

An HCI Principles based Framework to Support Deaf Community

Fazal Qudus Khan

Department of Information Technology
FCIT, King Abdulaziz University
Jeddah
Saudi Arabia

Asif Irshad Khan

Department of Computer Science
FCIT, King Abdulaziz University
Jeddah
Saudi Arabia

Mohammed Basher

Department of Information Technology
FCIT, King Abdulaziz University
Jeddah
Saudi Arabia

Abstract— Sign language is a communication language preferred and used by a deaf person to converse with the common people in the community. Even with the existence of the sign language, there exist a communication gap between the normal and the disable/deaf person. Some solutions such as sensor gloves already are in place to address this problem area of communication, but they are limited and are not covering all parts of the language as required by the deaf person for the ordinary person to understand what is said and wanted? Due to the lack of credibility of the existing solutions for sign language translation, we have proposed a system that aims to assist the deaf people in communicating with the common people of the society and helping, in turn, the disabled people to understand the healthy (normal people) easily. Knowing the needs of the users will help us in focusing on the Human Computer Interaction technologies for deaf people to make it further more a user-friendly and a better alternative to the existing technologies that are in place. The Human Computer Interface (HCI) concept of usability, empirical measurement and simplicity are the key consideration in the development of our system. The proposed Kinect System removes the need for physical contact to operate by using Microsoft Kinect for Windows SDK beta. The result shows that the It has a strong, positive and emotional impact on persons with physical disabilities and their families and friends by giving them the ability to communicate in an easy manner and non-repetitive gestures.

Keywords- Human Computer Interaction, Design, Human Factors, Deaf, Sign Language Synthesis, Kinect Devic.

I. INTRODUCTION

It has been noticed that deaf person face many difficulties when they try to communicate with their community by sign language especially in public places like hospitals, hotels, and markets. That is what usually forces them to accompany individuals to help them to get their needs. For this reason, this research focuses on suggesting an HCI based solution

using technology that may give them a chance to depend on themselves and get out from their solitude. This research work target for people with disabilities (deaf people). Al-Amal Institute that deals with such cases were contacted to test the proposed system and study the impact of the application on deaf people. A system which is based on HCI Framework is proposed to help the deaf community people to understand and use it easily. This system would provide:

- Easy-to-use user interfaces activated by hand motions.
- An economical IT solution for deaf and dumb people rather than accompanying individuals to help them get their needs.
- An efficient IT solution rather than other IT technologies and solutions such as the sensor gloves that are used in translating sign language. KTDP system can work in any public place at any time as well as it can recognize the motion of the head, arms, hands, and fingers.[9]

The proposed solution for the previous problem is developing a system consists of Kinect device as shown in figure 1, PC, and a database. Kinect is Microsoft's motion sensor add-on for the Xbox 360 gaming console. The device provides a natural user interface (NUI) that allows users to interact intuitively and without any intermediary device, such as a controller in the middle [1,3].

The database contains a set of images that represent a dictionary for the sign language; each image is stored with a particular meaning. By programming the Kinect device, it will be responsible for capturing and detecting the motion of the deaf then sending the captured scene or image to the PC.

The PC then will compare between the captured image and the images in the database using the appropriate matching algorithm. After matching the images, the PC will show the text of the captured image. Subsequently, this system is considered as a

mediator between healthy people and those of special needs. As mentioned previously, we called this system Kinect Technology for Deaf People (KTDP).

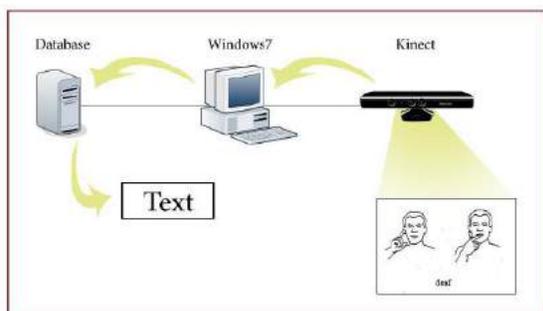


Figure 1. Shows how the system works.

This paper is organized as follows: section I introduction of the paper, sections II discusses the related work, section III the proposed Framework, section IV result of implementation and testing of the framework, and lastly section V the conclusion and future work.

II. RELATED WORK

Michelangelo in one of his famous brainy quote said "If people only knew how hard I work to gain my mastery, it would not seem so wonderful at all" from this saying we realize that hard work makes us prepared to face adverse situations. Hard work helps any student to become extraordinary. It's the key to success. Hence, we believed that we need to keep working patiently and hardly till the goal of our project is achieved. To us, we need to "see a little further into the sea" that we had to search efficiently, read deeply a lot of scientific websites and books of different fields and topics to get the best information and reference required for the project to keep going on.

Yasir Niaz Khan and Syed Atif Mehdi in their article describe the use of a device known as the Sensor gloves. This device is made of cloth attached with sensors. While they suggest that the utilization of a device called "the data glove" is a better choice over the utilization of a traditional camera, the reason being that the user has the flexibility of free movement which is dependent on the length of wire connecting the glove to the computer. However, when the camera is used, the user should stay in position in front of the camera. The gloves performance is not affected by any disturbing factor i.e. electromagnetic fields, the light or any other disturbances.[8]

In total seven (7) sensor glove of 5DT Company were used in their system out of which five (5) sensors are used for the fingers and thumb. While among the two sensors left, the sixth sensor is used to

measure the tilt motion of the hand and the seventh or the last sensor is used for the rotation motion of the hand. While the flexure of the fingers is measured by the Optic fibers which are placed on the gloves.[8]

The project under discussion uses only postures because of the reason that the glove can only capture the shape of the hand and not the skeleton/shape or motion of any other part of the body. Among the Signs, there are two letters for which the signs are ignored and those are for letters "j" and "z" because they involve moving gestures. Only Two special signs i.e. Space between words () and Full stop (.) were added to the input set. There was no compulsion in doing so but they have been added as to perform Basic English sentence writing functionality.[8]

