User Experience in Training a Personalized Hearing System

Gabriel Aldaz[1], Tyler Haydell[1], Dafna Szafer[1], Martin Steinert[2], Larry Leifer[1]

[1] Mechanical Engineering, Stanford University
{zamfir,thaydell,dszafer,leifer}@stanford.edu

Abstract. In this paper, we introduce Awear, a context-aware hearing system comprising two state-of-the-art hearing aids, an Android smartphone, and a body-worn Streamer to wirelessly connect them. Awear aims to improve the sound quality perceived by individual hearing aid wearers by learning from their stated preferences. Users personalize, or "train," the system by performing several *listening evaluations* daily. The Awear app features two types of user-initiated listening evaluations, the A/B Test and the Self-Adjustment Screen. After a longitudinal (6-week) study in which hearing impaired participants (n = 16) used Awear, 10 of the participants stated a preference for training their system using the A/B Test, 3 preferred using the Self-Adjustment Screen, and 3 stated No Preference. Of the 10 who chose the A/B Test, 7 named simplicity or intuitiveness as the primary reason for this preference. We also found a strong correlation between user level of functionality and listening evaluation preference, and a supplemental interview (n = 24) verified this correlation. Lastly, we discuss the most important aspects of the user experience: cognitive, functional, and psychological dimensions.

Keywords: User experience, hearing aids, smartphones, personalization, mobile apps, listening evaluations.

1 Background

According to the World Health Organization, 360 million people (over 5% of the world's population) have hearing loss [17]. In the United States, hearing loss affects 34.25 million people, more than 10% of the population; surprisingly only 1 in 4 Americans with hearing loss uses hearing aids [12]. In Germany, France, and the UK, hearing aid adoption is only slightly higher, perhaps due to government incentives [10].

For those who can benefit from hearing aids, have the financial means, and are willing to take action, the first step to better hearing is often to undergo a thorough evaluation by a hearing care professional. A fundamental part of the hearing care professional's audiological assessment is to use an audiometer to generate an audiogram, a representation of the softest sounds a patient can hear at different frequencies in each ear. Another important consideration during the audiological assessment is choosing a fitting rationale. A fitting rationale is a prescriptive formula

A. Marcus (Ed.): DUXU 2014, Part III, LNCS 8519, pp. 3–14, 2014.

describing the electro-acoustical characteristics of the hearing aids in response to given inputs.

The audiogram and fitting rationale only partially address the fact that hearing loss is an individual experience. How the brain processes sensory information varies from person to person, and two people with similar audiograms and fitting rationales may experience different levels of satisfaction in the long run [5]. Therefore, most fittings require multiple office visits to fine-tune the parameters correctly.

A person wearing advanced hearing aids in their daily life benefits from the devices' real-time digital signal processing, implementing a selected fitting rationale and employing sophisticated algorithms to automatically select settings to enhance speech or suppress transient noise. Hearing aids have many automatic features, such as turning Directionality and Noise Reduction on and off, as well as classifying the current sound environment (quiet, noise, speech in noise, etc.). Nonetheless, a hearing aid has very limited sensor inputs, relying entirely on its two on-board microphones to collect information about incoming sounds. Furthermore, current adaptive algorithms in hearing aids lack the ability to improve performance over time in response to sensor inputs.

In 2007, Edwards described future hearing aids that could allow fine-tuning to be done automatically outside of the clinician's office and that would have the ability to learn, making them "intelligent." Edwards predicted a greater industry shift from uniformity of patients and universal treatment to individuality of patients and therapy [6]. While the relatively slow processing speed, small storage capabilities, and limited user interface of hearing aids have largely prevented Edwards' vision from becoming reality so far, these limitations may be overcome by regarding the smartphone as part of an intelligent hearing system.

The smartphone has the potential to revolutionize the way users interact with their hearing aids, providing unprecedented personalization and increased satisfaction levels. Smartphones provide a powerful mobile computing platform, with formidable sensing, processing, communication, and memory capabilities. While the hearing aids will continue to perform the real-time sound processing, the smartphone opens new possibilities for an additional layer of processing that takes into account factors that change on a much longer timescale – such as the user's sound environment, location, or even her intentions – known in the field of computer science as context awareness [15].

2 Related Work

Although Awear is, to our knowledge, the first context-aware hearing system, there are a number of hearing aid-related apps on the market today. These may be segmented as enabling users to interact with their hearing instruments through their smartphone, akin to a remote control, or as transforming a smartphone itself into a "personal amplification device."

