

## Design of a Multimodal Hearing System

Bernd Tessendorf<sup>1</sup>, Matjaz Debevc<sup>2</sup>, Peter Derleth<sup>3</sup>, Manuela Feilner<sup>3</sup>, Franz Gravenhorst<sup>1</sup>, Daniel Roggen<sup>1</sup>, Thomas Stiefmeier<sup>1</sup> and Gerhard Tröster<sup>1</sup>

<sup>1</sup> Wearable Computing Lab., ETH Zurich  
Gloriastr. 35, 8092 Zurich, Switzerland  
{lastname}@ife.ee.ethz.ch

<sup>2</sup> University of Maribor  
Smetanova ulica 17, 2000 Maribor, Slovenia  
{firstname.lastname}@uni-mb.si

<sup>3</sup> Phonak AG  
Laubisrütistrasse 28, 8712 Stäfa, Switzerland  
{firstname.lastname}@phonak.ch

**Abstract.** Hearing instruments (HIs) have become context-aware devices that analyze the acoustic environment in order to automatically adapt sound processing to the user's current hearing wish. However, in the same acoustic environment an HI user can have different hearing wishes requiring different behaviors from the hearing instrument. In these cases, the audio signal alone contains too little contextual information to determine the user's hearing wish. Additional modalities to sound can provide the missing information to improve the adaption. In this work, we review additional modalities to sound in HIs and present a prototype of a newly developed wireless multimodal hearing system. The platform takes into account additional sensor modalities such as the user's body movement and location. We characterize the system regarding runtime, latency and reliability of the wireless connection, and point out possibilities arising from the novel approach.

**Keywords:** multimodal hearing instrument, assistive technology

### 1. Introduction

Recent studies show that hearing impairment is increasingly affecting people worldwide [18, 27]. The World Health Organization estimates that in 2005 the number of people in the world with hearing impairments was 278 million, or about 4.3% of the world's population [38]. Permanent hearing loss is a leading global health care burden, with 1 in 10 people affected to a mild or significant degree [7].

The demographic change in the European Union (EU) [30] leads to a strong increase in the number of hearing impaired people. At the age of 40 years, the auditory perception begins to deteriorate and beyond 60 more than 50% of the people perceive deterioration of their hearing ability. A report prepared by the Action on Hearing Loss estimates that in 2005 more than 81.5 million

adults in the EU had hearing problems and that this number will increase to 90.6 million by 2015. This figure indicates that more than 14% of adults in Europe will experience hearing problems [26]. Consequently, new technologies to assist people with hearing impairment have emerged. They include: Digital hearing instruments (HIs) [1], cochlear implants [8], implantable hearing devices [31], and more advanced assistive technologies such as the frequency modulation (FM) system [11].

Context-aware HIs automatically adapt to the estimated user's current hearing wish by switching hearing programs, e.g. switching HI programs in quiet or noisy environments, in face-to-face conversation, traffic, or music [2, 6]. By obviating the need for manual selection, these HIs avoid drawing the user's and other's attention to the hearing deficit. This is especially useful in situations where constant change of hearing programs is necessary. Currently, automatic adaption is based on computational auditory scene analysis (CASA [36]), which analyses the acoustic environment for music, conversations, or relative silence. However, in the same acoustic environment an HI user can have different hearing wishes that require different behaviors from the HI. We refer to this as the *ambiguity problem* [32]. In such scenarios, the audio signal alone contains too little contextual information to determine the user's hearing wish. This especially applies for complex hearing situations with multiple sound sources or different possible activities or movements of the HI user. Therefore, there is a strong need to have additional sources of information available for hearing impaired people to support them in these complex scenarios.

**Paper Scope and Contributions** In collaboration with an HI manufacturer, HI acousticians and HI users, we developed a wireless multimodal hearing system. To improve the HI adaption in acoustically ambiguous situations, we consider sensor modalities in addition to sound analysis. In our approach we do not need to equip the user's environment with additional hardware. Instead, we developed a miniaturized wireless head movement sensor attachable to commercial HIs. The user's head movement data and the sound feature data from the HI are transferred wirelessly to a smartphone. In a future generation of HIs the accelerometer will be integrated into the HI obviates the need for any additional devices. A dedicated smartphone application fuses the information together with sensor information from the smartphone itself (e.g. GPS or phone acceleration) to derive an improved estimation of the user's hearing wish. We characterize the system regarding runtime, latency and reliability of the wireless connection, and point out possibilities arising from the novel technology.

## 2. State of the Art in Supporting Hearing Impaired and Review of Additional Modalities to Sound

### 2.1. Hearing Instrument Technology

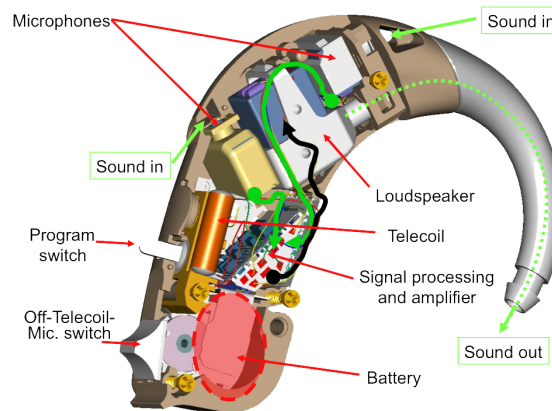
Most HIs use digital signal processing to process sounds from the acoustic environment and can be fitted to suit the HI user's individual hearing impairment

