Ms Leia - A Reading Tutoring Agent

Mayoor Rao
Georgia Institute of Technology
Atlanta, USA
mrao41@gatech.edu

ABSTRACT

Reading is a basic and one of the most essential skill to succeed in school. A child who can read before kindergarten would go to school with high confidence. Ability to read and comprehend also determines who well the students will succeed in other subjects. The schools have started adopting technology to impart reading skill in preschool and kindergarten. There are many popular mobile based applications which teaches reading. Giving mobile phones and tablet may not be a comfortable choice for every parent. The paper discusses an alternate approach of building reading applications which can be extended to teach other subjects as well. In this approach a smart speaker is paired with a television to facilitate tutoring. The student can engage in a conversational learning method with the intelligent tutoring system in this setup. The setup also provides a great control for parents as there is no way for the student to wander out of the teaching application. To the best of my knowledge such a teaching approach is unique.

Author Keywords

Intelligent Tutoring System (ITS), Smart Speaker, Interactive Television based tutoring

ACM Classification Keywords

Human Computer Interface -> *Natural language interfaces*, Artificial intelligence -> Natural language processing -> *Speech recognition*, Applied computing -> Education -> *Computer-assisted instruction; Interactive learning environments; Collaborative learning; Learning management systems*

INTRODUCTION

Teaching reading begins with reading books to the child. She is then introduced to print concepts such as reading from left to right, top to bottom, beginning of the book to the end, etc. Once familiar with print concepts, she can be taught letters, phonemes, sight words, decoding of words, reading short sentences and finally reading with intonation and expression. There are many mobile device based educational apps that teach literacy skills. Mobile devices provides a great interactive experience to the students. The mobile based applications provides a way to give a personalized learning experience using artificial intelligence. The student engagement can be increased by gamification and use of rich multimedia. The other popular form of teaching is using television. There are many educational content available either using streaming channel subscriptions or using DVDs. The challenge of a television based content is that it cannot be personalized and provides no interactive capability.

One of the main concern raised about a mobile devices is its effect on health. Since it is a new form of consumption method, there may not be enough studies to definitely conclude there may not be any health issues related to posture. The other major concern is the parental control. It is often very easy to step out of the app for the student and use applications or watch content that is not age appropriate. It is easier to implement parental control using television. Televisions provide a good viewing experience without having to struggle to maintain a good body posture.

Smart speakers are new form of interactive device that are widely available. Smart speakers allows user to engage in verbal conversations. Ms. Leia is an attempt to create an interactive intelligent tutoring system using smart speakers and television. The main object of such a tutoring system is to address concern of parents who might not be comfortable in giving mobile devices to children of young age fearing health and security issues.

RESEARCH

Mobile devices based tutoring applications are very popular. The school in my neighbourhood (Sunnyvale school district) makes use of tablet based apps in Kindergarten to enhance the reading skills. These devices enable to scale research backed education methods. Sight word training, where students are trained to recognize list of high frequency words and those which do not follow regular phonetic rules, helps improving the reading skills [5]. Sight words apps make use of (Incremental Rehearsal) IR and SIR (Strategic Incremental Rehearsal) methods to make the sight words training more effective [6,7,14]. The technology can provide adaptive tutoring, addressing the need for special attention to students who might need it [9]. An approach called blended learning was studied by JJen Elise Prescott, Kristine Bundschuh, Elizabeth R. Kazakoff & Paul Macaruso. In this approach the online learning was used as a complementary learning method to the classroom teaching. The study showed that there was great improvement in reading assessment. The effect was particularly high among students below grade 2 [10].

While technology is very effective, there are fears about its effect on young children. Educators fear that the computers can result in poor social skills resulting from less interaction with peers. They would spend less time in social problem solving which are very critical in the early years[11]. Children get more attached to the devices than their peers.

Light from Blue wavelengths can have effect on the sleep pattern which has other serious side effects [12]. Night mode is now a feature being adopted in all the mobile OS [15,16]. Study finds that the prolonged usage of the mobile devices causes issues with postures such as rounded shoulder, forward neck and can affect the breathing patterns[13]. Forward Head Posture (FPA) is one of the most common abnormalities which predisposes an individual to conditions such as neck pain, headache, etc. Studies have shown that prolonged usage of computer can cause FPA especially while working with smaller screens. This study also concludes that the mobile usage can cause FPA.

