Vol 6. No. 1, March 2013
African Journal of Computing & ICT

www.ajocict.net

An Experimental Comparison of Speech and DTMF for VoiceXML-Based Expert Systems

Oyelami Olufemi Moses & Akinyemi Ibidapo Olawole4
Department of Computer and Information Sciences,
Covenant University, Ota, Nigeria
femi.oyelami@covenantuniversity.edu.ng, ioakinyemi@covenantuniversity.edu.ng

Uwadia Charles Onuwa
Department of Computer Science,
University of Lagos, Akoka, Lagos, Nigeria
couwadia@yahoo.com

Akinwale Adio Taofeek3
Department of Computer Science,
University of Agriculture, Abeokuta, Nigeria
aatakiwale@yahoo.com

ABSTRACT
Comparisons of DTMF and speech modalities for interacting with diverse dialogue systems for different tasks, among different user populations have led to different design recommendations for different user populations. This paper reports the results of the experimental comparison of these input modalities in a new context of VoiceXML-based diseases diagnosis expert system among a new user population - Nigerians. The results show that DTMF was more satisfying than speech for system satisfaction. Modality wise, speech was more satisfying than DTMF. Speech was also more natural than DTMF. DTMF was preferred by the majority and was more effective and efficient than speech. For diseases diagnosis expert health dialogue systems in Nigeria, DTMF is recommended for effectiveness and efficiency. It is also recommended for satisfaction. Speech is recommended for modality satisfaction while both modalities are recommended for entertainment purpose. Speech is advocated for modality naturalness. However, a platform that incorporates the two modalities will provide the benefits of the two, and allow the users varieties of choices that best suit their needs.

Keywords- DTMF; speech, expert system; VoiceXML; diagnosis.

African Journal of Computing & ICT Reference Format
Oyelami Olufemi Mose, Uwadia Charles Onuwa, Akinwale Adio Taofeek & Akinyemi Ibidapo Olawole (2013). An Experimental Comparison of Speech and DTMF for VoiceXML-Based Expert Systems. Afr J. of Comp & ICTs. Vol 6, No. 1. Pp 87-94

1. INTRODUCTION
VoiceXML as an extensible mark-up language was designed for human-computer audio dialogues that feature synthesized speech, digitized audio, recognition of spoken input, Dual Tone Multi Frequency (DTMF) key input, recording of spoken input, telephony and mixed initiative conversations [1,2]. The language does not, however, support sophisticated dialogue systems that are artificial intelligence-based. In addition, users of VoiceXML-based systems have the options of either interacting with them using speech or DTMF. Various research work have compared the two input modalities in different contexts and among different user populations [3, 4, 5] and have come out with different results and design recommendations. There is a need to study the characteristics of different user populations as relating to dialogue systems. One of the reasons for that is that, different user groups speak with different accents and exhibit different characteristics.

It has also been found out that user preference for either DTMF or speech depends on the nature of task performed [4]. In a previous research work, the authors have used component orientation approach to develop VoiceXML-based expert system - Health Dialogue Expert System (HDES) that acts as a physician to diagnose some selected types of fever rampant in Nigeria - malaria, lassa, yellow and typhoid with a view to augment healthcare [6] as the ratio of physicians to citizens in this part of the world is abysmally low [7]. HDES incorporates Java expert system shell (Jess) into VoiceXML-based system as the expert system component using component orientation. In developing HDES, five physicians were interviewed as to the symptoms and how to diagnose the fever types handled by HDES. Literature was also consulted. The same physicians tested HDES and approved its reliability.

As at 2007, Nigeria had an estimated population of 140 million, but the ratio of doctors to the population was about 1 to 3,333 [8]. The density of physicians per 10, 000 population was 4, that of dentists was .5, 16 for nurses and midwives, and 1 for pharmaceutical personnel. According to the World Health Organization (WHO) 2010 report, in Nigeria, current life expectancy is 49 years. The most common causes of death in rank order are as follows: malaria, diarrhea, other diseases, pneumonia, prematurity, birth asphyxia, neonatal sepsis, HIV/Aids, congenital abnormalities and injuries [7]. As can be seen, malaria is actually the number one killer disease in Nigeria. There is therefore a need to fight the disease and other types of fever with all armories. Currently, the government is encouraging the populace through radio broadcasting to see their doctors for proper diagnosis instead of assuming they have malaria and self-medicate.