Among the problems mentioned by the authors of the article, One of the problems was that some letters were left out of the domain of the project as they involved dynamic gestures and may not be recognized using this glove. The use of two sensor gloves was not tested out and another problem was that some gestures require the use of both hands which was not discussed in the project.[8]

The article "Multiperspective Thermal IR and Video Arrays for 3D Body Tracking and Driver Activity Analysis" by Shinko Y. Cheng, Sangho Park, Mohan M. Trivedi focused on the body part movement for driver alertness. They develop the system to determine the used multi-perspective (i.e. four camera views) multimodal (i.e., thermal infrared and color) video-based system for robust and real-time 3D tracking of important body parts and to track some other things like head and hand motions during the driving. So their focus was on tracking in a noisy environment to avoid accidents while adopting the said system for the sign language is not a cost effective solution, as many sensors and furthermore, a lot of image processing is involved, which makes the system more complicated and less efficient.[11]

A. Comparing methods to find information First approach; using the internet

Strength

- The availability of internet all the time (24 hours a day, 7 days a week). [2]
- Cheap resource.
- A resource that has a huge amount of rich and useful information.
- The existence of search engines helps us a lot.
- Getting latest and up-to-date of information.

Weakness

- We face some difficulties to find the information that we need especially when we are looking for a small point.[3]
- There is much wrong information on the internet so we cannot trust everything we read on the internet.
- We must verify the reliability of information source before using them.

B. Second approach: using the interviews

Strength

- Making interviews help us to get information from experts.
- The ability to get more details in specific knowledge area or field.

Weakness

- Facing difficulties to make appointments with some responsible people.[3]
- Hiding some correct information could happen during the meeting from some people.[3]

III. THE PROPOSED FRAMEWORK

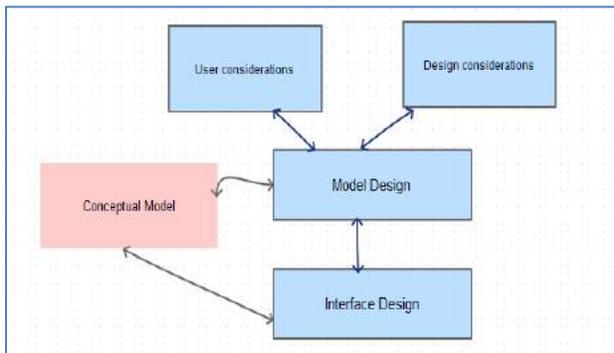


Figure 2 HCI Principles based Framework for the proposed system.

The proposed system as shown in figure 2 is based on the HCI Principles theories exist relating different design and user consideration. The user requirements consideration of the system is based on how the user communicate with the Kinect device and how it is interpreted? Here we consider Several elements that influence the way the model is developed, tested and maintained. At the Deaf users' side: their understanding of the sign language, as they will give input to the system. At the Normal Person users' side: their understanding of the Arabic language, who will see the interpretation of the gestures.

Design considerations:

This system is developed for the especially needed person, a user-friendly application is proposed that should help the users to form the correct productive rational interpretation on the screen of the given gesture by the deaf person. Common design methods that are considered in our design include the following factors:

- *Simplicity in Gestures interpretation:*

Frequently and commonly used gestures by the deaf person are to be focused so that the interpreted function should be easily readable and understandable by a normal person. commonly used simple gestures by the deaf person will be considered for interpretation rather than asking the deaf person to memorize a new gesture for the system requirement, So that the system should be easy to understand and simple to use and transparent enough for the user to concentrate on the actual meaning or message of the sign that is used.

- *Familiarity with the sign language:*

As this framework is built upon concrete requirement determination, it is very important to use this fact in designing a system. Relying on sign language the deaf people are familiar, this system is designed in such a way so that the familiarity factor within our system must be addressed and prioritized.

- *Availability of options:*

Since recognition is always better than remembering of the options available, Our system is efficient enough in terms of the user friendly interface and the options that we are providing, that it always suggest help in the form of animations and visual elements to ease the user to recall the functionality of the system.

- *Flexibility:*

The user will be able to use any hand, in the same sequence that the sign language use and the system will be flexible enough to handle it, at any time.

- *System Feedback:*

The Kinect device reads the movement of the gestures continuously and gives feedback through the system, through the action/hands movement of the deaf user. We understand that the prompt feedback helps to assess the correctness of the sequence of actions and that is the reason this feature is of high priority to us in our proposed framework.

3.1 SYSTEM DETAILS

A. Managing the database

Description: This system is mainly depending on its database in which it is impossible to accomplish the main functionality of the system without it. After creating the database, the administrator must be able to add, delete, and manage the elements of the database.

Priority: this function has a very high priority due to its role in the system.

Requirements:

- 1- Database to store the data.
- 2- Enough storage capacity.
- 3- Administrator to manage the database.

B. Making gestures

Description: special needs person has to stand in front of the Kinect device and make some motions or gestures that can be translated to text according to sign language rules.

Priority: this function has a high priority because it is considered as an essential pillar of our system as well.

Requirements:

- 1- Special needs person to make the motions.
- 2- Kinect device to detect the motion.
- 3- A program to translate the motion into text.

C. Reading gesture translation

Description: after detecting a motion by Kinect device, the appropriate text will appear on the screen to show the meaning of that motion to normal people. Priority: this function has a moderate priority because the system can work and accomplish its main functionality without the existence of the normal people to read.

Requirements:

- 1- Kinect device to detect the motions.
- 2- A program to translate the motion into text.
- 3- PC screen to show the translated text.
- 4- Normal people to read the translated text.

As shown in Figure 3, the use case diagram of the system, there are three main actors / stakeholders of the system, the Deaf person can make a gesture which is translated by the system into a plain English text and the normal person (with whom the disable person wish to communicate) can see the message on the screen. The developer of the system can add more gestures in future to enhance the system. Currently, only few gestures with basic needs are programmed.

