

Table of Contents

1. LITERATURE REVIEW	1
1.1 RESEARCH METHODOLOGIES DEFINITION.....	1
1.2 RESEARCH STATEGY FOR MY RESEARCH.....	2
2. SECONDARY RESEARCH BACKGROUND OF AI	3
2.1 AI DEFINITION.....	3
2.2 TYPES OF AI.....	3
2.3 AGENTS	4
3. AI SPEECH RECOGNITION IN LEARNING ENGLISH.....	5
3.1 AI SPEECH RECOGNITION	5
3.2 EFFECTIVENESS OF SPEECH RECOGNITION IN LEARNING LANGUAGE	5
3.3 CURRENT CHALLENGES OF AI.....	5
3.4 GROWTH OF AI SPEECH RECOGNITION	6
3.5 HYPOTHESIS.....	9
4. PRIMARY RESEARCH PLAN	9
4.1 PURPOSE OF PRIMARY RESEARCH	9
4.2 QUANTITATIVE RESEARCH PLAN	9
4.2.1 <i>Purposes of the survey:</i>	9
4.2.2 <i>List of questions:</i>	9
4.3 QUALITATIVE RESEARCH PLAN.....	10
4.3.1 <i>Interview</i>	10
4.3.2 <i>Purpose of interview</i>	10
4.3.3 <i>List of question:</i>	11
5. PRIMARY RESEARCH ANALYSIS.....	11
5.1 SURVEY RESULT ANALYSIS	11
5.2 INTERVIEW RESULT ANALYSIS	11
APPENDIX	11
RESEARCH PROPOSAL	11
RESEARCH PLAN	12
PRIMARY RESEARCH RAW DATA.....	13
1. <i>Survey result</i>	13
2. <i>Interview result</i>	16
REFERENCE.....	17

1. Literature review

1.1 Research methodologies definition

Secondary research:

According to Do (n.d), Secondary Research is 'a common research method; it involves using information that others have gathered through primary research'. It is a kind of research in which researchers use already completed studies of other researchers for their own situations. Commonly, secondary research is implemented by looking for studies via the Internet, book, articles, etc. and summarizing them all together to support your research purpose.

There are a lot of advantages of this kind research methodologies. First, it is easy and cost-effective because all studies have already existed and usually, are available via Internet.

Moreover, secondary research helps clarify the research question, thereby supporting the implementation of primary research.

Primary research:

According to Driscoll (2011), primary research can be defined as a methodology where researchers will collect raw-data first-hand instead of gathering already studies via public resources such as Internet, book, etc. and summarize to analyze that data.

There are multiple techniques to perform primary research which helps researcher to gain accurate and diverse data from individuals such as interview, survey, etc.

Primary research is the most common research method used to answer provided questions of the research.

1.2 Research strategy for my research

Secondary research:

Recent years have witnessed many breakthrough in AI technologies, which means secondary research can be conducted with the support of various resource such as scholarly journals, digital journals, journals and books. By using some specific keywords such as digital english learning, machine learning, AI technology, together with Google Scholar, I can find a lot of useful information

Primary research:

The primary research will be conducted by 3 different method: Survey, Experiment and Interview.

Survey: Survey is conducted by handing out paper for people. All the data collected will be analyze and used for further research.

Interview: Interview 3 professional people in area being studied by this research.

Experiment: Use some test to check out the functionalities of AI speech recognition.