However, the use of mobile phone in the country has been on the rise. As at October 2012, the number of connected lines stood at 138, 029, 637 for mobile (GSM and CDMA) and fixed wired and wireless while the total number of active lines stood at 109, 499,882 [9] out of the population of 140, 431, 790 [10]. This provides an avenue to reach a vast majority of the populace for healthcare which has hitherto been inadequate. This article reports the results of the experiments conducted to compare the use of DTMF and speech as input modalities for interacting with HDES in order for callers to diagnose their types of fever.

2. A REVIEW OF RELATED WORK

Aditi et al. designed a speech dialogue system for the provision of health information to caregivers of HIV positive children in Botswana, Southern Africa. They compared touchtone (DTMF) and speech input modalities in the context of low literacy users and a health information service [3]. They hypothesized the following:

a. DTMF is likely to be more acceptable in the developing world because general numeracy is less common; and
b. (b) DTMF-based systems are much easier to develop than natural language systems and are consequently more attainable in the resource-constrained environments that typically characterize the developing world.

In developing the system, doctors, nurses and nutritionists were interviewed as to the contents provided by the system. Further requirements were also got from printed materials. "All the health contents and prompts were translated using a registered translation service into the dialect of Setswana spoken in Botswana". The study was a within-subjects comparison, and after the completion of each modality's trials by the 33 caregivers involved, a post questionnaire was administered verbally to each of them.

The results obtained showed that there were no significant differences between task completion rates though speech performed slightly better that DTMF. 59% of the users preferred DTMF while 19% preferred speech. This is in contrast to studies carried out in the developed world where users preferred speech. The results, however, correlate with the fact that simple DTMF is generally not viewed as favourable. The subjects that preferred speech did not as in the studies in the developed world comment that speech is more entertaining or enjoyable but rather on the utility of speech as being more accessible for older people or faster" [3].

Neil et al.[4] compared speech and dialed input voice user interfaces for farmers in Gujarat, India. They "designed Avaaj Otalo ("voice-based community forum"), a Gujarati language application allowing farmers to access agricultural information over the phone. Navigational nodes in the application were limited to two or three options, and only directive-style prompts were used in order to avoid command ambiguity". They "partnered with Development Support Center (DSC), an NGO in Ahmedabad, Gujarat, to conduct a joint needs-finding exercise, based on which the three system's features were identified and implemented. Both isolated word speech and DTMF versions of Avaaj Otalo were implemented. Prompts were recorded in a professional studio by one of the DSC radio program's popular female voice personalities. Barge-in input was disallowed for both speech and DTMF.

Avaaj Otalo was built and deployed using IBM Research India's WWTW platform. Gujarati commands were converted to lexicons using the American English phoneme set for the speech recognition. The system performed with a recognition accuracy of 94% and was tested with 45 participants recruited from ten districts throughout rural Gujarat and they were all farmers. 87% of the subjects reported to have never used a PC". The experiment was, however, not a within-subjects experiment design. "Input modality (speech vs. DTMF) was randomly assigned to each user, but was anonymously corrected to maintain balance across age, education and gender. Testing sessions were led by a DSC staff member who had experience communicating with the target user group.

African Journal of Computing & ICT

www.ajocict.net

Each participant completed three tasks with Avaaj Otalo corresponding to its three features (listening to announcements, listening to archived radio program recordings, and posting questions), ordered by increasing difficulty. 38 participants were tested in a quiet office using a landline phone, with only the DSC staffer and two researchers as observers while the remaining 7 were women and were tested in their homes using a mobile phone because of travelling difficulty. The prototype application was instrumented to log task completion, errors and call duration to measure performance. During the test, two researchers noted points of difficulty, facial expressions, and comments made during the call. A post-test questionnaire with Likert scales was administered to measure user satisfaction, ease of use and learnability"[4].

