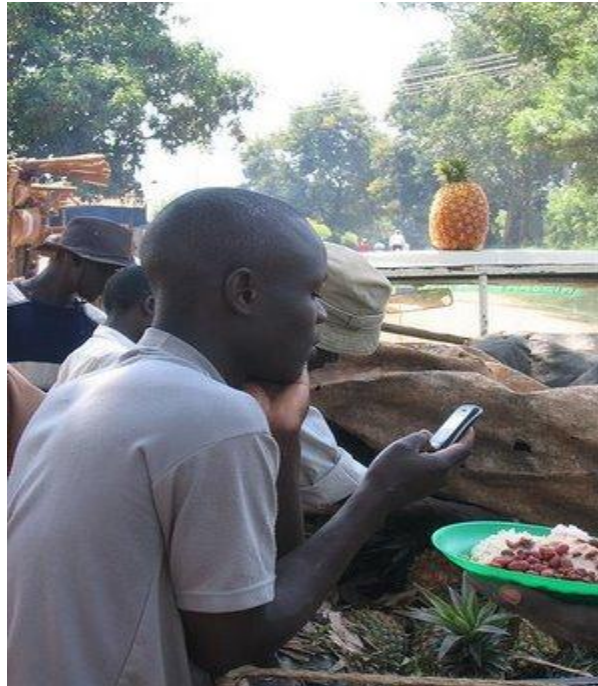# Voice User Interface Design for m-Event Organizer

Re-greening Africa Project

Master Thesis
MSc Information Science
VU University Amsterdam
Faculty of Science

Student: Deepak Chhetri
Student ID: 2205409
Main supervisor: Dr. V. De Boer, Vrije Universirty
Co-Supervisor: Prof. Hans Akkermans, Vrije University

Centre for International Cooperation (CIS)
vrije Universiteit amsterdam

SAHEL ECO

SEVENTH FRAMEWORK PROGRAMME

WORLD WIDE WEB FOUNDATION

The Network Institute

Abstract

Information and Communications Technologies provide a lot of options for information access by literate users but there is much less opportunities for information access by illiterate users. Implementing effective Voice User Interface in rural community of developing world requires addressing a number of important challenges. The recent market penetration of mobile phone and its impact in developing regions has gained lots of interest from Voice User Interface Designers.

In this research, I will present literature reviews on Voice User Interface Design with reference to low literate users in developing regions of the world. I will also describe the knowledge and experience I gained from each review and its implication to my research. Based on these literature reviews, I will present a framework that can be used to check the most important aspects required in designing an appropriate VUI for low literate users. The research ends with successful paper prototype VUI for m-Event Organizer. This research will be helpful to other designers during the process of designing a VUI prototype for low literate users.

Contents

1   Introduction

Information and communication technology (ICT) can play a major role in achieving the Millenium Development Goals (MDGs) by improving the access to information and by enabling communications in the developing world [18]. Despite the world being a similar and a globalized, there is still a huge digital-divide between the developed countries and the developing countries. Information and Communication Technologies for Development (ICT4D) have proved to offer enormous potential to bridge this huge digital-divide[18]. ICT has been already a significant part towards developing world's economy. An important factor in the choice of technology for an ICT4D project is that radio and telephone have the strong benefit of being language and literacy independent, given that the literacy rate of many developing countries is in the order of 50% [20].

"There is a widespread agreement that ICT services, especially mobile ones, have the potential to play a major role in furthering social and rural development in developing economies."[4] In the developing world, very few have a computer and access to the internet. "The market penetration and rural community adoption of basic mobile telephony and services has been extremely rapid in recent years."[4] This adoption happens primarily because basic mobile is easy to use, affordable, and suitable for illiterate users. The greatest impact of mobile communications on access to communication services – in other words, increasing the number of people who are in reach of a telephone connection of any kind – can be seen in developing countries[25]. Taking the advantages, the organizations with a global development have turned their attentions to the mobile phone as a potential platform for delivering development services, in sectors spanning education, finance, health, agriculture, and governance [14]. Therefore, there is huge potential for mobile ICT services.

Interestingly, low literate people in the developing world cannot read text but would be very much comfortable with voice. Voice can be a powerful tool for use in human computer interaction because it is the fundamental means of human communication [5]. So the Voice User Interface (VUI) makes human interaction with computers possible through a voice/speech platform in order to initiate an automated service or process [21]. With the rapid growth of wireless communication and advance in technology, VUI has become more familiar and need of good VUI application has increased rapidly. In recent years, many researchers in the area of human computer interaction (HCI) have attempted to enhance the effectiveness and efficiency with which work and other activities are performed using voice based interfaces[5]. Various research on voice user interface has resulted in a number of systems in several domains such as desktop, telephony and smartphone. But there has been very little research done on VUI for illiterate users in the developing world.Therefore, a well design voice user interface is an important factor in the success of voice applications in the developing world.

---

[4]VOICES, http://mvoices.eu/, reviewed 28th november 2012

1.1 Context to m-Event Organizer

The Web alliance for Re-greening in Africa (W4RA[1]) is a joint initiative from the Web Foundation[2] (Geneva, Switzerland), VU University (Amsterdam, Netherlands) and Sahel Eco[3] (Bamako, Mali). "W4RA envisions a world where all people are empowered by the Web: One Open Web, Everyone regardless of language, ability, location, gender, age or income will be able to communicate and collaborate, create valued content, and access the information that they need in order to improve their lives and communities"[1] . W4RA creates a communication platform where farmers can create, access, and share knowledge about market information via simple mobile and community radio. While by the initiative of W4RA1in West Africa, farmers are encouraged to use farming techniques more in greener ways.

"One of the projects under W4RA initiative is VOICES. The VOICES Project (Voice based Community-centric mobile service)"[4] is funded by the European commission and it is associated with 12 partner organizations. "The project started 1st January, 2011. The VOICES project is aimed at supporting remote communities by creating spoken web–content"[4]. In the VOICES Project, there are two use cases, which are m-Agro and m-Event organizer.

In the m-Event use case, Sahel Eco[3] organize events among farmers in Tominian area (Mali[5]) to create awareness about regreening. Figure 1 below shows the overview network configuration for the the m-Event organizer and it represents present scenario.

[1]W4RA, http://www.w4ra.org/en/about, reviewed 1st November 2012
[2]Web Foundation, http://www.webfoundation.org, reviewed 5th November 2012
[3]Sahel Eco, http://www.saheleco.net/, reviewed 12th november 2012
[4]VOICES, http://mvoices.eu/, reviewed 28th november 2012
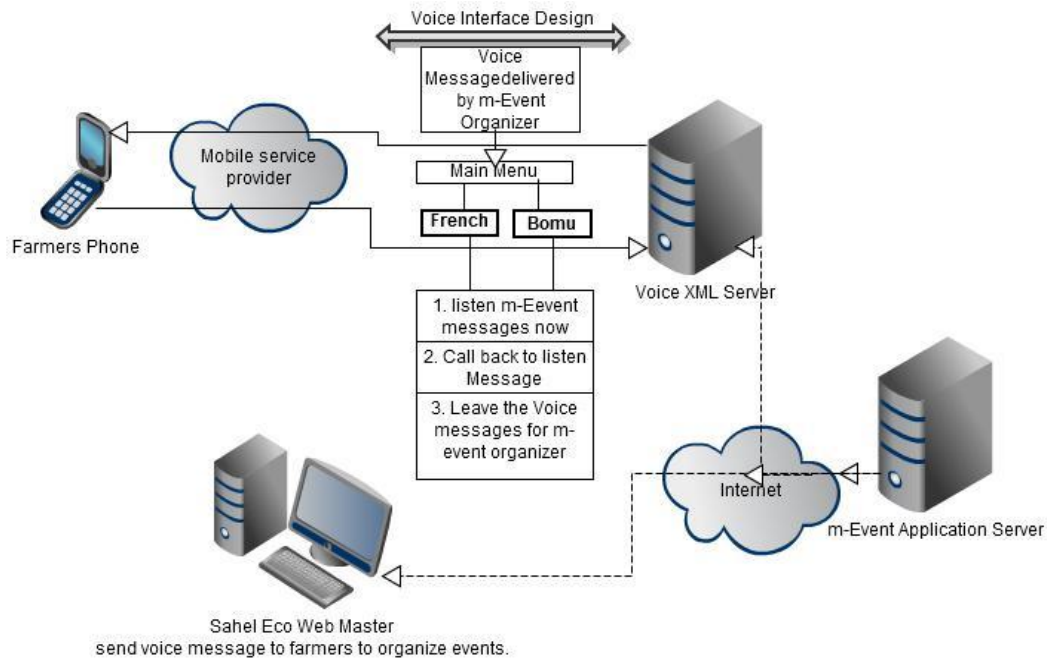[5]Mali,https://maps.google.com/maps?hl=en&q=tominian+area,+mali&ie=UTF-8, Reviewed 3rd November,2012

Figure 1: Overview of m-Event Organizer.