for different hearing wishes. The most common types of HIs on the market are behind-the-ear (BTE), in-the-ear (ITE), in-the-canal (ITC) and completely-in-the-canal (CIC). Frequency modulation (FM) systems are used to transmit distant sound directly to an HI user. An approach to overcome difficulties on the phone is to use magnetic induction with a dedicated coil (telecoil, T-coil), which allows different sound sources to be directly and wirelessly connected to the HI regardless of background noise. Figure 1 depicts the components of a BTE HI. Current systems include a digital signal processor, two microphones to enable directivity and conversion of sound, a miniature loudspeaker (receiver) a telecoil, and a high-capacity battery. The sound is conveyed acoustically via a tube to a custom ear mold (omitted in Figure 1). A HI performs the audio processing function of the HI encompassing audio pickup, processing, amplification and playback. The HIs at the user's ears can communicate with each other to stream sound and configuration data. They can also integrate a variety of accessories such as remote controls, Bluetooth or FM devices to form wireless networks, so called hearing instrument body area networks (HI-BANs) [4]. This motivates and supports our investigation of additional sensor modalities for HIs that may eventually be included within the HI itself, or within the hearing system of which the HI is one component. The automatic hearing program selection estimates the user's hearing wish based on the acoustic environment of the given situation and adjusts the sound parameters of the HI from among a set of hearing programs [16]. The classification is based on spectral and temporal features extracted from the audio signal [6] with regard to the audiometry data of the hearing impaired [17]. The hearing programs are optimized for different hearing wishes and selectively use advanced signal processing such as adaptive noise canceling, directivity ("beam forming") or multiband compression. Most current high-end HIs distinguish between four hearing programs: natural hearing (*Clean Speech*), speech intelligibility in noisy environments (*Speech in Noise*), comfort in noisy environments (*Noise*), and pleasure listening for a source with high dynamics (*Music*). Each hearing program represents a trade-off, e.g. between speech intelligibility and naturalness of sound. The automatic selection of hearing programs in HIs according to the user's current acoustic environment allows the hearing impaired to use the device with little or no manual interactions, such as program change. This also avoids drawing attention to the user's hearing impairment. Users consider automatic adaption mechanisms for changing the hearing programs as beneficial [6].

## 2.2. The Acoustic Ambiguity Problem

State-of-the art HIs which implement the automatic program selection based on auditory information only show intrinsic limitations. They select the most suitable hearing program according to the user's acoustic environment based on computational auditory scene analysis (CASA) [2, 6, 10, 36]. This approach performs well as long as the acoustic environment and hearing wish are directly correlated, e.g. when listening to direct speech in quiet environments. This assumption does not hold in all cases and leads to a limitation we call *Acoustic*

B. Tesselndorf et al.



**Fig. 1.** Components of a behind-the-ear (BTE) HI [25].

*Ambiguity Problem* [32]. Specifically, in the same acoustic environment a user can have different hearing wishes that require different hearing programs to be active. A sound-based processing cannot distinguish between these different hearing wishes. For example, when there is a person reading a newspaper in a busy train, the HI senses speech in noise. Solely based on this acoustic information, it is not clear whether the HI should optimize the sound processing for speech intelligibility or the user desires a hearing program that provides comfort in noise. Usually, HIs favor to optimize for speech, as social interactions, conversations in particular, are important for HI users. Unfortunately, the hearing impaired person's hearing wish in this case is not to listen to the passengers next to them. However, in a similar situation the passenger could actually favor to participate in a conversation. The HI detects the same acoustic environment and, thus, cannot select a suitable hearing program in both of the cases. Therefore, it is important to not only analyze the acoustic environment but to also assess the relevance of auditory objects [28]. The challenge here is not the effectiveness of the dedicated hearing programs but rather how to automatically adapt the hearing program to the user's specific hearing wish. Other typical situations in which state of the art HI program selection algorithms tend to fail include listening to music from the car radio while driving [13], participating in street traffic as a pedestrian [35], conversing in a cafe with background music [13], and watching TV [35].

### 2.3. Sensor Modalities Additional to Sound in Hearing Instruments

We can extract contextual information to support automatic hearing program selection using the following approaches:

**Head Movements and Mode of Locomotion** Head movements carry nonverbal cues during conversations [15]. Hadar et al. [12] found a relation of timing, tempo and synchrony in the listeners' head movements as responses to conversational functions. In our previous work [32] we confirmed head movements and head acceleration in particular, to be a relevant additional sensor modality to recognize the HI user's current hearing wish. We compared accuracies for hearing wish recognition between different on-body sensor positions in an office scenario and found the highest accuracy for the head position [32]. Besides the characterization of conversations, movement patterns at the head can as well be used to recognize head gestures or the user's mode of locomotion. Each of this additionally unveiled contextual information that supports the automatic hearing program selection. Atallah et al. demonstrate a triaxial accelerometer in their study that was placed inside an HI-shaped housing and was worn behind the ear to perform gait analysis [3]. Different activities such as reading, walking, lying down, walking slowly, and running fast could be detected.

**Smartphones** Smartphones could be used as user interfaces for hearing program selection, for configuration of manual contextual information, for user feedback to improve and personalize automatic program selection or to use the wirelessly connections, and most important as a rich sources of sensors. In a survey among 80 HI users we investigated the availability of smartphones and the users' acceptance for using them to improve HIs [35]. A share of 28% of the respondents always carries their smartphones, another 24% most of the time and we found a clear trend that younger age groups carry the smartphone more often than elderly. About 64% don't have concerns to leverage their phones for the HI, 24% are not sure and demand more information to decide, and the remaining part doesn't want the phone to communicate with the HI. According to this survey smartphones represent promising devices to enhance HIs.

**User Location** In our previous work [33] we investigated the potential of the user's location to improve automatic hearing program selection. It was evident, that the combination of the user's location with the mode of locomotion reveals significant correlations with the user's current hearing wish. A smartphone can be used to capture the user location via GPS or via wireless network fingerprints. Through the smartphone's connectivity, the internet can be used to associate the raw location to currently ongoing events which may also impact the user's hearing wish. Through these location-aware services, the automatic switching algorithm can consider if the user is listening to an open air concert or he is just walking in a park.

**Tagging** Tagging refers to putting dedicated beacons to objects. In a study by Hart et al. an attentive HI based on an eye-tracking device and infrared tags was proposed [14]. Wearers can "switch on" selected sound sources such as a person, television or radio by looking at them. The sound source needs to be

augmented with a tag device that catches the attention of the HI user. This way, only the communication coming from the sound sources, which are looked at, are heard.

