Methodology and guidelines for the evaluation of accessibility of public terminal devices by people with visual or hearing disabilities: Sound, audio and speech design considerations

T. Vasiliadis\textsuperscript{a} and P. Angelidis\textsuperscript{b,*}
\textsuperscript{a}Department of Electrical and Computer Engineering, Telecommunications Laboratory, Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece
\textsuperscript{b}Southeastern Europe Telecommunications & Informatics Research Institute, Thessaloniki, Greece

Abstract. This paper presents guidelines and methodologies for testing the accessibility of card-reading or other terminal devices installed in public places. The testing area of our concern is sound, audio and speech. In all cases, every effort has been made in order to meet the "design for all" concept. This work was funded by CEN/EC and has influenced the activities of CEN TC224/WG6.

Keywords: Design model, interface design, social design, user participation, design for all

1. Introduction

In this paper, we present the results of our research on the acceptable levels of ambient noise in the vicinity of card-reading or other terminal devices, which are installed in public places. The applications taken into consideration include human-device interactions involving speech and/or sound, as well as telephony applications. In close conjunction with this case, lies the study of the intelligibility of spoken messages against different acoustic levels of background noise.

This study recommends an adequate methodology and means, by which background noise measurements should be performed. The goal is to define the thresholds over which an acoustic shielding of the card-reading machine is considered necessary.

Finally, some basic recommendations and design guidelines are proposed for a formal definition of the level, form, type and distance requirements for the case where audio signals are used to help the visually impaired to locate a card-reading device.

The report is organized in Sections. Section 2 includes up-to-date background work in the area. Section 3 presents the main core of the results, i.e. the proposed methodology for testing on audio aspects. Section 4 describes the two templates used for testing. Further extensions on the present work are suggested in Section 5, whereas the paper concludes in Section 6.

This work was funded by CEN/EC and has influenced the activities of CEN TC224/WG6 [2,3].

\textsuperscript{*}Address for correspondence: Pantelis Angelidis, Telecommunications Engineer, Southeastern Europe Telecommunications and Informatics Research Institute, 9th Km Thessaloniki-Thermi Road, P.O. BOX 60080 GR-57001 Thessaloniki, Greece. Tel.: +30 2310489386; Fax: +30 2310489387; E-mail: pantelis@vidavo.org.
2. Standards review

2.1. Documentation

An up-to-date investigation of all relevant standards (currently in-force) has been carried out. The area of research was focused mainly in telecom and telematics related documents. A list of all standards, recommendations or technical reports that have been reviewed is given in Table 1.

2.2. Hearing disability and functional limitations

Hearing impairment can affect the whole range or only part of the auditory spectrum. The important spectral area for speech perception is between 250 and 4000 Hz. The term deaf is used to describe people with profound hearing loss, while hard of hearing is used for those with mild to severe hearing loss.

It should be pointed out that although many of the 80 million hard of hearing people in geographical Europe will have problems, for example, using a public phone in a noisy location, they will however not necessarily consider themselves as disabled or be registered as such [5].

2.3. Volume control

The public terminal should, if possible, be located in “low-noise-level” environment with good architectural acoustics. The telephone (if the terminal under consideration is used for telephony applications) should be equipped with an adjustable volume control, which can be identified and located visually, and by touch.

Volume controls, whether they are contained within the handset or integrated into the terminal, should be capable of increasing the volume within the range of 12 dB minimum and 18 dB maximum above the non-amplified state. The 18 dB maximum should not apply where an automatic reset (i.e. on-hook) to the non-amplified state is provided. But even in the case of an automatic reset capability, the level should never exceed 20 dB above the non-amplified state. However, careful should be taken that under no circumstances should the maximum attainable amplification be high enough to cause either hearing damage or acoustic feedback. It is preferable for the volume to be always reset to the non-amplified state when the handset is returned to the cradle or after a short time-out [8,9].

As hearing impaired people do not necessarily have elevated thresholds of loudness discomfort, some form of output limitation will be required. Recent work has indicated that Automatic Gain Control (AGC) can provide a better automatic means of limitation than peak clipping. In addition, it is recognized that the frequency response to give maximum intelligibility to some hearing impaired people may require shaping.

