Dialogue design for the OpenPhone system

TJ NDWE, A SHARMA, M KGAMPE, C KUUN
Meraka Institute, CSIR, PO Box 395, Pretoria, 0001, South Africa
Email: JNdwe@csir.co.za – www.csir.co.za

INTRODUCTION

OpenPhone is a healthcare interactive voice response (IVR) system that seeks to provide relevant information on care-giving issues to caregivers for HIV/Aids-infected children in Botswana. The intended users, the caregivers, can access this information over a telephone line in their local language of Setswana. The system lowers accessibility barriers to the illiterate and to the highly marginalised community of the blind. The caregivers are characterised as mostly females, ranging from semi-literate to illiterate adults with numerical literacy, who are familiar with the telephone for communication purposes. As the caregivers are not accustomed to using an IVR system for accessing information, the technology is entirely novel to them. These characteristics make this user population unique from other similar studies that have been done. This uniqueness necessitates the design of a system to meet specific concerns and decisions, which will ensure its usefulness and usability to its intended users.

PRESENT COPING MEANS

The caregivers presently rely on lectures from the Botswana Baylor Children's Clinical Center of Excellence (BBCCCoE) where they receive medication and consultancy service. It has been observed that the caregivers are passive listeners during the lectures and consequently forget what they have been told. The IVR system is an intervention that is complementary to the present method of ineffective lecturing, which remotely instructs on, reminds of and consolidates knowledge on care-giving issues for the caregivers, as written material cannot be supplied.

THE DESIGN

This encompasses various design constraints of the specific user characteristics. Careful design of the system dialogue is essential: It is the part of the system that the users interact with and directly experience. It is also the part on which users evaluate the success or failure of the system. How skilfully the system dialogue is designed and written, largely determines the nature and effectiveness of telephony systems [3]. Dialogue must be effective and satisfactory to users, both in context and in content.

Context design

The context is the structure and the layout of the information in the system and how this information is delivered by the system to the user and from the user to the system.

An IVR system puts demand on short-term memory resources through a technique of learning called chunking. Studies have shown that humans have a capacity of chunking, which ranges from five to nine chunks of information at a time [4]. Based on this premise, the researchers took a design decision to include no more than five items per menu. The number of items per menu can be viewed as the breadth of the menu. The breadth gets smaller as the dialogue progresses from one level to the next. It was also decided that the depth of the menu would be no longer than three levels deep as illustrated in the dialogue design in Figure 1. These decisions were taken in order to minimise the users' cognitive load during the use of the system.

The keywords used as user utterances or commands were picked according to some attributes that were part of the design decisions. The commands had to be:

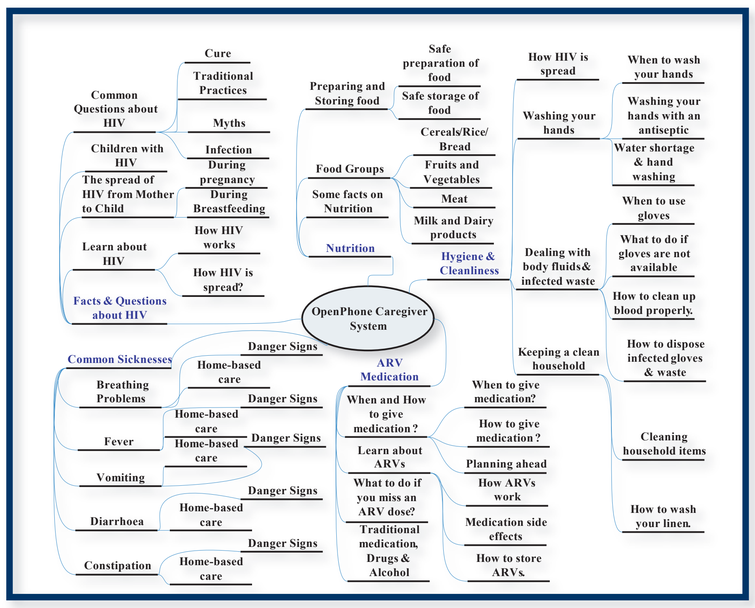


Figure 1: Dialogue design structure of the OpenPhone system

1. Short and precise - most commands to be given in the system menu with a single noun only. This decision was made so that the users only had one or two words to remember as commands. The more constraining a prompt, the smaller the active vocabulary at that point, and consequently the higher the recognition accuracy obtained and the fewer errors that will occur [2].
2. Intuitive - the command is a keyword that is clearly connected to the effect that it is supposed to produce. For example: *System: To learning about hygiene and cleanliness, say 'HYGIENE'.*
3. Acoustically dissimilar - keywords that sound similar were not used in the same menu list.

Information is provided with incremental levels of detail: When a user has not responded, a more detailed and longer explanation of the menu is given by the system. This was specifically done at the root menu for correct routing in this hierarchically structured system design. If the user gets misrouted at the root level, more effort is needed to complete the task.

Items in each menu were listed and given to the user in their descending order of importance. During the user needs solicitation meeting, the caregivers were asked to prioritise the items that they wanted the system to address by their order of importance. The order of menu presentation reflects this prioritisation by the caregivers: the more important issues are mentioned first in the menu.

Figure 2: BBCCCoE in Gaborone

A barge-in facility was introduced to enable the users to interact with the system at anytime without having to wait for the menu list to be concluded. This feature presents a more natural dialogue between the system and the user, similar to conversation. This barge-in feature will also be convenient for future advanced users when the system is fully operational. Barge-in allows users who are accustomed with the system to navigate swiftly through the system dialogue menus to improve the time taken to access information, thereby improving the efficiency of the interaction. Barge-in also supports recovery from errors as the user can barge-in and recover from an erroneous input.

Global keywords, for use anywhere and anytime during the interaction, have been introduced into the design, and the user is frequently reminded about the availability of the global keywords. For example, 'START' prompts the system to go straight to the root menu whenever the user utters the keyword/ command. Throughout the dialogue, it is frequently reiterated that the user who wants to go back to the beginning, can say 'START' at anytime.

Content design

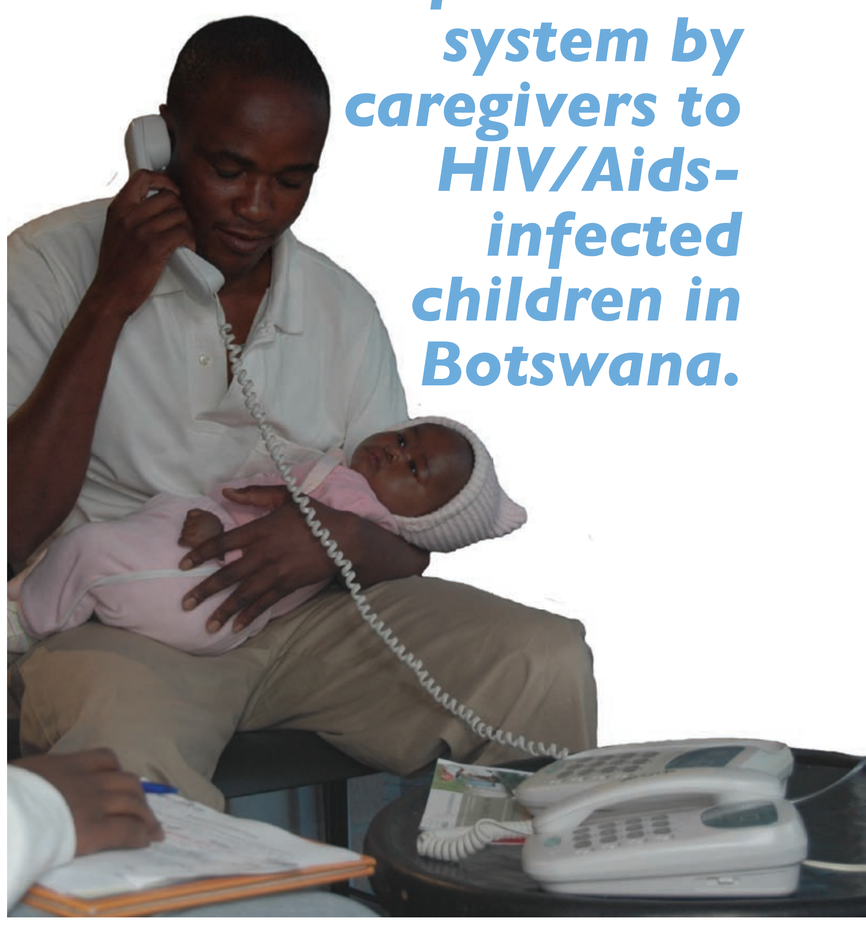
The content of the dialogue script was brought about from various sources; the major stakeholder, the BBCCCoE, had the most influence in deciding on the final dialogue script. The bulk of the content data was mined from the lecture notes used by the BBCCCoE staff members during the lectures delivered in the clinic every morning. The BBCCCoE staff members consulted with individual specialists within the clinic, such as a nutritionist, paediatrician and social worker. The caregivers contributed by alerting the research team to the needs to be addressed by the system and then prioritising the information.