In their study Choudhury and Pentland propose body-worn IR transmitters that are used to measure face-to-face interactions between people with the goal to model human networks [9]. The success of IR detection depends on the line-of-sight between the transmitter-receiver pair and all partners involved in the interaction needed to wear a dedicated device.

**Auditory Selective Attention** Capturing the user's auditory selective attention helps to recognise a person's current hearing wish. Research in the field of electrophysiology focuses on mechanisms of auditory selective attention inside the brain [29]. Under investigation are event-related brain potentials using electroencephalography (EEG). In a heart rate analysis, done by Molen et al. the influence of auditory selection on the heart rate was investigated [23]. However, the proposed methods are not robust enough yet to distinguish between hearing wishes in mobile real-life settings.

None of the mentioned approaches for additional sensor modalities have managed to reach integration into off-the-shelf HIs yet because either the performance is too poor or support for deployment in mobile settings is missing. Based on the review above we consider head movements and user location as promising modalities to be integrated into a multimodal hearing system.

#### 2.4. Actuator Modalities Additional to Sound in Hearing Instruments

Besides sensor modalities, research is ongoing considering actuator modalities to enhance HIs. The region behind the ears at the mastoid bone is one of the most sensitive head regions for vibrotactile stimulation [24]. In previous studies we investigated bilateral vibrotactile feedback, integrated into HIs for localisation [34]. An advantage of using vibrotactile feedback is that there is no interference with the sound from the acoustic environment. Moreover, tactile reaction time can be faster than auditory feedback [22].

In the study from Borg et al. a pair of glasses were enhanced with 4 vibrators and 3 microphones [5]. Sound source angles were located through the integrated microphones and translated to vibration patterns for the visually or hearing impaired wearer of the glasses. The approach did not focus on integration into HIs, instead the user needs to wear the proposed enhanced goggles. The results show an average share of correctly detected sound source angles of about 80% in a sound-treated room.

Weisenberger et al. present a vibrating device to be placed inside the ear mold to transduce sound into vibration [37] with a vibration frequency of 80 Hz for low frequency acoustic signals and 300 Hz for high frequency acoustic signals. Subjects have been tested in three tasks: sound localization, sound identification, and syllable rhythm and stress. Overall, the ear mold vibrator system

showed promising results as an actuator modality additional to sound, especially for aiding sound localization.

### 3. Multimodal Approach

By complementing sound with contextual information from additional information channels we provide the means to improve the automatic HI adaption in acoustically ambiguous situations. Based on our review of additional modalities to sound in the previous section we introduce a newly developed wireless multimodal hearing system, which takes into account the user's head movements and location. Previous studies have shown that additional modalities, head movements in particular, improve automatic hearing program selection [32, 33]. Furthermore, the integration of the head movement sensor allows the users to control their HIs using head gestures.

#### 3.1. Architecture of the Multimodal Hearing System

Figure 2 depicts an overview of the architecture of the multimodal hearing system and its communication. Commercial HIs are extended with a miniaturized triaxial acceleration sensor and communicate sound and acceleration data to the user's smartphone. The user wears a commercially available vendor-specific relay (shown here: Phonak iCube) around his neck to establish a wireless communication between the HI and the smartphone. The commercially available relay translates the bidirectional communication between the proprietary wireless protocols used by the HIs to a commonly used protocol, e.g. Bluetooth. The relay can stream sound from a phone or TV or play music from portable devices.

The protocols used for the system communication are denoted in brackets. The smartphone invokes an updated hearing program based on the analysis of the multimodal information. The system architecture is:

- *opportunistic*, i.e. the system falls back to a working stand-alone HI if the user's smartphone is currently not available, and
- leveraging a *smartphone* to collect and process sensor data and user interaction with the smartphone,
- *modular and scalable*, it is not limited to the selected set of modalities but can be extended to further modalities using standard protocols like WiFi, ANT+ or Bluetooth,
- *backward-compatible*, i.e. older HIs that support a relay or remote control can be upgraded with this technology.

Detailed descriptions of the architecture's building blocks are given in the following sections.

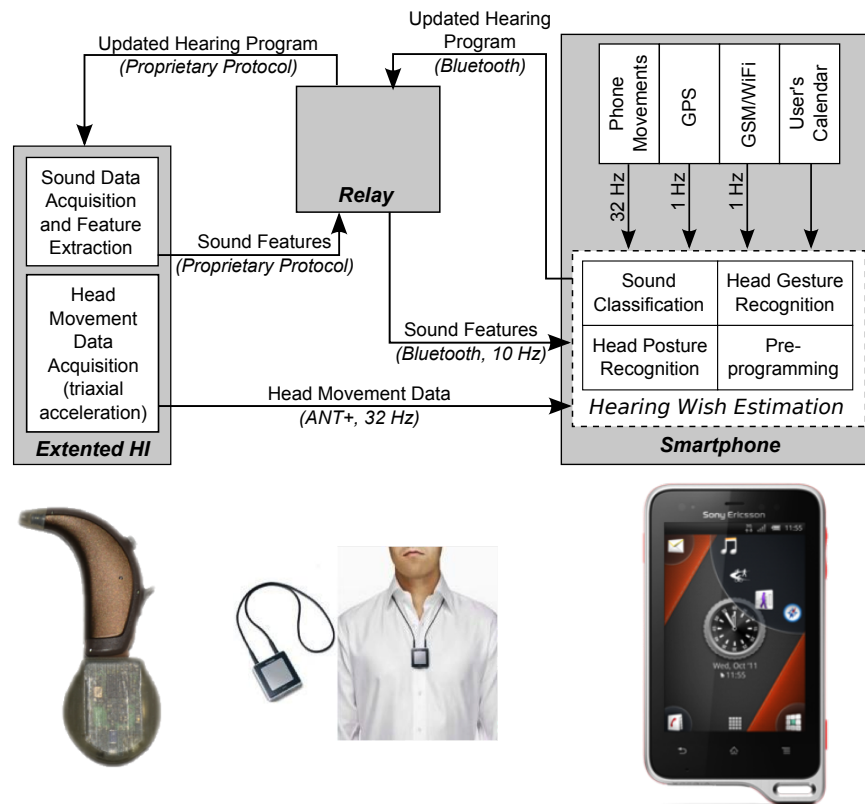


Fig. 2. Architecture of the developed prototype of a wireless multimodal hearing system.