2.1 Remote Controls

Today's advanced digital hearing aids can store several preset programs in their memories, each created for a different listening situation. There may be a general-purpose program, one for listening to music, and one for understanding speech in noisy environments. Program changes, along with volume control, are the primary mechanisms for users to make adjustments to improve sound quality. Presently, users may change program and volume via buttons on the hearing aids themselves, on gateway devices, or on dedicated remote controls.

ReSound [13] and Starkey [16] are two manufacturers that have launched apps that enable remote control of hearing aids via a smartphone. The ReSound Control and Starkey T2 Remote (Figure 1) have simple and straightforward interfaces that allow a user to control her device's volume and program on her smartphone. Although Control and T2 Remote do not offer any novel features, they open the door to a world of possibilities by giving users a much higher degree of convenience and discretion in controlling the settings of their hearing devices.

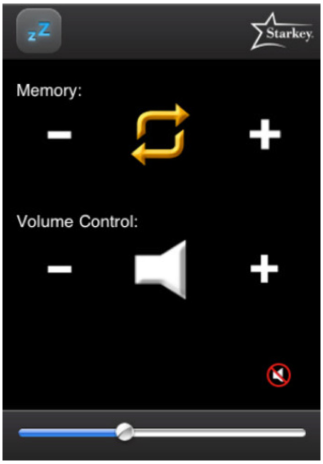

Fig. 1. Starkey T^2 Remote User Interface

2.2 Personal Amplification Apps

These apps use a smartphone's built-in microphone to pick up sound, amplify and adjust different qualities, and transmit the resulting sound using an off-the-shelf earpiece. Their primary advantage of these personal amplification apps is that they eliminate the need for users to purchase expensive hearing aids. However, software running on smartphones lacks the sophisticated digital signal processing of hearing aids and are often used without consulting a hearing care professional [2]. Furthermore, earpieces such as headphones and ear buds are not meant to be worn all day.

BioAid [1], created by a group at the University of Essex in England, aims to make sound amplification more accessible to those who cannot afford the steep cost of real hearing devices. The app offers six basic settings that the user can fine-tune using six additional screens. Finally, a "Noise Gate" allows the user to adjust background noise between 0-100%. This complicated interface raises the question of whether or not the average user will know how to interpret and make use of moderately technical language, or if users will ever even explore all twenty-four possible settings instead of settling for one of the presets. The same holds true for other available apps that try to mimic hearing aids without the actual devices. Apps including the SoundAMP from GingerLabs [8], HearYouNow from ExSilent [7], and others allow the user to control sound quality by adjusting the volume of individual frequencies, the volume of each ear, and even in some cases the controls for a dynamic compressor.

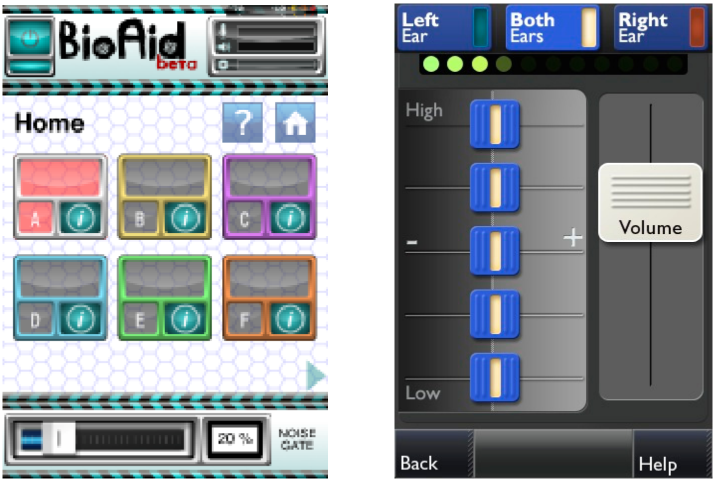

Fig. 2. BioAid (left) and SoundAMP (right)

3 Awear: A Context-Aware Hearing System

3.1 Hardware

The Awear off-the-shelf hardware (Figure 3) comprises a pair of state-of-the-art hearing instruments, an Android-based smartphone, and a gateway, a body-worn device with a built-in microphone that wirelessly links the hearing aids to a mobile phone. In the future, we expect technological advances to enable direct 2-way communication between the hearing aids and the smartphone.