Many educational apps use concept of gamification to promote student engagement. Gamification is mostly implemented in the educational apps in form of virtual stars or points. The student gets motivated to earn more stars and hence stays with the app completing the exercises. Though gamification may seem to be a perfect solution to keep the student motivated, as discussed by Andrade, F. R., Isotani, S., & Mizoguchi, R., there is a bright side and a dark side to gamification [4]. Gamification can force competition among students and students who are lower in the score compared to their peers may feel low in confidence or just lose interest. Gamification can cause addiction among some students. It's effect can be that they could learn only if there are gamification features else they don't feel interested in the activity.

Benefits of the educational apps are widely seen and the schools are enthusiastic about adopting them. However, there are issues such as parent's anxiety about the possible effects of these methods, posture related concerns and psychological effects due to gamification which need to be addressed. The work discussed in the paper tries to address these issues by providing an alternate ecosystem to provide quality education which is equally effective as iPad or mobile based apps.

DESIGN

A high level overview of the component layout is shown in Figure 1.

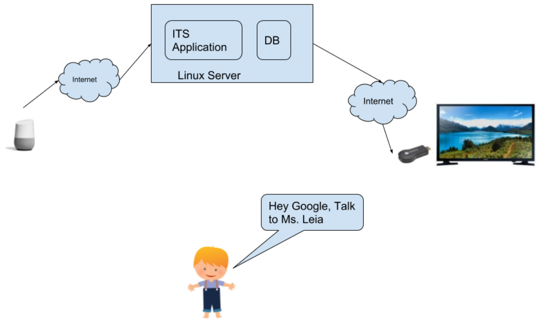


Figure 1: Component Layout

In the current implementation chromecast is used to enable internet on television and Google home as the Smart speaker. The choice of these two devices was based on what was already available. These devices can be replaced by any other device as long as they can achieve following objective -

Device	Expected Feature
Smart	● Ability to discover intent

Speaker	within the user utterance ● Ability to call a webservice at the end of the user utterance
Streaming Device	● Provide a way to bring up an HTML5 based page ● Support for Websockets

Table 1: Component Requirement

TECHNOLOGY

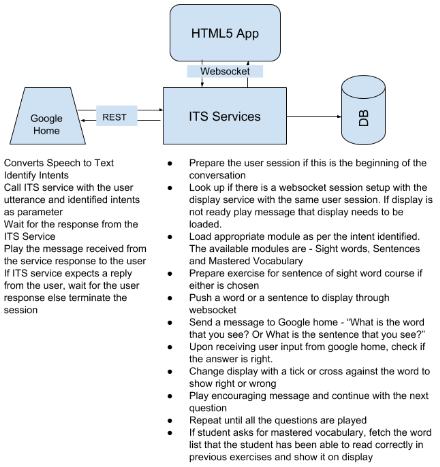
Smart Speaker: Google Home is a great voice based assistant which is very popular. The speech to text has a high degree of accuracy. *Dialogflow*[8] is a developer tool which helps to design experience around natural language based interfacing. It helps to train the google assistant with the expected user utterances and identifying the user intent from the utterances. Google home with *Dialogflow* makes speech to text and Natural language processing very easy to use in an application.

Streaming Device: Chromecast is a popular streaming device. It was first of its kind where a simple dongle could connect to the USB and HDMI port of the TV and make any TV internet enabled. It comes with chrome engine, which gives a good support for HTML5 based application. A web application can be launched on a chromecast enabled TV by using a cast option from a chrome browser running on a computer or mobile devices. Alternately the cast options are available through Android and iOS based application. The cast option lets you bring up the content on the TV.

ITS Services: Ms Leia's backend services are built using python based framework called Tornado. Tornado provides an event based web framework, which means it is non-blocking. The communication with the display page on the TV and the backend service is using Websocket. SockJS provides websocket plugin for Tornado.

Database: PostgreSQL is a popular open source relational database. Any relational database could have been used, Postgres was chosen as it is open source and comes with no cost.