### 3.2. System Communication

The communication paths between the system components and the corresponding communication protocols are shown in Figure 2. The smartphone communicates via Bluetooth with a vendor-specific relay (in our case a Phonak iCube as shown in the middle of Figure 2). The relay communicates over a proprietary wireless protocol with the HIs. Head movement data is sent to the user's smartphone. The acceleration sensor wirelessly communicates via the ANT+ protocol with the smartphone. The ANT+ protocol is a low power protocol designed for reliable transfer of data between sensors and display devices such as watches, heart rate monitors and bike computers. It ensures interoperability to guarantee seamless digital wireless communication in the 2.4 GHz license-free band. The transmitted data can be secured with a private network key. The adjustable sensor sampling and transmission rates are 32 Hz, which is sufficient for most activity recognition tasks [21]. The two HIs (HI model Phonak BTE Ambra 2012) worn at both ears can communicate with each other to syn-



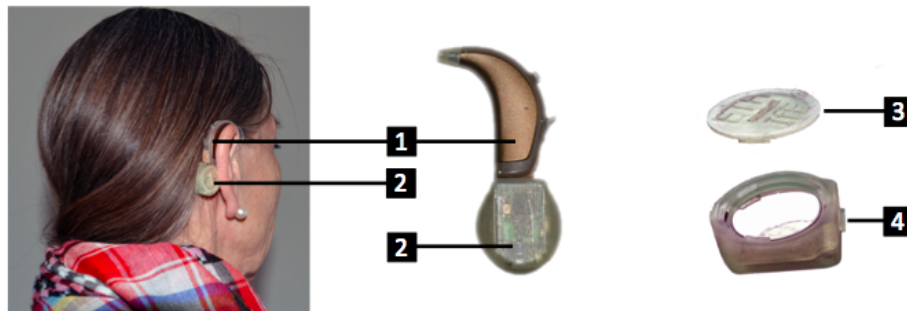
chronize manual hearing program switches and stream sound data. They wirelessly receive hearing program change requests invoked from the smartphone via the relay. In turn, the HIs send sound features, e.g. the sound level, to the smartphone. The smartphone acts like an automatic remote control. Our opportunistic approach ensures an automatic fallback to the original functionality of the HI in case the smartphone or the other additional system components are not available.

### 3.3. Extension of Hearing Instruments with a Head Movement Sensor

To produce the housing of the modular HI extension shown in Figure 3 we used a 3D CAD rapid prototyping method based on an acrylic photopolymer material. The head movement sensor has the dimensions of 22.6 mm × 21.6 mm × 10.3 mm and weighs 5.1 g (a typical HI weighs 4.7 g) and is attachable to commercial HIs. HIs to be used with the head movement sensor need to feature a slide mechanism at the lower end of the HI housing to mount the device. Most of the commercial HIs have this slide mechanism available to attach for example accessory FM receivers (replacements for the battery compartments for different types of HI are offered, which have the additional slide mechanism). The newly developed head movement sensor is based on the BodyANT platform [19]. It integrates a triaxial acceleration sensor (Bosch SMB380) and is powered by a 140 mAh CR1632 coin cell battery. The battery is placed in a battery compartment and can be replaced by using a coin to turn and open the battery cover. Acceleration can be measured with a bandwidth of up to 1.5 kHz in ranges of  $\pm 2\text{g}/\pm 4\text{g}/\pm 8\text{g}$  corresponding to a resolution of 4.0 mg/7.8 mg/15.6 mg. A stable clock cycle is provided using a 16 MHz crystal as clock source for both the radio transceiver and microprocessor. When active, the microprocessor periodically reads sensor values and sends messages to the radio transceiver according to the ANT message protocol. The transceiver continuously broadcasts the messages at a predefined message rate. If not activated, the microprocessor and radio transceiver are kept in power save mode. As mentioned before, the sensor allows the system to capture the user's head movements which is beneficial for improving HIs [32]. Due to the progress in the miniaturization of microelectromechanical systems (MEMS) and the reduction of power consumption of MEMS this technology manages to meet appropriate comfort requirements demanded by HI users [35].

To capture the HI user's head movements we opted for a modular solution, which is attachable to most of the state-of-the-art commercially available HIs. The design decision for an additional piece of hardware compares to integration of an acceleration sensor into the HI itself as follows:

- *Availability*: The modular solution is available now for all compatible state-of-the-art HI models; integration into the HI itself takes at least the time of an HI product development cycle for each single HI model we want to support.
- *Production costs for low volumes*: Our rapid-prototyping solution is cost-effective compared to the complete production of a next generation HI.



**Fig. 3.** The wireless multimodal hearing system comprises conventional commercial HIs (1) and the attachable head movement sensor (2). On the right the battery cover (3) and the slide mechanism (4) that sticks out to be attached to the HI, are shown.

- *Research platform:* Large scale real-life evaluations with HI users are needed before HI manufacturers decide to invest in this technology. These kinds of studies are feasible with our proposed platform. The modular approach to the multimodal HI extension allows to evaluate the benefits and limitations of multimodal HIs before a later integration of the additional modalities into the HI housing itself.
- *Backward compatibility:* The head movement sensor can be attached to any older HI that features a slide at the bottom, which is the case for most of the HI devices on the market. It represents a way to upgrade previous HI product generations.
- *System interoperability:* Only the communication with the HI relay is vendor dependent and needs to be adapted for different brands of HIs, the remaining system is vendor independent.

