Understand My World: An Interactive App for Children Learning Arabic Vocabulary

Zeyad Ali
Dept. of Computer Science & Eng.
College of Engineering
Qatar University
Doha, Qatar
zali@qu.edu.qa

Moutaz Saleh
Dept. of Computer Science & Eng.
College of Engineering
Qatar University
Doha, Qatar
moutaz.saleh@qu.edu.qa

Batoul Khalifa
Department of Psychology
College of Education
Qatar University
Doha, Qatar
zali@qu.edu.qa

Jihad M. Al‘a’am
Dept. of Computer Science & Eng.
College of Engineering
Qatar University
Doha, Qatar
jaam@qu.edu.qa

Dominic Massaro
Dept. of Psychology
Social Science II
University of California
California, USA
massaro@ucsc.edu

Samir Abou Elsaoud
Department of Computer Science
College of Science
British University in Cairo
Cairo, Egypt
Samir.Elsoud@bue.edu.eg

Abstract—Covid19 imposed a new paradigm in education especially in elementary schools. Children are no longer able to go regularly to schools as normal. They should then rely mainly on themselves to learn and acquire knowledge. Understand my world is a new technological application that allows children to learn the Arabic language interactively and independently using their smart devices (i.e., ipad, iphone). They can use their device’s camera and microphone to explore the world, understand spoken words, and read written language properly. Images captured by the camera are labeled and defined in both speech and writing. Children can also speak into their device to record their speech or ask questions. The app then presents the dictated words or answers to questions in both spoken and written form. Using the required Internet connection, the interface automatically provides exceedingly accurate image and speech recognition, all while preserving the participants’ anonymity. The application is designed to be intuitive and seamless to use, which makes it very attractive to children to learn Arabic vocabulary through multimedia.

Keywords—Arabic language, educational app, independent education, learning.

I. INTRODUCTION

From birth and even prenatally, children are immersed in spoken language but are essentially shielded from written language. Although there is evidence for analogous processes in understanding spoken and written language, written language is believed to require formal instruction and schooling whereas spoken language is seamlessly acquired from birth onward by natural interactions with persons who talk. Covid19 changed the way of teaching and disseminate knowledge. In fact, schoolchildren are constraints to learn distantly through the Internet. They cannot go regularly to schools and learn in a face-to-face manner. The fundamental principle now emerging is that there are analogous processes at work in speech perception and reading [7], [12], [14] (Massaro, 1998; Massaro & Jesse, 2005; Movellan & McClelland, 2001). There is now a solid evidence that many aspects of language processing (a) involve a form of pattern recognition, (b) are influenced by multiple sources of information, and (c) have the capacity to be quantitatively described by the Fuzzy Logical Model of Perception [7] (Massaro, 1998).

This same processing appears to occur in language acquisition, not just in accomplished language users [3], [4] (Fennell & Waxman, 2010; Hirsh-Pasek, Golinkoff & Hollich, 2010). Our objective in this research work is aimed at whether the learning difference between spoken and written language is necessary. It is possible that reading might also be acquired without instruction if sufficient immersion in written language is implemented. Our goal is to create an interactive app, Technology Assisted Reading Acquisition (TARA), to allow children to acquire literacy naturally and independently as their presence in schools is very limited. Language acquisition is well-described by sensory adaptation, pattern recognition based on multiple sources of information, and statistical learning [21] (Werker et al., 2012).

Previously, we sought to create an environment that would seamlessly provide written language. The goal was to build a smart room for the acquisition of literacy. We obtained a US Patent, Method And System For Acquisition Of Literacy [8], [10] (Massaro, 2012a, 2015a), that set the stage for this initial work. The goal was to create a smart room that monitors child/caregiver normal interactions to embed the child in written language. The invention exploited current knowledge and developments in behavioral science and technology to provide methods for automatically recognizing the meaning inherent in the child’s experience, and describing that experience in a written language form that is appropriate to the perceptual, cognitive, and linguistic capabilities of the child.

The logic of this learning pedagogy is intended to extract the relevant meanings embedded in a child’s experience and to describe some of that experience in written language. It uses two methods to recognize and convert the child’s experience of spoken language into written language. The first method is to recognize the speech that is being spoken during the experience. The second method is to recognize the scene, the objects, and the actions it contains. The outcomes of the speech and scene recognition are then presented in writing to children who are learning to read.