COMPONENT SERVICE DETAILS



- Converts Speech to Text
- Identify Intents
- Call ITS service with the user utterance and identified intents as parameter
- Wait for the response from the ITS Service
- Play the message received from the service response to the user
- If ITS service expects a reply from the user, wait for the user response else terminate the session

- Prepare the user session if this is the beginning of the conversation
- Look up if there is a websocket session setup with the display service with the same user session. If display is not ready play message that display needs to be loaded.
- Load appropriate module as per the intent identified. The available modules are - Sight words, Sentences and Mastered Vocabulary
- Prepare exercise for sentence of sight word course if either is chosen
- Push a word or a sentence to display through websocket
- Send a message to Google home - "What is the word that you see? Or What is the sentence that you see?"
- Upon receiving user input from google home, check if the answer is right.
- Change display with a tick or cross against the word to show right or wrong
- Play encouraging message and continue with the next question
- Repeat until all the questions are played
- If student asks for mastered vocabulary, fetch the word list that the student has been able to read correctly in previous exercises and show it on display

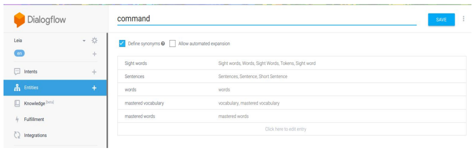
Figure 2: Component Service Details

Figure 2 lists the high level functional details of each component. The HTML5 app serves as a display only component. There is no user input expected on display component. The user interaction is only through Google Home. ITS application receives user utterance in a text format as converted by Google Home and publishes contextual content on display. The ITS service also generates contextual verbal responses to be played on Google Home. The types of verbal response include - greetings, asking for course type, asking questions, motivating students, encouraging students and providing clues.

GOOGLE HOME SERVICES

Dialogflow is a developer studio provided by Google to create and manage the Natural language based interface for an application. There are 4 main parts to setting up a project in Google Home -

1. Create a Project and associate an invocation phrase. Here the invocation phrase is set to *"Ok Google, Talk to Ms. Leia"*
2. Define Entities - Entities are a set of words that denote an action.



command

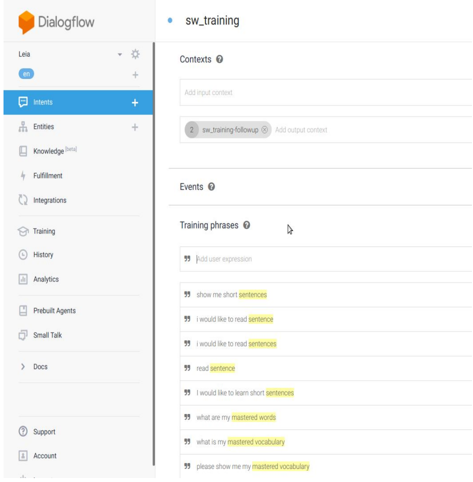
☑ Define synonyms □ Allow automated expansion

Sight words	Sight words, Words, Sight Words, Tokens, Sight word
Sentences	Sentences, Sentence, Short Sentence
words	words
mastered vocabulary	vocabulary, mastered vocabulary
mastered words	mastered words

Figure 3: Entity Setup

Figure 3 shows the entity configuration on **Dialogflow** studio. Here I have setup an entity called command and associated following words with the entity - Sight words, mastered vocabulary, sentences, mastered words

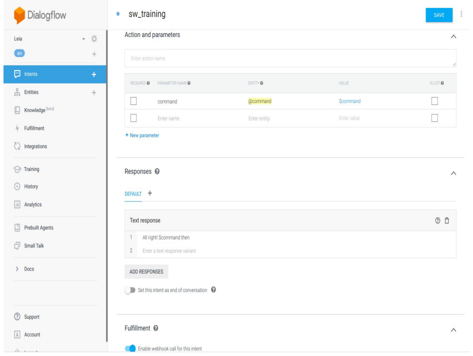
3. Define Intent - Intent is finding the entity and corresponding value from a user utterance. Eg. If the student says - *"I would like to learn Sight words"*, the entity is "command" and value is *"sight words"*. Note that, *"command"* is the only entity that is defined for this project. In order for the agent to correctly identify the intent, we need to train the agent with expected variants of the phrases from the user. The studio provides easy interface to provide training samples for training the agent.



show me short sentences
i would like to read sentence
i would like to read sentences
read sentence
I would like to learn short sentences
what are my mastered words
what is my mastered vocabulary
please show me my mastered vocabulary

Figure 4: Training Intent

Figure 4 shows the list of phrases that were trained to discover the intent.