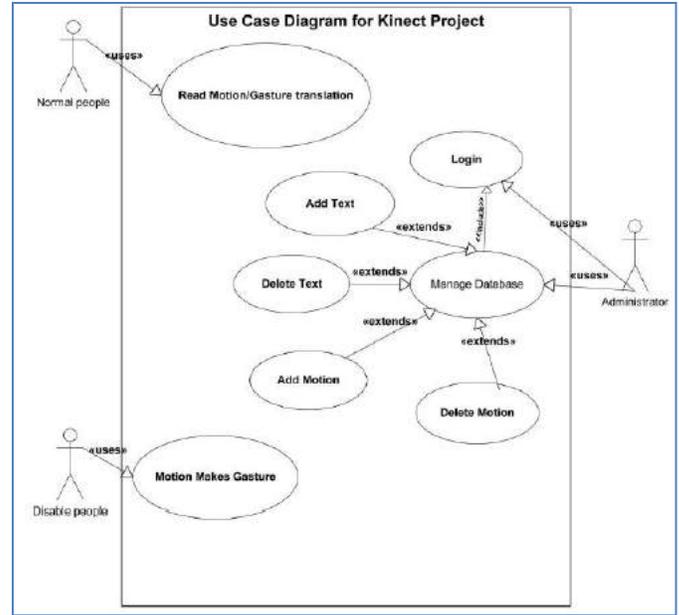


Figure 3. Shows high view of the main system requirements.

3.2 NON-FUNCTIONAL REQUIREMENTS

The non-functional requirements are not fundamental like functional requirements. They represent the qualities of our system. The KTDP system requires the following non-functional requirements to be filled in:

- Modifiability: the admin need to add more gestures, so modifiability is required as to enhance the system in future. We will make admin rights for the code to be modified.
- Usability: the system is useful if it can help the Deaf community, For usability, the device need to be installed with the modifications that we do in the using the Kinect SDK.
- Response time: a quick response time is required, this can be achieved if the device is in range.

A. Data requirements

After making several meetings with instructors and doctors that have a good experience in our project fields, and after looking in similar projects and literature searches, our team concludes what data should be stored in the database. The database should include the following data:

- The Sign Language gesture's data.
- The Meanings of each gesture's data.
- The Administrator's data to add and delete gestures and its meaning.

3.3 DESIGN AND METHODOLOGY OF GESTURES RECOGNITION

A. Methodology of Gestures Recognition

Gesture recognition in our system as shown in figure 4, can be explained in three steps; getting joints positions, comparing the position of some joints with the other joints positions, and generating a value for each gesture depending on the comparison step. [4]

Using the skeleton tracking feature that the Kinect device provide, we can track up to 20 joints in the user body who stands in front of the Kinect. In our system, we are interested in 14 joints which are in upper half of the body. [9]

By tracking these joints, we can get the position of each one as X, Y, and Z coordinates. After that, the system will go through 25 conditions (IF statements) that we built in the code to compare the positions of right hand, left hand, right elbow, and left elbow with each other and with the rest of the 14 joints. Each condition will result in one bit (1 or 0) and add it to a variable, let us call it "value_word". [7, 10]

Finally, the value_word variable will be a bit pattern consists of 25 ones and zeros, and this pattern could not be repeated for more than one gesture as each gesture differs at least in one joint position from others. This value is converted from binary to decimal before being saved in the database. By this way, the system can recognize up to 2 different gestures. Each gesture represented by one unique value. The following figure (figure 8.1.1) shows how the system compares joints positions then generates a value for the gesture. [7, 10]

For adding a new gesture to system purpose, we only need to store the generated value in the database as primary key with its title and id. On the other hand, recognizing the gestures and showing their meanings after they being added can be done by comparing the values generated while the deaf user is making gestures with the stored values in the database. If one value produced by deaf user gesture matches one stored value, then the title will appear on the screen.

3.4 ANALYSIS DONE FOR THE REQUIREMENT

There are many useful ways to collect data from different sources. One of them is making interviews and meetings with experienced persons who could give us valued information about our project. Besides, questionnaires could also be necessary to get feedbacks and opinions from large samples of different classes of the society. As well, searching and browsing the internet is a very important

technique that makes us able to find huge amounts of both practical and theoretical information that is useful for our project. This way of collecting information could be much better in saving efforts, time, and money than going to libraries and searching for specific books.[2]

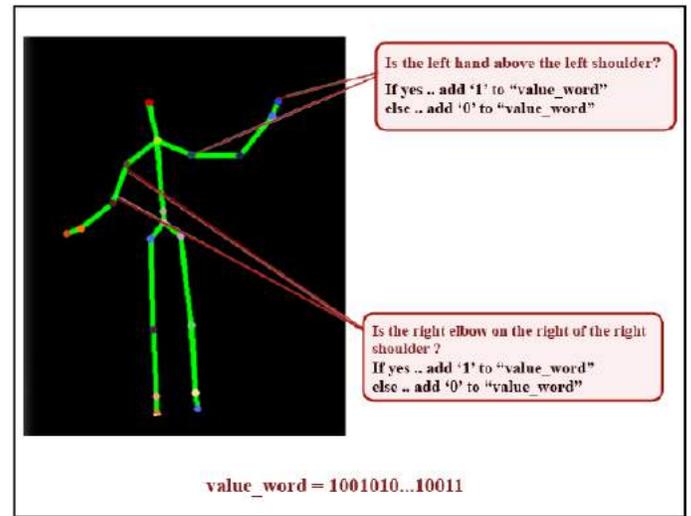


Figure 4. Gesture recognition

A. Questionnaire Analysis

a survey was conducted online, and the results were analyzed. In the following section, we list the requirements confirmation based on the questionnaire results.

Results of main questions:

As can be seen in figure 5, the statics of questionnaire showed good result that prove the importance of our project i.e. 94% of normal people don't understand the sign language. Also, 89% of them think that deaf people live in isolation of the society. Based on the previous percentages, we have verified on the importance and the benefits of our system for the society.

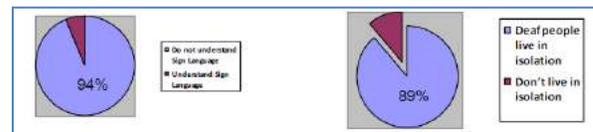


Figure 5. Statics of the questionnaire

as can be seen in figure 6, shows that the most of the respondents i.e. about 72% of people see it 's hard to communicate with deaf people and 89% of them prefer to use the technology in translating the sign language so that it could be a moderator between them and deaf people.

