Abstract: In this paper the concept and architecture of a monitoring and assistant system for people with hearing deficiencies will be described and a first user study will be presented. Core of this concept is the implementation of algorithms that are usually used in hearing aids at a central Home Information and Communication (HIC) platform. This platform integrates formerly separated devices like TV, telephone, and home automation. Thereby, the TV becomes the central human machine interface of the household linking together most of the acoustic inputs listened to by the hearing impaired in his/her home environment. Using the central HIC platform the acoustic inputs can be amplified and adapted according to the specific hearing impairment and the individual thresholds of hearing. These thresholds will be estimated during the initial diagnosis at the audiologist and transmitted to the HIC platform. Supportive Audio Signal Processing (SASP) algorithms are used to fine-tune the amplification according to the acoustic situation of the living room. These fine-tuning, frequent tests of speech understanding and monitoring of the individual adjustment of volume and sound controller are used as a feedback to the audiologist. In this way the transition from mild to moderate or severe hearing loss can be supervised, the use of hearing aids can be motivated and the adjustment of hearing aids can be improved and adapted to the individual situation of the patient. The ideas of this sound processing and alerting system at home, the so called Hearing at Home System (HaH) were presented to 62 hearing aid users and non-users (“young-old” to “old-old”; mean age: 69yrs.) of three countries (Netherlands, Sweden, and Germany) with a different degree of hearing loss (from slight to moderate). The first results of the user requirement study showed, that most of the users will benefit from the HaH environment system, and the results are very useful for further research and development, also in the realm of tele-monitoring applications.

This project has been funded by the European Commission (FP6-2005-IST-6 project 045089).
1 Motivation

The increase in the number of persons having hearing deficiencies is strongly connected to the demographic change in the European Union. The sense of hearing begins to degrade from the age of 40 onwards and estimates indicate that more than 50% of people over the age of 60 have some degree of hearing loss. The demographic change and age related hearing losses will result in an increasing number of hearing-impaired people in the European Union. Estimates of the number of people in the EU with various hearing disabilities in the early 1990s report a number of 6% of hard of hearing and 0.1% of deaf people. A study of the Institute of Hearing Research (UK) estimates that 81,536 thousand adults will have a hearing loss in Europe by 2005. By 2015 the figure will increase to be 90,588 thousand. This means that more than 14% of adults in Europe will have hearing problems. This makes people with hearing disabilities one of the largest groups facing the challenge that communication is mainly audio-based.

Hearing aids are supposed to support the hearing-impaired but the acceptance and use of such devices is influenced as much by psychosocial factors as by the performance of the hearing aid itself. Statistics show that on average people with a hearing deficiency tend to wait seven years before they search for assistance and 75% of hearing-impaired who could profit from a hearing aid don’t own one. This behaviour cannot be explained with reasonable arguments. Unlike eyeglasses, hearing aids are not considered a fashionable article and the general nonsensical preconception imputes “intelligence” to people wearing glasses and “stupidity” to people with a hearing problem. The latter is the effective outcome of the disability when contributing to a conversation: giving the impression of being an unintelligent individual. Unfortunately this stigma prevents the hearing-impaired from acquiring an ALD (Assistive Listening Device), which would abolish this impression.

The remaining 25% of users who already own an ALD don’t necessarily use it. Users regularly complain that hearing aids are “fragile devices that often do not work properly.” This has multiple causes: Hearing aids need to be fitted. Fitting is the process to parameterize a hearing aid to “fit” a single end user’s personal needs for his unique hearing disability. This process usually does not deliver satisfactory results right away resulting in several visits at the hearing aid dispenser for readjustments. Depending on the circumstances (severeness of hearing disability, quality of hearing aid, expertise of hearing aid dispenser, user’s readiness to cooperate, distance to travel, etc.) a user may come to the early conclusion, that the device is useless or faulty by design.

2 Approach

The used approach aims at integrating hearing support technologies within common digital TV/STB-like Home Information and Communication (HIC) platform devices and connecting this device with a residential gateway for external communication and integration of home appliances. Using this way, the acceptance barrier is lowered to a minimum. At the same time the available TV screen can be used to support speech intelligi-
bility by visual support. The description of the approach is divided into two parts: (1) the support system and (2) the tele-monitoring system.

2.1 Support System

The support system integrates all relevant audio sources and other information sources to enhance the pure audio signal and extend it by video output.