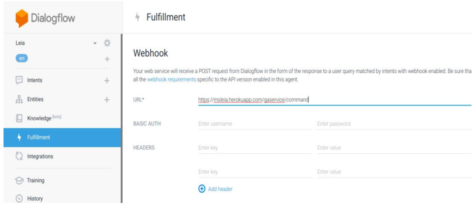
REQUIRED	PARAMETER NAME	ENTITY	VALUE	IS LIST
□	command	@command	$command	□
□	Enter name	Enter entity	Enter value	□

Text response
1 All right! $command then
2 Enter a text response variant

Figure 5: Intent to Entity Mapping

Figure 5 shows entity to intent mapping. Webhook integration is enabled for the intent so that the agent makes the rest call to ITS backend service with the identified entities and the values.

4. Once the agent identifies the type of course that is chosen by the student, we need a follow up intent, which will pick any user utterance and just send it to the ITS backend service. This is typically response to the question posed by the ITS, such as - *What is the word that you see?* Or *What is the sentence that you see?*
5. In order to receive the parsed output from Google Home, we need to configure the REST endpoint that needs to be invoked.



Your web service will receive a POST request from Dialogflow in the form of the response to a user query matched by intents with webhook enabled. Be sure tha all the webhook requirements specific to the API version enabled in this agent.

URL* https://msleia.herokuapp.com/gaservice/command

Figure 6: Webhook Configuration

Figure 6 shows the webhook configuration.

ITS SERVICES

ITS Services is a REST based application. It contains 4 main modules -

Google Assistant Handler: This module manages all the interactions with the Google home assistant. The handler receives payload from Google home, which contains the user utterance. The response is parsed and then passed to the Teacher module

Teacher Service: The teacher service loads the appropriate course module as per the student request. Once the course is initialized, it draws a question from the course and sends it to the HTML5 App that is loaded on television. It creates an appropriate verbal message that is sent to Google Home. Eg. In case of Sight words exercise, it will load sight words course, and then draw a word from the course exercise and send it to the display module to display on television and then plays a message - "What is the word that you see?" on Google Home. Upon receiving a response from the student, it compares with the answer. If right, it draws next question. If wrong, it repeats the question with the message - *"Can you try that again?"*. A message is sent to the display service to show a cross or check mark based on the correctness of the answer. If a student makes two mistakes, teacher service picks a clue associated with the word and plays that to the student. Eg. If the student reads “can” wrong two times, a clue (which is configurable through Admin API) such as *“It starts with Ka sound”* is played. If a student reads the word wrong thrice, then it moves on to the next word. Every response is recorded so that we can draw insight about the mastered vocabulary. When there are no more questions left, the teacher checks if there is any reward that is configured. A reward can be configured through Admin service, which is a message that will be played after completion of n number of exercises as configured. An sample reward message could be - *“Mommy is so proud of you! She has a surprise for you. Go and collect it!”*

Courses: All courses are implementation of Course class. There there two courses that are implemented in the current release - Sight words and Short sentences. Sight words course generates a flash card from the list of words that are setup through the Admin API. Flash card generation follows Incremental Rehearsal (IR) approach. The IR approach goes as follows -

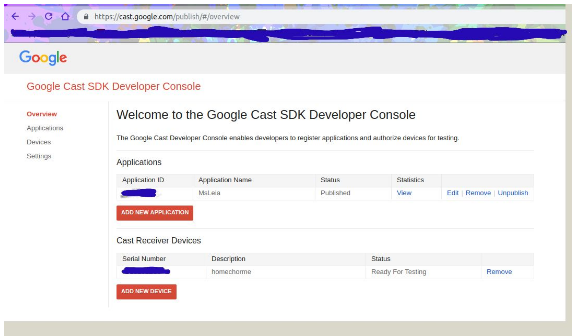
1. Create a deck of n words with 1 unknown word and n-1 known words. I have set n as 6 in current release.
2. The service reads out the first word so that the student learns the word and asks her to repeat it.
3. Once the student repeats the word, a known word is shown.
4. Show the sight words in following order -
    a. Unknown word followed by 2 known words
    b. Unknown word followed by 3 known words
    c. …..
    d. Unknown word followed by n-1 known words
5. When all the words in the deck are shown, the exercise is complete

Short Sentences exercise is generated based on mastered vocabulary. A sentence repo is created using Admin APIs. When sentence course is loaded, it generates an exercise by picking sentences that contain all the words mastered by the student. In the current release, a sentence exercise contains maximum of 6 sentences.

Admin: Admin module provides APIs for setting up sight words, sentences, rewards and encouraging sentences. The list of sight words are sourced from Dolch sight word list [3].