It is estimated that with the provision of additional amplification to the levels recommended above, possibly up to 80% of hearing impaired users would benefit even when not using their hearing aids to couple to the telephone set. If a hearing aid is worn in addition and inductive coupling is also provided, then the pro-
portion of hearing impaired users who will be able to have satisfactory telephone conversations will further increase [9].

2.4. Hearing aids

Public terminal devices that facilitate voice communications should be equipped with handsets that are hearing-aid compatible through inductive or electric coupling. These handsets generate a magnetic field to which hearing aids may couple [2]. All necessary requirements for magnetic field intensity for telephone compatibility with hearing aids can be found in [6,7] and [9].

2.5. Visual indication and other guidelines

Some general guidelines regarding acoustic output format and other aspects of the overall system design are the following [5]:

- Always provide alternative visual indication for any acoustic signal (lights and/or LCD displays).
- All tone signals should include high and low frequency components.
- If possible, provide user selectable sounds with different pitch.
- If pitch cannot be selected by user, high pitch sounds should be avoided.
- A connector for external earphones should be provided as alternative to loudspeakers.
- Pay-phones should be able to adapt to user profiles on Smart Cards (EN 1332-4). The technology that is developing around smart cards enables a user to store their own preferences on the memory chip of a smart card. Smart cards based public terminals allow the user’s card to instruct them in order to make specific adjustments. These could possibly affect sound quality, volume, typeface sizes and language preferences.

3. Methodology of tests

The recommended methodology of the proposed tests is discussed here, in order to evaluate the adequacy of the terminal’s design. We focus on two main areas of research, (a) ambient noise measurements and (b) acoustic-aided locating of the terminal device. As defined in the scope of this article, ambient noise measurements will investigate the need for additional acoustic shielding of the terminal, as well as the intelligibility of spoken messages for different levels of background noise.

3.1. Ambient noise measurements

The proposed methodology for the measurement of the maximum permissible background noise level is illustrated in Fig. 1.

In all cases, there should be a distinction between indoor and outdoor applications, since there are different design considerations that should be taken into account when the device is to be used outdoors, compared to an indoor environment.

Moreover, we consider the following sub-categories, which apply in both indoor/outdoor cases as illustrated:

1. The message under consideration is user originated. An example of such a message is a user prompt (command) that is passed to the machine in applications incorporating speech recognition technologies.

2. The prompt is machine originated. The majority of voice prompts in all applications are coming from the terminal device. This case draws the most of our concern. Different types of prompts should be tested:

   a. Pre-recorded voice prompts. In this case, voice messages (or text-to-speech messages) should be tested in terms of audibility and intelligibility for different ambient noise levels.

   b. Single frequency tones. In applications having simple man-machine interface, some prompts shall have the form of single-frequency tones or combination of simple tones. In such cases, an examination of the frequency ranges for best audible tones should be carried out for different groups of people.

3.1.1. Measurement setup

The measurement setup for the lab trials shall contain the following equipment:

- Acoustic noise generator (white/pink noise).
- Audio amplifier.
- Microphones.
- Sound meters.
- Loudspeakers.
- Headphones.
- Computer.

At first, for the measurements to be realistic, some recordings of typical background noise levels shall be performed in sites where the card-reading devices under investigation are to be installed. To the extend possible,
worst case solutions should be chosen, for example, next to a city road in extreme traffic conditions (outdoor application) or in a crowded corridor of a public building (indoor application). These recordings should be used as noise samples in the measurements (in addition to the acoustic-noise generator).

Three different cases shall be examined in detail during the lab trials. For either indoor or outdoor applications, all sub-categories presented earlier (regarding the type of the audio signal) should be investigated.

In special applications like public telephone terminals, there's a different approach since the user has the ability to operate the device with a handset. This case will be examined separately.

The measurement setup for applications that incorporate speech recognition technology is illustrated in Fig. 2. In these cases, the ambient noise source is simulated by an acoustic-noise generator connected to a loudspeaker. On the other hand, the card-reading device is simulated by a personal computer connected to a loudspeaker(s) and a microphone for all necessary speech communication between the user and the device.