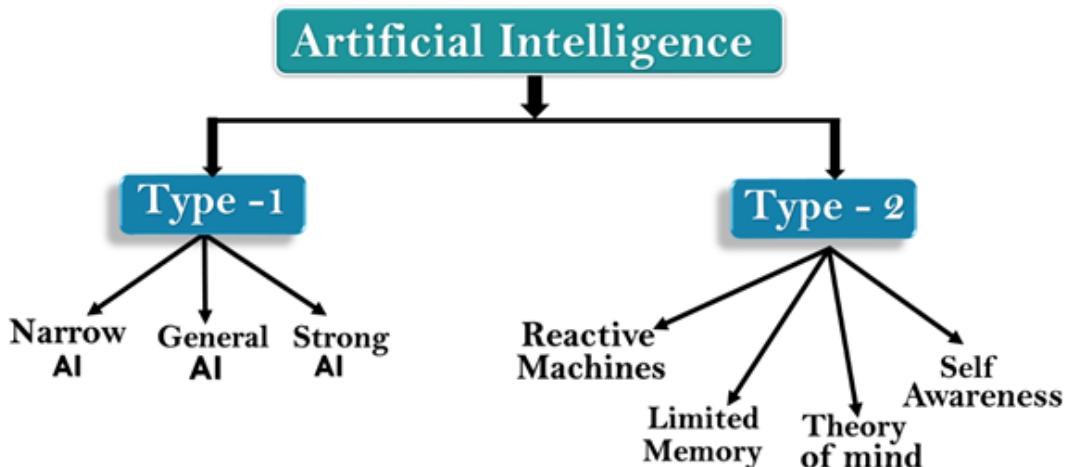
2. Secondary research background of AI

2.1 AI definition

AI (Artificial Intelligence) is used to describe a branch of computer science where we can create machines that have human based skills such as learning, reasoning, and problem-solving. With AI, we do not have to preprogram for every action that machine will do, instead, we can use algorithms which make that machine work with its own intelligence (AI Introduction, n.d.).

2.2 Types of AI

Artificial Intelligence is divided into a wide range of types, but there are two main categories of AI which depends on capabilities and functionalities. According to Types of Artificial Intelligence (n.d), these are types of AI agents



- Type I – Capabilities:
 - Narrow AI: It is currently the most common AI that only performs within its trained specific task. Some examples of Narrow AI that only work with pre-defined range of functions: Apple Siri, self-driving car, speech recognition, image recognition, playing chess, etc.
 - General AI: “General AI is a type of intelligence which could perform any intellectual task with efficiency like a human.”
 - Super AI: this type of AI can perform tasks beyond human intelligence. Main properties of this strong AI include think, solve problems, plan, judge and communicate by itself.
- Type 2 – Functionalities
 - Reactive machine: No stored memory and past experiences for future actions. Choose the best possible action for current situation. For example, IBM’s Deep Blue System and Google’s AlphaGo.
 - Limited Memory: Have memory but only store for a short time. One example is self-driving car

- Theory of mind: This technology has not been developed yet. It is meant to understand human emotions, and interact with human
- Self-awareness: This technology has not been developed yet. It is even smarter than human and has its own emotion and self-awareness.

2.3 Agents

An AI system is composed of an agent and its environment. The agents act in their environment. An agent perceives its environment through sensors and act upon on their environment through actuators

There are five types of AI agents:

- Simple Reflex Agent:
 - take decisions on the basis of the current percepts and ignore the rest of the percept history.
 - works on Condition-action rule, which means it maps the current state to action
 - Not adaptive to changes in the environment
- Model-based reflex agent:
 - These agents have the model, "which is knowledge of the world" and based on the model they perform actions
 - Updating the agent state requires information about: How the world evolve. How the agent's action affects the world.
- Goal-based agents:
 - Goal-based agents expand the capabilities of the model-based agent by having the "goal" information
 - These agents may have to consider a long sequence of possible actions before deciding whether the goal is achieved or not. Such considerations of different scenario are called searching and planning, which makes an agent proactive.
- Utility-based agent:
 - These agents are similar to the goal-based agent but provide an extra component of utility measurement which makes them different by providing a measure of success at a given state.
 - Utility-based agent act based not only goals but also the best way to achieve the goal
 - utility function maps each state to a real number to check how efficiently each action achieves the goals.

- Learning agent:
 - learn from its past experiences, or it has learning capabilities and adapt automatically through learning

3. AI Speech Recognition in Learning English

3.1 AI speech recognition

According to Velde (2019), speech recognition software can analyze the sounds you make by filtering what you say, digitizing it to a format it can “read”, and then analyzing it for meaning. Then, based on algorithms and previous input, it can make a highly accurate educated guess as to what you are saying. It gets to know the speaker’s use of language.

3.2 Effectiveness of speech recognition in learning language

Support teacher with correcting pronunciation errors: provides language learning potential by highlighting the worst word in an utterance and giving scores on pronunciation. These pronunciation problems are easily found by AI software rather than human teacher. Therefore, it helps tutors to find the necessary information to correct that student’s mistakes (Hincks, 2014)

Help students understanding lectures in various languages: In a study about Artificial Intelligence in The Classroom, Sam McNeil (2018) illustrates that any Presentation Translator is a free plug for PowerPoint to create real-time subtitles of what teacher is saying, displaying them below the presentation. This Translator is extremely useful for classes that have students coming from various nations. (Hincks, 2014)

Improve strong accent: AI software is beneficial for strongly accented students. (Hincks, 2014)

3.3 Current challenges of AI

Misrecognized words: Students must be alert for errors that go unrecognized by the program (words misunderstood by the software). This problem results from homophone phenomenon. Homophones are words that are pronounced the same but have different meaning and orthography. This is one of the biggest challenges for speech recognition because it needs large context to realize those words (Speech recognition for learning, 2010).