Figure1shows the m-Event scenario. It is showing a scenario whereby a Sahel Eco representative has a computer with internet connection in his office and access to the m-Event application. Then he/she presses the m-Event button and enter the farmer name and phone number. There is an option in the application to choose the number of languages that Sahel Eco representative wants to record. Then he/she records the voice message and send to registered farmers mobile phone. In the current implemented m-Event application, the voice messages are sent in only two languages, either in French or Bomu. The farmers receive the voice messages in a similar way as they receive the call and they have options to choose languages as per their preference. The voice message that the farmer receives has three options to choose from. The options are "will attend the meeting, press 1", "Are you not attending the meeting, press 2" and "Leave your message, press 3". The m-Event systems also logs whether the messages are delivered or a farmers' phone is switched off. The farmers who are not registered in m-Event systems can call the Sahel Eco number and listen to voice messages posted on the system. In this case the m-Event system will record the new farmers number and his/her voice message with a name. Finally, the m-Event organizer organize the m-event.

1.2 Aim of Research

The aim of this research is to do a literature review on voice user interface design. Based on the literature reviews, I will create a framework to check the most important aspects that are required in designing an appropriate voice interface for developing countries. In this research I will also propose a prototype voice interface design for m-Event organizer, which will be more simple and convenient for the farmers in Tominian area, Mali.

2   Problem Statement

2.1 Research Challenges in Design of Voice User Interface

Designing a Voice User Interface is different than designing other interfaces because of its transient and invisible nature. Designing speech applications, particularly speech-only ones, is substantially different from designing graphical applications [10, 22]. It is mentioned in [6] that VUI is a comparatively new experience for developers. There are very few good VUI guidelines available today mainly because the VUI design is so new and requires ample usability testing [6]. The most research paper on voice user interface doesn't present about designing voice interface, rather it focuses on VUI implementation. This makes it hard for new researchers to come up with new voice user interface designs for low literate users in the developing world.

In order to implement a successful voice application in the developing world, a number of major challenges faced by these communities will have to be addressed. One such challenge is that 40% of the population in the least developed countries is not literate [16] and even the literate are novice users of ICTs. It is a known fact that low literate users face great difficulty in accessing information and it is a difficult task to implement ICTs to improve access to information. Access to important information such as in the field of health care, can often mean the difference between life and death [7]. There were many researches done on voice user interface for low literate users and found that users avoid complex functions but use phones for synchronous voice communication only [15]. The question this brings to any researcher is that how we can design mobile phone interfaces in a way that can benefit low literate users.

3   Literature Review

I have done a literature review of 9 research papers. Most of these research papers present the VUI with reference to the developing world. From those papers, I will try to understand how design prototypes were presented and implemented. I will review on aspects from each paper, which are essential when designing the voice user interface for low literate users. I will use these aspects in my framework, which will guide me to design appropriate VUI for the m-Event organizer. The literature review will also consist of each research paper aim and research outcome. If paper presents the pilot studies or usability test, then I will review on its finding or issues in VUI. These finding and issues are very significant for the new VUI designer and it will lead to better design prototype.

3.1 Designing Interactive Voice Response (IVR) Interfaces for low literacy users [1].

Dual Tone Multi Frequency (DTMF) refers to push-button telephones for tone dialing. Interactive Voice Response (IVR) is a technology that allows a computer to interact with humans through the use of voice and DTMF tones input via telephone keypad or speech recognitions [26]. In [1], the authors address the issues of how to design IVR interfaces for the developing world. The typical design issues are recognition errors, auditory and cognitive challenges, user control of the interaction and ambiguities. The authors propose a dialog design model shown in Figure 2, which should be considered valuable for designing optimal IVR interfaces for low literacy users in the developing world. The design model consists of three parts, which are Get input, Error-recovery, and Play result.
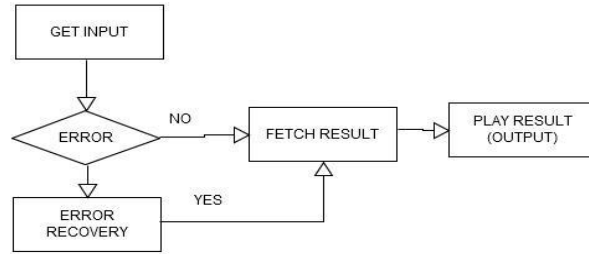
Figure 2: Dialog design model for IVR interfaces [1].

In [1], the authors looked into mobile usage culture, Error recovery and Multiple language support. I used these aspects in my framework. In the Get input section of the design model for IVR interface, [1] looks at what sort of input mode should be considered in IVR interfaces for non-literate users in the developing countries. The authors agree that Dual Tone Multi Frequency (DTMF) is preferred with a natural language interface.

The following point is mentioned in favor of DTMF over natural language interfaces.

1. Firstly, There is limited funding to build natural language with regards to linguistic knowledge, technical expertise, computing infrastructures in developing world.
2. Secondly, mobile usage culture; where there is a multiple user with the single ownership model. For an example like Kiosks in Africa, where people resell mobile minutes. In this case there will be speech recognition issues from multiple users, noisy environments and interruptions. From this point, it's clear that designers should keep in mind about mobile usage culture when designing the VUI for low literate users. Therefore, I use mobile usage culture as one of aspect in developing my framework.
3. Thirdly, the issue of privacy. For example, an Aids patient would want to get information via Healthline Services. He/she prefers to use DTMF rather than natural languages, because he/she can get the information by pressing phone keypad. Therefore, people around the phone user don't have idea, which information he/she in listening.

The writers postulate that error recovery in IVR interfaces for the developing world should be based on clear instructions such as explicit confirmation were needed [1]. I used the error recovery aspects in the framework that I developed. It is mentioned that the unfamiliarity of low literacy users with IVR systems can lead to more errors but this may be only in the initial stage. The availability of robust error management is an important element of audio prompt design for low literacy users. It is very essential to provide clear error recovery feedback for low literacy users else it will be very difficult for users to learn and will result in an annoying experience. This could discourage the use of IVR and users will have a negative impression towards the technology. The important aspect of error recovery is to provide the users with help at each point in the dialogue. All error feedback messages should be clear and context-dependent so that users have unambiguous instructions on exactly what they can say/press at any point [1].

At the end of the paper, it mentions how information or results are presented to low literacy users. It is very important in developing countries that voice output should be in local languages in IVR interfaces. The attention should be paid to translation, persona and user interface metaphors in prompt design.

8

3.2 Experiences Designing a Voice Interface for Rural India [2]

In [2], the writers describe their experience while designing, implementing and evaluating a mobile voice interface for farmers in Gujarat, India. "The authors outline their design process from initial contextual inquiry to a formal user evaluation, and use this discussion to motivate research guidelines for others designing voice interfaces in developing regions" [2]. The authors highlight three guidelines for researchers designing the voice interface for rural communities in the developing countries. I used these three guidelines as aspects from this paper in my design. These aspects are discussed below.