![Software architecture of the HIC platform.](image)

Figure 1 shows the software architecture of the HIC platform. It combines audio technologies like the Jack framework [JA07] and the Master Hearing Aid (MHA, [GR06]), the Synface visual lip reading support [BE04], a multi-media framework (FFmpeg, [FF07]), a controller running within an OSGi framework [OS07], and a residential gateway. The controller observes the environment (input audio/video channels, remote control, signals form home appliances through the residential gateway) and derives control commands to the other modules depending on the classified scenario (see chapter 3). A datapoint abstraction layer is used to convert the sensor inputs from different modules to a unified format and to convert control commands and parameter settings to the device dependent format. Depending on the detected scenario the video de-multiplexer prioritizes a certain input channel (streams coming from digital TV, DVD player, radio or telephone) and divides it if necessary in separate audio and video streams.
The audio stream will be processed by the MHA to improve the audio signals in terms of noise reduction and audio scene classification (Global-, G-SASP) and according to the individual hearing impairment of the user (Individual-, I-ASAP). The output of the G-SASP module is transmitted to the Synface subsystem. Because pure audio signals (e.g. telephone conversation) do not deliver extra visual information, lip-reading is impossible. Even video and TV streams, like news documentaries with background voice-over, do not provide this support in general. Synface is an animated artificial face showing natural lip movements based on speech input signals. This face will be synchronized with the original video stream and displayed at the TV screen.

Next to audio input from TV, radio, video and phone, home event detection is audio based, too. A ringing doorbell or finishing alarms of microwave ovens or washing machines are events of this kind. Especially the latter being quite often high-pitched beeping signals are often missed by the hearing impaired. By the integration of the residential gateway such events can be detected on the HIC platform and appropriate visual and audio signals can be generated.

2.2 Tele-monitoring System

The tele-monitoring system is used for the initial parameterization of the supportive audio signal processing (ASAP) algorithms, especially the individual SASP, the adaptation of the specific home environment and the generation of feedback to the audiologist. Aim of this approach is the optimal fitting of the audio processing algorithms according to the individual hearing impairment in the preferred environment of the patient. In this way the communication abilities of the patient and the general acceptance of hearing aids will be improved. It is assumed that while the hearing loss is mild to moderate the acceptance of the described home-based support system will be higher than a typical hearing aid.

The three major parts of the tele-manipulation system will be realized as follows (see Figure 2)

1. **Initial Configuration**: After initial diagnosis of the individual hearing loss using an audiogram the audiologist selects usually specific chains of audio processing algorithms and fits the parameter of this device like volume, frequency-specific amplification, compression, etc. These parameters will be transferred to the HIC platform either manually or by a web interface and corresponding paths of signal processing blocks within the MHA will be configured and the parameters will be set.

2. **Adaptation/ Tuning Wizard**: At home the user can evaluate the configuration of the processing chain and different parameters sets by testing the different scenarios. Different audio-visual test methods will be implemented and accompanied by graphical and video support. In this way, this fine-tuning of parameters takes into account the individual room acoustics, the selected modality (e.g. switching the parameters if switching from watching TV to make a telephone call) and the a priori unknown transfer functions of headphones or loudspeakers.
3. **Feedback**: Apart from using the results of the test methods for the fine-tuning at home, the results will be stored for analysis and further adaptations by the audiologist. Using a web interface and given access by the patient the audiologist can read the data and will be supported by trend analysis (e.g., increased volume adjustment or changing of sound control adjustments over a longer time period can be interpreted as a sign of increasing hearing loss).

![Feedback Diagram](image)

Figure 2: Schematic diagram of the tele-monitoring system and the underlying Master Hearing Aid (MHA) framework. Several signal processing blocks (Pn) are combined to form processing paths.

### 3 Scenarios

In the first step three different scenarios are addressed, which represent typical home living situations with the need of massive communication. All scenarios will be based on the core SASP system embedded in the HIC platform:

1. **Phone communication scenario**: Input from different communication channels (phone, VoIP, Intercom) will be processed using the Supportive Audio Signal Processing (SASP) and output to the communication device (e.g., headset or phone receiver). In parallel, an animated synthetic face is rendered for supporting lip-reading.

2. **Entertainment scenario**: Digital (MPEG) stream is being decoded and separated into video and audio stream. SASP enhances audio to the personalized needs of the user and extracts feature tokens to control the Synface application. In parallel the video stream is being delayed to be synchronized with the enhanced audio stream.
(expected latency about 200 ms). A Picture in Picture (PIP) engine combines original video stream with visual hearing support.