It became apparent that a smart room would require a great expense because it would require a dedicated space, cameras, microphones, and large display monitors. This type of implementation would be beyond the practical and financial situation of most families and they cannot afford the heavy
cost. In addition, government set strict measures on schools to control the health of children and instructors. Luckily, however, with the recent miniaturization of technology, it is possible to achieve the same learning environment by simply giving the child nothing more than a mobile phone or a tablet with the necessary software. Thus, children would always have the adequate technology to carry with them and learn independently at any time and from any location. Furthermore, the child can acquire literacy without the constant input of teachers, which might be particularly valuable during social distancing under COVID19 pandemic.

To this end, we have implemented a literacy app in a child-friendly application that can be employed on a low-cost smart device with an Internet connection. The proposed app, Understand My World (2020) [1], automatically labels and defines objects and scenes, records and transcribes a child’s speech, and even answers questions in a large variety of domains, while simultaneously giving written language experience.

The hypothesis to be tested is whether this experience with spoken and written language will allow the child to acquire literacy and the mechanics of reading seamlessly and without direct instruction like in traditional education at schools. The app can serve different categories of learners including those with Arabic learning difficulties (LD) and reduce their isolation from education. In fact, the population of children with LD is noticeably increasing across the globe. It is possible that these children can learn successfully with different methods than used with their normal peers. Thus, it is highly important to understand the difficulties that these children may face and develop a learning curriculum to accomplish remedial learning and put them on a productive path of new learning. Multimedia technology can improve the learning of the language for children with LD as we have shown in our previous work. In fact, multimedia can demonstrate the new concepts using multiple appropriate modalities such as visual, haptics, auditory and gesture. It keeps the children engaged for a longer time and takes into consideration their different levels of difficulties to learn new concepts.

The app provides an engaging multimodal learning environment that immerses the child in cognitive challenges and linguistic support. There is impressive evidence that early reading has a positive outcome throughout the lifespan. Learning to read earlier appears to have long-term benefits. In an early study, [2] Durkin (1966) found that the earlier the child was taught to read, the better she/he could read six years later (see also [19] Titzer, 2012). Ritchie & Bates [17] (2013) carried out a long-term analysis of reading skill at age 7 and Social Economic Status (SES) 35 years later on a large, population-representative sample from the United Kingdom. They found that reading ability at age 7 was positively associated with SES 35 years later and accounted for a complete standard deviation of SES status 35 years later.

Children who have the opportunity to learn to read early in life have an advantage in reading to learn more complex language. We have analyzed [9], [11] (Massaro 2012b, 2015b) two different spoken language databases and an assembled picture book corpus for their vocabulary and their linguistic and cognitive complexity. The results indicated that print was more complex than speech. The likelihood of observing a word not contained in the most common 5,000 words in English was 3 times as great in picture books than in child-directed speech. Five measures of reading grade level gave an average grade level of 4.2 for picture books and 1.9 for child-directed speech. These results indicate that written language has the potential for valuable language experience and the earlier children learn to read the more they can benefit from this source of knowledge.

Although the emphasis here is on the benefits and possibilities to be derived from immersing children in written language earlier and more supportively than has been possible in the past, the proposed research does not challenge the view of reading that has emerged concerning the centrality of decoding and does not advocate a “sight word” or “whole language” approach. These previous proposals neither considered the potential power of introducing written language immersion long independently of schooling, nor did they consider the affordances possible with the newest technologies. There is a significant amount of orthographic structure in alphabetic writing systems, which the child is capable of learning, and thus is not dependent on “sight word” reading [12] (Massaro & Jesse, 2005).

The present work also builds on brain science, which has documented so-called critical periods in audition, vision, and language. These critical periods are crucial for development. In contrast to later in life, the brains of young children are especially malleable. Deprivation of sensory or linguistic input during this time can diminish neural cell growth, produce cell loss, and reduce the number of dendritic connections among neural cells. This can result in a substantial deficit in the functions of sensory and language systems of the child (Huttenlocher, 2002). Similarly, limited written input early in life can place the child at a disadvantage in learning to read when schooling begins (Mayberry, 2010).

Technology assisted reading acquisition (TARA) would also be of particular value for deaf, hard-of-hearing children, and non-native Arabic speakers since it will allow reading acquisition without requiring a mastery of spoken language. The proposed app provides an opportunity for them to learn written language in parallel with the learning of sign language and/or spoken language (Mayberry, 2007; National Reading Panel, 2000; Shanahan & Lonigand, 2013; VL2, 2012).