The authors worked together with Development Support Centre (DSC) in Ahmedabad, Gujarat and designed Avaaj Otalo. Avvaj Otalo is the Gujarati language application which allows farmers to receive agriculture information via phone. The writers did perform the several rounds of design iteration, testing with farmers and other aspects before implementing a prototype version of Avaaj Otalo. After implementing the system, it was found that dialed input performance is far better than the speech input while considering completion rate. The writers mentioned the following guidelines, which are necessary for researchers when designing the voice interface for developing world.

1.   Enhance Existing Systems

There is a long list of past computing projects for the developing world that have not sustained beyond pilot deployments [2]. Usually, the information in the developing world is shared through mobile phone, television, radio and human interaction. The authors found that community based radio is most effective to share information such as agriculture topics. Therefore, it is mentioned in this paper that they integrated their system with the existing radio program, so that the farmers can contribute questions and comments on the conversation being aired on radio.

2.   Involving the users.

On this point the writers focus on farmers participation in the early stage of Avaaj Otalo. The authors included the farmers in the design process and came to know the insight of farmers interaction ability with voice interface. In the implementation phase, it is benefiting from developing simple features early and testing with farmers where they are already comfortable with technology. The authors learnt that farmers, Local NGO's and DSC are helpful in validating author's design ideas and advising alternative. They also discoverd that involvement of farmers in the design stage will lead to less error occurrence.

3.   Evaluate Design Choices Empirically (Dialogue Flow)

In this guideline, the design space dimensions for voice interface such as input modality, prompt style, call flow structure and use of sound effects are mentioned [2]. The writers stated that it is necessary to compare quantitative and qualitative alternatives within those dimensions.

While designing the Avaaj Otalo, the writers experienced many practical lessons. "The Writers experience designing Avaaj Otalo revealed some practical lessons that may extend beyond voice interfaces to the design of other information systems for the developing world" [2] . The paper concludes by explaining the writers experiences in designing the Avaaj Otalo, which is the voice interface for farmers in Gujarat, India to share and access the information via phone.

3.3 Voice User Interface Design Patterns [3]

[3] presented the voice user interface design patterns and the challenges with the audio. From this paper, I use the Persona, Multiple language selections, Menu Hierarchy (navigation) and Dialogue flow aspects in the framework for designing the voice interface in the developing world. The writers stated that designing the voice interface is difficult because of their transient and invisible nature.

The language is an important area in the VUI design process. Therefore, the authors stated that the selection and development of an appropriate dialog strategy is the first step designers must take when building a VUI interface [3].The writers emphasize the challenges with audio when designing the voice user interface. The design of efficient speech interfaces is considered by many to be more of an art than an engineering science [3].

The paper presented the following Dialog strategy.

1. Dialogs have to to be efficient and short
2. Dialogs have to be clear and structured
3. Novices have to be guided. They need to know what kind of information to provide
4. Experienced users know what to say and need a fast way to enter the data

Menu Hierarchy plays a key role in keeping the users' concentration on the system. According to the writers, it is important to guide users towards their interest and the data should be provided to users via the selection processes. "The options to be presented to the user are interrelated in a hierarchy, or can be made to appear that way"[3]. It is mentioned that the hierarchical structures are easy to understand, but the limitation of short-term-memory, authors like Miller [12] said that any human being can be capable of remembering 7+/-2 information chunks only. The writers conducted a user study and found that more than 6 options are rejected by users in most of the cases. "Therefore, the length and the depth of such a menu tree should be as flat as possible, and should also not exceed the magical number 7"[3].

The paper presented more about Persona, which is a look and feel for voice based applications. The writer stated the following forces with regards to Persona [3].

1. The interests of the target groups are different.
2. Interaction with voice based application should provide an underlying coherent and stable personality.
3. The system's responses must be consistent. New characters distract the user.
4. Dialogs should appear more natural.
5. Users know that they are talking to a machine.
6. The character must match the user's mental image of the application.

The voice application is affected by a number of objective factors such as speed, noise and attention but language remains as the biggest barrier [3]. It is very obvious that users of VUI always want to access the information in their native language. The paper presents consequences such as allowing users to easily recognize an option, Enable users with little foreign language skills access the system and Listening options in many different languages can be confusing [3].

In this paper, the writer aimed to provide the design patterns, which helps the voice interface designer to understand the nature of problems and how to solve those issues. This enables the designers of voice user interface to provide a successful voice interface. The Writers concluded that Persona and Structured Audio helps to design the system response.

3.4 Efficient Voice Information Services for Developing Countries [4]

The author aims to understand people's need for information via an easily accessible device like mobile phones in the developing world. The author studies impact of those services in developing regions, where most of people are illiterate and use simple mobile phones. The author demonstrates voice based services in Kenya and it is called Banana Information Service. They built a pilot Kiswahili system in local languages.

The paper presents much less with regards to design part of voice services. The Author writes more about ICT in developing world and the implementation of Banana Information Service in Kenya. The framework developed in this thesis uses the Navigation and Error aspects.

In order to provide an easy navigation in Banana Information Service, all the information was well structured into different categories. The navigation overview is shown in Figure 3. While designing the voice user interface for Banana Information Service, proper navigations play a vital role. The author stated that they provided the users with the option of pressing 0 to repeat the last information that had been played to them, and to press * to go back to the previous menu [4]. These options are essential for novice users and it will help the users when they get lost in navigation.
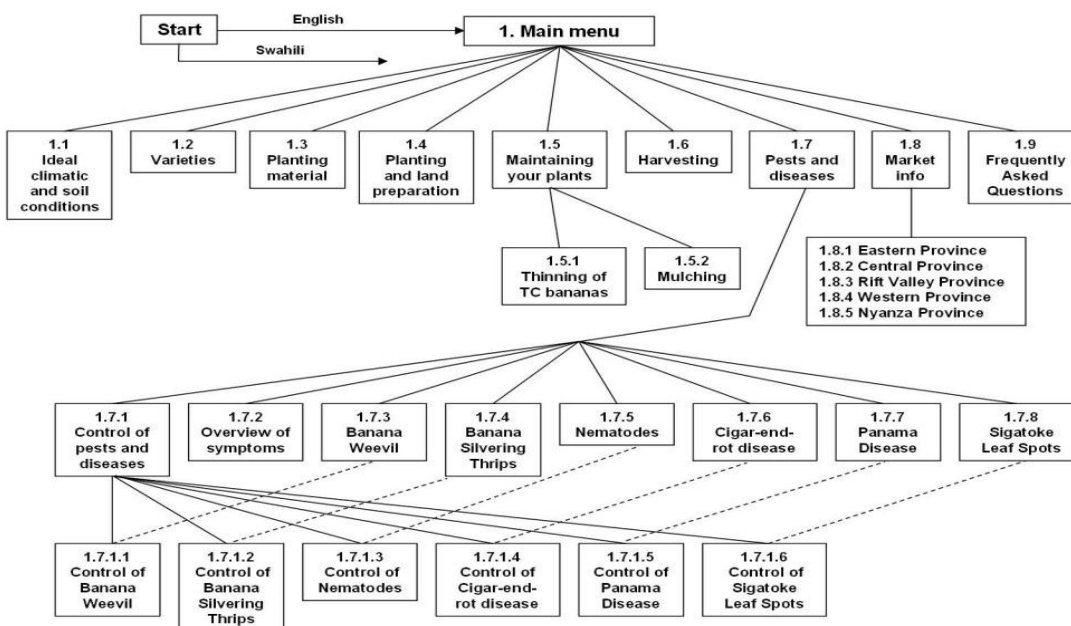

Figure 3: Overview of navigation in the Banana information service [4]

The author performed the usability tests and chose rural Kenyans as participants for the test. They contacted the banana farming community in Kibirigwi, Kirinyaga and selected ten participants to test performance measurements. To find an error component, the author tested the accuracy of using the

system. They calculate the error rate by dividing the average number of wrong answers by the total number of questions and found it to be 36 % [4]. The major reasons for the errors were following.

1. The participants could not recognize the spoken output produced by the text-to-speech system.
2. The information was played only one time in each section. The participants could have understood more, if they were able to listen twice.
3. Most of the participants chose English rather than Swahili and voice was with a British accent.
4. The participants stated that the voice was speaking too fast.