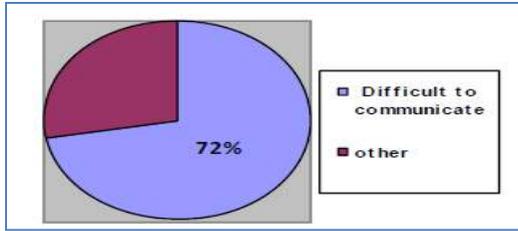


Figure 6. Statics of the questionnaire

B. Interview Analysis

Results of main questions:

The following is a list of the most important points that we conclude from the conversation with him:

- Deaf people live in isolation everywhere they go in their community.
- The Arabic Sign Language has not unified completely yet.
- The most of the sign language gestures can be made by only one hand.
- Deaf people can read and understand simple words. On the other hand, they cannot read or write long, complex sentences and paragraphs.
- Al-Amal Institute will help us anytime we need help even if we need a sign language translator to work with us.

3.4 MODELING OF THE SYSTEM

The static model (class modelling) of the proposed system is shown in figure 7.

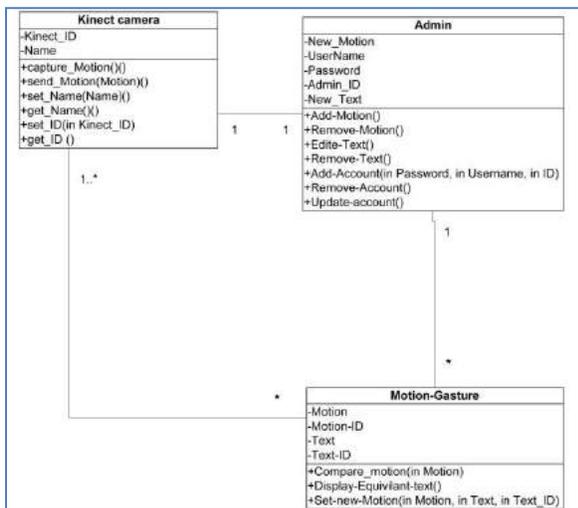


Figure 7. Shows class diagram of the system

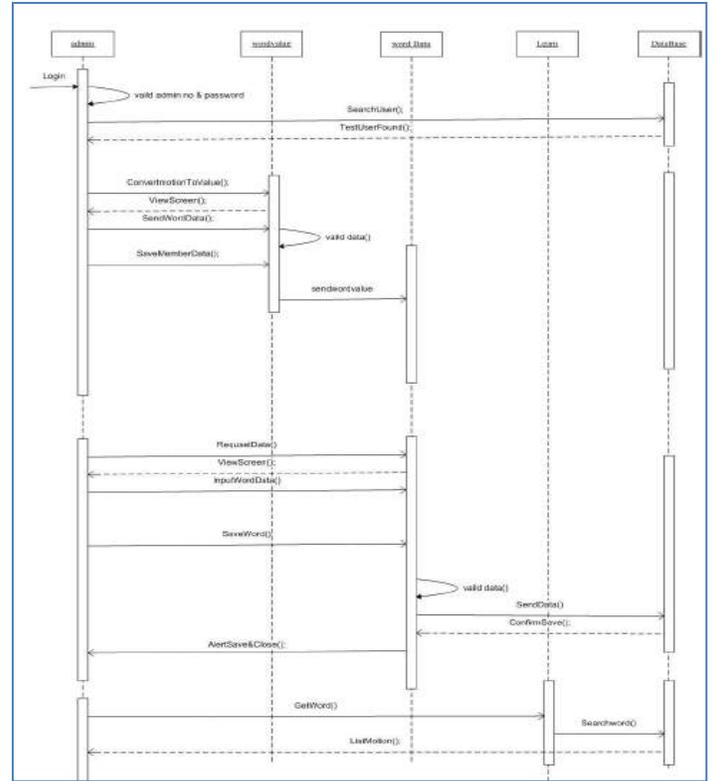


Figure 8. Shows sequence diagram of the system

The sequence diagram is necessary to show the sequence of steps that occurs while using the system. In the generic sequence diagram illustrated in figure 8, first, the registration and login process take place to confirm whether the user is a normal user or an admin. The Administrator can add gestures and its subsequent meaning in the system. While once the normal user is logged into the system after that different gestures that are performed by the disabled person are translated and delivered to the screen where the normal user can see them in plain English text.

While the HCI based User Interface is shown in figure 9.

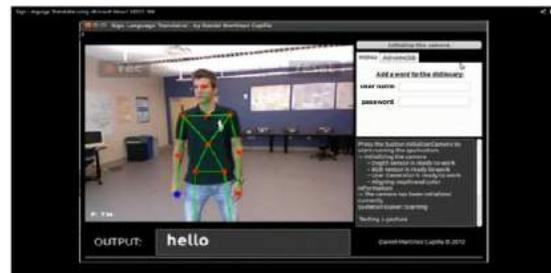


Figure 9. HCI based User Interface

IV. IMPLEMENTATION AND TESTING RESULTS

After implementing the system, the following are the interfaces that the user will interact with:

The main interface

When the user runs the program, the following interface as can be seen in figure 10, appears



Figure 10. Main interface

It contains 7 buttons. As default, 5 buttons of them are enabled and the other two are disabled. The buttons are :

- **About** button: This button moves the users to another window which show information about the authors who builds the system. This button is enabled for all users as default. If the users press this button the following window will appear
- **Login** button: This button is used by admins to login to the system as admin user. This button is enabled for all users as default. After clicking on this button another window will appear that asks the user to enter the username and password to authorize the admins. After clicking the **Login** button the following window as shown in figure 11, appears



Figure 11. Admin Log-In interface

If the user is successfully authorized, the system will move the user to the main interface again but all buttons will be enabled as can be seen in figure 12 below.