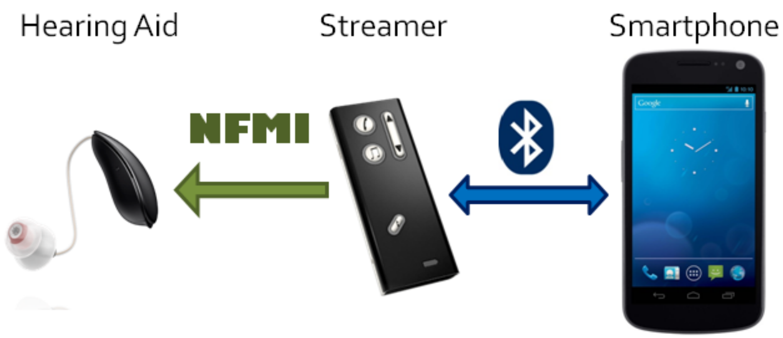

Fig. 3. Hardware components of the Awear system

3.2 User Interface

Although the Awear user interface works as a remote control – allowing the user to switch programs and change volume on the hearing instruments – its primary purpose is to solicit current user preferences for hearing aid microphone Directionality (on or off) and Noise Reduction (on or off)[1]. These interactions take place via user-initiated *listening evaluations*. There are two kinds of listening evaluations: the A/B Test and the Self-Adjustment Screen.

A/B Test. The A/B Test (Figure 4) is a standard method of comparing two variants applied here in a novel way to compare two hearing aid settings, A and B. The participant listens to both settings and gives a subjective, relative evaluation (“A is better,” “B is better,” or “No difference”). The settings corresponding to A (for example, “Directionality on and Noise Reduction off”) and B (for example, “Directionality off and Noise Reduction off”) are randomized during each evaluation.

Fig. 4. A/B Test Screen Sequence

[1] Hearing aid microphones can be configured to pick up sounds uniformly from all directions (Directionality off) or primarily from the front (Directionality on). With Noise Reduction on, the hearing aids attempt to reduce amplification of non-speech signals while preserving the amplification of speech signals. Turning Noise Reduction off simply disables this digital signal processing step.

Self-Adjustment Screen. The self-adjustment screen (Figure 5) allows users to turn Directionality on/off and Noise Reduction on/off until they find their absolute preferred setting. Whereas the A/B test conceals the names of the settings, the Self-Adjustment Screen explicitly labels the settings, making them visible to the users.

Fig. 5. Self-Adjustment Screen (Directional Listening has just been turned on)

4 Method

Between July and December of 2013, we conducted a longitudinal (6-week) study of Awear with 16 participants (10 men, 6 women, mean age = 55.5 years) from the San Francisco Bay Area, enrolled by various private audiology clinics. All participants were at least 18 years old, had sufficient cognitive ability to successfully operate a smartphone and Awear software, and had moderate to severe hearing loss.

To understand which type of listening evaluation (A/B Test or Self-Adjustment Screen) benefited users more, we used a within-subjects study design. In the pre-interview, we asked participants about their background and demographics. Most of the data collected was self-explanatory, with the exception of "Experience" and "Functionality." Experience refers to that person's familiarity and use of computer and mobile technology. Table 1 is adapted from the typology in [9].

Table 1. User Experience Levels with Computer and Mobile Technology

Experience	Typical Assets	Typical Actions
High	Laptop Tablet Smartphone	Use computers for programming, creative work, or other advanced tasks Use mobile devices for calling, texting, information, sharing, and entertainment
Medium	Laptop Cell Phone	Use computers regularly for information, sharing, and entertainment Use mobile phones for calling and texting
Low	Desktop Cell phone	Use computer to explore internet and stay in touch with friends Use mobile phones for calling and texting
Zero	Landline phone	No online access

Functionality indicates how deeply a person delves into the features of a product. For example, a Low Functionality user may prefer to use point-and-shoot cameras, whereas a High Functionality person might want to explore the camera's technical features. Table 2 illustrates typical attitudes of these user types with regard to hearing aids.