Speaker variability: speech recognition performance can be degraded when it is recognizing accented or non-native speech. Also, speakers have various ways of pronunciation and emphasis. These might be affected by social, emotional and personal factors. Therefore, it is such a challenging problem that speech recognition has to deal with (Speech recognition for learning, 2010).

Channel variability and noise: There are two main reasons behind this phenomenon External noises and microphone used to receive sound are two main factors that can affect the quality speech signal. This phenomenon, called channel variability, can be handled by applying filter or using high quality microphone (Speech recognition for learning, 2010)

3.4 Growth of AI speech recognition

According to Boyd (2018), The earliest advances in speech recognition focused mainly on the creation of vowel sounds, as the basis of a system that might also learn to interpret phonemes (the building blocks of speech) from nearby interlocutors. However, it was not until the 1950s that this line of inquiry would lead to genuine speech recognition. Up to this point, we see attempts at speech creation and recording, but not yet interpretation. Audrey, a machine created by Bell Labs, could understand the digits 0–9, with a 90% accuracy rate. IBM Tangora, released in the mid-1980s and named after Albert Tangora, then the world's fastest typist, could adjust to the speaker's voice. It was only in 1997 that the world's first "continuous speech recognizer" is capable of understanding 100 words per minute.

In recent years, with Machine Learning, Google has made remarkable improvements on the accuracy levels of previous speech recognition technologies (Boyd, 2018). There are multiple applications appearing on mobile phones such as Elsa, Duolingo, etc. to support learning languages. However, there are some people prefer to be taught by human tutor rather than AI speech recognition system (Devlin, 2016). Moreover, based on the download record on Appstore and Google store, these applications are not widely used.

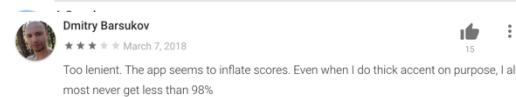
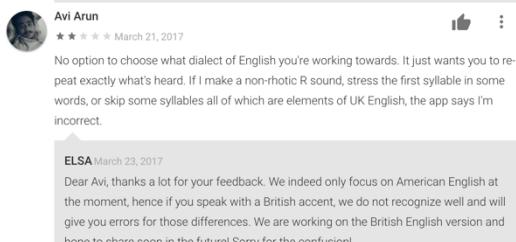
3.5 Review the most popular English learning applications using AI speech recognition

There are multiple applications currently in application store for mobile users like Duolingo, Babbel, Busuu, etc. but the most popular one that uses AI speech recognition is Elsa – English Language Speech Assistant.

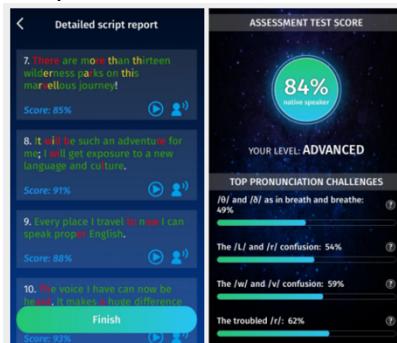
Unlike Duolingo which only use AI in increasing/decreasing level of difficulty in tests based on or Babbel that boost learner with the aid of pictures, Elsa uses AI speech recognition to adjust students' pronunciation. Elsa works by listening to users' voices, then it compares with the correct American English pronunciation. "The speech recognition technology out there is slightly different [from ELSA]. They try to guess what you're saying, whether you're saying it right or wrong. It's obviously very forgiving to mistakes. What we want to do is the exact opposite," Van (2018). Moreover, the app is very supportive when it can make suggestions as to how the speaker can improve on sound, for example, how to adjust their mouth shape or move their tongue. Therefore, as we can see while other applications only focus on teaching grammar and vocabulary, ELSA also uses AI to help with pronunciation. Furthermore, it has a dictionary, definitions, pictures and example sentences.

In Vietnam, language education is currently using textbook and non-native English teacher to educate students. However, this poses some issues. Firstly, most of English teachers at primary and secondary levels are usually disqualified because they had little chance to study in a native speaking country (Hoang, 2018). This results in high probability in pronunciation mistakes and low influence in English communication. Second, current curriculum of all levels use sets textbooks are mainly grammar-based (Hoang, 2018). Third, many students find it hard to study new vocabulary by traditional methods (learning by heart and writing many times) (Hoang, 2018). Therefore, Elsa's appearance promises to solve above problems.