The results showed that "task completion rate with DTMF was significantly higher than with speech (74% vs. 61%; $p< 0.05$). There was no significant difference in user satisfaction. In both groups, over 80% of users reported that they found it easy to access information from the system. Over 75% of both groups said they would "definitely" use such an application if it was made available. The users were asked to state the difficulty level of a particular task as either "difficult" or "very difficult" on a five-point Likert scale. Across all tasks, the percentage of such responses was 49% for speech and 30% for DTMF ($p < 0.05$). For difficulty faced when providing input to the system, 81% of DTMF users answered "no" or "definitely no", compared to 38% for speech users ($p< 0.01$)" [4].

Kwan and Jennifer also reported "an experiment that critically tested user's preference for an input modality (speech vs. DTMF) in a phone-based message retrieval system using a fully functioning natural language system. The experiment was a within-subjects design. All participants used both the speech-only and the DTMF-only modalities and they all completed identical tasks for both input modalities. Sixteen participants (8 women and 8 men) were recruited from the IBM T. J.Watson Research Center located in the state of New York. The participants were in a wide range of ages (from early 20s to over 60). None of the participants had any prior experience using mobile assistant, and they were all native speakers of English. The participants took the study one at a time in a usability lab. Upon arrival at the lab, the participants were seated and given a booklet with an instruction on the first page. The instruction page described the purpose of the study and had a list of required tasks. It also had a phone number they should call to reach the system, the name of the test account they should use, and a password. In the case of the DTMF condition, the instruction page had the function mapping and a sample interaction.

In the case of the speech-only condition, this page had sample phrases that could be spoken to the system. The participants used a regular wired office telephone on the table. The speech output from the system was played through a set of speakers as well as the handset so that the system's output could be captured via a video camera in the usability lab. In the study, all three elements of usability of telephony applications according to the European Telecommunications Standards Institute's guidelines on usability in telephony applications were examined. The system effectiveness was measured by calculating a success rate for each user task. The amount of time to finish each task was used as a proxy measure for system efficiency. An evaluation of user satisfaction was carried out through a series of questions asked immediately following the use of each modality" [5]. Their results indicated that "(a) DTMF was more effective and efficient for linear tasks, whereas speech was better for nonlinear tasks; (b) speech was preferred to DTMF by a majority of users; (c) speech was judged as being more satisfying, more entertaining and easier to use than DTMF; and (d) user preference for a particular modality was better predicted by user performance in nonlinear tasks rather than linear ones" [5].

3. METHODOLOGY

HDES was developed using VoiceObjects Desktop for Eclipse 9 and Voxeo Prophecy 8 was used as the implementation platform. XLite softphone was used to call HDES for testing. All these tools allow for developing and testing dialogue systems on a PC without having to deploy them on any telecoms service provider's network. HDES could not be hosted as there is currently no single voice service provider that provides such service in Nigeria.

3.1. Dialogue Flow of HDES

Once HDES is called, the system responds by welcoming the caller and introduces the caller to the services it provides. The system then lists all the symptoms the caller can choose from, as well as their corresponding keys (Table 1). The caller has the option of making the system repeat all the symptoms as many times as possible until he is familiar with them. After this, the caller can say three initial symptoms or press the corresponding keys on the telephone keypad of the symptoms he has noticed he has. The caller would then need to confirm the symptoms supplied before diagnosis is carried out. If all the symptoms are correct, the system goes on to diagnose, otherwise, the caller can re-supply or correct any of the initial three symptoms supplied. If the symptoms supplied are sufficient to diagnose the caller's disease, the system responds by telling the caller the disease being suffered from. Otherwise, the system would need to ask a series of questions in order for it to determine the kind of ailment the caller is suffering from.