The main advantage of integration of the sensors into HIs over the modular solution are shared hardware resources, in particular the micro controller and transceiver for wireless communication. This way a saving in power consumption and form factor could be realized. Thus, both approaches have advantages and represent parallel solutions.

### 3.4. Smartphone Application

Smartphones are becoming the central computer and communication device in people's lives [20]. We leverage a smartphone as a component of the multimodal hearing system for the following reasons:

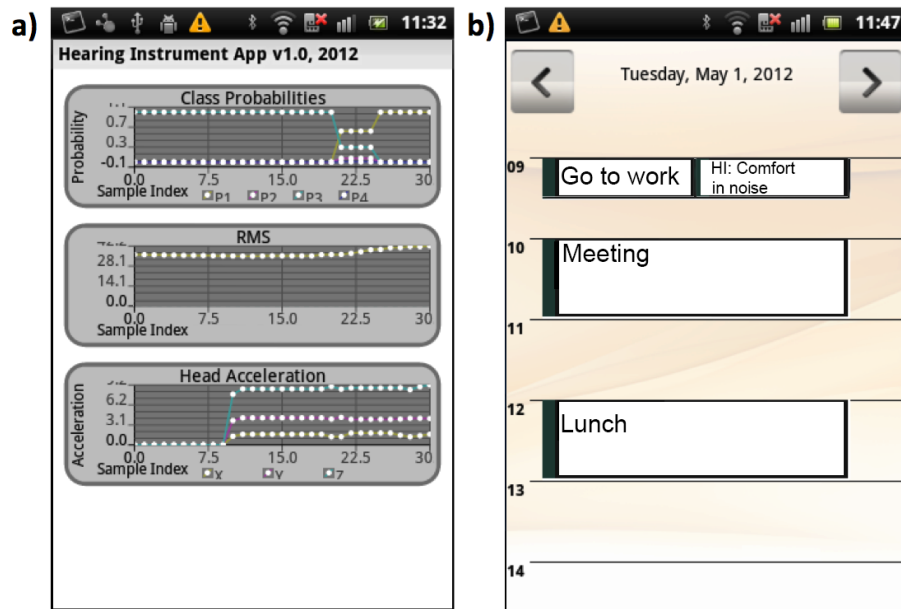
- *Processing power:* With an uprising trend modern smartphones offer processing resources up to 2000 MIPS, e.g. to execute complex context recognition algorithms.

- *Availability*: In previous work we identified smartphones to be available and accepted by their users [35].
- *Sensors*: Smartphones provide a rich source of sensor information such as an accelerometer, digital compass, gyroscope, GPS, microphone, WiFi, Bluetooth, ANT+, and camera.
- *Connectivity and scalability*: Smartphones provide internet access for cloud connectivity, access to the user's calendar, and support standard wireless protocols to extend the system with additional sensors in a modular way.
- *Extensibility*: Smartphones are programmable and additional applications can be developed, leveraging crowd sourcing and community driven software development.
- *User interface*: The smartphone can provide the user with a GUI to change more complex settings than possible with the buttons of the HI.

The newly developed smartphone application runs on any Android based phones that support the ANT+ protocol (we used a Sony Ericsson Xperia active smartphone). We opted for an Android-based software approach because it provides open source software development tools. The smartphone application performs the following tasks:

- receive and process sound features and acceleration data,
- provide a visual real-time data presentation of the sensor data,
- process the smartphone's local sensors,
- read the user's calendar and activate HI settings based on calendar entry, calendar entries that start with the special tag *HI:* are parsed by the smartphone application.
- use the user's location to activate room specific HI settings, e.g. reverberation characteristics of a concert hall; the application is prepared to download location-specific HI settings from a database from the cloud. This database can be populated by HI users to form a virtual HI community to share HI data, e.g. users can label their hearing wish in a specific place,
- perform classification of the sensor data,
- allow for data annotation, which is useful for HI developers and for HI end users also to train the HI using machine learning algorithms to let the HI automatically adapt to new context situations,
- enables to remotely log into the smartphone for debugging (with considerations for data anonymization and usage according to privacy laws and user agreement).

Figure 4 depicts screenshots of the smartphone application showing a data visualisation of sound features (hearing program class probabilities calculated within the HI, root mean square (RMS) sound level) and head acceleration, and a GUI that allows the user to program the HI using the smartphone's calendar. In this example the user programmed the hearing program "Comfort in Noise" to become active when traveling with public transport to his workplace, because he usually reads a newspaper and does not want the HI to optimize to the conversations around him. The smartphone application parses calendar entries



**Fig. 4.** Screenshots of the smartphone application: (a) Data visualisation of sound features and head acceleration, and (b) a GUI to program the HI using the built-in calendar.

that start with the special tag “HI:” and sets the hearing program accordingly at the start time of the entry.

## 4. System Characterization

### 4.1. Power Consumption

Battery lifetime is a critical factor for HI users [35]. The battery runtime for the head movement sensor with a 140 mAh CR1632 coin cell battery is more than 17 hours when the sensor sampling and transmission rates are 32 Hz. It increases to more than 4 days when the rate is reduced to 16 Hz, which is still sufficient for many activity recognition tasks [21]. The battery lifetime does not directly correlate with the power consumption due to the nonlinear discharge curve of the battery. Figure 5 shows the power consumption of the head movement sensor for different sampling rates. The ANT+ transmission rate was the same as the sensor sampling rate for the measurements. The runtime for the smartphone application is more than 16 hours with a 1200 mAh Lilon battery (3.7 V) when no other additional applications are being executed. Figure 6 shows the share of different components for the power consumption of the smartphone. We obtained the values by measuring the current from the

battery when having the different components individually activated. The runtime is sufficient for everyday use when it is recharged overnight. The battery lifetime of the relay we used is up to a 10 hours.