However, as can be seen in Google Play, one out three users who downloaded this application are not fully satisfied with this application (1/3 rate less than 5 stars).

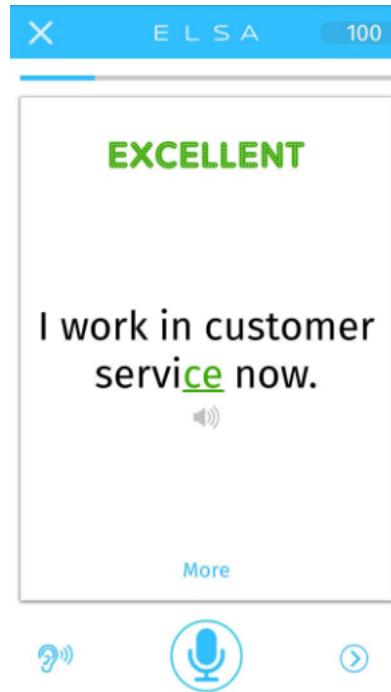


Some users figure some defects of this application. According to Eltplanning (n.d), Elsa is lack of focus on connected speech:



'Looking at the detailed script, it looks like some of the parts I got 'wrong' (like in number 8) were probably just contractions. I think the next stage for the developers is to work out how the app can recognize and accommodate features of connected speech. Maybe I'm just bitter that I'm only 91% accurate...'

Moreover, the application only focuses on the target sounds, which means the apps still returns **Correct!** if learners pronounce the targeted sound correctly but others isolated sound wrongly.



3.5 Hypothesis

Although AI speech is proved to be very useful and helps language learners to improve their problems but there is a small number of people who use it. My hypothesis is that AI speech recognition is not popular because users find it not enough to improve their skills and hard to use.

4. Primary research plan

4.1 Purpose of primary research

Although AI speech recognition has been applied to educations for many years and proved to be very useful in learning languages, it is showed that its application is not really popular. This survey aims to figure out how helpful the AI speech recognition actually is in education and why it is not a common choice of people for learning.

4.2 Quantitative research plan

I will ask about 30 users in the school (15 Female, 15 Male), 20-25 years old, almost all can speak English. They will answer some questions to figure out following questions:

- How useful Elsa application is in helping student learning
- What features do they like/dislike about Elsa?
- Would they choose Ai speech recognition rather other methods of learning?

4.2.1 Purposes of the survey:

This survey is meant to show the percentage of users frequently use AI speech recognition to learn, how helpful people about AI speech recognition when studying English, what they like/hate about it. As a result, we can figure out the reason behind the frequency of using AI speech recognition

4.2.2 List of questions:

1. How often do use Elsa?
2. How satisfied you are when you use Elsa to learn English?
3. Is learning with Elsa easier than with book or real tutor?
4. Does Elsa help to fix pronunciation problems?
5. Do you memorize new vocabulary faster by Elsa?
6. Elsa is hard to use?
7. Which is the thing you like about Elsa most?
8. Which is the thing you hate about Elsa most?
9. Do you think that Elsa can replace old teaching method (class, book, etc.)?
10. Do you find Elsa lack of interaction?

4.3 Qualitative research plan

4.3.1 Interview

These three people have all experienced Elsa and tried applying it when they teach their students. There are three sets of questions of each interviewee.

Interviewee	Time	Address	Occupation	Gender	Age	Objective
Nguyễn Hà Linh	9AM 8/7/2019	The coffee House	CEO of Zim academy	Female	37	To find out whether AI speech recognition is frequently used to support students to learn IELTS
Nguyen Thuy Duong	10AM 9/7/2019	The coffee House	Headmaster of An Duong Secondary Schoo	Male	40	Find out whether AI speech recognition is applied at school for students and its helps them to improve their English
Le Thanh Trung	10AM 11/7/2019	The coffee House	10 years of teaching literature and English	Male	35	To find out whether AI speech can help increase student reading and writing skills

4.3.2 Purpose of interview

This interview aims to answer following questions: how AI speech recognition is applied at school and does it really helpful? What is the cause of its unpopular/popular now? What functions that people expect AI recognition to have in the future?

4.3.3 List of question:

This interview aims to answer following questions: how AI speech recognition is applied at school and does it really helpful? What is the cause of its unpopular/popular now? What functions that people expect AI recognition to have in the future?

5. Primary research analysis

All the raw data is included in appendix

5.1 Survey result analysis

5.2 Interview result analysis

As mentioned in secondary research, traditional methods of teaching English in Vietnam results in many problems in speaking and learning new vocabulary so appearance of AI speech recognition should have been used commonly by Vietnamese and applied by many schools and English center. However, based on the interview results, there are still many issues existing in this application.