Table 2. User Functionality Levels with Hearing Aid Technology

Functionality	Typical Assets	Typical Actions
High	Hearing aids with multiple programs and volume control	Actively change programs, use each one in its intended listening situation. Change volume several times per day
Low	Hearing aids with only one program, no volume control	Put on the hearing aids and forget *Laissez-faire* attitude

During in-situ use, participants were instructed to wear the system as often as possible during the test period of 6 weeks. Participants were instructed to do approximately 8 listening evaluations daily, especially whenever they encountered an interesting or challenging listening situation. At the conclusion of the test period, we asked the participants open-ended questions about hearing situations that they experienced and the performance of the context-aware hearing system.

In January 2014 we conducted a brief follow-up interview where we presented the two types of listening evaluations, in random order, to an additional 24 hearing impaired persons. The purpose of this interview was to confirm or refute the qualitative listening evaluation preference results of the previous study.

5 Results

The 16 participants completed a total of 3,754 listening evaluations (5.5 per person per day). At the conclusion of the test period, we asked each participant to give us his or her subjective assessment of the two types of listening evaluations: A/B Test, Self-Adjustment Screen, or No Preference. To find the key predictors of listening evaluation preference, we compared participant responses to pre-interview data (Table 3).

As Table 3 indicates, 10 participants preferred the A/B Test, 3 preferred the Self-Adjustment Screen, and 3 stated No Preference. Figure 6 lists the reasons given for choosing the A/B Test. Simplicity and intuitiveness, with 7 responses, was by far the most common response.

Table 3. Test Person Characteristics and Listening Evaluation Preference

	Attributes						
Person	Age	Gender	Years with Hearing Aids	Experience	Functionality	Smartphone Owner	Evaluation Preference
1	21	M	11	High	Low	No	AB
2	34	M	3	High	Low	Yes	AB
3	41	F	7	Medium	Low	Yes	AB
4	46	M	15	High	High	Yes	SA
5	46	M	4	Medium	High	No	SA
6	48	F	<1	Medium	Low	Yes	AB
7	51	F	<1	Medium	Low	Yes	AB
8	54	M	5	Medium	Low	Yes	NO PREF
9	54	M	<1	Low	Low	No	AB
10	61	F	<1	Medium	Low	Yes	AB
11	64	F	15	Medium	Low	Yes	AB
12	68	M	15	High	High	Yes	SA
13	73	F	<1	High	Low	Yes	AB
14	73	M	8	High	Low	Yes	NO PREF
15	76	M	23	Low	Low	No	NO PREF
16	79	M	4	Medium	Low	Yes	AB

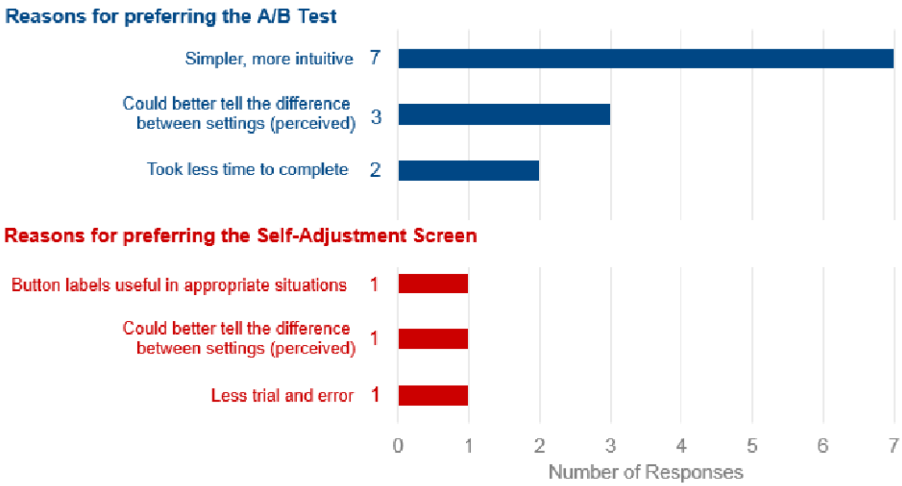

Fig. 6. Participant-given reasons for listening evaluation preference

A simple statistical analysis of the data in Table 3 indicated a strong correlation between Functionality and Listening Evaluation Preference. Of the 13 participants with Low Functionality, 10 preferred the A/B Test and 3 stated No Preference. On the other hand, all 3 participants with High Functionality preferred the Self-Adjust Screen.

Since 16 respondents was inadequate for generalizing the result of this investigation, we conducted a brief follow-up interview where we presented the two types of listening evaluations, in random order, to an additional 24 hearing impaired persons. With a sample size of to 40 (19 men, 21 women, mean age = 63 years), we computed the chi-square test for independence of each attribute with respect to listening evaluation preference.