3. **Home automation scenario**: Home appliances are being connected via a home area network (HAN) to the residential gateway. Acoustic events signalled by the home appliances (e.g. microwave, washing machine, etc.) will be transferred to the HIC-platform which will act depending on the user-defined fitting: acoustic alarm signal via the HIFI-system, visual notice on TV-screen, control of home automated light sources.

All scenarios are designed based on the results of questionnaires (end-users, organizations representing hearing-impaired people and professionals in the field, see chapter 4) in order to ensure practical relevance of the technical work. After the development and implementation the demonstrators will be carefully evaluated with end-users in order to ensure usability. Figure 3 shows the combination of the entertainment scenario with the home automation scenario. While the user is watching TV, signals from the home appliances may incorporate the time-shift of the TV stream, plays a signal tone and displays a message at the screen.

![Figure 3: Combined entertainment and home automation scenario based on the Home Information and Communication (HIC) platform.](image)

### 4 User requirement Study

The general objective of the end-user study is the definition of user needs, use case scenarios and assistive communication applications in a home environment focusing on TV, radio, telecommunication and home-automation applications. This work will be carried out with representative users from the Hearing at Home (HaH) target population, i.e.
elderly people who have problems with their hearing but do not like using a hearing aid, or who do use a hearing aid but are not satisfied with its performance in the home environment. The results of the user requirement study in combination with other not here reported research methods (e.g., focus groups) will form the basis for the development and quality assessment of the proposed system.

4.1 Methods

Description and recruitment of the sample: Altogether 62 interviewees, 33 male (53%) and 29 female (47%) were found and participated in individual interview sessions (duration: 1 to 1.5 hour) in Germany (Hörtech, Oldenburg), Sweden (KTH, Stockholm) and The Netherlands (Viataal, Sint-Michielsgestel). Almost one quarter of our interviewees lives alone (n=16). Of the other 46 persons, 12 have a partner who experiences hearing problems as well. From our sample we can deduce, therefore, that about one fifth of all 62 potential users live with another person with hearing problems.

The first selection criterion was the age: The mean age was 69.0 years (SD=10.4 years); they were divided into three age groups, as follows: “young-old” (50-64 yrs.): 22 persons, mean age 57.9 (SD=4.5), “middle-old” (65-75 yrs.): 19 persons, mean age 69.2 (SD=3.6), and “old-old” (75+ yrs.): 21 persons, mean age 80.6 (SD=4.4).

The second selection criterion was the status of hearing impairment: subjects with a slight hearing loss (27 persons with an average PTA of 38 dB, SD=4.5 dB) and moderate hearing loss (35 persons with an average PTA of 56 dB, SD=9.4 dB).

The third recruitment criterion was the usage of hearing aids: 61% of our interviewees were using hearing aids; 76% of them were bilaterally aided. It was important that the users, as a potential target group for the HaH system, were not satisfied with their hearing aids.

The questionnaire and procedure: The questionnaire contained 137 questions. Most questions were 5-point scales or yes/no answers, a small number required open answers. The questions were divided into the following six categories: 1.) personal, incl. hearing problems [ME05] and hearing aid use [CO02], 2.) TV use, 3.) radio/hifi use, telephone use, 4) home automation and attitude towards, and experience with technology following [MO04], and first-approach of a HaH solution.

4.2 Results

Subjective reported hearing abilities: comparison with reference group: The hearing abilities of the interviewees were checked with the Oldenburger Inventory-R [KO04], which contains questions about how well people manage in three situations: (Q) hearing in quiet, (N) hearing in noise/reverberation, and (L) localisation of sound. The scale is defined from 0 to 100% subjective hearing ability. We compared our results with those of a sample of a reference sample [ME05]. Apart from two minor significant differences, there seems no clear difference between our interview group and the reference group.
regarding hearing loss, and status of provision with hearing aids, indicating that we do not have a systematic bias in our sample.

The repeated measurement factor hearing dimension (hearing in quiet vs. in noise vs. localisation) showed a strong main effect: hearing in noise (estimated mean: 48.1) was significantly worse than the other dimensions of hearing ability (mean hearing in quiet: 71.5, mean localisation: 69.3; F(2,114)=39.0, p<0.01). A first conclusion is that hearing in noise is the biggest problem for our target group, implying that improvement of signal-to-noise ratio would be a promising solution for the HaH system.

Experience with technology: A cluster analysis of these statements yielded two clear clusters of technology experience and attitude towards technology. The first cluster of the sample (69%) represented a positive attitude towards technology and the second cluster (31%) a more or less negative opinion about technology (Figure 4). However, the clusters did not account for any of the variance we found in the interview results, indicating that the assessment of the features of the HaH system are dependent from the general attitude towards technology. Interesting detail is the finding that the persons in the non-technical cluster responded much more positively to the statement "I learned / wished to have learned the handling of a PC" than to the other statements. In this respect they behave similarly to the technically oriented persons. This finding is in contrast to results from Mollenkopf & Kaspar [MO04, database from 1999].

