

Let voice be your guide

Using a mobile phone while you yourself are mobile typically means glancing back and forth between the screen and your surroundings to make sure that your fingers and your feet are both navigating you to where you want to go. In addition, users with poor eyesight or poor literacy will have difficulty in operating their phones. Developed at Nokia Research Center, the Voice aid application uses speech compression and speech synthesis technology to guide users in basic phone functions eyes-free. **By Janne M. Vainio, Sakari Himanen, Hannu J. Korhonen, Juha Marila, Hannu J. Mikkola, Tuomas Vaittinen, Toni Nieminen, Henri Toukoma**

A S **MOBILE PHONE** functionality and user interface complexity increase, even the simplest needs are becoming more difficult to fulfill. Despite these trends, phone usage is still mostly related to basic communication: finding a name from contact list, calling, answering missed calls, and sending and receiving messages [3, 4].

Natural voice feedback is the best way of replacing or supplementing visual feedback. Users do not have to learn what different sounds mean. Voice prompts explicitly contain the information about the function being used. There is just the requirement for the speech prompts to be clear, simple, and easy to understand.

More than a screen reader needed

For vision-impaired users, special screen reading applications do exist. They operate

When voice makes a difference

DRIVING A CAR OR A BICYCLE are the most obvious cases for eyes-free usage of the phone. Drivers have many visual objects to observe, in addition to the road view. Sodhi et al describe the visual demands of a driver: mirrors, road view and car controls [1]. There is no room for extra tasks without compromising traffic safety. Although speaking on the phone and generally concentrating on anything other than driving is itself harmful for driving performance [2], diverting the eyes from the road is still far more dangerous. While driving, the user is also not willing to carry out very complicated operations. Thus, a simplified user interface with reduced set of functionalities will fit well into the use case of driving. Additionally, the performed tasks are not time critical and hence the driver can suspend the phone task and fully concentrate on driving when needed.

When walking, it may be the most natural behavior to stop before trying to make a phone call or browsing a contact list. However, walking is quite automatic in nature, so using basic phone functionalities should be easier than using a phone while driving a car.

People with poor eyesight or completely blind people have to operate a phone in an eyes-free or almost eyes-free mode all the time. The screen provides very little or no information to the user.

Elderly people may also belong to the group with poor eyesight, but they will also benefit from the simplified user interface itself. When only the most essential communication features are visible, it is easier to learn and handle the interface and carry out simple tasks.

For illiterate people, voice prompts are one way to lower the barrier to using the phone and reduce the dependency on having to read text from the display.

Small children even if they are not yet able to read, may be able to operate the phone with voice feedback. Hearing from the phone words “dad” or “mom” will make it possible to select the right person to be called.

There are also some other **special conditions**, in which it is more feasible to operate the phone without looking at the display. One example is very cold weather, when it is more convenient to use your headset and keep your hands in your pocket while operating the phone. Moreover, the phone is protected from the fierce weather conditions, such as heavy rain. Headset and voice prompts also help to use your phone unobtrusively when it is not appropriate to take a phone visible and use it publicly.

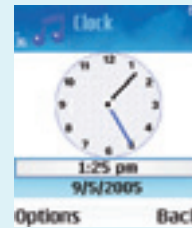
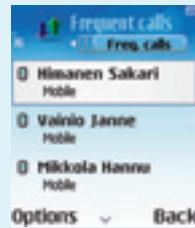
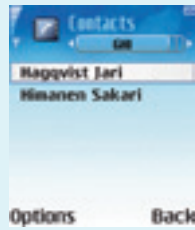
in the background and echo user interaction and visible screen items to the user. However, the screen contains quite a large variety of components and text, thus reading out everything seen in the screen makes voice prompts long and difficult to follow. Combined with a highly visual UI, the result is very clumsy and confusing. Voice prompts are most effective when they are short and simple. Although converting graphical UI to voice prompts makes it possible to use a phone, it does not necessarily make it easy or convenient.

In the case of the Series 60 UI, the design has been optimized for a small screen, but there is not much provisioning for having no display at all. The UI strictly relies on the screen to give the status of the application, making it possible for the user to navigate between different functionalities and views. >>

Listen up!

Users can browse the **contact list** by listening to the names. To assist in finding the right name, the contact list is divided into tabs according to the letter groupings on the keypad. Each tab is accessible via the corresponding key or by navigating to the left and right.