First, this application only works for intermediate learners who have pretty good vocabulary. Elsa can help them to fix some minor mistakes in their pronunciation. In contrast, Elsa is not beneficial for beginner because its AI speech recognition system only is not enough to push learner's memorizing words and practice many times. It should only be used to support teachers during their classes, not fully replace them because English after all requires interaction between human and human.

Second, some students love using this application, other may not. With students who are active and tend to communicate, they find learning with an AI machine boring. In contrast, others who are pretty reserved, love self-studying, and scare of judgements, they feel fine with learning Elsa. Therefore, we can see that Elsa must be developed more features to satisfy wider range of users.

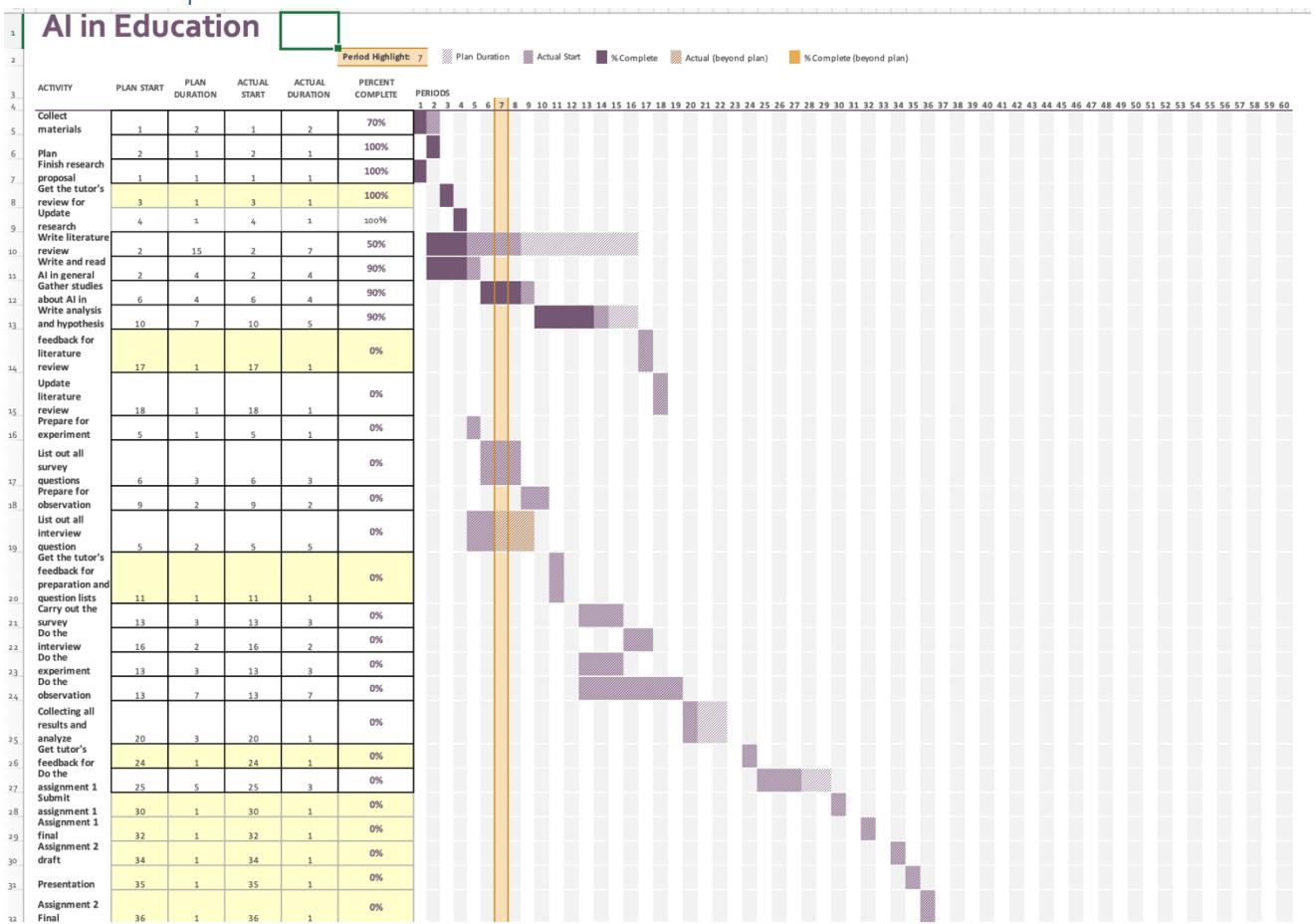
Third, there are still many errors of this systems when it comes to recognizing connected speech of full sentences, only focus on targeted sounds, etc. Therefore, when students think they are pronouncing correctly but Elsa still recognize it as false, students easily lose trust at this application and stops using it.