In all other cases, communication between the user and the device is to be carried via a conventional keypad. The measurement set-up for this particular case (except telephony applications) is illustrated in Fig. 3, where the microphone has been replaced by a keyboard.

The acoustic noise generator shall be able to generate white noise in different sound pressure levels (SPL) that can be adjusted externally (the use of an audio amplifier is recommended). Normal transactions between the user and the device are simulated using the computer while manually increasing the ambient noise level. The maximum permissible SPL is recorded for each case. The SPL of the background noise shall be measured either at the position of the card-reading device or, alternatively, at the position of the user relative to the device (usually at a distance of approximately 0.5m from the device), no matter where the noise generator’s loudspeakers are located during the tests.

3.1.2. User originated messages

If the application uses speech recognition techniques during the user interaction, the setup in Fig. 2 shall be deployed in the lab according to the following guidelines.

The recognition engine on the terminal device shall be configured and optimized using echo cancellation and noise reduction.

Echo cancellation improves the quality of a speech signal by diminishing any echo that might have been introduced by the telephone line. To support barge-in,
the application should also support echo cancellation. Otherwise, the recognition engine cannot provide accurate results because the echo from the played prompt is often mistakenly assumed to be the user’s utterance.

For increased recognition accuracy and efficiency, it is critical for the system to distinguish any leading or trailing background noise and/or silence from the actual utterance, before sending it to the recognizer. Modern recognition engines have algorithms to reduce incoming steady-state background noise. The engine enhances the user-originated message and effectively filters out noises such as tones, buzzing, humming and hissing.

However, this mechanism was not designed for speech recognition in non-steady-state background-noise environments, such as other voices behind the primary speaker. As noted earlier, the same measurements should be carried out also by using pre-recorded noise samples from real installation sites, instead of a noise generator. In the case of speech recognition applications, this should be taken into serious account, since the noise generator produces steady-state noise (white spectrum) and a potential noise-reduction mechanism in the recognition engine could filter out all incoming noise, hence giving us erroneous or misleading results.

The use of noise reduction mechanisms in such applications increases the CPU usage for the recognition engine. There is a trade-off between noise reduction efficiency and computational cost (Fig. 4). Therefore, this mechanism should be used depending on the overall required speed of application.

Taking into account the aforementioned configuration aspects, tests shall be performed by increasing the ambient noise level and recording the recognition results. At the same time, some tuning of the recognition engine should be performed in terms of VAD thresholds. Increasing this threshold can result in worse recognition performance for constant user utterance level and background noise level. The overall acceptable background noise level (SPL) shall be retrieved according to the application’s recognition requirements (e.g., the confidence levels for specific VAD threshold using echo cancellation and noise reduction mechanisms).

All tests shall be performed using pre-recorded typical voice prompts played back as input to the recognition engine. This ensures that the speech level and characteristics from the user during the trials is consistent and uniform at all instances. The playback shall be performed with a loudspeaker placed in front of the device and at typical distances from it (height about 1.80 m – 0.5 m distance from the machine). A typical list of recommended speaker prompts for some applications is given below:

- Digits (all digits shall be tested).
- “Yes”/“No” (confirmation dialogs).
- “Abort”/“Cancel”.
- “OK”/“Enter”.
- “Increase volume”/“Decrease volume”.
- “Money Withdrawal”.
- “Money Transfer”.

The testing procedure is summarised in Section 4.2.1.

3.1.3. Machine originated voice messages

Regardless if the user input is via speech commands or by a standard keypad, this test is carried out to examine the audibility and intelligibility of pre-recorded voice prompts that the device is playing back to the user (see Figs 2 and 3).

The user is asked to verify the spoken-out message. The process is repeated until the user fails to do so, at which case the SPL is noted (see Section 4.2.2). Messages are transmitted at random order. The threshold level is verified by repeating all messages.

Two sub-cases can be distinguished regarding the quality and source of the voice prompt. This pre-recorded message can be either a high-quality digital voice recording (up to 44100 Hz–16 bit – stereo), played back in a set of loudspeakers, or a digital synthesized voice-prompt made by a TTS engine.

