*Asian Journal of Research in Computer Science*

*4(2): 1-11, 2019; Article no.AJRCOS.52404*
*ISSN: 2581-8260*

# Spoken Dialogue System: Its Applications in the Developing Countries and a Technology for Bridging the Digital Divide and Augmenting Scarce Services in those Countries

**Oyelami Olufemi Moses[1*]**

*[1]Department of Computer Science and Information Technology, Bowen University, Iwo, Nigeria.*

***Author's contribution***

*The sole author designed, analysed, interpreted and prepared the manuscript.*

***Article Information***

DOI: 10.9734/AJRCOS/2019/v4i230110
Editor(s):
(1) Dr. G. Sudheer, Professor, Deptartment of Mathematics and Computer Science, GVP College of Engineering for Women, Madhurawada, India.
Reviewers:
(1) Oladele, Mathias Omotayo, The Federal Polytechnic, Ede, Nigeria.
(2) David Lizcano, Madrid Open University, Spain.
Complete Peer review History: https://sdiarticle4.com/review-history/52404

***Mini-review Article***

***Received 13 August 2019***
***Accepted 28 October 2019***
***Published 11 November 2019***

## ABSTRACT

**Aims:** This article reports the various application areas of the spoken dialogue system in the developing world to determine if the system could be used to bridge the digital divide prevalent in these regions of the world. The work also aims to identify in which developing nations is the system currently being put to use.
**Study Design:** A survey of twenty articles on the subject matter was carried out and their domains of the application were identified. The different forms of the evaluation carried out on them were also identified towards determining their outcomes positivity for bridging the digital divide. Various comments made of the different evaluations were also considered in determining the suitability of spoken dialogue systems in bridging the digital divide.
**Place and Duration of Study:** Department of Computer Science and Information Technology, Bowen University, Iwo, Nigeria, between February 2013 and October 2019.
**Methodology:** The different domains of the works, the different forms of the evaluation carried out on the systems, the various comments consequent upon the testing of the systems by the participants and the developing countries where those works were carried out were identified. A position was now taken based on the results obtained.

**Corresponding author: E-mail: olufemi.oyelami@bowenuniversity.edu.ng;*

**Results:** Nine of the works are in the healthcare domain, three in agriculture, one in banking, one in aviation, one in secretarial work, one in the accuracy of recognition, one in education and three having multiple domains. The various comments and results from the evaluations all point towards the system's suitability for bridging the digital divide. The spoken dialogue system is currently being used in only six developing nations of the world.
**Conclusion:** Consequent upon the results obtained, it is clear that spoken dialogue systems can be used to bridge the digital divide in the developing world and that other application areas not yet covered could be explored for the benefits of the citizens of these regions, especially the digitally disadvantaged ones.

*Keywords: Spoken dialogue system; mobile phone; digital divide; ICT; developing the world.*

## 1. INTRODUCTION

Information and Communication Technology (ICT) has ushered in an era that makes the world a global village because of the ease with which information about virtually anything in the world and different services can be obtained anywhere and anytime. However, the benefits provided by the growth of ICT have not been fully enjoyed in the developing countries of the world. The reason is that ICT penetration in these parts of the world is still lower than that of the developed world. According to the International Telecommunications Union (ITU) 2018 global and regional ICT estimates, even though there are a continuous usage and access to information and communication technologies worldwide, the developing world is still lagging behind as compared to the developed ones. For Internet, usage had increased from 51.3% in 2005 to 80.9% in 2018 in the developed world while in the developing countries, there was a growth increase from 7.7% in 2005 to 45.3% in 2018. For computer ownership, 83.2% of households in the developed world possessed a computer in 2018 with only 36.3% in developing countries. However, according to the report, there is a more predominant increase in mobile cellular subscriptions in the developing world as compared with the developed one [1]. This means that this opportunity can be exploited to bridge the digital divide existing in developing countries.

One of the technologies that can be accessed via mobile phone is the spoken dialogue system (SDS), and it has been identified to be useable to bridge the digital divide and provide wide access to information and services in the developing world using voice services and short message service (SMS) [2]. In this work, the author reviews work in spoken dialogue systems in the developing countries and provide empirical evidences on the various ways this technology has been used towards positing that the technology can be used to bridge the existing digital divide and provide access to scarce-to-access services in these regions. The author is unaware of any other study that has made this contribution.

## 2. REVIEW OF RELATED WORKS

The prospects of a voice-enabled healthcare system which obviously show that people with low literacy level, those who are blind, the visually impaired and those that are not computer literate will be empowered to access health resources that are available on the Web in graphical user interface form in voice user interface form were identified and discussed in [3]. Also, people with poor reading habit will be able to access health information through voice using their phones and that immediate access to healthcare will be provided to people in places where there are no health facilities or where the facilities are distant from.