These errors that were found provide vital information for any new designers and guide to design better VUI. The authors review error feedback on design prototype and further development of voice is in progress. The paper concludes with successful implementation of Banana Information System, which provides information about growing bananas through an interactive voice response system.

3.5 Designing SpeechActs: Issues in Speech User Interfaces [5]

In [5], the authors do not present with reference to the low literate user in the developing parts of the world but the issues mentioned and discussed in this article are applicable to any VUI design. Implementing a usable conversational VUI involves overcoming substantial obstacles [5]. The authors investigated the principles and challenges of VUI design and pointed out the limitation of the current technology.

The authors did the survey and a field study, which is very important for any preliminary VUI design. After the field study, the working prototype design was finalized and they conducted the usability study, which resulted in several iterative redesign prototype.

After the pilot usability study, the following challenges which were essential in VUI design were discovered. The framework presented in this paper, was also adopted in my research framework.

1. Prosody: It was found that many of human utterance cannot be applied in VUI. For example, speech like hmm or huh , cannot reproduce exactly the same in other languages. Therefore, any new designer in VUI has to look for a new idea or technology to resolve this issue. But pilot usability studies in this paper showed that most of the speech output was understandable. While, the participants did complain that the voice sounded a bit electronic[5].

2. Information Flow (Dialogue Flow): "Just as one way of organizing information can be clear on the screen and confusing when spoken, so it is with information flow"[5]. This paper presents a number of experiments on information flow. For example, in the case with the exit dialogue, the authors treated any invalid input as an implicit "no"[5].

3. Recognition Errors: It is mentioned in this paper that recognition errors are frustrating because of the recognizers inconsistent responses. "It is common for the user to say something once and have it recognized, then say it again and have it misrecognized" [5]. There are many circumstances which can lead to recognition errors and it is divided into 3 categories. These are;

   a. Rejection Error: This error occurs when the recognizer has no hypothesis about what the user said [5].

b. Substitution Error: Involves the recognizer mistaking the user's utterance for a different legal utterance, as when "send a message" is interpreted as "seventh message" [5].

c. Insertion Error: in this case, the recognizer interprets noise as a legal utterance [5]. For example, if people are talking in another room or someone is in telephone.

4. Speed and Persistence (Voice Command): As we know that speech is easy to produce but it is hard to consume for a normal human being. The authors stated that written information can be absorbed faster than verbal information. So,when designing the VUI for low literate users, voice command plays a vital role in the implementation of the design.

The challenges like information organization and Ambiguous silence are also mentioned but the framework designed in this thesis does not take these challenges into consideration. This paper presents the broad view on the challenges faced by the VUI designers. It is very important to do field a study, which is helpful in designing preliminary VUI and consequently usability study for design prototype.

3.6 Voice User Interface Design for a Telephone Application Using VoiceXML [6].

VoiceXML (VXML) is the W3C's standard XML format for specifying interactive voice dialogues between a human and a computer [13]. "It is designed for creating audio dialogs that feature synthesized speech, digitized audio, recognition of spoken and DTMF (Dual Tone Multi-Frequency) key input, recording of spoken input, telephony, and mixed initiative conversations" [6]. The voice interface design is a very important part in developing the VoiceXML application. "In this paper, the VUI for a VoiceXML 'Cinema Service' telephone application is designed and a number of experiments are undertaken to help the design of the VUI" [6] .

The authors performed series of experiments with the users in order to understand user experiences. The authors then improved the initial VUI and users helped to provide valuable information for the future VUI development. The authors stated that Speech recognition technology is still imperfect, and users encounter failure for various reasons [6]. Therefore, it is very necessary to include error recovery during the design process. In addition, the author's emphasized on navigation as it makes it easier for users to use services, if the Voice User Interface Design is simple, short and clear dialogs.
The writers performed the following experiments and the framework presented in this thesis used some of the aspects.

1. Navigation
   Its aim was to find the user experience in navigating through the application. It was found that the users in the beginning are a bit slow and at a later stage, users become more comfortable with the application.
2. Memory and Age Group
   In this experiment, it is used to test how long and how many voice prompts can be spoken before the user get confused or forget instructions. It was found that there is no correlation between age groups and most menus are easy to remember.
3. TTS (Text-to-speech) setting and Pre-recorded Prompts
   Its aim is to find the user preferred application setting. For example, male or female voice, text to speech or pre-recorded human voice prompt. It was found, most preferred female text to speech.

The writers concluded that there should be two different user interfaces so that the user with different level of experience can choose the different interface. The experiment also proved that multiple, long and descriptive prompts make navigation difficult [6].

13

The example of VoiceXML.

```
<?xml version="1.0" ?>
<vxml version="2.0">
<form id="greeting"><block>Welcome to the Cinema service.
<goto next="#main"/></block>
</form>
<menu accept="approximate" dtmf = "false" id = "main">
<prompt>You can book movie tickets, find session times, ...
What would you like to do?</prompt>
</menu>
<form id="book">
<block> Booking movie tickets. <goto next = "book.vxml" />
</block>
</form>
...
</vxml>
```

3.7 Speech vs. Touch-tone: Telephony Interfaces for Information Access by Low Literate [7].

In this paper, the authors developed spoken language interface prototypes aimed at low literate users, and tested them with community health workers in Pakistan [7]. The authors chose the healthcare domain, which is very important in terms of information access in developing regions. The paper also presents the comparison between Speech and Touch–tone.

A number of pilot user studies were conducted with community health workers in Pakistan, where most of the workers are illiterate. This paper presents following finding after pilot studies and these were adopted in the framework presented in this thesis.

1. The information presented orally needs to be short.
2. Low literate users were less likely to have ever used a phone.
   The pilot user studies show that low literate user were very reluctant to pick of the phone and need to ask for permission. The phone usage culture differs in rural regions of the developing world and it has implication on the design of voice interface for low literate population. Therefore, mobile usage culture is vital consideration when designing VUI with reference to low literate users.
3. The national language is not always optimal
   The authors stated that initially their local NGO and partners told them that Urdu is a most common language spoken in Pakistan. So Urdu was the acceptable choice of the system. But pilot studies show that Urdu was not understood by the 50% of participants in Umarkpot and 66% those in Dadu [7]. Therefore, language selection support is an important element in the design phase of voice user interface.
4. The regional language is also not always optimal.
   The authors conducted pilot studies in Dadu District and most people there speak Sindhi. But studies show that many of the participants belong to migrant communities from Balochistan and they speak Balochi without any written text. Thus, it is very necessary to involve the target population in pilot studies before coming up with a voice interface design.
5. Training and working with local facilitators is essential

The authors have conducted a number of pilot studies and performed various testing on spoken language interface prototypes, which leads to various design decisions. From this paper, I take these pilot studies and testing, which are essential to researchers designing the voice user interface focused on low literate users in the developing world. The authors conclude by stating that for some target population, a well-

designed speech interface can significantly outperform a touch-tone interface for both low-literate and higher literate users [7].

3.8 Designing an Architecture for Delivering Mobile Information Services to the Rural Developing World [8].

The authors mentioned that there are many limitations in current mobile software platforms and it makes it difficult to implement in the developing world. The authors stated that web-based mobile applications are hard to use, do not take advantage of the mobile phone's media capabilities and require an online connection [8]. In order to overcome those types of limitations, the authors came up with CAM- a new framework for developing and deploying mobile computing applications in the rural developing world [8]. CAM is very much amended for low literate users because of its minimal support, paper-based navigation, a simple script programing model and offline interaction.

The paper does not provide related information to be used for the framework in my research except Navigation, which will guide the users in the right directions. But there is more information which can be helpful for the voice user interface designers. The information such as environmental and user challenges can be useful in my research because it gives an idea about the rural community in developing world which helps to design better voice interface. The Authors also highlight environmental challenges such as Intermittent Power, Intermittent Connectivity, Long Travel Times, Variable Population Density and Lack of secure storage. User challenges like Limited education, Underemployment and Limited Disposable Income are also stated. The above mentioned environmental and users' challenges were not explained in this thesis as these were not included in the framework designed for this thesis but only used as guidelines when designing the VUI in the developing world.