These messages shall have the form of a potential prompt for any application. Typical examples are given below:

\[ \text{Fig. 3. Lab trial measurement setup (general case).} \]
Fig. 4. Relations between configuration parameters and acceptable ambient noise level for applications with speech recognition enabled.

Fig. 5. Parameters affecting the intelligibility of spoken messages.

- “Please insert your card”.
- “Invalid card!”.
- “Please enter your PIN code”.
- “Please take your card”.
- “Would you like a receipt?”
- “Amount of money for transaction?”
- “Please make a selection”.

The testing procedure is summarised in Section 4.2.2. The parameters affecting the intelligibility of spoken messages are schematically represented in Fig. 5.

For the case of pre-recorded messages, taking into account that normal speech signals consist of spectral components about as high as 8000 Hz, we consider that the minimum quality of pre-recorded prompts shall be at least of telephone quality (high frequency components at 4000 Hz). Lower quality is considered inadequate for such applications. In general, a worst-case scenario can be adopted and all tests shall be performed using pre-recorded messages sampled at 8000 Hz—8 bit—mono.

In some applications, text-to-speech technologies can be used for the spoken messages. For a TTS using a modest phoneme database, the resulting synthesized audio file could be very hard to understand in extreme ambient noise conditions. The intelligibility of messages produced by such a system increases when the prosody of speech has been carefully designed. In some cases, there should be different intonation curves applied to some prompts. In general, intonation shaping is recommended when investigating the intelligibility of spoken messages in noisy environments.

In both of these cases (recorded and synthesized messages), many different prompts shall be used in order to achieve high accuracy. Moreover, different types of voices shall be used for the recordings, including some variants of specific male and female voices.

3.1.4. Machine originated tone messages

For simple man-machine interfaces where the prompts have the form of single frequency tones or combination of simple tones, the following guidelines apply.

The spectrum of audible tones is limited by human anatomy of the ear. The minimum audible frequency tone is about 16 Hz and the maximum is about 20 KHz. The upper boundary is gradually decreasing with age to about 10 KHz.

The field of hearing for any human ranges from 100 Hz to 8 KHz for normal speech sounds. In general, high frequency components (greater than 8 KHz) shall be avoided in all cases. Simple frequency messages are also considered inappropriate since there are humans that suffer from severe hearing loss in specific spectrum areas. Therefore, combination of tones is highly recommended for such kind of messages.

Moreover, early research in electroacoustics shows that the human ear can’t distinguish easily two single tones of the same level but slightly different in frequency. For that reason, the decision for the tones must
Fig. 6. Masking of a tone by a louder tone.

Fig. 7. Definition of mean levels \( L_{eq} \) in the noise profile chart (example graph).

meet specific requirements in order for the tones to be equally audible and distinguishable [10].

Another critical consideration that should be taken into account is the masking effect. According to acoustics theory, a single frequency tone at a specified sound pressure level can mask all tones in a wider spectrum area (towards the highest components) of a level specified by the appropriate masking curves [1]. The masking effect is illustrated in Fig. 6.

Therefore, the decision and design of frequency tone messages for card-reading devices has to meet the aforementioned requirements. Taking into account these design aspects, the intelligibility and audibility of this kind of messages can be tested according to the methodology previously mentioned in Section 3.1.3. All messages are played to the user while increasing the background noise level (using both noise generator and pre-recorded noise samples) until the intelligibility of the message is not acceptable.

The testing procedure is summarised in Section 4.2.3.

3.1.5. Machine originated prompts – Statistical aspects

In both cases, when investigating machine originated prompts (either voice or tone messages), the maximum permissible ambient noise levels are extracted by testing the intelligibility of spoken messages by the end-user.

Every effort shall be made to involve people from various population groups, in order to validate the simulation results and generalize to an extent the main conclusions extracted from the trials. The main concern shall be to record feedback from people with hearing disabilities, since this group is expected to pose stricter requirements in audio interacting design for card-reading or other terminal devices. The population sample (in a general case) has to be large enough to be considered representative in statistical terms.
expected range for the age of users is about 18–75 years (this can be extended some years further). However, there is no need for complete representation of this whole range, since problems in the usage of terminal devices are expected only in special population groups. Therefore, for audio and speech applications, the users' sample needs only to include elderly people (which are hard of hearing due to ageing) and people with hearing impairments of any age. These groups are considered the worst case for the audio and speech trials presented in this work.