In [4], it was observed that speech production changes in children as they develop introduces temporary variations in their speech production at specific ages and as such, create difficulties in developing spoken dialogue systems for them. The authors identified the speech corpora available for children and the languages in which they were available. The factors responsible for the decrease in the recognition of children's speeches were also identified. Furthermore, application areas of spoken dialogue systems for children like in reading as pronunciation aids and for mastery of words by preschool-age children are also highlighted. Children with special needs like the deaf and those with cleft-up palate were also identified to benefit from spoken dialogue systems. Guidelines for good user experience of spoken dialogue systems were presented also. The work, also, presented issues related to emotions by children as they interact with spoken

dialogue systems and how emotions can best be recognized. Lastly, they presented the application of spoken dialogue systems in toys and robots.

In [5], areas where speech emotions have been utilized and the reasons why emotions identification is difficult were identified. The work also considered the design criteria to be followed when designing emotional speech corpora. Furthermore, the work identified the available emotional speech databases in different languages and noted that the majority of them are adult-focused with an exception of two which are BabyEars and KISMET. The authors were also able to identify the problems besetting the existing databases of speech emotions and the features ideal for recognizing emotions in speeches locally and globally. Lastly, they were able to identify the best speech characteristics from which emotional contents can best be extracted and stated that recognition of emotions in speeches is a two-stage activity of extracting the appropriate features from the speech and then the classifying of the type of emotion of the speech. Consequently, the work discussed the various methods of classifying emotions with their merits and demerits but concluded that the Hidden Markov Model (HMM) is mostly used.

In [6], it was emphasized that spoken dialogue systems were formerly used to make the interaction between man and the computer easy by using them to carry out simple tasks, but currently being used for more complicated tasks. The work carried out a review of the components of an SDS, initial efforts at building systems that produce sounds, the state of the art of the technology, its merits, application areas and identified opportunities for research in Automatic Speech Recognition (ASR), Spoken Language Understanding (SLU), Dialogue Management (DM), Natural Language Generation (NLG) and Text-to-Speech synthesis (TTS).

Recent advances in spoken dialogue systems were reviewed in [7] by discussing how developments in big data and deep learning techniques have contributed to the advances in task-based and non-task-based dialogue systems. The structure of the pipeline-based task-oriented dialogue system is discussed with the different approaches employed in its four components of language understanding, dialogue state tracker, dialogue policy learning and natural language generation. For non-task-based systems, different approaches such as neural generative models of building data-driven systems were discussed and their shortcomings and areas of improvements highlighted. Retrieval-based methods for generating responses like single-turn response matching were also discussed. Lastly, hybrid methods combining neural generative models and retrieval-based models to improve performance were discussed also.

## 3. REVIEW OF SPOKEN DIALOGUE SYSTEM APPLICATIONS IN THE DEVELOPING COUNTRIES

In [2], the design, development and testing in South Africa, of a spoken dialogue system to provide information about employment, social grants and health to the community workers in Lwazi and the members of the community at large using the eleven official languages of South Africa was reported. Pilot studies of the spoken dialogue system were carried out in six places across the country from June 2009 to June 2010 and the results obtained showed that “that automated telephony services can be used as an effective means to communicate and disseminate localized information and enhance government service delivery in rural communities”. It is to be noted that the six places where the pilot took place were a combination of semi-rural, rural and urban areas.

The resources needed to develop a spoken dialogue system in an inexpensive manner, to allow illiterate farmers in Tamil Nada (a predominantly illiterate community), a village in India access information about the weather conditions, prices of produce and rainwater collection so as to assist them in their work in the domain of agriculture were gathered in [8]. The developed spoken dialogue system running on a laptop was tested among thirteen participants from three districts of Tamil Nadu and with varying levels of literacy using a phone and a laptop microphone. The results obtained show that the participants that used the system were willing to use it. They also showed excitement in using it to obtain accurate information and believed that the time they spent interacting with the system was not a waste. Their excitement prompted them to also put forward useful suggestions that could make the system better in serving the people. They also glossed over the few errors of recognition by the system because they found the system useful. Furthermore, those living far from the market were also willing to pay for accessing the services if the information provided by the system was accurate. Even

though the illiterates among the villagers had some difficulty in restricting their inputs to recognizable forms, however, they were sure that if trained, they would be able to operate the system.