#### 4.2. Packet Loss

We measured the packet loss occurring during the transmission from the sensor to the smartphone for the user wearing the smartphone on different locations on the body: left and right front trouser pocket, left and right back trouser pocket, and left and right upper arm, attached with a strap that was shipped with the smartphone. The HI with the head movement sensor was worn at the left ear. For each phone location the user performed activities of daily living including sitting, standing, and walking. We calculate the packet loss as the rate of data packets that were not received at the smartphone using Equation 1:

$$\text{Packet Loss} := \frac{\#Sent\ packets - \#Received\ packets}{\#Sent\ packets} \quad (1)$$

Table 1 shows the measured packet loss values. Packet loss is low for all smartphone locations and renders the wireless communication suitable for application in multimodal hearing systems. The largest packet loss value was on the back right trouser pocket, where the user's body is in the line of sight of the transmission path, this way damping the signal. We did not observe any packet loss from the smartphone to the HI.

#### 4.3. Latency

Latency refers to the time between the onset of a head movement and the time the system reacts to it. The total latency is below 100 ms and is comprised as follows: The communication delay between head movement sensor and smartphone is below 6 ms. Acceleration data is usually evaluated using block processing with a sliding window. The main influence on the latency is the window

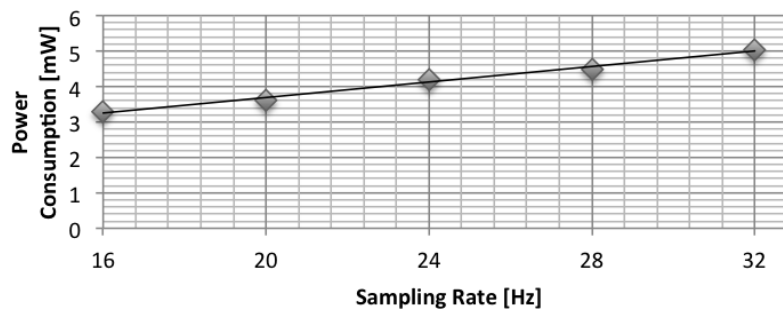
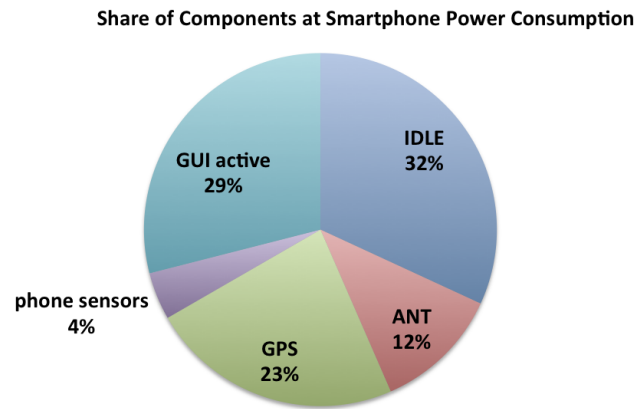


Fig. 5. Power consumption of the head movement sensor for different sampling rates.



**Fig. 6.** Share of different components for the power consumption of the smartphone.

**Table 1.** Packet loss for the system communication between head movement sensor and the smartphone on different locations on the body.

| Smartphone Location | trouser front pocket |       | trouser back pocket |       | upper arm |       |
|---------------------|----------------------|-------|---------------------|-------|-----------|-------|
|                     | left                 | right | left                | right | left      | right |
| Packet Loss [%]     | 0.04                 | 0.16  | 0.71                | 1.07  | 0.10      | 0.07  |

size used and is in the range of the duration of the head movement to be detected. The roundtrip communication time between the smartphone and the HI via the relay is 60 ms to 90 ms. Based on the above results we conclude the wireless communication of the system is fast enough to transfer sound features and head movement data and set HI programs.

#### 4.4. Potential of the Multimodal Hearing System

We designed the multimodal hearing system as a flexible platform to pave the way for use in a broad spectrum of applications. Its pervasive usability, small size, flexibility and possibility for long-term deployment allows us to bring the efforts from various research domains into daily use. Besides passive HI control through recognition of the user's context and active HI control, e.g. through user head gestures, we see further applications of the platform at least in the following research domains:

- activity recognition and pervasive computing,
- human computer interaction (HCI),
- computational social science,
- long-term behavior monitoring, and
- self-adapting and -learning systems.

#### 4.5. Limitations

A crucial aspect to make the system usable and reliable is to ensure a low complexity. The prototype allows for both automatic and manual control of the HI. It has to be mentioned that additional manual interaction with the smart phone contributes to high user interaction costs. For the automatic hearing program selection, however, the context is derived from the user's movement and the user does not have to change his behavior or learn to use the system. In a future generation of HIs the accelerometer will be integrated into the HI and there will be no need to carry additional devices such as the mobile phone to benefit from improved automatic hearing program selection. Additional functions such as programming the HIs using the phone's calendar, or using head gestures require more interaction and training of the user. The amount of required training will depend on the technical background of the user and the actual user interface for these features. This has to be studied and optimized in additional experiments.

The head movement sensor is implemented as an additional device attached beneath a HI, it is not yet integrated inside the HI itself. Therefore we face some additional limitations:

- Users of the prototype system need to carry the relay and phone as additional devices with additional weight and battery maintenance to benefit from any of the new functionalities.
- Our presented setup requires an Android-operated smart phone which features the ANT+ protocol. Up to now, there are only a very limited number of phones which support ANT+. We opted for ANT+ since this technology is consuming less power than conventional bluetooth transmission and might be widely established in the near future. However, the setup can easily be adapted to work with bluetooth to ensure compatibility with older phones. Alternatively, the Bluetooth low energy protocol could be used. In any case the used wireless protocol should be standardized across HI and smart-phone manufacturers, and could be integrated into the HI's relay.