Table 4. Chi-square test for independence p-values per user attribute

Attribute	***p* (*n* = 16)**	***p* (*n* = 40)**
Functionality	0.000335	0.000230
Gender	0.0561	0.254
Years of Experience with Hearing Aids	0.221	0.485
Age	0.735	0.574
Experience	0.768	0.744
Smartphone Owner	0.837	0.861

With both the small and larger sample sizes, we see that Functionality is the only attribute where the p-value is less than the significance level (0.05). Thus, we conclude that there is a relationship between Functionality and listening evaluation preference.

Overall, 23 respondents preferred the A/B Test, 10 preferred the Self-Adjustment Screen, and 7 claimed No Preference. Reasons for preferring the A/B Test were "simpler, more intuitive" (18) and "took less time to complete" (4). Reasons for preferring the Self-Adjustment Screen were "button labels useful in appropriate situations" (8) and "less trial and error" (2).

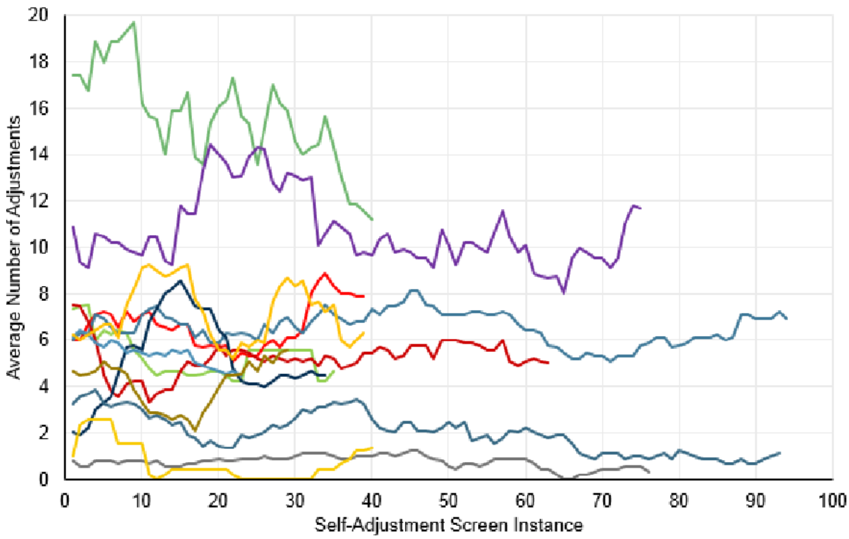

Fig. 7. Moving Average for Number of Adjustments by Participant

For the Self-Adjustment Screen, we wondered if participant engagement, measured by number of adjustments, would lessen over time. We recorded every time the participant pressed either the Directional Listening or the Noise Reduction button during a particular instance of a Self-Adjustment Screen. Figure 7 shows a running average of the number of adjustments each participant made over time. In general, we see that each participant maintained a constant level of number of adjustments. This indicates that the participants remained engaged and continued to try to find the best setting in each listening situation during the entire training phase.

6 Discussion

We conclude with a discussion of the most relevant user experience dimensions regarding training a personalized hearing system: cognitive, functional, and psychological.

Cognitive. The process of selecting participants for the longitudinal study revealed that presently there is a significant portion of the hearing impaired population for which smartphone apps are not appropriate. The average age of our study participants was 55.5 years, compared to the average age of a hearing aid user, which is around 70 years [11]. Two major exclusion criteria were low cognitive skills and Zero Experience (did not own computers or cell phones, were never online), which are more prevalent among the 80+ age group. In the follow-up interview, only 4 out of 9 (44.4%) in the 80+ age group felt hearing aid personalization could be of value, while 100% in the <80 group expressed it could be useful. However, in the future, we can expect an increasing percentage of hearing impaired people to be familiar with smartphones and apps.

To gain a better understanding of the participants' mental models [3], we asked them during the post-interview what they thought the Awear app did and how it worked. Participant 15 had the most difficulties. He had no recollection of the Self-Adjustment Screen, although the data showed that this screen did come up regularly and he did not make any adjustments. Regarding the A/B Test, he stated that, "Program A is best for speech comprehension and not bad for music. Program B is best for music listening – sounds are clearer and expanded and subtly muffled for speech." Although Settings A and B were randomized, that did not prevent Participant 15 from creating his own, very detailed mental model.