In [9], it was identified that in Kenya, access to computers and the Internet to obtain information was very limited as very many people were disadvantaged in that they could not afford the device and could not access the Internet due to illiteracy. However, the majority of the people had access to telephones. The work, therefore, exploited this opportunity to provide speech-based access to information by developing in Kiswahili, a system that provided all information about banana farming to farmers in Kenya. The information provided also included market information. The system's usability test was carried out through a questionnaire filled by ten farmers in Kibirigwi, Kinrinyaga, who used the system. The participants were rural dwellers. These dwellers consisted of illiterate ones or those with little or no education at all, or those who had little or no knowledge of English. They consisted of three female farmers and seven males. Six of them were from age forty-one and three between thirty-one and forty while one was between eighteen and thirty. Three of the participants used a telephone daily, five every week and two rarely used a telephone. Eight of them had never used a computer before, one used a computer often while another one had a little use of the computer. There was only one illiterate among them.

The results of the usability evaluation show that:

i. the system was accepted by the farmers that participated in the evaluation as 100% of them said they liked the system;
ii. the system was easy to learn as eight out of the ten involved in the study attested to it even though the majority had never used a computer before and just three of them used the telephone daily;
iii. 100% of the participants were satisfied with the system as they preferred to use it rather than the Internet and books;
iv. The error rate was 36% and this showed that the interface needed to be improved on;
v. all the participants (100%) believed that the system would be useful for the banana farmers and that it would be beneficial if the system could be enhanced to give information about other crops and livestock;
vi. all, but one of the participants believed that this kind of system would be very useful especially where extension workers are inadequate and/or are not available for consultation and to confirm advice given by them;
vii. The visually impaired farmers would find the system beneficial as revealed by a blind computer teacher that listened to the system and also offered useful suggestions for its improvement. Kenya Society for the Blind (KSB) also approved of the system;
viii. The National Agriculture and Livestock Extension Program (NALEP) also accepted the system.

In Pakistan, a developing nation, telephone-based access was provided to access health information by the less literate community health workers in a bid to enhance their work. The information provided was in Urdu language and information on diarrhea, pneumonia and important messages for health workers was provided. This was done to provide them with easy and fast access to the health information contained in the large health manual which was beset with the difficulty of being carried around [10]. The participants that took part in the study had the lowest level of literacy, lived in the rural areas and belonged to the Pakistan Initiative for Mothers and Newborns (PAIMAN). The results of the evaluation of the prototype system by six health workers with low literacy level using six dimensions on a questionnaire on a 5-point Likert scale show that the system was useful as presented in Fig. 1 below:

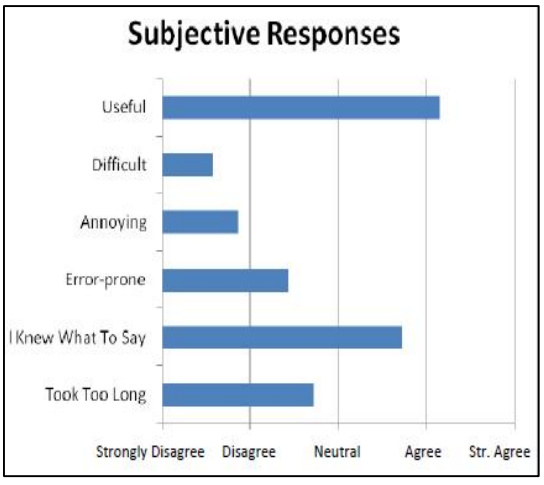


**Fig. 1. Telephone-based health information usability [10]**

In [11], it was posited that the illiterate population in the developing countries were disadvantaged in accessing the contents of the Web which are majorly in text and English language. In bridging this gap, the work provided Audio Wiki, which provided speech-based access to the corresponding text-based contents of the Web. The system allowed the creation, editing and listening to the contents in the audio format created by others on any topic of interest. A user can search for topics by speaking a keyword. If the topic does not exist, he can enter the new keyword using the phone's keypad and its equivalent audio is also recorded to serve as a keyword. If the topic searched for, is already existing, the user can listen to the contents on the topic, add his comments on the topic, edit his comments already existing and delete his comments. The Audio Wiki with English interface was accessible with feature phones prevalent in the developing world due to the cost implication of owning smartphones. The prototype system with a Hindi interface was also being developed in parallel to be deployed in India.

In [12], needs assessment for information among rural dwellers in different villages of South India was carried out for seven years. Ten information needs were identified from which only four were used to develop VoiKiosk, a spoken dialogue system that provided information to these villagers. The system enabled farmers to perform the following tasks:

i. As an avenue to seek for experts' opinions about problems relating to their crops;
ii. Access for the villagers to get information about their health conditions and the schedules of doctors' visits;
iii. Access to daily news items and information about the distance education programme being run across villages;
iv. Recording of advertisements they had and job seekers were able to browse available jobs.