African Journal of Computing & ICT

www.ajocict.net

Table 1: Symptoms and Their Equivalent DTMF Keys

Key	Symptom
1	Vomiting
2	Headache
3	Fever
4	Chills
5	High fever
6	Muscle pain
7	Tiredness/Fatigue
8	Sore throat
9	Diarrhea
10	Loss of appetite
11	Stomach pain
12	Rash
13	Bloody stool
14	Nose bleed
15	Low consciousness
16	Low heartbeat
17	Constipation
18	Bleeding from anus
19	Mouth bleeding
20	Joint pain
21	General discomfort
22	Abdominal rash
23	Chest rash
24	tongue discoloration
25	Excessive thirst
26	Black stool
27	Internal heat
28	Back pain
29	Dry cough
30	Cough
31	Tiredness
32	Weight loss
33	Chest pain
34	Blood pressure changes
35	Hypertension
36	Swollen neck
37	Swollen face
38	Swollen eyes
39	Ringing in ears
40	Blood in urine
41	Change in heartbeat
42	Shaking
43	Nausea
44	Dizziness
45	General pain
46	Sweating
47	Fall in temperature
48	Paleness
49	Shortness of breath
50	Bitter taste
51	Low urine
52	Stress
53	Anxiety
54	Skin redness
55	Red eyes
56	High heartbeat
57	Low backache
58	Red tongue
59	Bad breath odour
60	Bone pain
61	Blood vomiting

African Journal of Computing & ICT

www.ajocict.net

3.2. Experimental Design
A paper list containing the symptoms as well as their corresponding DTMF keys was given to each participant. This was necessary because of memory loss associated with dialogue systems with many menu options and correlates with Kwan and Jennifer's approach. Barge-in was therefore allowed once a caller knew what key or symptom to supply, so that there would not be any need to wait till all the symptoms have been listed before responding. A within-subjects design was adopted in the evaluation as in the case of Kwan and Jennifer. All the subjects tested both DTMF and speech separately. They completed the same task for each of the input modalities. Once a subject completed a task using a particular input modality, he/she would fill the questionnaire for that input modality. Once that was done, he/she would use the other input modality, after which the questionnaire on that input modality would be filled.

3.3. Participants
Twenty one subjects participated in the evaluation. This number is more than 16 used by Kwan and Jennifer. The participants consisted of undergraduates of Covenant University, Ota and University of Agriculture, Abeokuta, all in Nigeria. Others are doctors from Victory Medical Centre, Eleyele, Ibadan and State Hospital, Jericho, Ibadan, all in Oyo State, Nigeria. The doctors were five in number. The subjects also included professionals in teaching and health maintenance organizations. Nine of the subjects are males while ten are females. Two did not indicate their gender. Five of the subjects are in the age range of 31- 40, four in the range of 21 – 30 and ten in the range of 15 – 20. Two did not specify their age range.

3.4. Apparatus
3.4.1. DTMF Modality
The subjects were instructed about the functionality provided by the system and how they could interact with it. They were informed that their mode of interaction with the system in this modality was through the pressing of the keys on the telephone keypad. The keys representing the symptoms were to be pressed, and all other interactions with the system should be through the keys instead of speech. For other interaction with HDES, they were asked to listen to the system for guide at each stage of the dialogue. Like the speech modality condition, the system output was synthesized lady's speech.

3.4.2. Speech Modality
In the speech modality, the subjects were instructed to use only speech for interacting with HDES. The list containing the symptoms served as a guide for them as regards the symptoms. They were, however, also asked to listen to the system for guide for other conversations with it.

3.5. Procedure
The subjects tested the system one at a time in an office. The tools used for the implementation allow the testing of the system in a standalone mode. They were taught how to initiate a call to HDES using XLite soft phone. The subjects were informed that the purpose of the study was to examine usability issues related to the use of a phone-based health dialogue expert system for diagnosing diseases. Once each subject arrived at the venue where the test was carried out, they were given a detailed instruction about what the system does and how it could be interacted with. A paper list containing the symptoms and their equivalent DTMF keys was also given to them. The subjects were told that their task was to interact with HDES using the soft phone in order to diagnose their ailments. They were asked to pick specific symptoms representing the ones they assumed to have and say or press them on the telephone keypad. They were informed that they would need to fill a questionnaire each for each input modality. The speech output from the system was played through a set of external speakers connected to an Acer laptop running Windows Vista on which HDES resided.