As Figure 6 illustrates, 3 of the participants who preferred the A/B Test stated that it was because they could better tell the difference between settings in the A/B Test than in the Self-Adjustment Screen. On the other hand, 1 participant who preferred the Self-Adjustment Screen claimed the opposite. The settings were inherently the same in both types of listening evaluations, but individual participants started to find patterns where perhaps none existed. Nonetheless, 14 out of 16 participants reported positive feelings of clarity, competence, and mastery of both types of listening evaluations.

Functional. All 16 participants understood the purpose of the listening evaluations and the overall app functionality. We uncovered that, of the user attributes elicited, listening evaluation preference correlated highly with the Functionality user attribute. At least in the field of hearing, one option for future apps would be a layered approach: a simple, intuitive UI (such as the A/B Test) as the default, while offering affordances for a more in-depth, technical UI (such as the Self-Adjustment Screen) that only the High Functionality users will choose to access.

Psychological. We left the choice of when they wanted to complete their listening evaluations entirely up to the participants. This participant-triggered self-reporting contrasts with the traditionally time-triggered Experience Sampling Method (ESM) used in psychology experiments [4]. Although participants averaged 5.5 listening evaluations per day – below the target of 8 per day – they did so with no reminders or further encouragement from us. Of the 21 participants who started the longitudinal study, 5 dropped out within the first few days for various reasons. The remaining 16 participants completed the 6-week study, with only 2 being asked to retain the system longer to accumulate more data. This shows the extreme level of participant motivation when involved in co-creation and personalization in an issue as fundamental to well-being as hearing.

7 Conclusion

Advances in wireless technology are quickly eliminating the need for intermediate body-worn devices such as the Streamer used in the Awear experimental setup. A new generation of Personal Sound Amplifier Products (PSAPs), which can be sold directly to consumers as electronic devices and are exempt from government regulations, connect directly to smartphones [14]. Hearing aid manufacturers are following suit. The user experience for these apps, as well as those designed by the hearing aid companies, are going to shift from simple remote control operation to incorporating increasingly sophisticated levels of personalization.

We have found that participant-triggered listening evaluations allow the user to become part of the hearing device fine-tuning process, thus leveraging to some extent the positive motivational effects of co-creation and participatory design. As customary, no single user experience design is right for every user. Designers in the arena of hearing-related apps must take into careful account the cognitive, functional, and psychological aspects of the user experience.

Acknowledgements. The authors are indebted to all personnel at Oticon A/S who supported this research, as well as the study participants. The project was funded by the Oticon Foundation.

References

1. BioAid, http://www.bioaid.org.uk
2. Coleman, M.: There's a hearing app for that. Hearing Journal 64(11), 12,14,16 (2011)
3. Craik, K.: The nature of explanation. Cambridge University Press (1943)
4. Csikszentmihalyi, M., Larson, R.: Validity and reliability of the experience-sampling method. Journal of Nervous and Mental Disease 175(9), 526–536 (1987)
5. Dalebout, S.: The Praeger Guide to Hearing and Hearing Loss: Assessment, Treatment, and Prevention. Praeger: Westport (2009)
6. Edwards, B.: The Future of Hearing Aid Technology. Trends in Amplification 11(1), 31–45 (2007)
7. ExSilent, http://www.exsilent.com
8. GingerLabs, http://www.gingerlabs.com
9. Hougaard, S., Ruf, S.: EuroTrak I: A consumer survey about hearing aids in Germany, France and the UK. Hearing Review 18(2), 12–28 (2011)
10. Horrigan, J.: The Mobile Difference, Pew Internet and American Life Project (2009)
11. Kochkin, S.: MarkeTrak VII: Hearing Loss Population Tops 31 Million People. Hearing Review 12(7), 16–29 (2005)
12. Kochkin, S.: MarkeTrak VIII: 25-year trends in the hearing health market. Hearing Review 16(11), 12–31 (2009)
13. ReSound, http://www.resound.com
14. Romano, T.: Better Hearing Through Bluetooth. New York Times (January 15, 2014)
15. Schilit, B., Adams, N., Want, R.: Context-Aware Computing Applications. In: 1st International Workshop on Mobile Computing Systems and Applications, pp. 85–90 (1994)
16. Starkey, http://www.starkey.com
17. World Health Organization Fact Sheet 300 (2013)