The evaluation of VoiKiosk was carried out for four months among the villagers using an initial thirty participants, the population of whom later grew. The results obtained revealed the following:

i. The users were able to understand the technology to the extent of being able to play with it;
ii. They were able to use the system innovatively;
iii. They were able to discover many things on their own;
iv. They patiently spent much time with the system;
v. Over 87% of the calls made to the system was for content accessing while around 13% was for creating contents.

The implication of the above as concluded by the authors is that the system was accepted by the users and that content accessing found more usage. It was also concluded that with the increasing usage of the system, speech-based systems could be used to provide information and communication technologies in rural areas.

In [13], a spoken dialogue system to provide health information to caregivers of children that were HIV positive in Botswana was designed and developed. The information provided included:

i. Answers to questions usually asked about HIV
ii. Information about its cure
iii. Sicknesses normally experienced
iv. Information about medication
v. Nutrition information
vi. Spread of HIV
vii. Issues about personal hygiene

The work also compared the use of touchtone also called Dual-tone Multi-Frequency (DTMF) key input and voice interface in accessing the system. In evaluating the system, thirty three (33) caregivers were involved in the study, however, twenty seven (27) tested the system via DTMF and speech while the remaining six did not carry out testing on the two systems as a result of the constraint of time or system set-up failure. Testing was done for five days in April of 2008 in the premises of Baylor, and a facilitator, an observer and a Wizard-of-Oz (WOZ) operator were present. The facilitators were local graduate students trained by the investigators to facilitate in Setswana, the local language of the people. The results show among others that the majority (59%) of the participants preferred DTMF while 19% preferred speech and 6% preferred both. The authors concluded that telephony applications could be used by the low literate as well as the semi-literate users and that spoken dialogue systems in local languages could serve as a very useful means of providing health education.

In [14], a framework for the design of speech-enabled self-care eHealth systems to provide the

same information accessed in text form on the Internet by the consumers to cater for the needs of the visually impaired and the blind was proposed. Based on the framework, a prototype application was developed. The contribution of this framework lies in the fact that it is speech-based and it takes care of the needs of the underserved people majorly domiciled in Africa.

In [15], a spoken dialogue system that assisted farmers with formal education to ask questions about farming, access questions other farmers have asked and listen to the responses given to these questions was developed. The system also served as a platform for information dissemination to the farmers about weather conditions, prices of farm produce in the market, government programmes, etc. Lastly, with the system, farmers could listen to missed radio programmes archived in the system. Fifty one participants were involved in the pilot study carried out in India. The farmers were all small-scale farmers of Gujarat. Of the fifty one subjects, one was a school teacher and another a businessman while the remaining forty nine were farmers and had an existing relationship with Development Support Centre (DSC), a non-governmental organization in Ahmedabad, Gujarat, India. These participants were drawn from four districts across the state of Gujarat and because it was difficult getting female participants, all of them were males. Demographic data were available for forty-five participants, but nineteen out of them had eight-grade education or less. Twenty had finished high school as at then while six had also completed their college education. Their age range was 18 - 60 and the median age was 29. All the subjects were native speakers of Gujarati, being their first language and all of them had had no prior Internet experience significantly. The results of the usability evaluation as concluded by the authors imply that the system provided relevant agricultural information promptly and served as a forum to interact with experts and share opinions with peers. The results also show that speech is a suitable means of creating online communities in rural areas of the developing nations.

In [16], an intelligent framework for usable speech-enabled e-health system was presented. The framework on which a prototype application was developed makes available health information people search for, on the Internet in text form, available in speech form so as to enable the visually impaired, the blind, the people with low literacy and the non-literate and those that are not computer literate access the same health information people search for, for self-care. The framework also incorporates rule-based reasoning into interactive voice response (IVR) systems for screening diseases to enable them to diagnose diseases thereby imitating physicians.

In [17], it was identified that the VoiceXML standard does not support the development of expert systems, knowledge-based systems, natural language understanding (NLU), natural language generation (NLG) and natural language knowledge representation. It only supports output of synthesized speech, output of audio files, recognition of spoken input, recognition of dual-tone multi-frequency (DTMF) key input and recording of spoken input. Although previous research efforts had integrated successfully, intelligent components for NLU into VoiceXML-based systems, the work investigated the possibility of integrating intelligent components for knowledge-based systems into VoiceXML-based systems by proposing the integration of Java expert system shell (Jess) into health dialogue systems to enable them to diagnose diseases using component orientation. The prototype system developed was used to diagnose common fevers prevalent in Africa and it was tested. The results show that it is viable to develop knowledge-based systems with VoiceXML as the integration was successful.