These challenges can be used as guidelines and provide an in-depth virtual view of rural community in the developing world. It is important for designers to keep those challenges in mind while designing voice interface for developing regions.

3.9 Speech Recognition for Illiterate Access to Information and Technology [9]

"Speech recognition has often been suggested as a key to universal access, but success stories of speech-driven interfaces for illiterate end users are few and far between"[9]. There are many reasons for a little success in speech recognition such as multilingualism, cultural barrier, dialectal variation, etc. The author presents a novel design for a spoken dialog system (SDS) that is scalable, modifiable, and designed to operate with limited linguistic resources [9]. This design for spoken dialog system addresses the social and economic challenges of the developing world.

The authors presented about user interface design for illiterate users but authors wrote more about the speech recognition. The authors conducted a number of case studies in rural Tamil Nadu, India and explore the possibility of speech driven user interface for low literate users. In the initial pilot studies, the author developed the Tamil Market, an SDS that provides weather, market prices for crops and agricultural information to low literate users who navigate by uttering one of 30 possible words in Tamil [9]. The paper presents the following findings after the user studies were done on Tamil Market.

1. The average time for interaction was 3 minutes and participants valued accuracy of information within average time. Therefore, when designing voice user interface for low literate users, dialogue has to be short and structured.
2. It was found that illiterate users are more hesitant to participate in the pilot studies. But writers suggest for training and familiarization of systems to low literate users. Therefore, it is very

essential for user involvement in pilot studies, when designing the voice interface in developing regions.

3. The writers injected 2% recognition error and it was not recognizable by participants because of their excitement towards the systems. The authors have tested the error in different methods and users found it interesting. This method can be implemented in my design for voice user interface but we need to find exactly what type of error is common.

The authors aim to find whether SDS like Tamil Market can be used by low literate users without formal training of the system. It was found that low literate users would likely perform better with more training and helps such as graphic output [9]. The paper presented very important findings after the pilot studies and these finding are also related to my framework. I use dialogue flow, involving with the users and error aspects in my framework.

4 Assessing Voice interface Design for developing countries

4.1 Framework for assessment Voice Interface Design for developing countries

Based on the literature review in the previous section, I created a framework that can be used to check the important aspects that are required in creating a voice interface design for developing countries. From the literatures, I adopted some important aspects that were raised in the literatures. These aspects are deemed to be necessary in designing the voice user interface for developing world. The aspects are generalized into three categories, which are Design Considerations, IVR Design and Implementation Aspects. Each categorization of aspects represents the different phase in the design part of VUI for low literate users.

1. Design Considerations: These are considerations that designers have to consider when designing the VUI for low literate users. The consideration such as involving the users during the design, referring to similar existing similar projecst and mobile usage culture. In this category, I put the emphasis on design considerations, which are vital for any new VUI designers. For example, If a designer X wants to design a VUI for low literate community in Kenya, the designer will have a comprehensive idea if he/she refers to similar project that has been implemented in such a context.

2. IVR: There are many issues with the design of the IVR interface with reference to developing world. In this category, I will present the aspects which are essential when designing the VUI. The enculturation of IVR interface in the developing world is different from the developed countries because of the influence of a different set of social-cultural, linguistics, and domestic challenges . "The enculturation of IVR interfaces suggests that such interfaces should be balanced against the existing social-cultural and domestic environment of the users, which may influence the usability of the applications in terms of its domain and nature of task" [1]. Therefore, the IVR design is an important category of VUI design. I have selected essential aspects from literatures and incorporated them in my framework.

3. Implementation Aspects: This refers to aspects which can lead to better design of VUI after the implementation preliminary prototype design for low literate users. This category also plays a vital role in design of voice user interface. For example, voice produce can mislead the users if it is not uniform. Also background sound can lead users to confusing state. Most of the literature reviews highlighted that implementation of a preliminary pilot user interface has a great effect on the final VUI design. Therefore, implementation aspects can lead to better VUI design with reference to low literate users.

The aspects, which are essential in designing the VUI for low literate users are presented below in table 1. The framework will help in identifying the most important aspects needed in designing the voice interface with reference to m-Event organizer. I elaborate on each aspect in section 4.2. The researchers in Table 1 referred to papers that I had reviewed and its column presents the respective reviewed paper references.

Aspect	Researchers
Design Considerations	
Enhance the existing system	2, 4
Involving with the user	1, 2,6,7,9
Mobile Usage culture	1, 7,8
IVR	
Different user interface for different user group.	6
Error Recovery	1,4,5, 6,9
Dialogue Flow: *Input modality* *Prompt Style and* *Use of sound Effect*	1, 2, 3, 5,6,9 *2* *2* *2*
Keep it Simple	2
Provide multiple language support	1,2,3,7
Implementation Aspects	
Navigation	3, 4,6, 8
Persona	1, 3
Prosody	5
Voice Commands: *Speed,orderly,brief,relevant and Persistence.*	5

Table 1: A framework consisting of aspects which are essential when designing VUI for low literate users.

4.2 Framework's aspects in details.

4.2.1 Aspects with reference to design consideration

1. *Enhance Existing System*: There is a long list of past computing projects in the developing world that have not sustained beyond pilot deployments [18]. The integration of existing systems in the design process would provide an advantage over a standalone application. There will be more positive effects on the new system such as saving time, users view towards system and many

other things . For example, the authors in [2] found that community based radio is most effective to share information such as agricultural products and they integrated their system with the existing the radio program. Therefore, I would recommend VUI designers to refer to the past similar project or existing systems that are similar to their project.

2. *Involving the users*: It is important to include the target population for pilot studies because most of these populations in developing countries are unfamiliar with the voice user interface. Involving them will help designers to get an insight of the systems end users. The authors in [2] state that their partner field agents were helpful in validating their design ideas and suggesting alternatives. The pilot studies always show the finding and it will lead better preliminary prototype design for voice interface. It will be a disastrous consequence if the findings were found after implementing a working prototype. I would suggest a new VUI designer to do a pilot studies and involve the target population in preliminary VUI implementations.

3. *Mobile Usage Culture*: In the developing world, mobile phone usage and its proprietorship is different compared to the developed world because of social economic disparities [2]. People share common mobile phones and pay to use the phone services. Therefore, the mobile usage culture has implications in the voice interface design. To illustrate with an example, suppose the system with speech commands will have challenges with speech recognitions because of the noisy environment and multiple users. This type of challenges and circumstance will lead to recognition error. Therefore, it is very important to consider this aspect, when designing VUI for low literate users.

4.2.2 Aspects with reference to IVR

1. *Different user interface for different user group*: In the paper [6], the authors conducted a series of experiments on voice user interface. It was found that two separate user interfaces should be available so that the users can choose the interface as per their experience. Only few past research on VUI with context to low literate user implemented this aspect but I would suggest to test it in the early implementation of the system. This would help low literate users in the beginning and as they get experience on the system, users can change to another interface, which have short navigations to get targeted information.

2. *Error Recovery*: The voice user interface is more error prone compared to the visual user interface. It is possible to eliminate the error with advances in technology but not totally. In the voice interface design, there should be clear error recovery feedback for low literacy in the developing world; if not, it will be a frustrating experience and users will have a negative attitude towards the technology. Help should also be provided at each point of dialogue. All error feedback messages should be clear and context-dependent so that users have unambiguous instructions on exactly what they can say/press at any point [4]. I would strongly recommend to VUI designers to have clear error recovery options so that the low literate users don't get lost or be lead to miscommunicate.

3. *Dialogue Flow*: The way information is organized can be clear on the screen and confusing when spoken [11]. The information flow has to be clear, be of high quality and well structured. Well-structured audio allows users to access information fast and easily. It will also help users to stay focused on the audio. In the VUI, the dialogue flow aspects such as input modality, prompt style, call flow structure and use of sound effects should be considered as important features by VUI designers. Therefore, the flow of dialogue plays a vital role in VUI and it should be considered as an important aspect.