One of the potential benefits of children learning to read early in life is that they can then read to learn. This innovative intervention would also help redirect financial resources where they will have the most impact. Although 90% of private and public educational spending is on children between the ages of 6 and 19, while 90% of brain growth occurs before age 6. Spending for nurturing children for literacy before age 6 will have the most impact in improving the quality of life, especially those children who currently reside on the wrong side of the digital divide.

The cost of illiteracy as well as the huge cost of formal literacy instruction is one of the major financial burdens on societies. In addition, many children who are considered literate still have difficulty in comprehension which may be related to delaying reading instruction until schooling begins. By embedding the child’s meaningful experience in written language, their learning to read becomes embodied in the same manner as learning spoken language. Our other research has shown that the language of picture books is more challenging than spoken language, so this exposure is valuable in the development of vocabulary, grammar, and language comprehension [11] (Massaro, 2015b). It has also been shown...
that reading ability at age 7 is a strong predictor of social economic status as middle age adults.

The remaining of the paper is organized as follows: In section 2 we discuss several apps for learning Arabic proposed in the literature. In section 3 we describe the Understand my Word app and its features. In section 4, we present the feedback from users on the app functionality, experience, and user satisfaction, while in section 5, we conclude the paper and propose new direction for its improvement.

II. ARABIC LEARNING APPS

Given the recent miniaturization of technology, Arabic-based learning apps give children the necessary technology to carry with them wherever they go and learn. In fact, all they need is a smart device like an iPhone or tablet with the necessary software and an access to the Internet. There are a lot of Arabic learning applications available nowadays in Google play and Apple store ranged from basics to advanced. For instance, the Ali/Bee app proposes a list of simple Arabic words with their direct pronunciation. The user selects a word and listens to its pronunciation. A list of matching exercises is then proposed where the user sees a word and try to match it with the corresponding sound. She/he can form the spelling letters of a vocabulary by clicking on the proposed letters. A variety of topics is proposed following the same structure. The app exists in two modes, free with advertisements and premium with a subscription fee. The app is intended mainly to English-speaking learners. The Learn Arabic Language by Drops is a new app to learn Arabic vocabulary on different aspects of life like, travelling, food, nature, animals, health, garden, and shopping. The app offers only 5 minutes of free usage per day. The tutorial consists of a falling drop showing an Arabic vocabulary with its representative image and its pronunciation. The user can listen to the word pronunciation and matches it with the corresponding image. For a longer daily usage, the user needs to pay a subscription fee. The Learn Arabic: course and language app by Mondly is a good software to learn Arabic vocabularies mainly for non-Arabic speaking learners. It offers tutorials for beginners, intermediate and advanced learners by showing words, phrases, images, and pronunciations with animations. The sounds seem to be pronounced automatically by a synthesizer and not recorded as in many other applications that we have evaluated. The app has a rich content on different topics and for different users. It offers a trial period of 7 days after which the user is charged for $9 per month. The learn Arabic app freely available on Google play is very basic. It contains a list of the most used English words with their corresponding spelling in Arabic along with their pronunciation. The user can click on the microphone icon to listen to the pronunciation. The learn Arabic 4 Kids app teaches the children the Arabic alphabet with their pronunciation assisted by images. It is also a very basic app with no tutorials or games. The smart teacher app is intended to teach learners the common words used in different domains including body parts, animals and birds, nature, fruits and vegetables, home appliances, etc. The principle of this app consists of showing an image with its name in English and a list of three words in Arabic. The user must select the proper Arabic name of the displayed object to pronounce it. The Rosetta Stone app proposes short Arabic tutorials with 30 minutes of daily usage. Its features include speaking, reading, and listening with full words and genres identifying (i.e., male and female). The complete version is accessible for $120 per year. The Memrise app teaches users the different shapes and structures of Arabic letters to form words. It has an annual subscription fee of costs $30. The Duolingo app is a very good app to learn Arabic vocabulary for non-native Arabic speakers. It offers a series of simple tutorials on different topics and aspects of life. The user can select the topic of interest and start learning. It allows to learn different shapes of Arabic letters by listening to their pronunciation, form words and phrases by a drag and drop plus many other options. The podcast: Arabic Pod 101 proposes simple pre-recorded tutorials for listening. The user can download, listen to them, and repeat broken phrases. The Arabic with Sam podcast offers Arabic tutorials explaining Arabic vocabulary occurring in sentences with details on pronunciations and grammars. It requires basic knowledge of Arabic before listening to the tutorials. It targets English-speaking learners. In all these applications, the user cannot change or add new contents as possible with our Understand my world app, where the user can interact with it, take images of objects from his/her surrounding and discover their names verbally and in written.