In [18], an interactive voice response (IVR) system that served as an auto secretary was developed. A caller has two choices of services: either to speak with a person or to make enquiries. To speak with a person, the caller will need to mention the name of the person and the incoming call will be routed to either the landline or the mobile line of the person in that order depending on availability. If the choice of the caller is to make enquiries, he can do this in any of the eleven official languages spoken in South Africa. The choice of the language will make the system route the call to an operator that speaks the language. The work's focal point was to optimize the acoustic model to allow the system route calls to the desired operator based on the choice of language and to enable the system to add new names. The results obtained using an independent testing corpus showed that the accuracy in recognition of named-language stood at 95.11% and for proper names, it stood at 93.31%. These results suggest that the system was a success.

In [19], drug/pharmaceutical counterfeiting was identified as a menace in both the developing and developed countries. Even though there had been different solutions proposed to deal with the menace, the work presented spoken dialogue systems as a novel way of detecting counterfeited pharmaceuticals/drugs to reduce the threats they pose to consumers, healthcare providers, drug manufacturers and governments. The work exploited the ever-increasing usage of mobile phones in developing countries as compared with any other digital device to safeguard the health of the populace. However, in [20], the author went further to carry out user satisfaction and acceptability evaluation of the system. Twenty-nine participants were involved in the testing and seventeen of them were females while twelve were males. Those in the age range of 15–20 were thirteen, eleven within 21–30, three within 31–40 while two were within 41–50. 10% of them were novices as per experience/skill in computer usage, 24% were average users, 52% were good users while 14% were experts. In the use of digital devices for work enhancement, 14% used cellphone/PDA, 69% laptop/notebook, 7% used desktop/PC while 10% did not specify any device. 93% owned mobile phone/PDA, while 7% did not respond. 21% had owned a mobile phone/PDA for six months, 7% for one year, 10% for two years and 62% for more than two years. 3% of the participants were making calls 1–2 times a day, 14%, 3–4 times a day, 17%, 5–6 times a day and 66% more than seven times a day. The results obtained from the twenty nine subjects that participated in the study show that the users accepted the system and were satisfied with it as 97% of them were in support of the system for healthcare delivery while 96.5% wanted National Agency for Food and Drug Administration and Control (NAFDAC) to introduce the system to enable the masses to identify fake drugs/pharmaceuticals. NAFDAC is the agency in Nigeria in charge of food and drug control and administration.

In [21], a comparison of DTMF and speech for interacting with a VoiceXML-based expert system developed for diagnosing diseases among the Nigerian populace is presented. The system provided two modalities: speech and DTMF to interact with it. The system was tested and the two modalities compared for interacting with the system. Twenty one participants were involved in the evaluation and they were undergraduates of Covenant University, Ota, Nigeria and University of Agriculture, Abeokuta, also in Nigeria. Medical doctors, five in number from Victory Medical Centre, Eleyele, Ibadan, Nigeria and from State Hospital, Jericho, also in Ibadan were among the twenty one that evaluated the system. Professionals in teaching and health maintenance organizations were also part of the participants. Nine of the participants were males while ten were females. Two did not specify their gender in the questionnaire filled after testing the system. Five subjects were in 31–40 age range, four within 21–30 while ten were within 15–20. Two did not specify their age range. The results show that each modality has its strengths and weaknesses as per the following dimensions: satisfaction, modality preference, naturalness, effectiveness, efficiency and entertainment. However, a system incorporating the two modalities was suggested for such systems so that the user can choose from either of the modalities for good user experience. The results imply that the system will be accepted once developed using the two modalities.

In [22], it is stressed that useable systems having small vocabulary could be developed for under-resourced languages with about 5-10 hours of transcribed speech in the targeted language and having recognition accuracies of 95% or higher of speaker independence. However, the work investigated if performance similar to the one recorded for small vocabulary systems could also be achieved for tasks that involve the usage of medium vocabulary using Woefzela and the tools developed in [23]. These tools run on Android-based mobile devices. The investigation employed five official languages in South Africa namely: Afrikaan, English, isiNdebele, isiZulu and Sesotho. The authors concluded that “recognition accuracies can be achieved on medium-vocabulary recognition tasks in low-resource languages using corpora collected with the new generation of smartphones-based data collection tools". This conclusion implies that since these tools are easier to use and most often not Internet-dependent and coupled with the low error rate, the prospects of building useable voice-based applications in under-resourced languages become enhanced to bridge the digital divide that exists between the developed and the developing nations.

In [24], it was observed that health information for self-care was present on the Internet in text form. However, because of the low literacy level and inadequacy of Internet access, only a small

percentage of the populace could access it to better their lot health-wise. The work developed a spoken dialogue system that provided health information about the following: malaria fever, lassa fever, yellow fever and typhoid fever. This system catered for the needs of those who were visually impaired, low literate people, computer illiterate and those with no Internet access. The system was tested for acceptability and user satisfaction among nineteen participants mainly students of Landmark University, Omu Aran, Nigeria. One of the subjects did not complete the questionnaire after testing the system. Nine participants were males while five were females. Four did not specify their gender. Ten of them were at most twenty years old while five were within the age range of 21–30. Three did not specify their age range. User satisfaction mean of 3.98, approximately 4 (recommended for a good usability study on 1-5 scale) was obtained and this suggested that the users were satisfied with the system. Also, 94% of the study participants approved of the use of the system in accessing health information. This implication of this result is that the system was accepted.