3.6. Measures
As reported by Kwan and Jennifer, according to ETSI, usability in telephony applications is defined as the level of effectiveness, efficiency, and satisfaction with which a specific user achieves goals in a particular environment. In this work, the three dimensions of usability were examined as in the case of Kwan and Jennifer. System effectiveness was measured by calculating the error rate for each of the input modalities. This was done by programming HDES to log errors: **Misrecognition** and **No Match/No Input**. Call duration was used as a measure of efficiency. This was extracted from a system log file after each user finished with an input modality. User satisfaction was evaluated through the items in the questionnaires administered after the use of each input modality.

In measuring user satisfaction, a questionnaire used in a similar study by Kwan and Jennifer [5] and user satisfaction survey used in Marilyn et al. [11] were adopted. "The measures used in the questionnaire have both face and content validities. In terms of face validity, all measures were constructed by experts who have more than 10 years of experience in usability tests of mobile and speech user interface (SUI) applications. In terms of content validity, the measures cover all dimensions of usability in telephony applications as defined by ETSI" [5]. The only modification done to the questionnaire by Kwan and Jennifer was the changing of certain adjectives to their simpler synonyms so as to aid the subjects' understanding. Two types of questions were administered – questions on modality evaluation and questions on system evaluation as adapted from Kwan and Jennifer [5].

"The first set of questions asked the subjects to evaluate the system regardless of their evaluation of the interaction modality". These were extracted from Marilyn et al. [12, 13]. For the second set of questions, the subjects were to respond to questions that evaluated their interaction modality with the system. It was speculated as done by Kwan and Jennifer that the evaluation of the interaction modality and the evaluation of the system could be different.

After carrying out the task, the subjects were immediately "asked to evaluate the system and the interaction modality by indicating how well certain adjectives described the modality and the system on Likert scales of 1 (strongly disagree) to 5 (strongly agree)". The adjectives – **comfortable**, **uncomfortable** (instead of **exhausting** used by Kwan and Jennifer), **frustrating** and **satisfying** were used to measure modality satisfaction index. Four adjectives – **boring**, **cool**, **fun**, and **entertaining** – were used for a modality entertainment index. The adjectives used for measuring naturalness index were- **natural**, **unnatural** (instead of **artificial** used by Kwan and Jennifer), **boring** (instead of **repetitive** used by Kwan and Jennifer and **nervous** (instead of **strained** used by Kwan and Jennifer [5]. Lastly, the subjects were required to provide answers to the question **Which of speech and DTMF do you prefer to interact with the system and why?**

4. RESULTS

As earlier mentioned, system effectiveness was measured by calculating the error rate for the task carried out using each of the input modalities. This was done by instrumenting HDES to log errors: **Misrecognition** and **No Match/No Input**. Call duration was used as a measure of efficiency. User satisfaction was measured through questionnaires. Subjects' DTMF satisfaction ratings were compared with their speech satisfaction ratings using a repeated measures t-test with modality as the repeated factor. "The within-subjects t-test, used for comparisons with a continuous dependent variable, is also known as the paired samples t-test (the SPSS term), the dependent samples t-test, correlated samples t-test, or the repeated measures t-test. It is used when the same person is in the study twice"[14]. t-test is used when the sample size is less than 30[15].

4.1. User Satisfaction

4.1.1. System Evaluation

The null hypothesis is that the mean difference between DTMF and speech user satisfaction is zero. The alternative hypothesis is that there is a mean difference between the two input modalities.

H_0: $\mu_d = 0$
H_1: $\mu_d \neq 0$

Satisfaction ratings for DTMF were significantly higher (M=37.1) than for speech (M=33) as indicated by a significant t-test, $t(18) = -3.45$, $t_{crit} = 2.1$, $p < .05$. Since -3.45 is less than -2.1, the null hypothesis is rejected. The finding thus indicates that DTMF was more satisfying than speech.

4.1.2. Modality Evaluation

For modality satisfaction, the null hypothesis is that the mean difference between DTMF and speech modalities satisfaction is zero. The alternative hypothesis is that there is a mean difference between the two input modalities.