III. UNDERSTAND MY WORLD APP

This app is intended to be totally free of charge and address mainly the school-age children with Arabic learning difficulties. It can be very useful during COVID19 restrictive period to learn. The goal is to evaluate it for improving the mastery of Arabic language. The app records all the interactions with a specific child to determine the exact history she/he has had with both spoken and written language. Using this history, we test the child’s ability to recognize and remember spoken and written words that they have experienced during interactions with the camera, statement, and question functions of the app. The children will acquire new vocabulary and improve their language skills seamlessly and without direct instruction. The child needs to simply click on its icon to run. She/he can start to take a photo of an object from her/his environment and start learning vocabulary. The app immerses the child in the world of experience combined with spoken and written language to support the acquisition of the Arabic language skills. It is designed to be intuitive and seamless to use. Intended to be held horizontally, the home screen displays statement, camera, and question icons on the bottom. The top icons show options for toggling selfies, flash, instructions, and settings. Figure 1 shows the home screen of the app with its available buttons.

![Fig. 1. Understand my world app home screen.](image)

There are three unique functionalities in the app: a camera, a statement, and a question. Figure 2 shows that when a picture is taken in the camera function, labels are presented on the screen. Figure 2 shows the image of a garden with the Arabic label “Hadeeqa” (garden) shown on the screen. The child can easily select the question mark (?) icon to learn about the
definition of the present label (e.g., “A garden is a planned space, usually outdoors, set aside for the display, cultivation, or enjoyment of plants and other forms of nature. The garden can incorporate both natural and man-made materials”), or select the next button (>) to be presented with a total of 5 labels pertinent to the garden image.

The statement function – depicted by a speaker icon – allows the user to say words aloud and have them repeated back, accompanied by their written presentation on the screen. The question function – depicted by a microphone icon – allows users to ask a variety of questions and have them answered by our intelligent algorithms. Many types of questions can be asked, ranging from “Tell me about…”, “Define…” and “What is…”, etc. In the camera function, images captured by the camera are labeled and defined in both speech and writing. The different modes of the understand my word app are shown in the following figure 3.

The technology for the camera functions are the image recognition tools of Clarifai available at (Clarifai.com). We choose the top five labels that Clarifai returns for a given scene. Definitions of these labels can be accessed and are provided by Houndify’s database of knowledge domains (Houndify.com). The spoken output is created by Apple’s AV text to speech synthesis. Figure 3 describes the process. The current obtained results are good. However, in some cases the scene recognition techniques can give wrong or irrelevant names. Therefore, we need to develop an approach to check the suitability of the provided names for children. We will build an educational corpus of most used Arabic vocabularies in education.

To eliminate the need for eye movements, the successive words are sequentially presented one word at a time in a large text at a single location on the screen. Literate persons can read this presentation mode just as efficiently and accurately as a typical document format ([6] Juola et al., 1995; [22] Yu et al., 2007). In fact, Kwon et al. (2007) found that reading speed using this method was significantly faster than traditional reading speed for 3rd graders through adulthood.

The statement function allows the user to record his or her speech, which then is transcribed and presented in both written and spoken form. The speech recognition is carried out by Houndify. Analogous to the camera output, the spoken output is created by Apple’s AV text to speech synthesis, and the written output of words is presented one word a time and is aligned with the spoken output.

The question function allows the user to ask questions, which are answered by accessing Houndify’s knowledge domains. The app repeats the question and provides answers to the question in both spoken and written form.

In the Settings screen, the speech rate and size of the letters can be easily controlled and adjusted across a wide range of options to accommodate all different types of learners. It also allows the user to select whether or not speech should accompany the presentation of the written words.

While this app makes use of a wealth of information, sometimes the program will not recognize what is asked, or it will not have the information requested. In these cases, the learner will receive the reply: “Sorry, I didn't get that.”

Using the required Internet connection, the interface automatically provides fairly accurate image and speech recognition, all while preserving the learners’ anonymity. This application ensures safety for all ages of users, as indecent or
offensive words will not be acknowledged in dictation or used in responses.