Fourth, Elsa cannot replace traditional teaching method for several reasons. AI is still a machine, not human, therefore, learning depends a lot on the awareness of students. Unlike real teacher who warn his students when they surf the Internet instead of studying, Elsa cannot manage what student are doing and is easily be turned off. Moreover, examinations in Vietnam still focus on grammar and writing, so Elsa plays insignificant role in supporting students in Vietnam.

APPENDIX

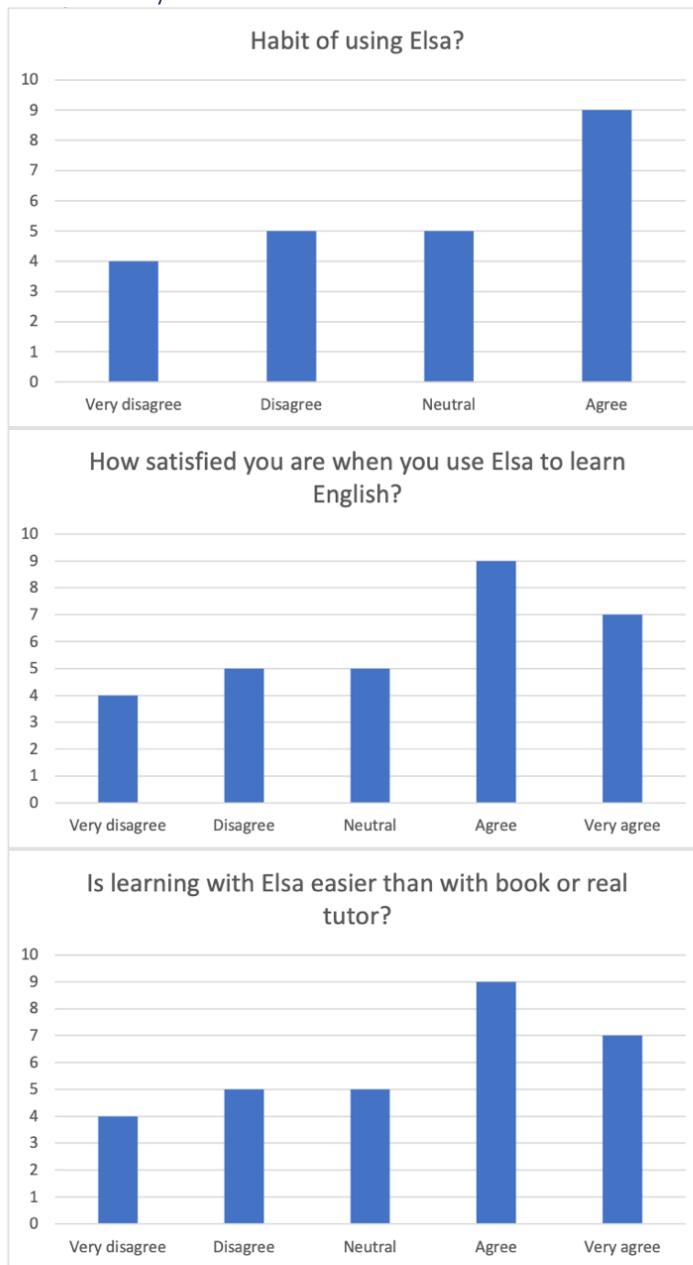