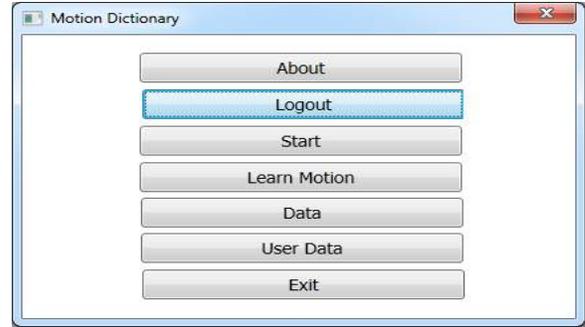


Figure 12. Main interface admin user

- **Start** button : This button is used to start recognizing the gestures made by the deaf user. This button is enabled for all users as default. After clicking on this button, another window will appear that contains Title ID, Title, and Word value fields that will show the id of the gesture, the text meaning of the gesture, and the value generated by that gesture respectively. Also, the window contains two squares. The left one used to show the skeleton of the user and the other one shows the complete sentence of group of gestures. After clicking the **Start** button the following window figure 13, appears [5,6]

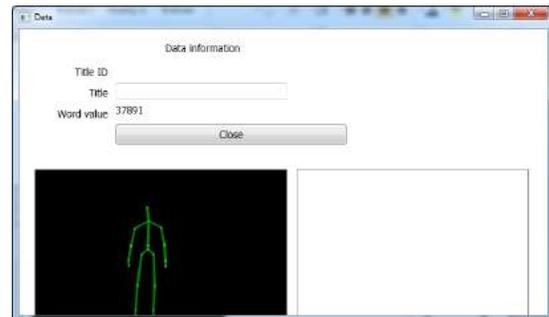


Figure 13. Data information interface

- **Learn Motion** button: This button moves the user to the system dictionary which has the gestures titles and their photos. This dictionary helps the users to learn how to make gestures. In the dictionary window, the user can search for a certain gesture by its title. This button is enabled for all users as default. After clicking on this button, another window will appear that contains a list of gestures and their values as well as an area for showing the photos of the gestures. After clicking the **Learn Motion** button the following window figure 14, appears

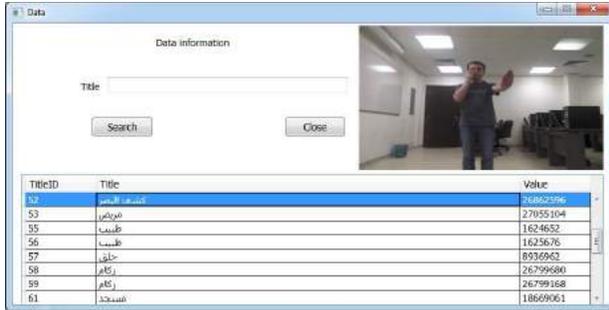


Figure 14. Dictionary interface

- **Data** button: This button is enabled only for admins after logging in the system. This button moves the admin to the window that makes him able to add, update and delete gestures. After clicking on this button, another window will appear that contains buttons, text fields, gestures list, area for skeleton view, and area for normal camera view. After clicking the **Start** button, the following window appears (Figure 15).

The **Clear** button clears the texts fields. **Add**, **Update**, **Delete** buttons are used to add, update, and delete system's gestures. **Start** button starts the skeleton view. **Close** button closes the window. **Find Value** button shows the value of a gesture. **Select Image** is used to choose an image for a gesture to be added to the dictionary. The rest two button are **Take Photo**, and **Capture**.

The first one shows the standard camera view, and the second one used to capture a photo using the camera. When the admin wants to add a new gesture, he should go through some steps respectively.



Figure 15. Data information(Admin interface)

First he should press the **Start** button. Then he should fill the text fields. After that, he has to select an image for the gesture. Finally, he can add the gesture to the database by clicking on **Add** button.

- **User Data** button: This button is enabled only for admins after logging in the system. This button moves the admin to the window that makes him able to add, update and delete admin users. After clicking on this button, another window will appear that contains a list of admin users and their information, some buttons, and some text field to enter admin

data. After clicking the **User Data** button the following window figure 16, appears.

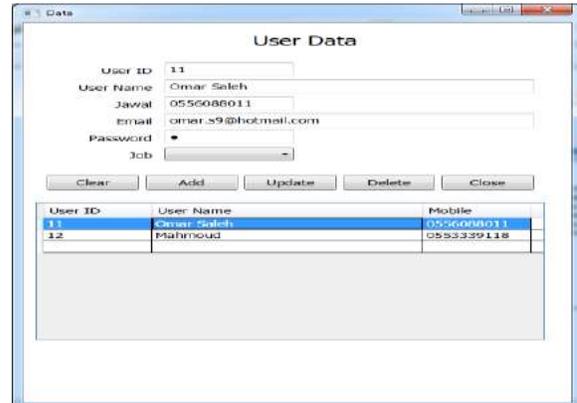


Figure 16. Admin information interface

- **Exit** button: This button is used to close the system.

TESTING THE SYSTEM

In testing the recognition of the system, we selected Ten (10) gestures randomly to be tested by some people and us.

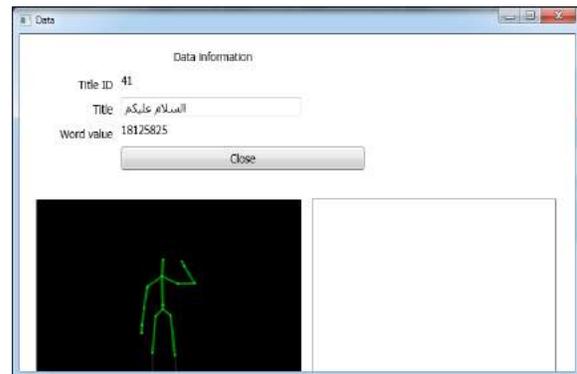


Figure 17. Gesture recognition testing(1)

The gestures are shown below Figure 17 till figure 26, with their skeleton views:



Figure 18. Gesture recognition testing(2)