It was observed in [25] that transacting with banks in Nigeria has become electronic to the extent that customers do not have to be physically in the banking halls to obtain information about their accounts balances, make deposits, transfer money from one account to the other and make withdrawals. However, in spite of the benefits provided by speech-based interfaces in terms of being cost-effective, widespread availability and the great prospects of bridging the digital divide with them, none of the existing solutions was speech-based. The authors, therefore, developed a spoken dialogue system for carrying out banking transactions and compared the use of speech and dual-tone multi frequency (DTMF) key input for accessing the services provided by the system. The system provided the following services:

i. Logging into the system using the account number and using a pin as the password;
ii. Account balance enquiry;
iii. Money transfer from one account to the other;
iv. Generating a new pin;
v. Updating of pin;
vi. Call routing to a customer relation officer.

The evaluation of the system was carried out among fifty Nigerian undergraduates. None of the undergraduate participants was a novice in the use of computer software. 20% rated themselves as average users, 68% as good software users while 12% rated themselves as expert users. None of the subjects was not using digital devices for work enhancement. 2% used the desktop computer to enhance their work, 2% also used the tablet, 48% used laptop/notebook while 48% made use of cell phone/PDA. All the subjects owned a mobile phone/PDA. 56% had owned a mobile phone/PDA for more than two years and 44% for two years. 0% was not making calls at all and 0% made calls 1 to 2 times a week. 38% made calls 3 to 4 times weekly, 54% 5 to 6 times a week and 6% for more than 7 times weekly. 2% did not specify how often they made calls and all of the participants supported the use of mobile devices for banking. The results obtained from the evaluation show that the system's satisfaction ratings for DTMF (M=36.18) were higher than that of speech (M=34.70) significantly as revealed by a significant t-test. The implication is that the system will be accepted in Nigeria if deployed, but with the users preferring DTMF for interacting with it.

It was identified in [26] that speech-based flight schedule information systems were prevalent in the developed nations of the world, but none was existent in Nigeria. The work reported the development of a spoken dialogue system for obtaining information about flight schedules in Nigeria and evaluated the acceptability and user satisfaction of DTMF and speech for interacting with the system. This was done to determine if the system would be acceptable and if so, what modality would best be suited to interact with it among Nigerians. This would eventually assist developers and people with stakes in the aviation sector to make an informed decision about its development and deployment in the country. Fifty subjects participated in the testing and 2% of them were novice user of computer software, 10% were average users, 63% were good users while 25% were experts. 82% made use of laptop/notebook to enhance their work, 14% used the tablet and 4% made use of cell phone/PDA. 52% had owned PDA/cell phone for more than two years, 18% for two years, 3% for one year and 27% for six months. The results of the evaluation show that the majority of the users (58%) preferred to interact with the system using DTMF while 42% preferred speech. Also, 99% supported the real-life deployment of the system while just only 1% did not support its development in real life.

In [27], a multimodal university admission requirements search system accessible in both audio and text forms using personal computers and phones (including feature phones for audio access only) was developed to provide access to candidates who do not own PCs and smartphones because of economic reasons. The user satisfaction and acceptability of the spoken dialogue component of the system was evaluated among twenty two subjects resident in Lagos, Nigeria. Lagos is home to all tribes in Nigeria because of its buoyant economy. 59% of the subjects were males while 18% were females. 23% of them did not specify their gender. 4% were novice computer software users, 41% were average users, 32% good users while 23% were experts. Those that used tablets for their work enhancement were 27%, those that used desktop were 18%, 9% used cell phone/PDA and 46% used notebook/laptop. 18% of the respondents had used mobile phone/PDA for a year, 5% for a year and a half, 9% for two years and 68% for beyond two years. Those that made calls one to two times a week were 4%. for three to four times weekly, 14%, from five to six times, 9% and for more than seven times, 73%. The results obtained show that 96% of them were in support of deploying the system in real life. This implies that the system was accepted by the majority. The mean value obtained for the user satisfaction obtained through a questionnaire scaled 1-5 was 3.6, approximately 4 suggested for a good usability study. This result shows that the users were satisfied with the system.

## 4. DISCUSSION

A total of twenty works as regarding spoken dialogue systems in the developing countries were reviewed. Out of the twenty, nine are in the domain of healthcare, three in the domain of agriculture, one in banking, one in aviation, one in secretarial work, three in multiple domains, one in the domain of recognition accuracy and the last in the domain of education.