In the geographical area of Europe there are about 100 million elderly people and 50 million who are disabled in some way [4]. Moreover, about 80 million people are hard of hearing compared to the 800 million of Europe's total population. The age distribution of the disabled people reveals that nearly half of them (51%) are older than 65 years, 19% are between 50 and 65 and roughly 30% of them are younger than 50 years. This clustering applied to the hard of hearing group of population provides a satisfactory estimate for the users' sample needed during the measurements.

Furthermore, taking into account the fact that the estimated population of Europe aged 65 and over (elderly people) is expected to reach 136 millions in the year 2030, it is clear that designing for the elderly and disabled people is often adequate criterion for the total population as well.

3.1.6. Telephony applications

In the case of a public telephone terminal, the user has the ability to operate the device using a handset. The methodology is the same as before, taking into account that the spoken messages by the device are played back into the handset and not on a loudspeaker. The tests can be carried out using the same guidelines as before.

Therefore, hearing aids can improve the overall susceptibility of the system to external noise sources. Vol-
volume control can be used without any problems to the surrounding environment. Furthermore, in telephony applications the use of acoustic shielding shall be considered inevitable in some cases, because of privacy reasons as well as for protection from ambient noise. All acceptable background noise levels that have been recorded for each of the above cases (if applicable) shall next be compared against each other and the worst-case scenario shall be adopted for the application. This level of ambient noise shall finally be compared with the average background noise measurements from the final installation site. This comparison shall guide the decision regarding the necessity of any acoustic shielding.

3.1.7. Design criterion for acoustic shielding

The measurement process described in the aforementioned paragraphs shall be used to extract useful conclusions about the application under consideration. Assuming that we have recorded the background noise profile of a real site (where a terminal device is to be installed) for a large period of time, we shall calculate the safe usage percentages of the device by using the information extracted from the lab trials.

The term ‘safe usage percentage’ is used to describe the percentage of people that will be served by the installed device without problems (caused by ambient noise) for a specified/guaranteed percentage of time (given that no additional acoustic shielding has been provided on-site).

Measurements for the background noise profile of a site shall be performed and the following statistical values shall be extracted from the profile:

- $L_{xx}$: noise level that has been exceeded for an ‘$xx$’ percentage of time in this site.

Figure 7 illustrates the definition of the mean levels for 2, 10 and 50% of the total observation time. The information behind $L_2$ (or $L_{10}$ for example) noise level is that these levels have been exceeded only for 2% (or 10%) of the total time. This guarantees that for the rest 98% (or 90%) of the time, the background noise level will be below that level at the specified site.

The results of the lab-trials shall be statistically processed and a graph of the distribution of acceptable noise levels for a user percentage level shall be created. Figure 8 illustrates, for example, the cumulative distribution of the results, i.e. the percentage of users that have accepted the specified ambient noise SPL. The absolute noise boundary is the maximum noise level for which none of the users could understand the spoken messages (i.e. people acceptance is 0%).

Using the measured data so far, we shall extract the safe usage percentages for the device under consideration (and for the specific installation site). The proposed process is depicted in Fig. 9.

Marking the appropriate mean noise level (in this example $L_2$) at the SPL axis, the distribution curve reveals the percentage of people that accepted this ambient noise level. Summarizing, this approach provides guarantees that for 98% (for the specified example—see Fig. 9) of the time, the device installed on that specific site will safely serve a percentage X% of the people, without any need for acoustic shielding.

Depending on the application’s requirements about these percentages, there shall be a decision about the necessity of additional acoustic shielding for this terminal device.