However, these limitations will become obsolete for future generations of HIs that integrate the accelerometer. When the accelerometer is integrated into the HI, we will face a reduction in the HI's battery lifetime. However, the additional power consumption cannot be quantified as long as the additional functionalities are not finally integrated into the HI. Possible optimizations concerning battery lifetime strongly depend on the actual implementation and applied power management techniques of the final integrated device. We expect the impact to be low due to the availability of low-power MEMS accelerometers (e.g. 250  $\mu$ A for the ST LIS331H accelerometer).

## 5. Conclusion and Future Work

We presented a newly developed wireless multimodal hearing system. It represents an enabling technology, which raises new possibilities for HI users, HI acousticians and HI manufacturers:

B. Tesselendorf et al.

- to improve automatic hearing program selection in acoustically ambiguous situations using additional sensor modalities,
- to implement and investigate the benefit of a gesture controlled HI,
- to introduce location aware support,
- to let the HI user schedule his daily routines and corresponding HI programs,
- to allow HI manufacturers remote debugging capabilities in the field to improve the product (with considerations for data anonymization and usage according to privacy laws and user agreement)
- to support the HI acoustician with fitting the HI to the user by providing multimodal contextual information from real-life situations, in which the user's appreciate modified HI sound settings

With a day of battery lifetime, reliable wireless connection and sufficiently small latency (below 100 ms), we found the system to be viable both as a research platform and as a working prototype for a potential product in the HI market. The head movement sensor can be used to upgrade previous HI generations and enables evaluations towards integration of sensors into HI itself.

In future work we plan to conduct long term real-life studies with HI users. Besides assessing the user acceptance, we want to confirm the benefit of multimodal hearing systems, already demonstrated in laboratory settings [32], for real-life situations. We further plan to assess the benefit of a head gesture controlled HI, particularly for elderly people.

**Acknowledgments.** This work was part funded by CTI project 10698.1 PFLS-LS. We especially thank Nadim El Guindi and Stephan Koch for valuable discussions and support with the relay framework.

## References

1. Abrams, H.: Digital hearing aids. *Ear and Hearing* 30(3), 385 (2009)
2. Allegro, S., Büchler, M., Launer, S.: Automatic sound classification inspired by auditory scene analysis. In: *Consistent and Reliable Acoustic Cues for Sound Analysis (CRAC)*, one-day workshop, Aalborg, Denmark, Sunday September 2nd 2001 (directly before Eurospeech 2001). Citeseer (2001)
3. Atallah, L., Aziz, O., Lo, B., Yang, G.Z.: Detecting walking gait impairment with an ear-worn sensor. *International Workshop on Wearable and Implantable Body Sensor Networks 0*, 175–180 (2009)
4. Biggins, A.: Benefits of wireless technology. *Hearing Review* (11 2009)
5. Borg, E., Ronnberg, J., Neovius, L., Lie, T.: Vibratory-coded directional analysis: Evaluation of a three-microphone/four-vibrator DSP system. *J. of rehabilitation research and development* 38(2) (2001)
6. Buchler, M., Allegro, S., Launer, S., Dillier, N.: Sound Classification in Hearing Aids Inspired by Auditory Scene Analysis. *EURASIP Journal on Applied Signal Processing* 18, 2991–3002 (2005)
7. Cavender, A., Ladner, R.: Hearing impairments. *Web Accessibility* pp. 25–35 (2008)
8. Chapman, R.: Cochlear implants. *Ear and Hearing* 29(3), 477 (2008)



9. Choudhury, T., Pentland, A.: Sensing and modeling human networks using the sociometer. In: ISWC. p. 216. IEEE Computer Society, Washington, DC, USA (2003)
10. Eronen, A., Peltonen, V., Tuomi, J., Klapuri, A., Fagerlund, S., Sorsa, T., Lorho, G., Huopaniemi, J.: Audio-based context recognition. *IEEE Transactions on speech and audio processing* 14(1), 321 (2006)
11. Fitzpatrick, E., Séguin, C., Schramm, D., Armstrong, S., Chénier, J.: The benefits of remote microphone technology for adults with cochlear implants. *Ear and hearing* 30(5), 590 (2009)
12. Hadar, U., Steiner, T.J., Clifford Rose, F.: Head movement during listening turns in conversation. *Journal of Nonverbal Behavior* 9(4), 214–228 (12 1985)
13. Hamacher, V.: Signal processing in high-end hearing aids: State of the art, challenges, and future trends. *EURASIP Journal on Applied Signal Processing* 18(2005), 2915–2929 (2005)
14. Hart, J., Onceanu, D., Sohn, C., Wightman, D., Vertegaal, R.: The attentive hearing aid: Eye selection of auditory sources for hearing impaired users. *Human-Computer Interaction –INTERACT* (2009)
15. Heylen, D.: Challenges Ahead: Head Movements and other social acts in conversation. In: AISB 2005, Social Presence Cues Symposium (2005)
16. Keidser, G.: Many factors are involved in optimizing environmentally adaptive hearing aids. *The Hearing Journal* 62(1), 26 (2009)
17. Keidser, G., Yeend, I., O'Brien, A., Hartley, L.: Using in-situ audiometry more effectively: How low-frequency leakage can effect prescribed gain and perception. In: *Hearing Review*. vol. 18, pp. 12–16 (2011)
18. Kochkin, S.: MarkeTrak VIII: 25-year trends in the hearing health market. *Hearing Review* 16(10) (2009)
19. Kusserow, M., Amft, O., Tröster, G.: Bodyant: Miniature wireless sensors for naturalistic monitoring of daily activity. In: *BodyNets* (2009)
20. Lane, N., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., Campbell, A.: A survey of mobile phone sensing. *Communications Magazine, IEEE* 48(9), 140–150 (2010)
21. Maurer, U., Smailagic, A., Siewiorek, D., Deisher, M.: Activity recognition and monitoring using multiple sensors on different body positions. In: *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. IEEE* (2006)
22. Mohebbi, R., Gray, R., Tan, H.: Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 51(1), 102–110 (2009)
23. Molen, M., Somsen, R., Jennings, J.: Does the heart know what the ears hear? A heart rate analysis of auditory selective attention. *Psychophysiology* (1996)
24. Myles, K.: Guidelines for Head Tactile Communication. Tech. rep., Army Research Lab Aberdeen Proving Ground Md Human Research And Engineering Directorate (2010)
25. Naylor, G.: Modern hearing aids and future development trends
26. Rohdenburg, T., Huber, R., van Hengel, P., Bitzer, J., Appell, J.: Hearing aid technology and multi-media devices. *ITG Fachtagung für Elektronische Medien* 13 (2009)
27. Shargorodsky, J., Curhan, S., Curhan, G., Eavey, R.: Change in Prevalence of Hearing Loss in US Adolescents. *JAMA* 304(7), 772 (2010)
28. Shinn-Cunningham, B.: I want to party, but my hearing aids won't let me! In: *Hearing Journal*. vol. 62, pp. 10–13 (2009)
29. Shinn-Cunningham, B., Best, V.: Selective attention in normal and impaired hearing. *Trends in Amplification* 12(4), 283 (2008)
30. Siltan, W., Schoenefeld, J.: An aging europe. demographic transition and economic implicatios (2010)