CHROMECAST AND DISPLAY MODULE

Display module is an HTML5 based application which is hosted separately from the python services. In the current release, the application is hosted through github by creating msleia.github.io repository. Github provides free hosting until certain limit and also gives HTTPS URL. Chromecast makes any TV with HDMI port, internet enabled. Once the chromecast is on, any device running a chrome browser can launch content on chromecast that is available in the same WiFi network as the device. Chromecast SDK provides a way to create a receiver module and a sender module. The receiver module is hosted on TV and uses Chrome Javascript based APIs to receive instruction from sender. A sender module when launched on chrome browser, shows a button called “cast”, which can be clicked to launch the app on TV. A developer needs to create a chrome developer account [2]. Once an account is created, we need to first register a test chromecast device and then register an application associated with the registered device.



https://cast.google.com/publish/#/overview

[redacted]

Welcome to the Google Cast SDK Developer Console

The Google Cast Developer Console enables developers to register applications and authorize devices for testing.

Applications

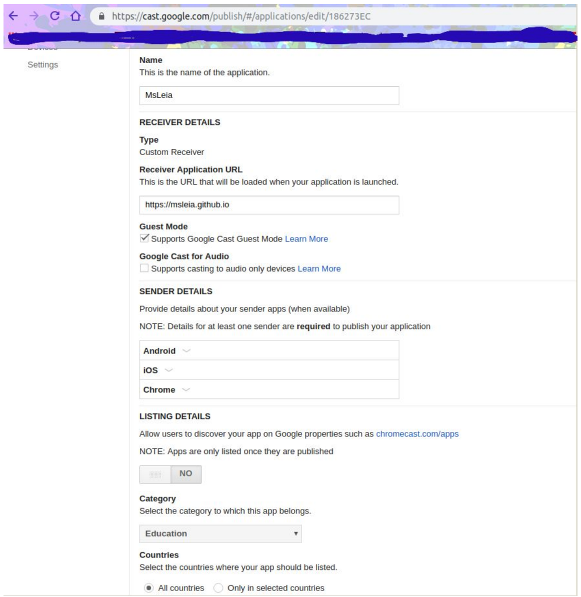
Application ID	Application Name	Status	Statistics	
[redacted]	MsLeia	Published	View	Edit | Remove | Unpublish

Cast Receiver Devices

Serial Number	Description	Status	
[redacted]	homechorme	Ready For Testing	Remove

Figure 7: Device and Application registration on Chromecast developer site.

The screengrab above shows registered application called Ms Leia and a registered chromecast device. Once the registration is complete we can publish an application that is to be launched. It is mandatory for the application endpoint to be HTTPS. The link provided corresponds to the receiver module for the chromecast.



https://cast.google.com/publish/#/applications/edit/186273EC

[redacted]

Settings

Name
This is the name of the application.

MsLeia

RECEIVER DETAILS

Type
Custom Receiver

Receiver Application URL
This is the URL that will be loaded when your application is launched.

https://msleia.github.io

Guest Mode
☑ Supports Google Cast Guest Mode Learn More

Google Cast for Audio
☐ Supports casting to audio only devices Learn More

SENDER DETAILS

Provide details about your sender apps (when available)

NOTE: Details for at least one sender are **required** to publish your application

Android
iOS
Chrome

LISTING DETAILS

Allow users to discover your app on Google properties such as chromecast.com/apps

NOTE: Apps are only listed once they are published

NO

Category
Select the category to which this app belongs.

Education ▾

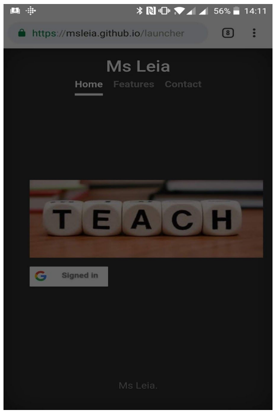
Countries
Select the countries where your app should be listed.

● All countries ○ Only in selected countries

Figure 8: Send and Receiver configuration for chromecast application

Figure 8 shows the configuration of the receiver module. The display service is a single page application written using React framework. As soon as the receiver module is launched it opens a websocket with ITS service. A callback is registered with the websocket which is called every time there is a message on websocket. The message received from the server contains the text that needs to be displayed and if the response type is right, wrong or none. If it is right, then a tick symbol is displayed along with the text, if the type is wrong, then a cross is displayed and if none, only the text appears.