![Cluster solution on the basis of attitude towards technology (1="not all true" to 5="very true")](image)

Use of television: 92% of the interviewees have a television set in the living room and most of them stated that the living room is the most important room where they watch television. Many people have a second television set, notably in the bedroom, study or a special TV room.
Our interviewees generally have a sloping audiogram, normally found in presbyacusis, and might therefore benefit by compensating for high tones. Despite this fact, hardly anyone ever adjusts the tone: 92% does this 'never' or 'seldom'. The reason for this is probably, as many interviewees mentioned that they do not know how to adjust it (or they were not even aware of the possibility).

Most people (93%) find it quite easy to operate the remote control (RC). Most of the people seem to use 2 to 5 functions, although the young-old seem to use relatively more than the other age groups: they are overrepresented in the ">5" category and underrepresented in the "2-3" category. However, the effect is not significant. Also, no statistical significant differences occurred concerning the clusters of technological experience. Only a minority (32%) owns a set-top box for digital television, but those who do, do not find it difficult to operate.

Most of the people find it important to watch the news, current affairs and documentaries (85%), and entertainment and sports programmes (69%). This result showed statistical significance ((F(1,59)=17.0, p<0.00); see Figure 5). There is a tendency for the “old-old” to find television watching more important than the other age groups, although they do not watch it more often. There is quite a range in the number of minutes of television watching: on average our interviewees watch 153 minutes a day, with a standard deviation of 84 minutes which is in accordance with an actual Dutch study [BR06].

![Figure 5: Importance of telecasts and age groups (1=not important to 5=very important)](image_url)

Interviewees having a partner (n=46) mostly watch TV with their partner: 98% does so regularly to always. About three quarters of the interviewees indicated that they cannot understand TV programmes well at normal volume: for that reason, they will put the volume up. The co-watcher, which is almost always the partner (if there is one), says the volume is too loud in 50% of the cases.
It was shown that not more than 28% of the interviewees use a headset (normal or special FM/infrared) while watching TV. Almost 68% is aware of the availability and possible benefit of headsets. During the interview we showed pictures of various headsets and explained their possible advantages. Even so, less than half of the interviewees said they would consider using it in future, indicating that the acceptance is not very high to use such devices.

**Use of telephony:** The phone is deemed important for social contacts (92% of the interviewees says so) and information gathering or making appointments (79%). There is a wide range in the number of calls per day. The mean is almost 5 with a standard deviation of 6.8 (the median is 3 calls a day or 25 minutes).

Our interviewees have an average of 2.25 telephone devices at home. Wireless phones are used a lot: 80% of the interviewees are using a DECT, cell phones are used by 78%, whereas 68% also uses the standard type, a fixed phone. Most people have phone calls in the living room (84%) but in more than a third of all cases the phone call is made from another room, mainly the kitchen or the study (37-45%). If the phone is answered while the television is on, 90% of the interviewees put the TV off or to mute.

**Home automation and alerts:** Over 40% of the interviewees reported that they have “regularly” to “always” problems recognizing the signals from household appliances and apparatus. Interestingly we did not find any statistical significant relation of these problems with hearing loss or hearing aid use. In general, people find it important to hear the door bell, telephone, fire alarm and (microwave) oven, but other home automation appliances such as dishwasher, washing machine, tumble dryer, intercom are not so important.

**First-approach opinion of HaH solution:**

**Sound processing:** First of all, we asked what people think of the sound processing, incl. home appliances' attention signals. Most interviewees scored in the answer categories 'moderately' to 'very much' useful. The usefulness of HaH's sound processing for TV watching (81%) was rated significantly higher (p<.05) than the usefulness of HaH for either radio listening (60%), making a phone call (56%) and perceiving attention signals (54%); see Figure 6. But we find no statistical significant relation between the acceptance of the HaH unit in sum and the degree of hearing loss. But, posthoc tests via t-test indicated that people with a moderate hearing loss differed significantly (p<.05) from people with a mild hearing loss: they think that the HaH is more helpful for phone calls. Additionally, no significant differences were found for a preference of the HaH system for technological friendly vs. sceptic participants. In answer to our question whether interviewees would like to control the sound manually or whether they preferred automatic control, we found two equally strong opinions: 52% prefers automatic control, 48% prefers a manual control.