The choice of language for commands and the information in the system has been a language that the caregivers understand in order to avoid technical terminology used by medical professionals. For example, the words 'immune system' do not have a direct translation in the Setswana language and an alternative explanation of the immune system is essential. This has been achieved through consultation with the nurses at the BBCCCoE who give the lectures to caregivers. They suggested that a term that literally translates to 'soldiers of the body' is well understood by the caregivers to mean the immune system.

A recorded voice was used instead of a synthetic voice; listening to and comprehending synthetic voices is more difficult, and a synthetic voice

Figure 3: A user testing two automated telephone systems

Skilful dialogue design in an interactive voice response system has facilitated the effective use of a telephone-based health information system by caregivers to HIV/Aids-infected children in Botswana.



requires more cognitive effort [1]. Connie Ferguson, a well-known actor in a popular local soap opera, was used as the voice talent for recording the voice. She is a female and her age is within the average age range of the caregivers. She speaks fluent Setswana in the local accent. One of the test participants recognised her voice and identified her without being asked. The decision to use a voice talent was made to get a voice that is familiar, in the local accent, female as most caregivers are, in the same age range as the caregivers and a person who is a professional. The voice talent offered her services for free.

The entire system dialogue follows a pattern of first introducing the effect and then the action that should be taken in order to bring about the effect instead of the other way around. For example:
System: For advice on washing your linen, say 'LINEN'; instead of:
System: Say 'LINEN', for advice on washing your linen.
Users tend to understand the critical information of the message better if it is in the final part of the sentence rather than in the initial part [1].

An introduction is provided under each menu item in the root menu so that the user is aware which topic was selected for discussion. This allows the user to know exactly the progression of the dialogue; unwanted topics can therefore be avoided by starting again. Instructions are also provided on how to have information repeated.

The system provides the user with an option to leave feedback and any comments or questions about the system. The feedback is recorded by the system. The user feedback will be used to improve future developments as it will be incorporated into the later versions of the system.

CONCLUSION

The design has enabled the researchers to start developing prototype systems that will be tested with real users in Botswana.

REFERENCES

1. Delogu, C., Conte, S., and Sementina, C. (1998). Cognitive factors in the evaluation of synthetic speech. In Speech Commun. 24(2). May, 1998. pp. 153-168.
2. Hone, K. S. and Baber, C. (1995). Using a simulation method to predict the transaction time effects of applying alternative levels of constraint to user utterances within speech interactive dialogues. In Proceedings of the ESCA Workshop on Spoken Dialogue Systems, P. Dalsgaard, L. Larsen, L. Boves, and I. Thomsen, Eds. ESCA, Vigso, Denmark, pp. 209–212.
3. Migneault, J. P., Farzanfar, R., Wright, J. A., and Friedman, R. H. 2006. How to write health dialog for a talking computer. J. of Biomedical Informatics 39, 5 (Oct. 2006), 468-481.
4. Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. Psychological Review, 63, 81-97.