SCREENFLOW



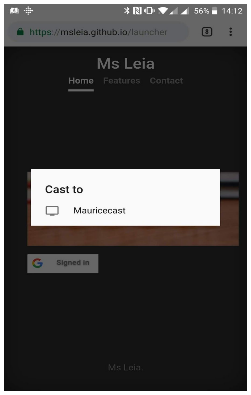


Figure 9.a: Home Screen **Figure 9.b: Cast Button**

Figure 9.a and 9.b shows how to launch the application on TV. Once the application is loaded and Teach button is clicked, the chromecast sender module displays a cast button with the list of chromecast in the local WiFi. Once the cast device is chosen the application is launched on television.

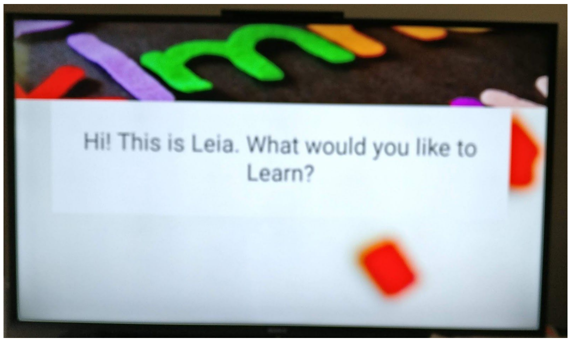


Figure 10: Application loaded on TV

Figure 10 shows Ms Leia launched on television after cast button was clicked. Once the student says - *"Ok Google, Talk to Ms Leia"*, the student hears a greeting - *"Hello, what can I do for you"*. The user can say *"I would like to learn sight words"* or *"I would like to see mastered vocabulary"* or *"I would like to read short sentences"*.

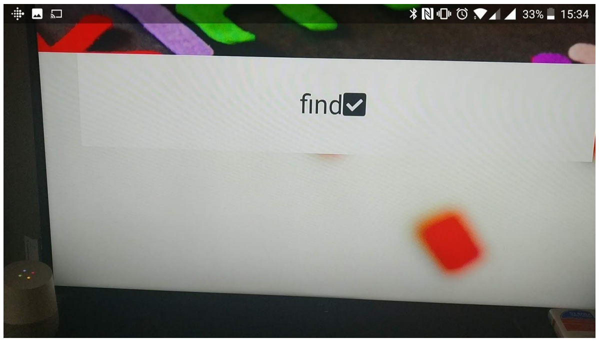


Figure 11: Sight word exercise with feedback

Figure 11 shows sight word and a tick symbol when a student reads the word correctly.

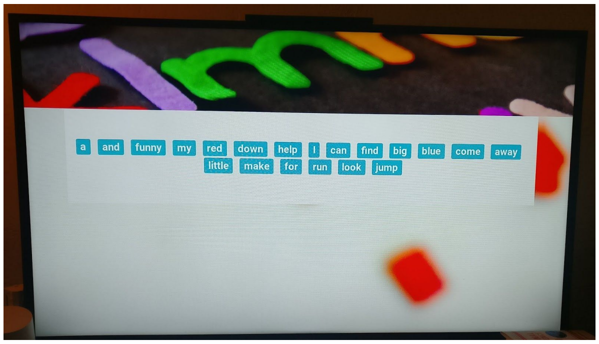
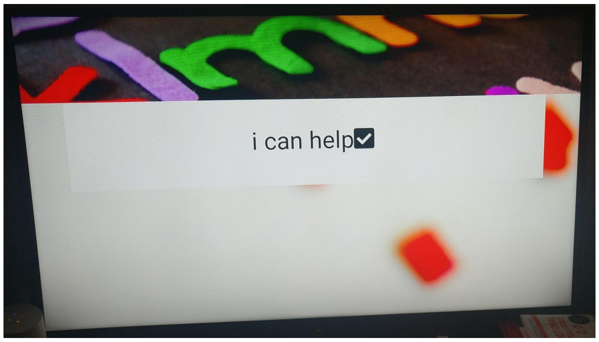


Figure 11: List of mastered vocabulary

The above screen shows the list of mastered vocabulary when the student prompts for mastered vocabulary.



i can help☑

Figure 10: Sentence exercise

The above screen shows a short sentence when the student engages with short sentence based exercise. It shows a tick symbol or a cross depending on whether the student read the sentence correctly or incorrectly.