Research proposal

Research plan

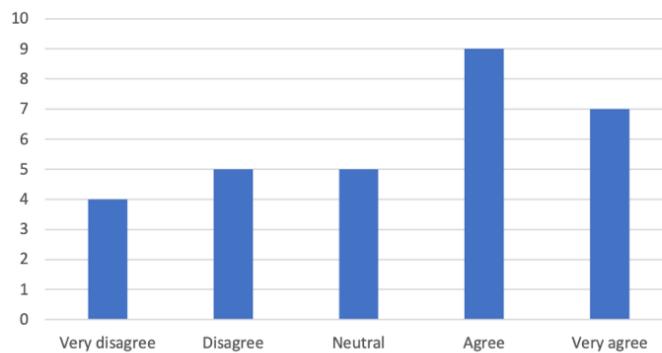


Primary research raw data

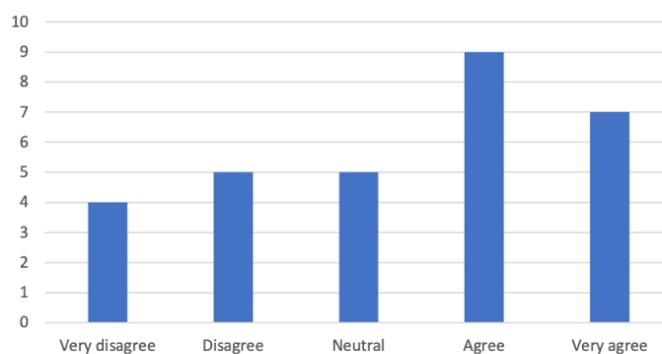
1. Survey result



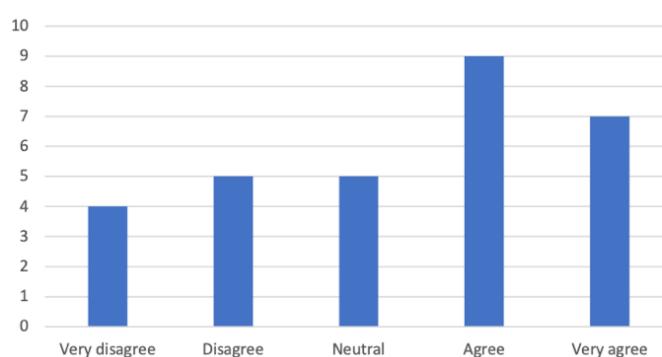
Does Elsa help to fix pronunciation problems?



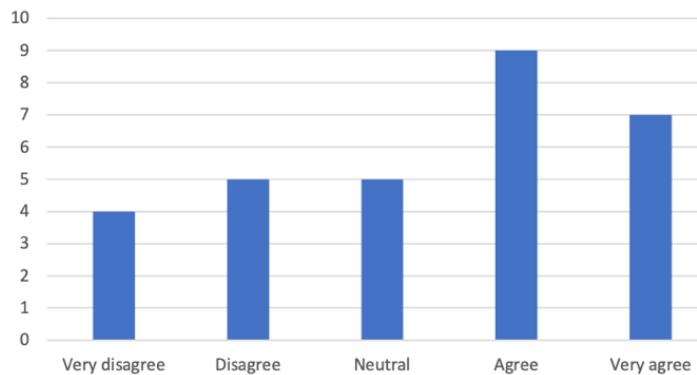
Do you memorize new vocabulary faster by Elsa?



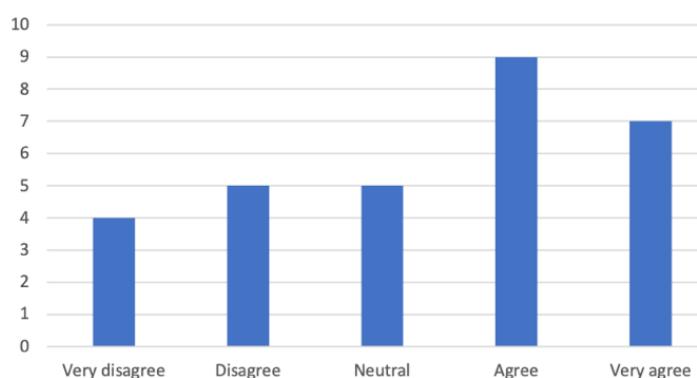
Elsa is hard to use?



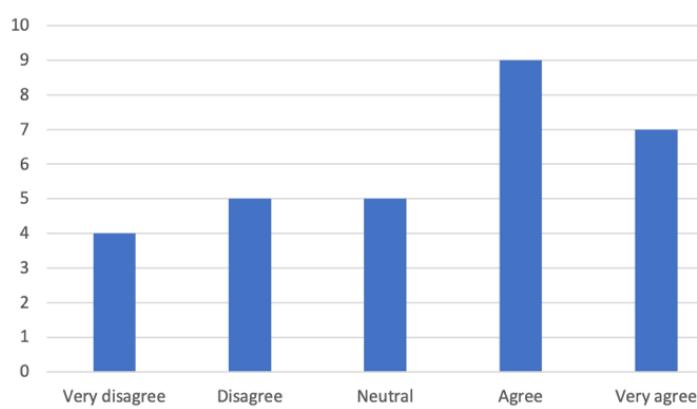
Which is the thing you like about Elsa most?