Out of the twenty, different kinds of evaluation like acceptability of the systems, comparisons of DTMF and speech input modalities, usefulness, user satisfaction, suitability, effectiveness, possibility of integrating intelligent components into spoken dialogue systems and recognition accuracy were carried out in fifteen, all of whose results were positive as to the dimensions measured. This means that spoken dialogue systems are capable of being used to bridge the digital divide. This is even more so considering the following comments from some of the works: "voice can be a suitable medium for online communities in the rural developing world" [15], "…automated telephone services can be used as an effective means to communicate and disseminate localized information and enhance government service delivery in rural communities" [2], "The increasing use of the VoiKiosk system for the different purposes leads us to believe that a voice-based mechanism for local content creation is a very powerful interaction modality to provide information and communication technologies in rural areas"[12] and "telephony services could, in fact, be easily used by semi and low literacy users, and that a spoken dialogue system in local language can be a powerful health education tool" [13]. Other encouraging comments like users glossing over few errors of recognition, users willing to pay to obtain services provided by spoken dialogue systems and the showing of excitement [8] all point to the fact that the system is capable of being used to bridge the digital divide. In addition, considering the fact that the systems considered were tested among low literate and illiterate people, the blind, the visually impaired, computer illiterates, underserved people and in rural communities where digital divide is most prominent (because of the low literacy level there and low or non-availability of other ICT facilities except mobile phones), it can be concluded and inferenced that this technology can indeed be used to bridge the digital divide.

From another perspective, spoken dialogue systems have been of dominant usage in healthcare from the works considered (nine out of twenty) and given the fact that the density of physician per 1000 of the population in the developing countries is low, the technology has been leveraged on, to augment healthcare provision. A lesson learnt from this is that spoken dialogue systems can be exploited to augment services not easily available in the developing world especially in the native languages of the people.

## 5. CONCLUSION

In this work, it has been hypothesized that spoken dialogue systems are capable of being used to bridge the digital divide existent in the developing nations of the world. Consequently, a review of the different ways through which the

system has been used in the said regions of the world to carry out tasks and services was made. The different countries that have applied the technology have also been identified. From the results of the evaluations of the systems in different sectors and from the reports that emanated from the users of the systems and considering the characteristics, spread and statistics of the participants of the different evaluations carried out, it is quite clear that the claim is valid and that spoken dialogue systems can of surety be used to bridge the observed digital divide and provide services that are not readily available in these parts of the world. It is suggested based on the results and the reports obtained that other areas could also be explored for the application of this technology. Also, the vast majority of the remaining developing countries can explore the usage of this technology. This will indeed engender more benefits to the regions and the people under study. Also, this will also pave the way for more research activities that are peculiar to the developing world to be conducted in this area.

## COMPETING INTERESTS

Author has declared that no competing interests exist.

## REFERENCES


1. ITU. Statistics; 2019.
   (Accessed 22 May 2019)
   Available:https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx
2. Grover AS, Barnard E. The Lwazi community communication service: Design and piloting of a voice-based information service. In Proc. WWW2011. Hyderabad, India; 2011.
   Available:https://doi.org/10.1145/1963192.1963357
3. Oyelami MO, Uwadia CO, Omoregbe NI. Prospects of voice-enabled healthcare system in the developing nations. In Proc. of the First International Conference on Mobile Computing, Wireless Communication, E-Health, M-Health & Telemedicine (MWEMTeM '08), LAUTECH, Ogbomoso, Nigeria; 2008.
4. Gerosa M, Giuliani D, Narayanan S, Potamianos A. A review of ASR technologies for children's speech. In Proc. 2nd Workshop on Child, Computer and Interaction, ACM, Cambridge, Massachusetts; 2009.
   Available:https://doi.org/10.1145/1640377.1640384
5. El Ayadi M, Kamel MS, Karray F. Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition. 2011;44(3):572–587.
   Available:https://doi.org/10.1016/j.patcog.2010.09.020
6. López-Cózar R, Callejas Z, Griol D, Quesada JF. Review of spoken dialogue systems. Loquens. 2015:1(2): e012.
   Available:https://doi.org/10.3989/loquens.2014.012
7. Chen H, Liu X, Yin D, Tang J. A Survey on Dialogue Systems: Recent Advances and New Frontiers; 2017.
   (Accessed 23 February 2019)
   Available: http://arxiv.org/abs/1711.01731
8. Plauché M, Nallasamy U, Pal J, Wooters, C, Ramachandran D. Speech recognition for illiterate access to information and technology. In Proc. International Conference on Information and Communication Technology and Development, ICTD2006, Berkeley, California, USA; 2006.
   Available:https://doi.org/10.1109/ICTD.2006.301842
9. Nasfors P. Efficient voice information services for developing countries. Master's thesis, Department of Information Technology, Upsalla: Upsalla University; 2007.
10. Sherwani J, Ali N, Mirza S, Fatma A, Memon Y, Karim M, et al. Healthline: Speech-based access to health information by low-literate users. In Proc. IEEE/ACM Int'l Conference on Information and Communication Technologies and Development, Banglore, India; 2007.
11. Kotkar P, Thies W, Amarasinghe S. An audio wiki for publishing user-generated content in the developing world. In Proc. HCI for Community and International Development Workshop at CHI 2008, Florence Italy; 2008.
    Available:http://groups.csail.mit.edu/commit/papers/08/kotkar-hci08.pdf
    (Accessed 15 March 2019)
12. Agarwal SK, Kumar A, Nanavati AA, Rajput N. Content creation and dissemination by-and-for users in rural areas. In Proc. IEEE/ACM Int'l Conference on Information and Communication Technologies and Development; 2009.