CONCLUSION AND FUTURE WORK

The chief purpose of this tutoring agent is to demonstrate a teaching method which is interactive, uses television as a medium to display and is conversational. One of the goals was to cater to the needs of parent such as myself who is not comfortable in giving mobile devices to the children who are too young. A television does not require the user to sit close to it as it has a large display. Hence, it helps the students maintain their posture through the learning session. The rewards managed by parent instead of all the virtual stars and scores, helps parents to incentivize learning in a manner that best suits their child. The rewards provided by the parents are tangible and gives parents a sense of participation in the learning exercise. The project in its current form meets all these objectives.

Future work and enhancements:

- ➢ This tutoring method can be applied to teach other skills such as math, analytical and logical thinking.
- ➢ Currently the agent supports only reading short sentences and sight words. This can be enhanced to add exercises which involve reading sentences mixed with pictures.

SOURCE CODE

All source code is available under Apache 2.0 licence.

Component	Source Repo
Dialogflow project	https://github.com/msleia/googlehome_agent
ITS Services	https://github.com/msleia/itsserver
Display App	https://github.com/msleia/msleia.github.io https://github.com/msleia/msleia-cast

Table 2: Source Code Inventory

Backend services are hosted on heroku[17]. The html5 is hosted on github using github hosting feature.

ACKNOWLEDGEMENTS

I would like to thank my mentor Tyler Bobbik whose feedback and encouragement made this project a success. I would like to thank Dr. Joyner for EduTech course which gave me an opportunity to explore this field and make contribution. I would like to thank my wife and kids for their continued support.

REFERENCES

1. Sarah Halzack (2015, November 24) The rise of the really-big-screen TV Retrieved from https://www.washingtonpost.com/news/business/wp/2015/11/24/the-rise-of-the-really-big-screen-tv/?utm_term=.c6da02c7b9b2
2. Google. Google Cast SDK Developer Console. Retrieved from https://cast.google.com/publish/#/overview
3. Sightwords.com. Dolch Sight Words. Retrieved from https://sightwords.com/pdfs/word_lists/dolch_kindergarten.pdf
4. Andrade, F. R., Isotani, S., & Mizoguchi, R. (2016). The bright and dark sides of gamification. Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9684, 176-186.
5. Hayes, Colleen, "The Effects of Sight Word Instruction on Students' Reading Abilities" (2016). Education Masters. Paper 327
6. January, SA.A., Lovelace, M.E., Foster, T.E. et al. J Behav Educ (2017) 26: 151. https://doi.org/10.1007/s10864-016-9263-2

7. Kupzyk, Sara, Daly, Edward J., III, & Andersen, Melissa N. (2011). A Comparison of Two Flash-Card Methods for Improving Sight-Word Reading. Journal of Applied Behavior Analysis, 44(4), 781-792.
8. Google. Dialogflow. Retrieved from https://dialogflow.com/
9. Hutinger, P., Bell, C., Daytner, G., & Johanson, J. (2006). Establishing and Maintaining an Early Childhood Emergent Literacy Technology Curriculum. Journal of Special Education Technology, 21(4), 39-54.
10. JJen Elise Prescott, Kristine Bundschuh, Elizabeth R. Kazakoff & Paul Macaruso (2018) Elementary school–wide implementation of a blended learning program for reading intervention, The Journal of Educational Research, 111:4, 497-506, DOI: 10.1080/00220671.2017.1302914
11. VanderScoter J, Ellis D, & Railsback J (2001) Technology in early childhood education: Finding the balance. Portland, OR: Northwest Regional Educational Laboratory. Wood E, Specht J
12. Harvard Medical School. (August 2018) Blue light has a dark side. Retrieved from https://www.health.harvard.edu/staying-healthy/blue-light-has-a-dark-side
13. Jung, Sang In et al. “The Effect of Smartphone Usage Time on Posture and Respiratory Function.” Journal of Physical Therapy Science 28.1 (2016): 186–189. PMC. Web. 15 Sept. 2018.
14. Elizabeth Mulvahill (2018, April 25). What Are Sight Words? Retrieved from https://www.weareteachers.com/what-are-sight-words/
15. Apple (2018) Retreived from https://www.apple.com/newsroom/2018/06/ios-12-introduces-new-features-to-reduce-interruptions-and-manage-screen-time/
16. Google (2018) Retrieved from https://support.google.com/families/answer/7103340?hl=en
17. Hekroku. Documentation. Retrieved from: https://devcenter.heroku.com/categories/reference