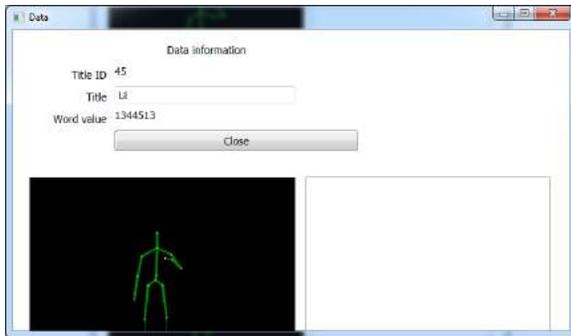


Figure 19. recognition testing(3)

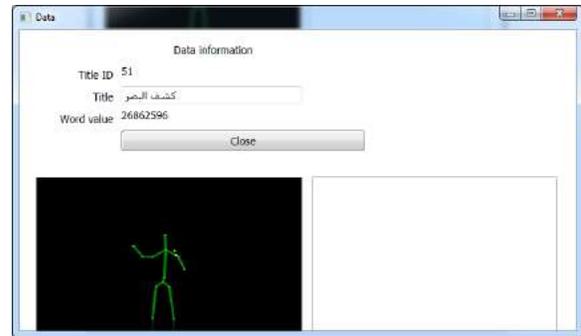


Figure 23. Gesture recognition testing(7)

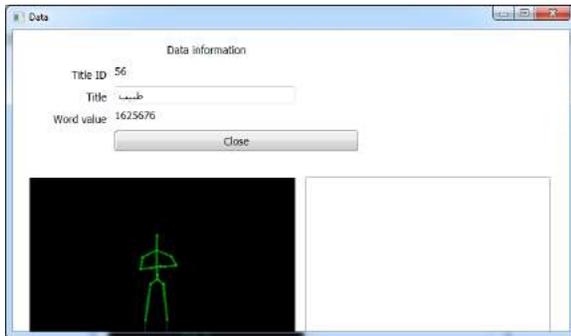


Figure 20. Gesture recognition testing(4)

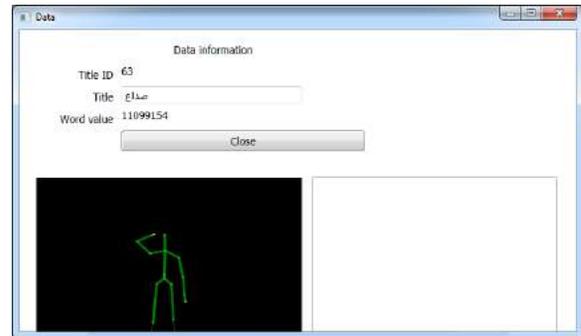


Figure 24. Gesture recognition testing(8)

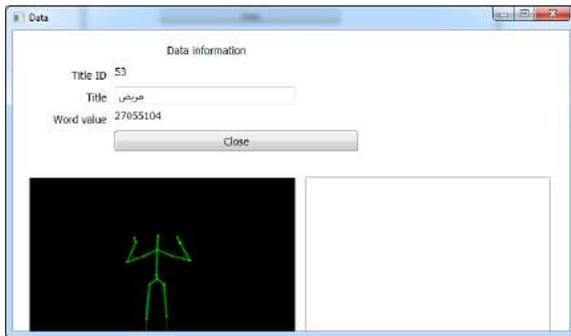


Figure 21. Gesture recognition testing(5)

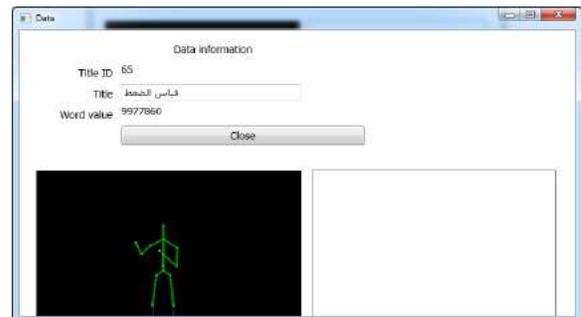


Figure 25. Gesture recognition testing(9)

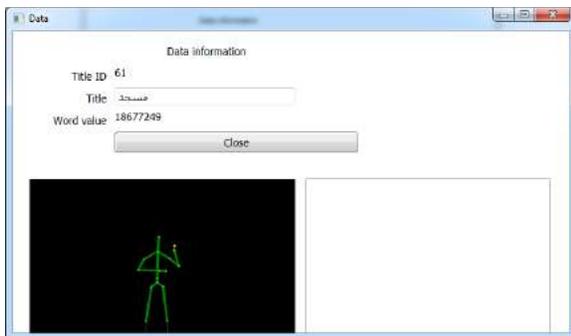


Figure 22. Gesture recognition testing(6)

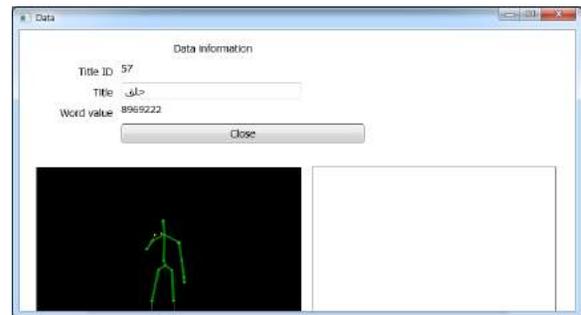


Figure 26. Gesture recognition testing(10)

Three other persons have tested this system. Each gesture of the ten gestures has been repeated five times by each. Appendix shows This system has been tested by 3 other persons. Each gesture of the 10 gestures has been repeated 5 times by each person.

V. CONCLUSIONS AND FUTURE WORK

The aim of this research study is to support and help deaf persons in communicating with their community by using a KTDP system. The system translates the body gestures made by deaf users into text.

The system has a manual for the gestures that can be recognized. Also, the system gives the admin the ability to add new gestures easily.

The HCI concept of usability, empirical measurement and simplicity are the key consideration in the development of our system. In future, the research will focus to include finger recognition to make it able to translate the whole official sign language gestures.