H_0: $\mu_d = 0$
H_1: $\mu_d \neq 0$

The subjects evaluated their interaction with speech as more satisfying (M=12.42) than DTMF (M=12.21) as indicated by a t-test, $t(18) = -.43$, $t_{crit} = 2.1$, $p < .05$, though the two systems did not differ significantly. Since -.43 in not less than -2.1, the alternative hypothesis is rejected. Thus, the two are modalities are equally satisfying.
For modality entertainment, the null hypothesis is that the mean difference between DTMF and speech modalities entertainment is zero. The alternative hypothesis is that there is a mean difference between the two input modalities.

H_0: $\mu_d = 0$
H_1: $\mu_d \neq 0$

An interesting result was got for modality entertainment; both speech and DTMF were rated equal (M=13.32) as indicated by a t-test, $t(18) = 0$, $t_{crit} = 2.1$, $p < .05$. Since 0 is not less than -2.1, reject the alternative hypothesis. Hence, the result shows that both speech and DTMF are equally entertaining. Finally, for modality naturalness, the null hypothesis is that the mean difference between DTMF and speech modalities naturalness is zero. The alternative hypothesis is that there is a mean difference between the two input modalities.

H_0: $\mu_d = 0$
H_1: $\mu_d \neq 0$

The subjects evaluated their interaction with speech as more natural (M=l0.63) than DTMF (M=10.12) as indicated by a t-test, $t(18) = -.66$, $t_{crit} = 2.1$, $p < .05$. The difference is not, however marginally significant. Since -.66 is not less than -2.1, the alternative hypothesis is rejected. Thus, DTMF and speech are equally natural.

African Journal of Computing & ICT

www.ajocict.net

4.1.3. Modality Preference

In response to the question **Which of speech and DTMF do you prefer to interact with the system and why?** 62 % of the subjects chose DTMF modality, whereas 38 % chose speech ($p < .05$). 67% of those who chose DTMF gave recognition inaccuracy of speech modality as being the reason why they preferred DTMF. Other gave reasons like convenience of using DTMF, higher speed of task completion with DTMF and that it is easier. For those that chose speech, the reasons given are "**It is inviting, interesting and exciting though not accurate**", "**It is easy to understand**", "**It reduces stress**", "**For clarity purpose**", "**It is more natural**", "**It is easier to use**" and "**It is natural though not accurate**".

4.1.4. Effectiveness and Efficiency

For effectiveness and efficiency, only data from subjects who were able to successfully interact with HDES using the two input modes were used. For modality effectiveness, the null hypothesis is that the mean difference between DTMF and speech modalities effectiveness is zero. The alternative hypothesis is that there is a mean difference between the two input modalities.

H_0: $\mu_d = 0$
H_1: $\mu_d \neq 0$

DTMF was more effective than speech as the errors generated by speech were significantly higher (M=8) than those generated by DTMF (M=2.5) as indicated by a significant t-test, $t(7)=-2.26$, $t_{crit}=2.36$, $p<.05$. The null hypothesis is therefore rejected.

For modality efficiency, the null hypothesis is that the mean difference between DTMF and speech modalities efficiencies is zero. The alternative hypothesis is that there is a mean difference between the two input modalities.

H_0: $\mu_d = 0$
H_1: $\mu_d \neq 0$

DTMF was also more efficient (M=1.14) than speech (M=1.97) as the completion time is less than that of speech as indicated by a significant t-test, $t(9)=-4.29$, $t_{crit}=2.26$, $p<.05$. The null hypothesis is therefore rejected.

4.1.5. Experiences with Mobile and Computing Devices

5% of the subjects rated themselves as **novice** in the use of computer software. 16% rated themselves as **expert**. 68% rated their skill as being **good**, while 11% rated their skill as being **average**.