B. Tessendorf et al.

31. Snik, A.: Implantable hearing devices for conductive and sensorineural hearing impairment. In: Auditory Prostheses, Springer Handbook of Auditory Research,, vol. 39, pp. 85–108. Springer New York (2012)
32. Tessendorf, B., Bulling, A., Derleth, P., Feilner, M., Roggen, D., Stiefmeier, T., Tröster, G.: Recognition of hearing needs from body and eye movements to improve hearing instruments. In: International Conference on Pervasive Computing (2011)
33. Tessendorf, B., Derleth, P., Feilner, M., Kettner, A., Roggen, D., Stiefmeier, T., Tröster, G.: Identification of relevant multimodal cues to enhance context-aware hearing instruments. In: 6th International Conference on Body Area Networks (Bodynets) (2011)
34. Tessendorf, B., Derleth, P., Feilner, M., Roggen, D., Stiefmeier, T., Tröster, G.: Improving game accessibility with vibrotactile-enhanced hearing instruments. In: 13th International Conference on Computers Helping People with Special Needs (IC-CHP), Linz, Austria (2012)
35. Tessendorf, B., Derleth, P., Feilner, M., Roggen, D., Stiefmeier, T., Tröster, G.: Survey on the potential of additional modalities for hearing instruments. In: Proc. of the 38th Annual Convention for Acoustics (DAGA 2012) (2012)
36. Wang, D., Brown, G.: Computational auditory scene analysis: Principles, algorithms, and applications. IEEE Press (2006)
37. Weisenberger, J., Heidbreder, A., Miller, J.: Development and preliminary evaluation of an earmold sound-to-tactile aid for the hearing-impaired. J Rehabil Res Dev 24, 51–66 (1987)
38. WHO: Deafness and hearing impairment, Fact sheet N°300, February 2012, <http://www.who.int/mediacentre/factsheets/fs300/en/index.html>
39. Wilson, B., Dorman, M.: Cochlear implants: a remarkable past and a brilliant future. Hearing research 242(1-2), 3–21 (2008)

**Bernd Tessendorf** received the Diploma degree in Electrical Engineering and Information Technology (Dipl.-Ing.) from RWTH Aachen University, Germany, in 2006. In 2007, he received the Diploma degree in Economics (Dipl.-Kfm.) from Fernuniversity Hagen, Germany. In 2008, he joined the Wearable Group at the Electronics Laboratory at ETH Zurich.

**Matjaž Debevc** received Ph.D. degree in computer science from University of Maribor in 1995. He is currently an Associate Professor in Computer Science at the Faculty of Electrical Engineering and Computer Science, University of Maribor. His research interests include human-computer interaction, e-learning, user interface design, adaptive user interfaces, internet applications, interactive TV, distance education and applications for disabled people.

**Peter Derleth** received Ph.D. degree in Physics from University of Oldenburg, Germany, in 1999. Since 2000 he is employed by Phonak AG, Switzerland. Since 2005 he is leading the Algorithm Concepts Group in the Research Department.

**Manuela Feilner** studied Electrical Engineering at ETH Zurich and received her Ph.D. degree from EPFL (Swiss Federal Institute of Technology Lausanne)

in 2002. Since 2003 she works for Phonak AG, Switzerland. Since 2008 she is working in the group of Advanced Concepts and Technologies.

**Franz Gravenhorst** received his Diploma (Masters) degree in Electrical Engineering and Information Technology at Karlsruhe Institute of Technology, Germany, in 2010. He worked at the Research and Development department of an automotive company in Detroit, USA. He joined the Wearable Group at ETH to start his PhD in 2010.

**Daniel Roggen** received PhD degree at the Laboratory of Intelligent Systems of EPFL, Switzerland, in 2005. In his PhD he developed bio-inspired electronic circuits with fault-tolerance, learning, and developmental capabilities. Since 2005 he is Senior Researcher in the Wearable Computing Lab at ETH Zurich.

**Thomas Stiefmeier** is a senior member of the research staff at the Electronics Laboratory at ETH Zurich. He received his master's degree in electrical engineering from the University of Technology Darmstadt and his PhD degree from ETH Zurich. Thomas Stiefmeier is CEO and a co-founder of Amphiro AG, a start-up company in the emerging field of smart water metering.

**Gerhard Tröster** studied electrical engineering in Darmstadt and Karlsruhe, Germany, earning his doctorate in 1984 at the Technical University of Darmstadt about the design of analog integrated circuits. During the eight years he spent at Telefunken (Atmel) Heilbronn, he headed various national and international research projects centered on the key components for ISDN and digital mobile phones. Since 1993 he directs the Electronics Laboratory at the ETH. In 1997 he co-founded the spin-off u-blox ag.

*Received: April 23, 2012; Accepted: November 23, 2012.*