**Signals:** The most favourable signals to be alerted from the HaH system are sounds (“beeps”) with 66%, followed by LED (52%), flashes (39%), and a text in the display (37%). Open-ended questions regarding to additional signals reveal that the participants suggested to using pictograms.
Synface Animation: We also asked the interviewees to give their opinion about the potential usefulness of Synface. Almost half of the interviewees think Synface might be useful for TV watching (47%) or making phone calls (50%). Smaller minorities thought Synface is useful for radio or intercom (both 35%). The valuations were in sum one scale point lower in comparison to the whole HaH system as a sound processing unit. There is a statistical tendency (p=.07) for the main effect hearing loss, indicating that the moderately hearing impaired interviewees rated the usefulness of Synface higher than the ones in the slight hearing loss group, indicating that this might be more a feature for moderate to severe hearing handicapped persons.

Use of headsets or loudspeakers in the HaH environment: The interviewees gave their preference for listening either to the loudspeaker or through a (special) headset: almost two thirds (65%) prefer the loudspeaker. Interesting, though, is the deviant response within the group of interviewees with a moderate hearing loss who do not use a hearing aid: 63% of them is in favour of listening through a headset. Of course, this would be the group for whom a headset is necessary, because of the relatively high gain that is required to make the sounds audible. Over 70% over the subjects thinks that free communication while watching TV with their partners is important.

5 Summary and conclusions of the user study

A number of questions in our interview addressed the actual hearing disabilities that our target group persons experience. We found that the majority regularly to often have difficulties with listening in noise, listening TV and radio at normal volume level, per-
ceiving attention signals from door bell, telephone, fire alarm and various home appliances.

The implication of all this is that there certainly is a potential use for an HaH application that processes sounds to make them more audible, thereby compensating the audiometric loss, and to increase the signal-to-noise ratio of television (and radio) programmes. The interviewees themselves seem aware of these potentials, especially for the purpose of watching TV.

People often watch TV with their partner, if they have one. They find it important to be able to communicate with their partner while watching TV. Only a minority listens to the television through a (special) headset; the loudspeaker is preferred, even for those who are aware of special headsets that provide better perception. The implication of this TV watching behaviour is that both processed and unprocessed sounds should be available simultaneously through a headset and a loudspeaker. When watching TV alone, the hearing disabled person may choose the sound output device. When watching with the partner, the hearing disabled person uses a headset, the partner will actually be free to choose, but will probably prefer a loudspeaker. The headset should be equipped with microphones for picking up the partner's voice.

Some interviewees have problems perceiving certain voices (unclear, high-pitched) through the telephone, but in general, phoning does not cause many problems within this group. The consequence may be that there is no great need for a visual support like Synface that makes lipreading possible. Nevertheless, some people say that they find the Synface option interesting, but they do not believe it would help them. What is the implication of this? We know that visual support will certainly help persons with a moderate hearing loss, although the benefit will probably be most clear with severe hearing losses (PTA over 70 dB HL). Also, the lack of interest for Synface may also be related to a feeling of being stigmatized. This is a point to take into account into future research, the focus group discussion, as we feel that in the interview setting, the benefits of Synface may have been underexposed.

Less than half of the interviewees have problems hearing attention signals from home appliances. They find it most annoying when they cannot hear the door bell, telephone, oven or microwave oven, but do not think the signals of other home appliances very important. An exception is the fire alarm, but it may be assumed that the sound output level of such an alarm will be enough to alert a person with a hearing loss up to 70 dB HL. When it comes to improving audibility of the attention signals, most people would be in favour of an acoustic signal or a LED. The implication may be that a well audible signal, possibly in combination with a LED, display or a pictogram on the TV screen, is required to make people aware of the home appliances when they are watching television or listening to the radio.

Whether people belong to the 'non-technical' cluster or not, they generally do not have problems operating the remote control or the set-top box and, as we found, do not keep away from a PC. We may therefore assume that absence of technical insight or interest does not preclude people from accepting or operating technical devices such as the HaH.
system per se. As long as the technique is inside and the user interface is easy, there should be no problem.

The reported results of the user requirement study built up the first step to formulate new assistive technologies and devices for hearing impaired persons in their home environment, also in the realm of user-oriented, acoustic input and output scenarios for tele-monitoring systems. It was shown that the proposed system and the scenarios are on the right track, because the presentation of the HaH environment, especially the sound processing applications, caused a great acceptance of our interviewees. For the next steps a fine-tuning of the use case scenarios and a final definition of the applications will be carried out.

Literature