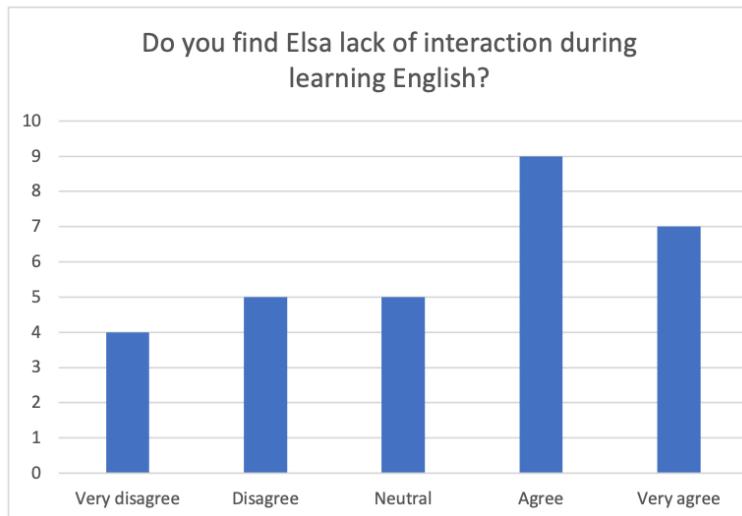


Which is the thing you like about Elsa most?



Do you think that Elsa can replace old teaching method (class, book, etc.)?





2. Interview result

2.1 For the first interviewee

1. How did you apply Elsa in teaching student at ZIM academy?

In speaking class, students are required to speak with teacher to check pronunciation. Teacher would note all mispronounced words and required students to learn that word in Elsa

2. How do students' scores change after using Elsa?

In general, there is not much change in pronunciation of student who are really bad at spelling and lack of vocabulary. However, with ones who have strong accents but are good at vocabulary, their scores improve a lot.

3. What do you think about using Elsa other than real tutor when studying English?

It is really helpful with speaking and learning new words. However, it is not enough for self-studying

2.2 For second interview:

1. How often does your school apply Elsa to teaching English?

We require student to learn Elsa at home and will check their studying in the class.

2. Do you find any improvement of students after using Elsa?

It is only beneficial for students who already have good background English and only need to fix some mistakes in pronunciation. However, with people who lack of English vocabulary, Elsa doesn't make any improvement, and they still need real human support to learn and practice new words.

3. What do you think about strength and weakness of Elsa?

Elsa is good for young students because they find Elsa interesting. Some students love learning with Elsa more than with teacher because they are scared of human teacher judges. However, with student who are active and want to talk a lot, they find Elsa boring.

4. Would you replace Elsa with a native teacher to teach student pronunciation?

It can only support teacher and it is not enough to fully fix student pronunciation because speaking requires conversation between human and human

2.3 For third interview:

1. What do you think about the impacts of Elsa on the writing and reading skills of students?

They barely have significant effect on student reading and writing skills but it helps to improve students with listening and speaking skills.

2. If AI speech recognition is developed, what do you expect it to have?

I want it to be easier to use. Moreover, it can talk to students like a real teacher and give more feedbacks and solutions to students' problems.

Reference

- Anon., 2010. *Speech recognition for learning*. [Online]
Available at: <https://www.readingrockets.org/article/speech-recognition-learning>
[Accessed 7 2019].
- Anon., 2019. *AI Introduction*. [Online]
Available at: <https://www.javatpoint.com/introduction-to-artificial-intelligence>
[Accessed 7 2019].
- Boyd, C., 2018. *The Past, Present, and Future of Speech Recognition Technology*. [Online]
Available at: <https://medium.com/swlh/the-past-present-and-future-of-speech-recognition-technology-cf13c179aaf>
[Accessed 7 2019].
- Devlin, H., 2016. *Could online tutors and artificial intelligence be the future of teaching?*.
[Online]
Available at: <https://www.theguardian.com/technology/2016/dec/26/could-online-tutors-and-artificial-intelligence-be-the-future-of-teaching>
[Accessed 7 2019].
- Do, L. T., n.d. *Design Research Techniques*. [Online]
Available at: <http://designresearchtechniques.com/casestudies/author/dee/>
[Accessed 4 8 2019].
- Driscoll, D. L., 2011. Introduction to Primary Research: Observations, Surveys, and Interviews. In: C. L. a. P. Zemliansky, ed. *Writing Spaces: Readings on Writing, Volume 2*. s.l.:s.n.
- Hincks, R., 2014. *Speech recognition for language teaching and evaluating*, s.l.: s.n.
- Velde, N. v. d., 2019. *How Does Speech Recognition Technology Work?*. [Online]
Available at: <https://www.globalme.net/blog/how-does-speech-recognition-technology-work>
[Accessed 7 2019].