4. *Keep It Simple:* The VUI design for low literate should be simple, which requires only basic knowledge of mobile phone. If the design is complicated, then users will definitely have a negative attitude towards the use of the applications. The author in [6] state that multiple, long

and descriptive prompts are difficult for users to remain concentrated on the information, eventually leading users to a frustrating experience. It is stated in [3] that menu hierarchy should not exceed the magical number 7. The authors in [9] did an experiment and found that the average time of interaction should not exceed 3 minutes else users can't concentrate on the information. Therefore, the VUI designer should always aim to come up with a simple design.

5. *Providing Multiple language support*: Language still remains the biggest barrier in the rural communities of developing regions. [7] states that national and regional language is not always optimal in language selections. Language plays a vital role in VUI, especially with reference to low literate users. The research in [2,3,7] shows that low literate users always prefer to access the information in their native language. Therefore, I would suggest doing a field study with the context spoken language before coming up with a design prototype. I would also like to recommend other designers to provide low literate users with multiple local language options in their designs and the selection of the preferred language should be easy as well.

4.2.3 Implementation Aspects

1. *Navigation*: It can be defined as a way to guide the user to the information of interest [3]. Navigation is crucial in keeping the users concentrated during the interaction. The proper navigation in voice interface design will lead to less error-prone for novice users. For example, if the low literate user gets lost within the interaction and if there is an option to go back to previous menu or listen to previous message, it will keep users on track [4]. The authors in [6] did an experiment on navigation and found that all the users are slow in the beginning but as users get comfortable with the application, their skill in navigation improve progressively. I would recommend other VUI designer to design simple and effective navigation.

2. *Persona*: It is the look and feel for voice based application [3]. Users of voice based applications build their own mental image of a personality or character that they infer from the application's voice and language; such mental image relates certain properties to the virtual dialog partner, where systems' responses should fall within a foreseeable range of possibilities [3]. The audience should be convinced with the application with regards to their feeling and impression. To get the user attracted to the application, it is very important to choose suitable nonverbal sounds and background music to increase the clarity of prompts and messages. The music and nonverbal sound should be consistent throughout the entire dialog. I recommend for persona to be considered as an important aspect in designing the effective VUI.

3. *Prosody*. This is the rhythm, stress, and intonation of the speech [23]. Prosody, or intonation, is an important element in conversations, yet many of the synthesizers available today do a poor job reproducing human-sounding intonation contours [5]. There is lack of prosodies and many participants don't understand the speech output and participants complained that voice sound electronic [5]. Prosody is an important aspect of speech output in VUI design and designers should come up with solutions to handle prosody in VUI applications.

4. *Voice Command*: Although speech is easy for humans to produce, it is much harder for us to consume [17]. The slowness or fastness of speech out play is a major factor in voice command. It is obvious that the written information can be absorbed faster than verbal information. The voice has to be persistent, so that it will not mislead the users. I would recommend to VUI designers to include only spoken speech that is relevant, brief, orderly, persistence and informative in their VUI interfaces.

5   Paper prototype voice interface design for m-Event organizer

In this section, I will describe the present m-Event VUI design. The present version of m-Event is very simple and it is very much convenient for users in developing regions. I will make a checklist of aspects against the present version of the m-Event organizer. The checklist will be presented in tabular format followed by an explanation. I will also present the improved version of the m-Event organizer in this section.

5.1 Present versions of m-Event VUI Design

Figure 4 shows the present version of m-Event systems. Voice messages are delivered in three languages, which are English, French and Bambara. The farmers receive the voice messages in their preferred languages with a welcome message. There were three options in voice message, 1-Attend the meeting, 2-not attending the meeting and 3-Do not know. Then the farmers have choice to record the message that they want to leave for m-Event organizer. The farmers also have choices to hear their recorded message for correction. Finally, m-Event design ends with call end. This VUI design is very simple and applicable to illiterate users in the developing world. But there are few design aspects, which were not present in the present VUI design of m-Event. Therefore, I will check the present version of m-Event VUI design against my framework and come up with new improve version of the VUI design of m-Event.
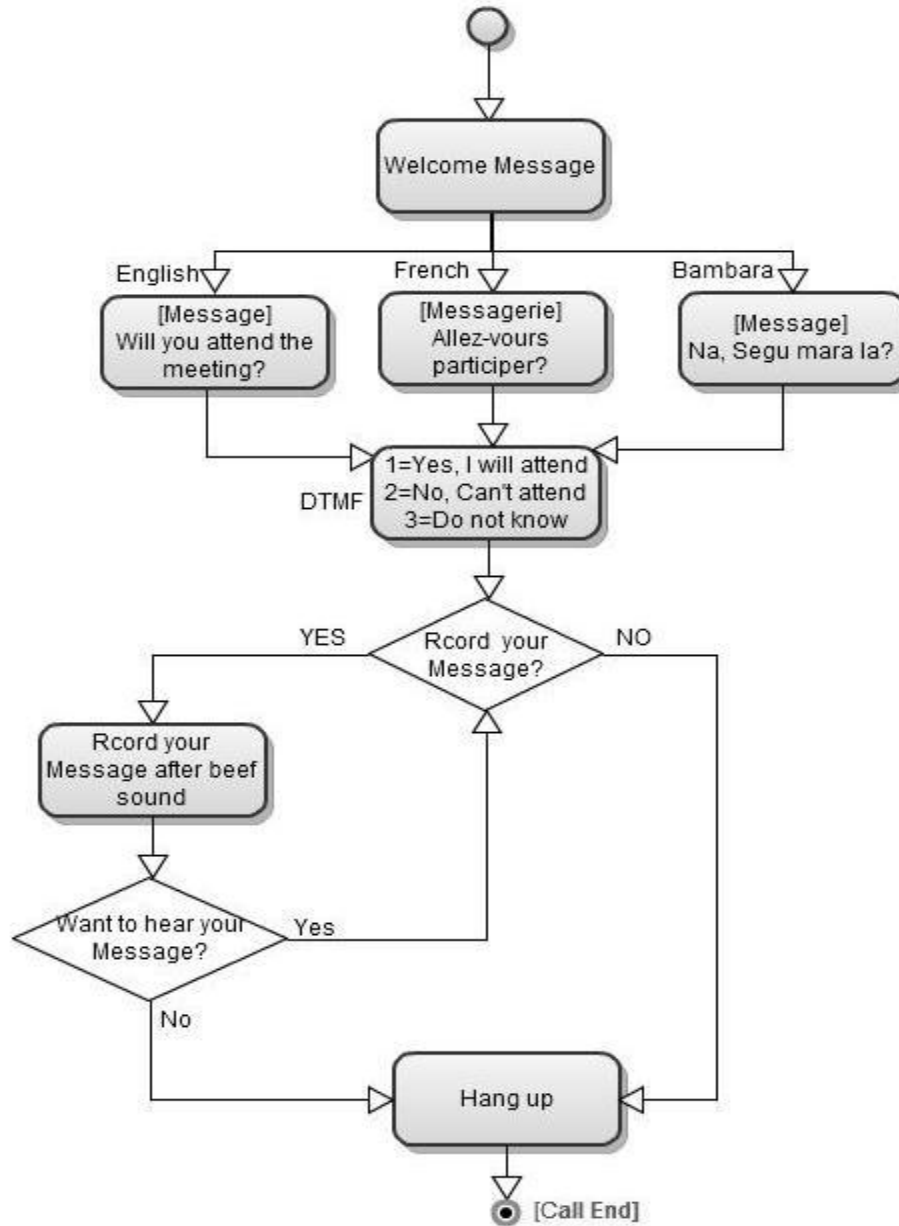
Figure 4: Present m-Event voice user interface.