3.2. Acoustic aid for locating the device

In the case of visually impaired users and not only, there shall be an audio signal helping them locate the terminal device. Two different sub-cases shall be examined, as far as the acoustic source is concerned:
### User originated prompt test

<table>
<thead>
<tr>
<th>Name</th>
<th>Date</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ASR engine used**
- **Echo Cancellation**: YES □, NO □
- **Noise Reduction**: YES □, NO □

**Initial VAD threshold**

<table>
<thead>
<tr>
<th>Speaker Prompts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

**Test Results**

<table>
<thead>
<tr>
<th>Prompt</th>
<th>Recognition status</th>
<th>VAD</th>
<th>SPL @ device</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Overall SPL** (minimum of successful recognitions) dB(A)

---

### 3.2.1. The source is located at the card reading device

The methodology for this kind of application is the same as before (Section 3.1.3). Many of the aforementioned principles apply in this case as well. The type of the acoustic signal shall be determined by reviewing the results of the previous lab-trials (pre-recorded voice message vs. frequency-tone message). The level of this audio signal shall be also adopted from previous tests.

### 3.2.2. The source is located at a handheld device carried by the user

This case shall be preferred because the audible signal can be at a lower level and thus it is less annoying for others moving at the surrounding environment. Moreover, the user shall be capable of adjusting the parameters of the audio signal in proportion to its own needs or even disabling this auditory feature (depends on the implementation). Optionally, the HHD shall be able to guide the user by providing navigational messages rather than just tones.

Although single-tone audio messages provide no information about the exact position of the card reading device, this kind of messages shall be used as alarms for informing the user about the existence of a device in his/her near vicinity. Such messages shall be used with simple hand-held devices with no extra hardware for more complicated informative messages.

Moreover, if the application allows this case, the HHD shall provide more complex audio messages (preferably voice prompts) with navigational instructions assisting the user to find the card-reading device. In the case of a more sophisticated device, there shall be an option for the card-reading machine to scan its surrounding environment for visually impaired users.
Table 3
Machine originated voice message test

<table>
<thead>
<tr>
<th>Name</th>
<th>Date</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Voice Prompt</th>
<th>Engine / Quality of recording</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTS synthesized</td>
<td>□</td>
</tr>
<tr>
<td>Pre-recorded</td>
<td>□</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Voice Message List</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompt</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompt</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Total number of participating users

---

(carrying special equipment – see below) asking for special assistance. This will provide all necessary information for the machine to instruct precisely the user towards the exact position.

In the case of users with other than visual impairments, there shall be an option for the terminal device to send to the user’s HHD a map with detailed instructions and exact position of the machine.

Two options shall be considered for this implementation.

a) The signal is user initiated.

In this case, the user knows that a card-reading device is in the near vicinity and asks the machine for additional help if possible. This help shall be about either the position of the device and/or (more specifically) the slot where the user should place his/her card. In the case of simple alarm signals (as described above), the user can ask if any card-reading devices are near him and a pre-defined alarm signal shall provide this information to him.

b) The signal is produced automatically.

In this case, the user shall carry a smart-card ‘tagged’ with an ID for visually impaired users which starts the audio signal (or the voice navigational messages) when detected by a card-reading machine in a specified distance. The distance shall be appropriate in order not to disturb neighbouring services (if any) and suitable for the user to hear the audible help. Furthermore, if an automatic system is to be implemented, there shall be an option for the user (in the case of a HHD)
to disable all automatic informative messages, if not wanted.

As illustrated in Fig. 10, the form and type of communication between the machine and the HHD carried by the user, in order to achieve this kind of service can be anything from infrared to novel/emerging wireless technologies. The exact description of this communication and the hardware implementations of the smart card or the HHD are out of the scope of this work.

4. Testing procedures

4.1. Testing report template (example)

This appendix presents some example report templates for each test that is to be carried out.

4.1.1. User originated prompt test
See Table 2.

4.1.2. Machine originated voice message test
See Table 3.

4.1.3. Machine originated tone message test
See Table 4.

4.2. Testing procedures

4.2.1. User originated prompts (Section 3.1.2)
1. The equipment to be used is listed in Section 3.1.1.
2. Deployment of measurement setup depicted in Fig. 2 (Section 3.1.1).
3. Preparation of testing parameters (see report template in Section 4.1.1).
4. Initialization of noise source's SPL (set to minimum).
5. Test user is seated in front of simulated machine (PC) at 0.5 m distance.
6. Test prompt is spoken by the user.
7. Recognition result is reviewed (SUCCESSFUL or not).
8. In case of successful recognition, proceed to step 11.
9. Noise SPL is measured at the machine using high-precision sound-meter.
10. If recognition is unsuccessful, the noise SPL is marked in test report (see template in Section 4.1.1). Proceed to step 12.
11. Repeat steps 6 to 10 for increased noise level (same prompt).
12. Repeat steps 4 to 11 for different test prompt (same user).
13. Repeat steps 4 to 12 for different test user.