REFERENCES

- [1] Kinect (n.d.). retrieved Janury 15th 2017, from <http://en.wikipedia.org/wiki/Kinect>
- [2] Yaqoub Sayed ikram and Ishaq Sayed Ikram "Complaints Reporting System", (2012) Graduation / senior Projects (CPIT 499), Faculty of Computing and Information technology, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia.
- [3] Kinect (n.d.). retrieved Feruary 5th 2017, from <http://searchhealthit.techtargt.com/definition/Kinect>
- [4] KinectSdk (n.d.). retrieved January 4th 2017 from <http://www.microsoft.com/en-us/Kinectforwindows/>
- [5] Microsoft Visual studio 2010 (n.d.). retrieved Feruary 5th 2017, from <http://www.microsoft.com/visualstudio/en-us/products/2010-editions/visual-csharp-express>
- [6] NET framework (n.d.). retrieved Feruary 5th 2017, from <http://msdn.microsoft.com/en-us/netframework/aa569263>
- [7] Kinetsdk (n.d.). retrieved Janury 6th 2017, from <http://research.microsoft.com/en-us/um/redmond/projects/Kinetsdk/about.aspx> (© Microsoft Corporation, 2011).
- [8] Mehdi, Syed Atif, and Yasir Niaz Khan. "Sign language recognition using sensor gloves." Neural Information Processing, 2002. ICONIP'02. Proceedings of the 9th International Conference on. Vol. 5. IEEE, 2002.
- [9] Near Mode: What it is (and isn't) (n.d.). retrieved Janury 10th 2017, from <http://blogs.msdn.com/b/Kinectforwindows/archive/2012/01/20/near-mode-what-it-is-and-isn-t.aspx>
- [10] Simon Lang, Raul Rojas, Marco Block-Berlitz, "Sign language recognition with Kinect", (September 2011), retrieved Janury 26th 2017 from <http://page.mi.fu-berlin.de/block/abschlussarbeiten/Bachelor-Lang.pdf>
- [11] Shinko Y. Cheng, Sangho Park, Mohan M. Trivedi "Multiperspective Thermal IR and Video Arrays for 3D Body Tracking and Driver Activity Analysis". 2nd Joint IEEE International Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum (OTCBVS'05) in conjunction with IEEE CVPR2005, San Diego CA, USA. June, 2005

AUTHOR'S PROFILE

Fazal Qudus Khan, M.Sc, is working as a faculty member in the department of Information Technology, FCIT, King Abdulaziz University, Jeddah, Saudi Arabia. He has over eleven years of experience (Two years Industrial experience with more than nine years in academia and research). Mr. Khan is currently a Ph.D. candidate at the University of Kuala Lumpur, Malaysia. He has received M.Sc. in Computer and Network Engineering from Sheffield Hallam University, Sheffield, UK & Bachelor degrees in Information Technology from NWFP Agricultural University, Peshawar, KPK Pakistan. He has published several research articles in leading journals and conferences; his current research interest includes Software Engineering with a focus on Component Based and Software Product Line Engineering. He can be reached at fqkhan@kau.edu.sa.

Asif Irshad Khan, Ph.D., is working as a faculty member in the department of Computer Science, FCIT, King Abdulaziz University, Jeddah, Saudi Arabia. He has thirteen years of experience as a professional academician and researcher. Dr. Khan received Ph.D. in Computer Science and Engineering from Singhania University, Rajasthan, India, and Master & Bachelor degrees in Computer Science from the Aligarh Muslim University (A.M.U), Aligarh, India. He has published several research articles in leading journals and conferences. He is a member of the editorial boards of international journals, and his current research interest includes Software Engineering with a focus on Component Based and Software Product Line Engineering. He can be reached at aikhan@kau.edu.sa.

Mohammed Basher, Ph.D., is an Assistant Professor and the Chairman of Information Technology Department at the Faculty of Computing and Information Technology in King Abdulaziz University, Saudi Arabia. He has fifteen years of experience as a professional academic. Dr. Basher received PhD in Computer Science from the School of Engineering and Computer Science at Durham University, UK. He received Master of Information Technology from Griffith University, Australia and Bachelor in Computer Education from King Abdulaziz University, Saudi Arabia. His current research interest is in HCI and E-learning. He can be reached at mbasher@kau.edu.sa.