5.2 m-Event checklist against the Framework

In the Table 2, I present the framework along with m-Event checklist, where I checked whether the aspects present in my framework is present in m-Event VUI design. I found that most of the aspects present in my framework were present in m-Event VUI design. But few aspects were not applied in the present version of the m-Event organizer. To check the m-Event checklist against my framework, I discussed with Mr. Nana Baah Gyan, who is an expert in m-Event organizer. We conclude that all the design consideration aspects were implemented in the present m-Event VUI. While the rest of aspects, I

21

checked and present in Table 2. Each checklist against the aspects are explained in details below the table 2. Finally, I came up with the improve version of m-Event VUI design and it is shown in Figure 5.

Aspect	Researchers	m-Event checklist
Design Considerations		
Enhance the existing system	2, 4	✓
Involving with the user	1, 2,6,7,9	✓
Mobile Usage culture	1, 7,8	✗
IVR Design		
Different user interface for different user group.	6	✗
Error Recovery	1,4,5, 6,9	✓
Dialogue Flow: *Input modality,* *Prompt Style* *Use of sound Effect*	1, 2, 3, 5,6,9 *2* *2* *2*	✓
Keep it Simple	2	✓
Provide multiple language support	1,2,3,7	✓
Implementation Aspects		
Navigation	3, 4,6, 8	✓
Persona	1, 3	✓
Prosody	5	✓
Voice Commands: *Speed,orderly,* *brief,relevant and Persistence*	5	✓

Table 2: Framework along with m-Event Checklist.

✓ : Aspects of my framework are present in the current version of the m-Event.
✗ : Aspects of my framework are not present in the current version of the m-Event.

The m-Event checklist against the framework is explained in details below:
Enhance the existing system: The m-Event designers did use the existing system to enhance the implementation of m-Event system. The designer used the Radio Marche system and it is integrated with the m-Event system.
Involving the users: The m-Event designers did involve the farmers in Tominian, Mali and considered users feedback to design better VUI.

Mobile usage culture: The designers did study the mobile usage culture of the target community in Mali and develop the m-Event system as per pilot studies done in Tominian, Mali.

Different user interface for different user group: In the present version of the m-Event, there is only one user interface. Which means that users have only one choice and experience users don't have alternative interface. But I feel, it is not necessary to have different user interface for different group of users for low literate users. Because, VUI should be very simple and there should not be complex navigation. Therefore, there is no use of different interface when design is very simple and short.

Error Recovery: There is no option available for error recovery in the present version of the m-Event. But Mr. Nana, who is an expert in m-Event told me that error recovery is available in VoiceXML technology. When users don't respond, VoiceXML automatically prompts and direct users in the right directions. Therefore, there is not necessary to use extra prompt for error recovery.

Dialogue Flow: Such as input modality, prompt style and use of sound effects are present in the current version of m-Event. I did hear the recorded audio of current m-Event system and found that information flow were clear, quality and well structured. Therefore, it is not necessary to make any changes in the design part of Dialogue flow.

Keep it Simple: The present version of m-Event is very simple and it is very much appropriate for the illiterate users in developing regions. I was told by m-Event expert that farmers are very much comfortable with simplicity of design present in m-Event. The farmers with the basic knowledge of mobile phone can use the system effectively.

Providing multiple language support: The current m-Event system provides the users with voice messages in preferred users' language. The preferred user language in recorded in user profile in the m-Event system. But based on the framework, I would prefer to have choices of language selection from the user side. Therefore, I will add language selection options in an improved version of the m-Event design.

Navigation: The present interface is very simple. The interface just have 3 options to choose and there is less probability that users get confused or lost. Complex navigation will have negative comprehension towards the technology but current VUI design of m-Event is appropriate for novice users.

Persona: It is present in the current m-event voice interface. I heard the voice recording of m-Event system and there is background music in the welcome message, which will have effects on users feeling towards the applications.

Prosody: Intonation is important element in VUI but synthesizers available in todays technology perform poor job. Therefore, it would be nice, if the m-Event designers come up with the solution for intonation, which is important with reference to low literate users.

Voice Command: I did hear the voice record of m-Event and I found that it is very much consistent, brief and orderly. Therefore, the voice command aspect is present in current m-Event system.

5.3 Improved version of m-Event VUI design

In the improved version of m-Event VUI design, I made a little change compare with present version of m-Event VUI design. Out here, I allow farmers to choose their desired language rather than the preferred language recorded on the m-Event system. Because, many farmers don't keep mobile phone with him/her and call may be received by the user relative, who may know other language better than the preferred language given in the m-Event system. Therefore, I present the option -1 for English, 2 for French and 3 for Bambara. While the rest of the design, I keep same as the present version of the m-Event. Because, I checked present version of the m-Event design against my framework and found that most aspects were presented.The new improve version of m-Event VUI design will be more effective and convenient for farmers in Tominian region, Mali. This improve version of VUI design can be a reference for other designers, who are developing an application with context to illiterate users in the developing world.
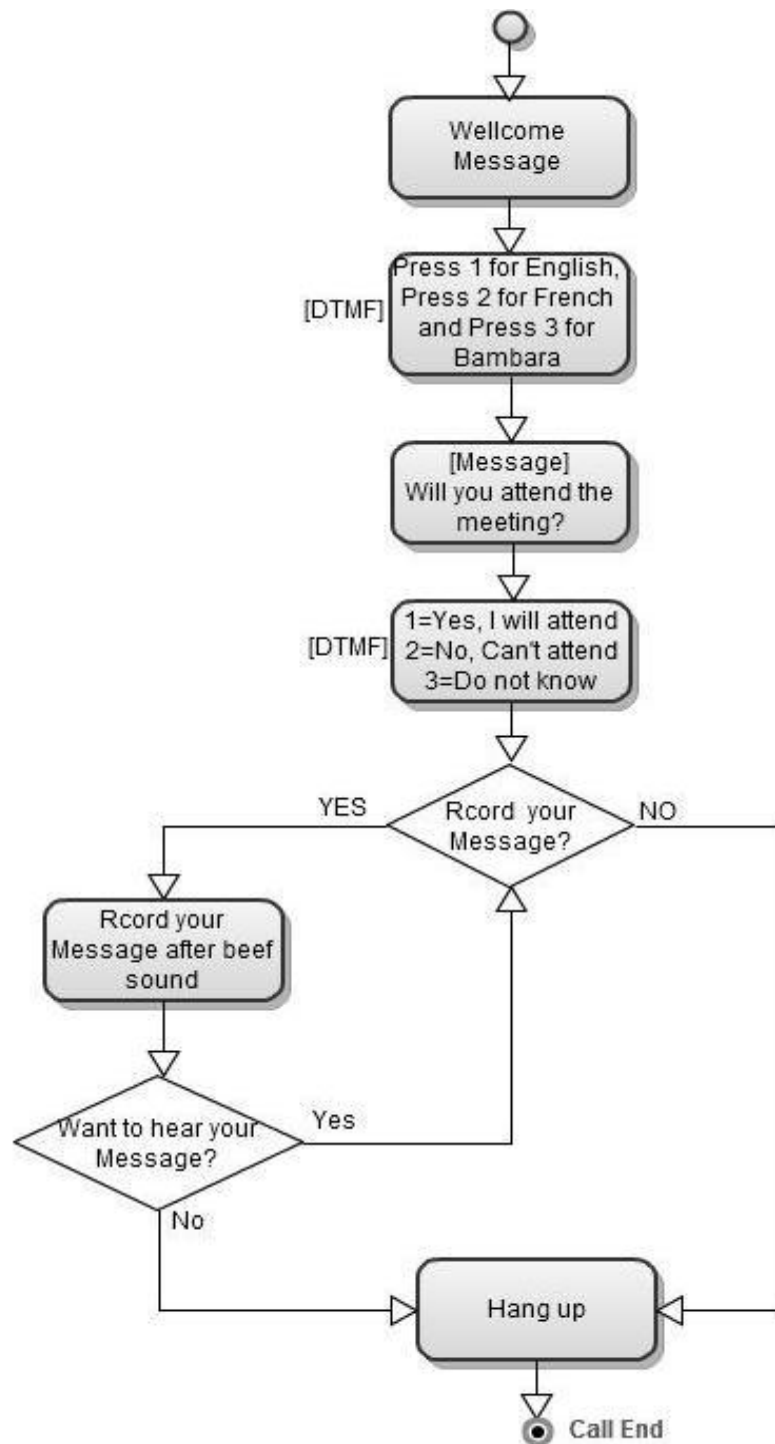
Figure 5: Improve version of m-Event VUI Design.