13. Grover AS, Plauche M, Barnard E, Kuun C. HIV health information access using spoken dialogue systems: Touchtone vs. speech. In Proc. 3rd International conference on Information and communication technologies and development, Doha, Qatar; 2009.
14. Oyelami MO, Uwadia CO, Akinwale AT. A Framework for the Design of Speech-Enabled Self-Care E-Health Systems. In Proc. 9th International Conference on Electronic Business (ICEB 2009), Macau; 2009.
15. Patel N, Chittamuru D, Jain A, Dave P, Parikh T. Avaaj otalo: a field study of an interactive voice forum for small farmers in rural India. In Proc. of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, Georgia, USA; 2010.
16. Oyelami OM, Uwadia CO, Akinwale AT. An Intelligent Framework for Usable Speech-Enabled E-Health System. In: Ajeeli ATA, Al-Bastaki YAL, editors. Handbook of Research on E-Services in the Public Sector: E-Government Strategies and Advancements. Hershey, New York: IGI Global; 2011.
17. Oyelami MO, Uwadia CO, Akinwale AT. Integration of expert system technology into VoiceXML-Based Systems. Journal of Computing. 2011;3(3):28–35.
18. Kleynhans N, Molapo, R, De Wet F. Acoustic model optimisation for a call routing system. In Proc. PRASA 2012, CSIR International Convention Centre, Pretoria; 2012.
19. Oyelami MO. Dialogue systems as a means of detecting counterfeited pharmaceuticals/drugs by consumers. In Proc.7th International Conference on ICT Applications, National Defence College, Abuja, Nigeria; 2012.
20. Oyelami MO. User satisfaction and acceptability of dialogue systems for detecting counterfeit drugs. In Proc. International Conference on Human-Computer Interaction, Amsterdam, The Netherlands; 2013.
21. Oyelami OM, Akinyemi IO, Uwadia CO, Akinwale AT. An experimental comparison of speech and DTMF for VoiceXML-Based Expert Systems. African Journal of Computing & ICT. 2013;6(1):87–94.
22. Barnard E, Davel MH, van Heerden CJ, Wet FDe, Badenhorst J. The NCHLT Speech Corpus of the South African languages. In Proc. SLTU 2014, St. Petersburg, Russia; 2014.
23. Hughes T, Nakajima K, Ha L, Moreno P, LeBeau M. Building transcribed speech corpora quickly and cheaply for many languages. In Proc. Interspeech, Makuhari, Japan; 2010.
24. Oyelami OM. An empirical study of user satisfaction with a health dialogue system designed for the Nigerian Low-Literate, Computer-Illiterate and Visually Impaired. In Amy Neustein, editor. Text Mining of Web-based Medical Content. Berlin, Boston: De Gruyter Inc; 2014.
25. Oyelami OM, Akinyemi IO. The acceptability and comparative study of DTMF and speech for banking transactions in Nigeria. British Journal of Mathematics & Computer Science. 2014;4(9):1290–1300. Available:https://doi.org/10.9734/bjmcs/2014/8741
26. Oyelami OM. Spoken dialogue systems for flight schedule information in Nigeria: Acceptability evaluation and user satisfaction comparison of speech and DTMF. In Proc. of 8th International Conference on Speech Technology and Human-Computer Dialogue, Bucharest, Romania; 2015.
27. Oyelami OM, Awolola A. Development, user satisfaction and acceptability evaluation of a multimodal admission requirements search system. International Conference on Inventive Computing Systems and Applications, Pattaya, Thailand; 2018.



*Peer-review history:*
*The peer review history for this paper can be accessed here:*
*https://sdiarticle4.com/review-history/52404*