Appendix: The following are the tests summaries



Fig. The first test

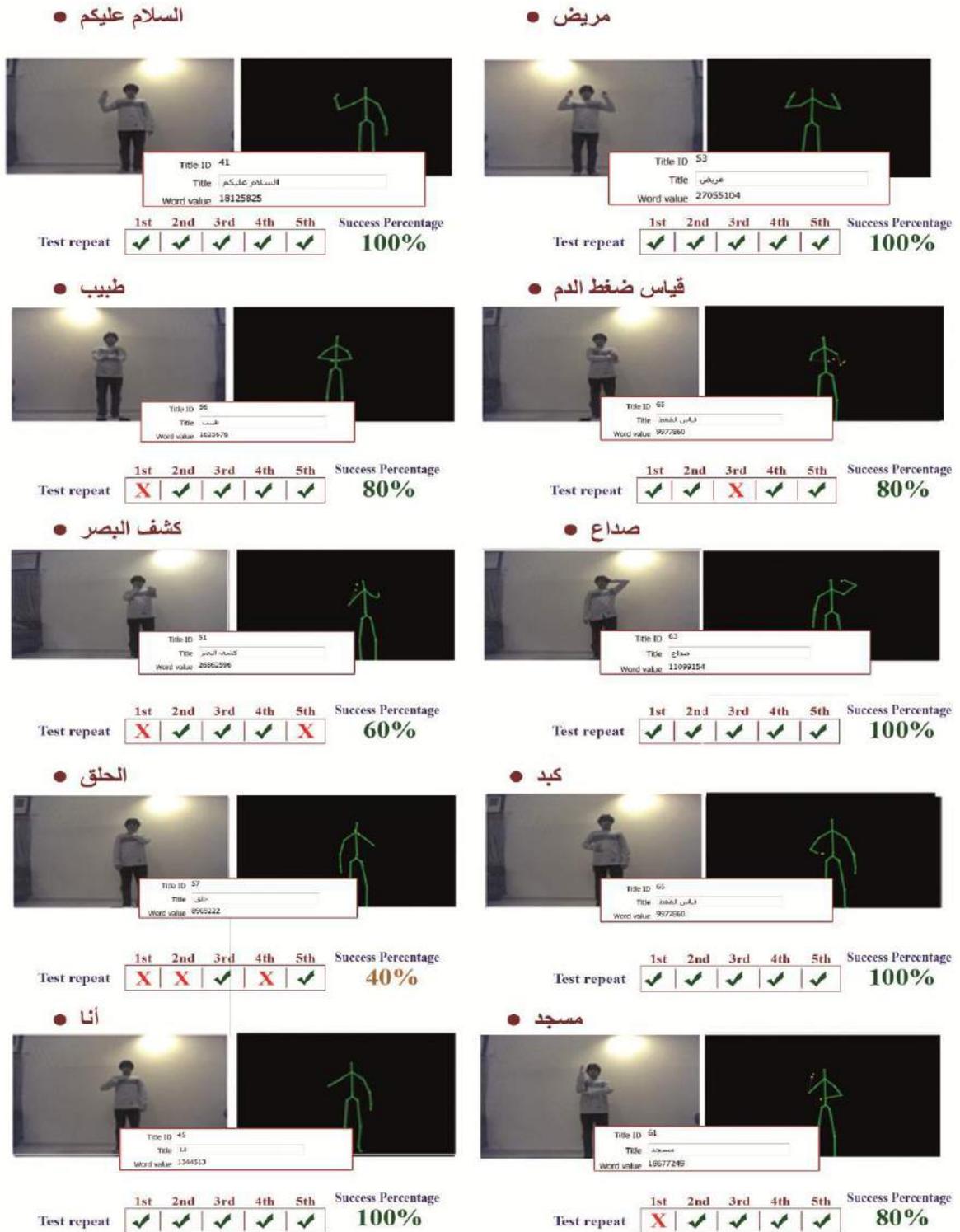


Fig. The second test

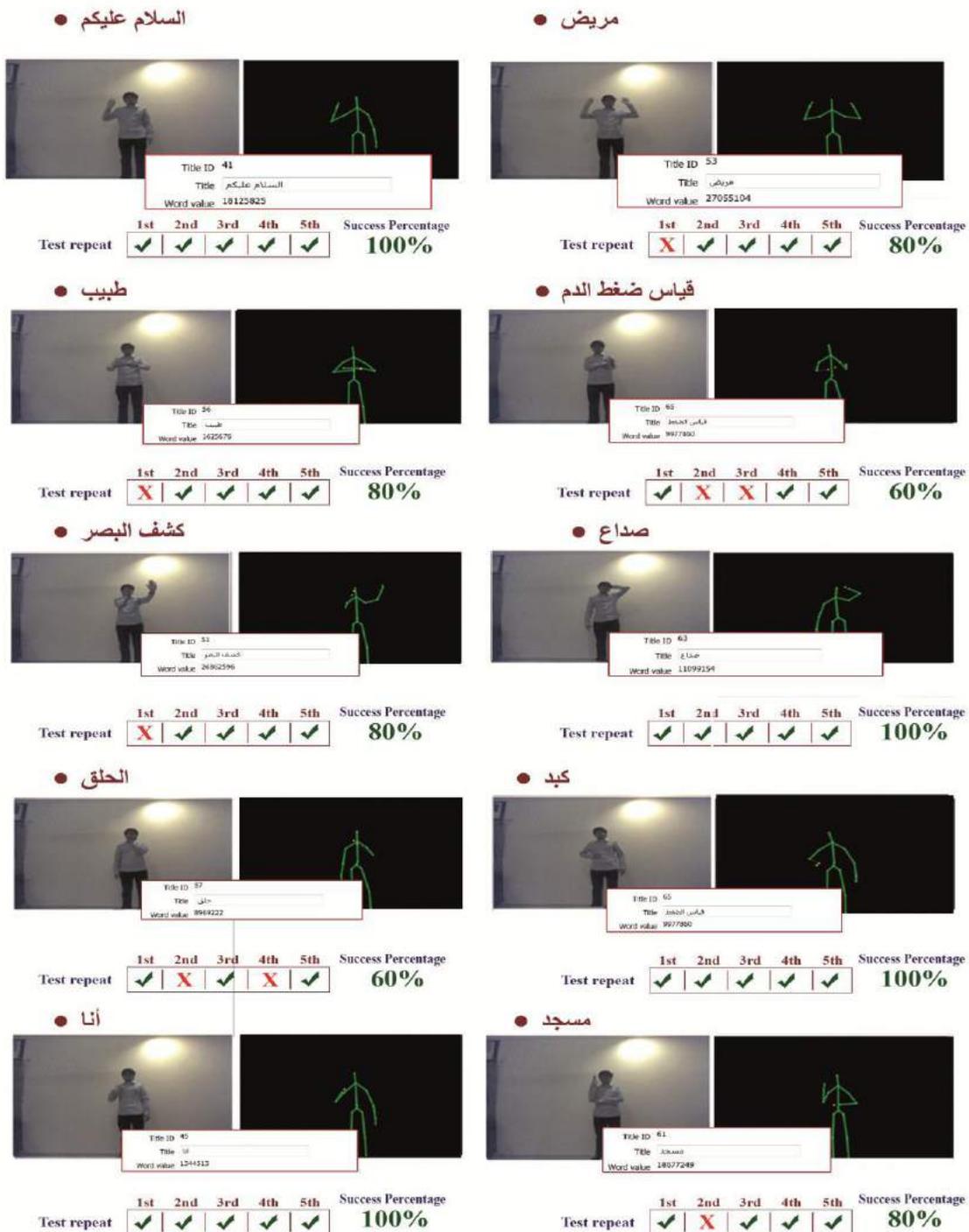


Fig. The third test



© 2017 by the author(s); licensee Empirical Research Press Ltd. United Kingdom. This is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license. (<http://creativecommons.org/licenses/by/4.0/>).