In the phone log the user can browse **missed, dialed and received calls** and make a phone call from these lists. Additionally, a new recent calls entry is introduced: **the most frequent calls list**. It makes it very easy to give a call to a person who the user is calling often, like your family or close friends. Other call lists are often littered with one-time calls, thus making it more difficult to find the frequently dialed entries.



It is also possible to make a call by **dialing the number directly** if no suitable contact exists in the phone book. This feature is implemented with special number roll, which can be operated by using just the 5-way navigation key and two soft-buttons.

It is also possible to get the **time and date** with voice feedback.

Additionally, some other useful functions exist, like easy **muting and unmuting** voice feedback and **volume** control of voice feedback. The volume and muting status of the Voice aid is independent from the actual phone UI.

Voice aids basic communication

» Voice aid is mobile phone software with a multimodal UI, with which you can operate your phone aided by voice feedback. This removes the need to see the display for people that are unable to, either temporarily or permanently. Still, the needed information is available on the display all the time, so the user doesn't have to rely on speech alone in situations where he or she can see the screen.

The current version of Voice aid supports the most essential phone functionalities [3, 4] – Contact list, Recent calls, Voice mailbox, Number dial and Clock (see inset box) – and follows most of the Series 60 conventions to make it easy to use even on the first try.

Voice aid uses pre-recorded voice prompts for permanent menu items, which have been compressed using technology developed at Nokia Research Center. This technology provides an effective very low bit-rate compression for narrow-band speech. It maintains sufficiently good speech quality while still squeezing the voice prompts into very small storage space, an important factor when several languages need to be supported simultaneously. It would also be technically possible to use text-to-speech (TTS) synthesis to generate the voice prompts, but, given the limited resources of mobile devices, the quality is not as good.

For non-permanent voice prompts like contact names, TTS synthesis must be used because those prompts cannot be pre-recorded. This TTS synthesis system has specifically been implemented and optimized for mobile applications in NRC.

42 languages supported

One of the big challenges of making an application with voice feedback is the localization effort. The Series 60 UI currently supports 42 languages, so the Voice aid application must support all of them in order to be delivered with the software of the phone product.

Though there are translations available for most of the Series 60 screen texts, all the voice prompts are translated exclusively to guarantee best fit for targeted use as audio. Furthermore, need for concatenation of voice prompts creates additional challenges in application development due to different grammatical rules among languages. Since the voice prompts are recordings of native speakers, any functional modifications causing changes in voice prompts require that all the speakers need to be invited back to record the new prompts.

The future looks vocal

The development of the Voice aid concept and accompanying software started in the autumn of 2003, and the first version of the software will be delivered to several product programs in the third quarter of this year.

The Voice aid concept is still under continuous evolution, in which more functions and different ways of interacting with voice are added. One example is voice recognition. With simple voice commands the phone can be used totally hands- and eyes-free. This is possible even when your phone is located out of your reach.

Also, for voice feedback there are things that need further development and investigation. It is not trivial to balance the features and functionalities of voice UI and to maintain

short and easily understandable voice prompts, which give good information about the status of the application and make navigation and operation as effortless as possible. Also, different user groups may need different considerations. ■

FURTHER READING:

- [1] M. Sodhi, B. Reimer, J. L. Cohen E. Vastenburg, R. Kaars, S. Kirschenbaum (2002) On-Road Driver Eye Movement Tracking Using Head-Mounted Devices. Proceedings, ETRA'02, New Orleans, Louisiana, USA. ACM 2002
- [2] Dave Lamble, Tatu Kauranen, Matti Laakso and Heikki Summala (1999) Cognitive load and detection thresholds in car following situations: safety implications for using mobile (cellular) telephones while driving. Accident Analysis & Prevention vol 31 (6) 1999, pages 617-623
- [3] Kari Virtanen (2004) Final Report, Smartphone 360, Nokia 6600, 5 Months: February–June 2004. Series 60 Platform / TP
- [4] Wireless World Forum (2005) MobileYouth 2005 study, page 108

FOR MORE INFORMATION, CONTACT:

Hannu Mikkola

Multimedia Technologies laboratory
Nokia Research Center, Tampere
hannu.j.mikkola@nokia.com

Janne Vainio

Multimedia Technologies laboratory
Nokia Research Center, Tampere
janne.m.vainio@nokia.com