5   Conclusion

I did a literature research on voice user interface for low literate users in the developing world and selected nine papers for literature reviews. The literature reviews consist of paper's aim, finding, aspects which are crucial when designing the VUI for developing regions. Based on the review, a framework was created to check the appropriate VUI design for m-Event use case.

It is mentioned in [6] that VUI is a comparatively new experience for designers and there is less work done with the VUI implementation in the developing world. Therefore, I had difficulty in finding the appropriate literature. Although there are many literatures on VUI but most are focused with the users in developed countries. Recent market penetration of mobile phones in the developing world had brought attention with VUI designers. Most of the literature that I had reviewed presents about the pilot user testing, implementation, finding and ICT rather than the design of VUI. Therefore, I have created a framework, which will guide new designer to design appropriate VUI for low literate users in the developing world. The framework consists of aspects, which are necessary when designing VUI. I generalized aspects in three categories after consultation with an expert in VOICES project.

There were limitation on the evaluation of existing VUI design of m-Event against my framework. For an example, to evaluate an aspect such as " Different user interface for different user group", we need target population for pilot studies. But I discussed with the experts on m-Event organizer and take their suggestion. I had taken the existing VUI design of m-Event and check against my framework. Then, I have suggested improved version of the voice user interface for m–Event use case. The framework provides the comprehensive checklist for VUI designers.

8   Reference

[1]   Aditi Sharma Grover, Osamuyimen Stewart and David Lubensky. Designing Interactive Voice Response (IVR) Interfaces: Localisation for low literacy users. 1Meraka Institute, CSIR, South Africa, 2IBM T.J. Watson Research Center, NY, USA.

[2]   N. Patel, S. Agarwal, N. Rajput, A. Nanavati, P. Dave & T. Parikh, A Comparative Study of Speech and Dialed Input Voice Interfaces in Rural India. ACM CHI 2009. April 4 - 9, 2009, Boston, Massachusetts, USA, 51-54.

[3]   Dirk Schnell and Fernando Lyardet (2006). Voice user interface Design Pattern: Telecooperation Group, Darmstadt University of Technology, Germany.

[4]   P. Nasfors. Efficient Voice Information Services for Developing Countries,   Master Thesis, Department of Information technology, Uppsala University, Sweden, 2007.

[5]   Kwang B. Lee, Roger A. Grice, The Design and Development of User Interfaces for Voice Application in Mobile Devices. 1 Samsung Electronics Co., Ltd., KOREA, 2 Rensselaer Polytechnic Institute, 2006.

[5]   Nicole Yankelovich, Gina-Anne Levow, Matt Marx, Designing SpeechActs: Issues in Speech User Interfaces. 1 Sun Microsystems Laboratories, Two Elizabeth Drive, Chelmsford, MA, USA 01824, 2 MIT AI Laboratory 545 Technology Square, Room 810 Cambridge, MA 02139, 1995.

[6]   Daniel Mecanovic and Hao Shi. Voice User Interface Design for a Telephone Application Using VoiceXML, School of Computer Science and Mathematics, Victoria University, Australia.

[7]   J. Sherwani, S. Palijo, S. Mirza, T. Ahmed, N. Ali, & R. Rosenfeld, Speech vs. touch-tone: Telephony interfaces for information access by low literate users, Proc. IEEE/ACM 3rd Int. Conf. on ICTD, Doha, Qatar, Apr. 2009, 447–457.

[8]   T.S. Parikh & E.D. Lazowska, Designing an Architecture for Delivering Mobile Information Services to the Rural Developing World. Proc. 2006 International World Wide Web Conference (WWW 2006), 791 – 800, Scotland,2006.

[9]   Plauche, Madelaine, Udhayakumar Nallasamy, Joyojeet Pal, Chuck Wooters, and Divya Ramachandran, "Speech Recognition for Illiterate Access to Information and Technology," International Conference on Information and Communication Technologies and Development (ICTD2006), Berkeley, 2006.

[10]   Nicole Yankelovich, Gina-Anne Levow, Matt Marx, Designing SpeechActs: Issues in Speech User Interfaces. 1 Sun Microsystems Laboratories, Two Elizabeth Drive, Chelmsford, MA, USA 01824, 2 MIT AI Laboratory 545 Technology Square, Room 810 Cambridge, MA 02139, 1995.

[11]   http://www.sigchi.org/chi95/proceedings/papers/ny_bdy.htm Accessed on 6th December. 2012.

[12]   George A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psycological Review, 63:81_97, 1956.

[13]   http://en.wikipedia.org/wiki/VoiceXML Accessed on 4th January, 2013.

[14]   DONNER, J. 2008. Research approaches to mobile use in the developing world: A review of the literature

[15]   CHIPCHASE, J. 2005. Understanding non-literacy as a barrier to mobile phone communication. http://research.nokia.com/bluesky/non-literacy-001-2005/index.html.

[16]   UNESCO. 2007. UNESCO Institute of Statistics Data Centre, regional literacy rates. http://stats.uis.unesco.org/unesco/TableViewer/tableView.aspx?ReportId=y201.

[17]   Schmandt, Chris. Voice Communication with Computers: Conversational Systems, Van Nostrand Reinhold, New York, 1994.
[18]   Fonseca, Rodrigo, J. Pal, "Computing Devices for All: Creating and Selling the Low-Cost Computer," International Conference on Information and Communication Technologies and Development (ICTD2006), Berkeley, 2006.
[19]   Indrani Medhi[1], Somani Patnalk[2], Emma Brunskill[3], S. N. Nagasenga Gautama[1], William Thies[1], Kentaro Toyama[3] "Designing Mobile Interfaces for Novice and Low-Literacy Users". [1]Microsoft Research India, [2]Massachusetts Institute of Technology , [3]University of California, Berkeley, 2011
[20]   Greenberg, A. (2005). ICTs for Poverty Alleviation: Basic Tool and Enabling Sector. Sida ICT for   Development   Secretariat,   Stockholm,   Sweden.   Online   version: <http://www.sida.se/publications>.
[21]   http://en.wikipedia.org/wiki/Voice_user_interface Accessed on 7th January, 2013.
[22]   Kamm, Candance. "User Interfaces for Voice Applications," in Voice Communication Between Humans and Machines. National Academy Press, Washington, D.C.,1994.
[23]   http://en.wikipedia.org/wiki/Prosody_(linguistics) Accessed on 20the January, 2013
[24]   Jahanzeb Sherwani , Speech Interfaces for Information Access by Low Literate Useras:Computer Science Department, Carnegie Mellon University,Pittsburgh, PA, 2009.
 [25]   ITU.   Mobile   overtakes   fixed: Implications   for   policy   and   regulation, 2003.http://www.itu.int/osg/spu/ni/mobileovertakes/Resources/Mobileovertakes_Paper.pdf. Accessed on 22 January, 2013.
[26]   http://en.wikipedia.org/wiki/Interactive_voice_response, Accessed on 6th Febrauray, 2013

Glossary

DSC	Development Support Centre
DTMF	Dual Tone Multi-Frequency
HCI	Human Computer Interaction
HTML	Hyper Text Markup Language
ICT	Information and Communication Technology
ICT4D	Information and Communication Technologies for Development
IVR	Interactive Voice Response
MDGs	Millennium Development Goals
NGO	Non-Government Organization
PDA	Personal Digital Assistant
SDS	Spoken Dialogue System
SMS	Short Message Service
TTS	Text-to-Speech
VoiceXML	A standard XML format for specifying interactive voice dialogues between a human and a computer
VOICES	Voice based Community-centric mobile service
VUI	Voice User Interface
W4RA	Web alliance for Re-greening in Africa
XML	Extensible Markup Language