In the use of devices for enhancing work, 5% replied that they do not use technology to enhance their work. 79% use laptop/notebook to enhance their work, while 16% use personal digital assistant (PDA)/cell phone. All the subjects reported that they owned a mobile phone or a PDA. 63% have owned a mobile phone/a PDA for more than 2 years, 16% for 6 months, 16% for 2 years and 5% for 1 year. In response to the number of times they make or receive calls a week, 5% do not make and receive calls. 11% make or receive calls 1-2 times a week, 16% 3-4 times a week, 5% 5-6 times a week and 63% more than 7 times a week. Lastly, 84% support the use of mobile devices for healthcare, while 16% do not. The 16% that do not support gave reasons such as speech misrecognition problem and that the use of mobile devices will not be as fast as with an encounter with a medical practitioner.

5. CONCLUSION

A within-subjects comparison between DTMF and speech for interacting with dialogue expert system for diagnosing diseases among Nigerian users has been presented. The results presented in section 4 have implications for developers and it is clear that if effectiveness and efficiency are the focus, DTMF is recommended. If satisfaction is the watchword for the overall system, DTMF is recommended. Speech is recommended for modality satisfaction while both modalities are recommended for entertainment purpose. Speech is advocated for modality naturalness, however, a well-designed diseases diagnosis dialogue system should strategically provide a platform that incorporates the two input modalities to reap the benefits of the two, and to allow the users varieties of choices that best suit their needs.

African Journal of Computing & ICT

www.ajocict.net

REFERENCES

[1]. R. José, "Web services and speech-based applications around VoiceXM," *Journal of Networks* ,vol. 2, February 2007, pp. 27-35.

[2]. W3C (2001). Voice Extensible Markup Language (VoiceXML) Version 2.0 http://www.w3.org/TR/2004/REC-voicexml20-20040316/. Accessed May 25, 2008.

[3]. S. G. Aditi et al., "HIV Health Information Access Using Spoken Dialogue Systems: Touchtone vs Speech,". Proc. 3rd International Conference on Information and Communication Technologies and Development, Doha, Qatar, 2009, pp. 95 – 107.

[4]. P. Neil et al., (2009). "A Comparative Study of Speech and Dialed Input Voice Interfaces in Rural India," Proc. CHI 2009, Boston, MA, USA, April 3 - 9, 2009.

[5]. M. L. Kwan and L. Jennifer,."Speech Versus Touch: A Comparative Study of the Use of Speech and DTMF Keypad for Navigation," International Journal of Human–Computer Interaction, vol. 19, 2005, pp. 343–360.

[6]. O. M. Olufemi, U. C. Onuwa and A. A. Taofeek "Integration of Expert System Technology Into VoiceXML-Based Systems," Journal of Computing, vol. 3, 2011, pp. 2151-9617.

[7].World Health Organization (2010). World Health Statistics 2010.

[8]. F. O. Ogunrin, O. Ogunrin and A. Akerele, "Motivating Nigerian Doctors for Improved Healthcare Delivery," International Journal of Health Care Quality Assurance, vol. 20, 2007, pp. 290-306.

[9]Nigerian Communication Commission (2012). Monthly Subscriber Data http://www.ncc.gov.ng/index.php?option=com_content&view=article&id=125&Itemid=73. Accessed December 20, 2012.

[10]. National Population Commission (2012). http://www.population.gov.ng/.Accessed December 20, 2012.

[11] M. Walker, D. Litman and C. Kamm, "Evaluating spoken language systems," Proc. American Voice Input/Output Society (AVIOS). May 1999.

[12]. W. Marilyn et al. "PARADISE: A framework for evaluating spoken dialogue agents," Proc. 35th Annual Meeting of the Association of Computational Linguistics (ACL 97), July 1997, pp. 271-280.

[13]. M. A. Walker, C. K. Kamm and D. Litman,. "Towards Developing General Models of Usability With PARADISE," Natural Language Engineering, vol. 1, 1998, pp 1-13.

[14]. Newsom USP 534 Data Analysis Spring 2009 http://www.upa.pdx.edu/IOA/newsom/ho_t-test%20within.pdf. Accessed March 31, 2011.

[15]. G. B. Alan, Elementary Statistics: A Step by Step Approach, New York: Mc GrawHill, 2004.