4.2.2. Machine originated voice message (Section 3.1.3)
1. The equipment to be used is listed in Section 3.1.1.
2. Deployment of measurement setup depicted in Fig. 2 (Section 3.1.1).
3. Preparation of testing parameters (see report template in Section 4.1.2).
4. Initialization of noise source’s SPL (set to minimum).
5. Test user is seated in front of simulated machine (PC) at 0.5 m distance.
6. Test prompt is played back.
7. User verifies the spoken message (TRUE or FALSE).
8. If TRUE, proceed to step 10.
9. If FALSE, proceed to step 13.
10. Noise SPL is measured at the machine using high-precision sound-meter.
11. Increase the number of total users who verified the played message for this particular noise SPL.
12. Repeat steps 6 to 11 for increased noise level (same prompt).
13. Repeat steps 4 to 12 for different test prompt (same user).
14. Repeat steps 4 to 13 for different test user.
15. Mark all data in test report (see template in Section 4.1.3) as derived from the above process.

5. Extensions – further testing requirements

The methodology described herein, regarding the ambient noise measurements, is directly applicable and can be immediately implemented, as soon as a desirable amount of test users is gathered (clustered as in Section 3.1.5). As pointed out earlier, the sample has to be large enough in order to achieve high accuracy in the extraction of results. Since the major conclusions that come out of this measurement process are highly dependent on the statistical properties that have been extracted by the lab trials, special care shall be taken into account when selecting the population groups for use in the tests (these groups define more or less the final statistical properties of the test results).

Special groups of people (i.e. with hearing disabilities) shall be weighted differently when selecting the users’ sample for the measurements. Special care should be considered for these people not only because they provide the worst-case scenario for our designing process, but also because they usually suffer from state analgesia in most of their everyday life activities. In fact, in order for this work to meet the “design for all” concept (and thus, ensure ‘access for all’ to terminal devices, including for people with disabilities) in audio and speech-related applications, the criterion for acoustic shielding (i.e. safe usage percentage) is derived only by groups of population who are hard of hearing.
Further extensions that shall be considered for the future are the categorizing of population groups according to specific applications. Surveys shall be carried out in the form of questionnaires in order to gather more information about the usage of publicly placed terminal devices by these groups of population. These surveys will be of much help in a later stage when the users' samples are built for our lab trials.

In the near future, there shall be an effort to perform as many trials as possible in order to acquire enough statistical data from the end users so that the design process is simplified.

As for the recommendations described in Section 3.2, it is clear that at the present moment the technological trends are favourable. HHDs, which can be used for testing, are already available in the market, as for example mobile phones equipped with GPS receivers and PDAs with mapping features. Modern communication protocols, such as GPRS, should be exploited for smart-card device-accessing.

A considerable effort towards additional development of such hardware is highly recommended. In addition, intense collaboration with the smart-card industry and the manufacturers of card-reading or other public terminal devices shall be considered in order to proceed with the necessary alterations to their current models.

Further collaboration with other service providers shall be taken into account in order to incorporate mapping services as described, for example, in Section 3.2.2. This service requires central management and cartographic databases interconnected to the terminal device infrastructure. There aren't any current protocols designed and tested for such applications.

carrying out lab-trials. The focus area of research was the investigation of the level of ambient noise that is acceptable before an acoustic shielding becomes necessary for card-reading or other terminal devices.

Taking into account aspects and recommendations presented in recent design considerations (that were investigated during the research work), the methodology contains a full-setup description, including all equipment required and testing procedures.

All available design guidelines that were extracted from previous standards and documents (related to telecom applications) have also been presented. Extensions based on current trends on communication technology are finally discussed.

6. Conclusions

In the present work, we have presented and drafted a measurement setup and complete methodology for