Unfortunately, the rich resources available for English are not available for Arabic. It was necessary then to use Google translate between English and Arabic to use the English resources. The English outputs of Clarifai and Houndify were directly translated to Arabic. The translated Arabic then was presented to the user in both spoken and written modalities using the Apple AV API for Arabic text to speech synthesis. In addition, the Arabic speech from the user in the statement and question functions had to be recognized by the Apple AV API before using Google translate to obtain the English equivalents to access Houndify’s knowledge domains. Then, of course, Houndify’s output had to be translated back into Arabic to provide the Arabic answer to the user.

Having successfully implemented the app in Arabic, we intend to test its functionality and usability with the potential users. Central to this goal, it was necessary to program a history of a given child’s use of the app. The history records the complete history of the three functionalities camera, statement, and question. The app was programmed to ensure that the child saw the text that was recorded in the history. When a word, or a definition, or an answer was presented, the screen was made unresponsive until its presentation was complete. Therefore, the child could not quickly continue pressing the icons so the words would flash by without being read or heard.

IV. EXPERIMENTS

For the time being we have tested the app with only five subjects of different ages mainly for its functionality, interface, and user satisfaction. The following figure 4 shows a girl and a boy using the app while at home to learn Arabic vocabulary.

Fig. 4. A snapshot image taken by the user smart device camera.

I need to be able to select any fruit object occurring in the image like for instance the watermelon, then the app highlights my selection and provide me with some detail about the object. This suggestion needs to improve the images analysis and recognition algorithms to extract iconic objects. Subject-2 said: “the app can teach vocabulary in an enjoyable and interactive manner; I like its simplicity and easy to use; I learnt few words quickly with good explanation”. Subject-3, “the app works relatively well, however, for some images I did not get a response.” Subject-4, “the app works only in a horizontal mode and on an ios platform only. In addition, the app closed several times after few minutes of usage”. Subject-5 suggested to add encouragement messages when learning properly a new vocabulary, like for instance, very good, excellent, well-done, etc. so that the user feel happy with her or his progress. Therefore, we should conduct several assessment surveys on likeability, usability, and user experience. Table-2 shows some questions included in the surveys to be analyzed for further improvement.

TABLE I. POTENTIAL QUESTIONS FOR THE APP ASSESSMENT

<table>
<thead>
<tr>
<th>Assessment Number</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>How much did you like the app?</td>
</tr>
<tr>
<td>2</td>
<td>Is the app easy to use?</td>
</tr>
<tr>
<td>3</td>
<td>Did you learn new vocabulary with the app?</td>
</tr>
<tr>
<td>4</td>
<td>Was the app engaging?</td>
</tr>
<tr>
<td>5</td>
<td>How much did you use the app per day?</td>
</tr>
<tr>
<td>6</td>
<td>Was the app useful for you?</td>
</tr>
<tr>
<td>7</td>
<td>Was the recognized objects correct?</td>
</tr>
<tr>
<td>8</td>
<td>Was the image description satisfactory?</td>
</tr>
<tr>
<td>9</td>
<td>Was it simple to understand the pronunciation?</td>
</tr>
<tr>
<td>10</td>
<td>Is the app easy to user?</td>
</tr>
</tbody>
</table>

V. CONCLUSION

We have proposed a new interactive app for children to learn Arabic vocabulary with fun through multimedia while being far from their schools during COVID-19 pandemic. The app is freely available on apple store to download and use on ios-based smart device. It is very useful to enhance the children knowledge in Arabic language. They can use it at any time and from any location. The app allows the children to learn independently and recognize the objects around them. The app needs extensive experiments with children especially for those with learning difficulties. We plan to link the app with an Arabic Corpus for adequate explanation. In fact, the app is currently connected with the open world with limited control on the contents it provides. In fact, some contents may be offensive for children and need then to be eliminated.

Subject 1, mentioned that the app is very useful to learn Arabic vocabulary, however, she said “I didn’t get all details about the objects occurring in the image snapshot that I took with the camera. I prefer that the app allows me to select an object from the image, highlight it by a circle or a given color, then analyze and give information about it”. For instance, when I took the following snapshot the app gives me the five following names as “fruit”, “food”, “apple” “healthy” and “dessert”.

Fig. 5. Student learn Arabic vocabulary at home through the app.
ACKNOWLEDGMENT

This work was made possible by NPRP grant #10-0205-170346 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES