Rethinking Hearing Aid Fitting by Learning From Behavioral Patterns

Benjamin Johansen
Tech. University of Denmark
Dept. of Applied Mathematics
and Computer Science
Richard Petersens Pl., bld. 321,
DK-2800 Kgs. Lyngby, Denmark
benjoh@dtu.dk

Per Sandholm
Oticon A/S
Kongebakken 9,
DK-2765 Smørum, Denmark
psan@oticon.com

Michael Kai Petersen
Eriksholm Research Centre
Rørtangvej 20,
DK-3070 Snekkersten, Denmark
mkpe@eriksholm.com

Jakob Eg Larsen
Tech. University of Denmark
Dept. of Applied Mathematics
and Computer Science
Richard Petersens Pl., bld. 324,
DK-2800 Kgs. Lyngby, Denmark
jaeg@dtu.dk

Niels Henrik Pontoppidan
Eriksholm Research Centre
Rørtangvej 20,
DK-3070 Snekkersten, Denmark
npon@eriksholm.com

CHI'17 Extended Abstracts, May 06-11, 2017, Denver, CO, USA ACM 978-1-4503-4656-6/17/05. http://dx.doi.org/10.1145/3027063.3053156

Abstract
The recent introduction of Internet connected hearing instruments offers a paradigm shift in hearing instrument fitting. Potentially this makes it possible for devices to adapt their settings to a changing context, inferred from user interactions. In a pilot study we enabled hearing instrument users to remotely enhance auditory focus and attenuate background noise to improve speech intelligibility. N=5, participants changed program settings and adjusted volume on their hearing instruments using their smartphones. We found that individual behavioral patterns affected the usage of the devices. A significant difference between program usage, and weekdays versus weekends, were found. Users not only changed programs to modify aspects of directionality and noise reduction, but also continuously adjusted the volume. Rethinking hearing instruments as devices that adaptively learn behavioral patterns based on user interaction, might provide a degree of personalization that has not been feasible due to lack of audiological resources.

Author Keywords
Hearing impairment; user behavior; health; aging; augmented audio

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction

The current practice of fitting a hearing instrument relies on a trained audiologist, and it takes on average two months, with 2-3 visits, to fit the hearing instruments [6]. Hearing instruments are rarely fitted optimally at the first consultation, as the amplification of specific frequency bands only explains part of the problems encountered when aiming to understand speech in noise. Postponing the first visit for a decade [3] from first experiencing hearing problems until acquiring a hearing instrument provides other challenges. During this period the brain may have started to rewire due to its inherent plasticity and consequently the ability to comprehend speech may have begun to degenerate [7, 9]. As a result it may be difficult for the wearer to separate voices in challenging listening environments. In many cases there may be a lack of audiological resources for optimally adjusting the device. A perceived bad user experience may result in the user giving up on adapting the settings or simply returning the hearing instrument to the clinic. Kjeldsen and Matthews [5] identifies two types of tests in the hearing instrument fitting: as a minimum identify the needs for amplification in the frequency bands affected by the hearing loss based on an audiogram and subsequently assess the user's ability to separate sounds in noisy environments in sessions with trained audiologists. It may be difficult for users to describe how they perceive sounds in words in order for the audiologist to adjust the settings. Furthermore, the listening experience is only simulated based on audio samples in the clinic, which may differ from the problems the user actually encounters in real life. Other papers within the HCI literature have addressed the issue of retrieving and describing a situation. Dahl and Hanssen [2], build a tabletop prototype, where the user could choose between predefined soundscapes, but such participatory approaches may require that an audiologist is present to be useful.

In this paper we investigate how a hearing instrument is used throughout the day. Meaning, rather than simulating listening scenarios in a clinic, we aim to infer the optimal settings based on how the user adjusts programs and/or volume as the context changes in real life situations. In the present study we focus on the temporal dimension of interaction patterns observed over hours and days within a 10 week period. To our knowledge, no other studies have investigated in situ temporal interaction patterns of hearing instrument users at this level of detail. Traditionally hearing instruments have been perceived as independent devices limited by memory size and processing power, and only recently been able to wireless connect with smartphones. Utilising the power of an Internet connected hearing instrument, we investigate how a snapshot in time, represents a situation where the hearing instrument performs suboptimal. In this scope, the hearing instrument is perceived as a device that augment a soundscape. Previous studies within HCI have described similar devices augmenting hearing, usually involving a pair of binaural microphones and a pair of head worn speakers[8] [10], however, they have not investigated adaptation patterns or user fitted hearing experiences.

We propose a different way of hearing instrument fitting, connecting hearing instruments with smartphones and the Internet, making them cloud connected devices. Based on data and user engagement, we generate new types of personalization of hearing instruments. We propose a paradigm shift where audiological best practice and interventions includes decisions making from user generated data reflecting everyday usage.

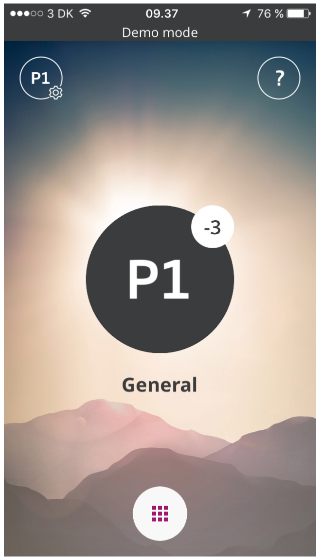


Figure 1: The interface of the Oticon ON iPhone app available from the App Store (iOS) to control the hearing instruments. To increase volume the user swipes up, and to reduce, swipes down. To change program the user taps the black circle and taps on a program to select it. The app then communicates directly with the hearing instruments via Bluetooth, and data is send via the iOS IFTTT app.

Method

Participants
6 participants volunteered for the study (6 men), from a database provided by Eriksholm Research Centre. The median age was *61.8 years (std. 11.1 years)*. All participants have more than a year experience using hearing instrument. The participants suffers from a symmetrical hearing loss, ranging from mild-moderate to moderate-severe as described by the WHO[12]. All have an iPhone 4S or newer. One participant was excluded due to missing data.

Apparatus
Each subject were equipped with two Oticon Opn™ hearing instruments, stereo Bluetooth low energy (BLE) 2.4 GHz. All subject used personal iPhone 4S or newer iPhone models with Bluetooth 4.0. The logged data consist of any user initiated program change or volume change through the Oticon ON iPhone app (see Figure 1), formatted as time series data, transferred using IFTTT (If-This-Then-That), stored in the cloud and shared via Google Drive. The hearing instruments were fitted with four programmes. The subjects were provided with a test user Google account prior to the experiment. The account was used for data collection, and the subjects had full ownership of the account and data.

Procedure

Subjects were fitted with OPN hearing instruments by an audiologist. The hearing instruments were fitted based on a unique frequency dependent volume amplification for each subject. Each subject was fitted with four programmes, through the Genie 2.0™ fitting software. The programs emulates different types of auditory focus, by increasing amounts of signal processing to enhance voices and reduce background noise when encountering challenging listening scenarios. These are trade offs between speech intelligibility, and background sound amplification. The four programs are:

- *P1*: Resembling an omnidirectional perception with a frontal focus. Sounds from the sides and behind the listener are slightly suppressed to resemble the dampening effect of the pinna.
- *P2*: similar to P1 but gently increasing balance and noise removal when encountering complex listening environments.
- *P3*: similar to P1 but increasing balance and noise removal even in simple listening environments.
- *P4*: similar to P3 with high sensitivity to noise increasing balance and noise removal in all listening environments.

Results and Discussion

In this section we first analyse the collected data to explore what differentiates the program usage based on the time of the day. Next we probe whether demands related to specific activities influence the behavioral patterns, by comparing program usage on weekdays (Mon-Fri) against weekends (Sat-Sun). Subsequently we discuss to what degree such learned behavioral patterns could sufficiently provide a foundation for adapting the device settings based on temporal aspects alone. For the analysis, only data collected between 8AM and 12AM is used, under the assumption that the hearing instruments would be switched off during the night. Data was collected between 12AM and 8AM, as the participants not always switched off the hearing instruments, introducing noise in the data set.

The difference between programs
Each subject shows unique interaction patterns when it comes to program usage. It should first of all be noted that the usage time for each participant varies between 3.5 to 8

	Average daily usage
S1	3.54 h
S2	7.21 h
S3	7.41 h
S4	6.66 h
S5	8.08 h

Table 1: Average hours of usage of the hearing instrument for each subject (S1-S5).

	P1	P2	P3	P4
S1	65%	17%	3%	15%
S2	80%	0%	0%	20%
S3	96%	0%	0%	4%
S4	67%	9%	10%	15%
S5	62%	9%	1%	28%

Table 2: Average usage of hearing instrument per subject. P1-P4 are programs, and S1-S5 are subjects. The average usage is in percentage of total usage of the device from 8AM to 12PM.

hours per day. The total usage can be observed in Table 1. To determine if there is a significant difference between the usage of the four programs an analysis of variance was performed. The mean usage of the four programs are: P1 18.4 minutes per hour (mph), P2 1.5 mph, P3 0.6 mph and P4 4.1 mph. Meaning, the difference in usage time related to the four programs was significant (F(3,4) = 23.1, p < .0001). The subjects have a preference for using P1, while P4 is second. The preference for P1, may reflect that it provides a frontal focus with a slight dampening of sounds from the back. This is similar to the acoustical characteristics provided by the natural shape of the ears and head. This suggest that P1 may provide adequate compensation in most of the listening scenarios encountered during the day. The three other programs offer increasing degrees of frontal focus and noise removal, where on average program P4 is preferred. However, from Table 2, subject 1 seems to prefer P2 which offers increased brightness facilitating speech intelligibility to P4. Based on the program changes alone it seems that at least two different auditory focus settings are needed. One program for less demanding listening scenarios allowing the user to shift the attention between several sound sources, and another program for challenging environments with multiple voices and background noise requiring more attenuation of ambient sounds.

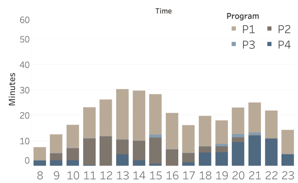
Table 2 shows the average usage of the four programs, P1-P4. Interestingly, we found that program P1 was preferred 74% of the time. This is significantly different from previous findings of respectively 33% [1] and 37% [11]. This could be due to manufacturer-specific noise reduction and gain reduction algorithms[4]. An interesting observation along the temporal dimension is illustrated in Figure 2. As an illustrative example, subject 4 uses P1 over the course of the day. However, the more supportive program P4 is primarily used between 11AM and 4PM and again between 7PM and 10PM. In Figure 2c patterns for the same two programs are shown for Subject 2. A notable difference appears for the usage of P4, who uses P4 from 9Am to 5PM, and then barely uses this program for the rest of the time period. The patterns thus seem highly individual and any design of algorithms for automatically adapting device settings would need to incorporate temporal aspects in regards to the individual preferences.

The different usage in weekdays compared to weekends

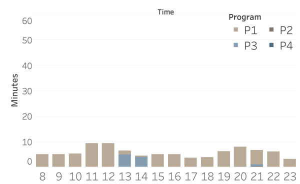
The next question to investigate is whether specific activities in weekdays and weekends change the behavioral pattern. The average use on weekdays are 7.8 hours per day, and 5.4 hours per day for the weekends. The difference between the aforementioned is significant ($F1,4$ = 17.0, $p <$ .02). To understand how the usage patterns varies between the weekdays and the weekend, a statistical analysis was performed on the four programs across participants. The different usage of the four programs was significant ($F_{3,12}$ = 23.1, p < .0001). However, the interaction between program and day was not significant ($F_{3,12}$ = 1.4, p > .5).

This indicates that the behavioral patterns vary over the course of a week. From Monday through Friday P1 is on average used 71% (of 7.8 hours) versus 80% (of 5.4 hours) Saturday to Sunday. Both the overall usage time and reduced selection of the P2-P4 programs, indicate that the user activities during weekends may represent fewer auditory challenges. In the light of this, we argue that any algorithms aiming to adapt the device settings according to behavioral patterns should also take these weekly patterns into consideration.

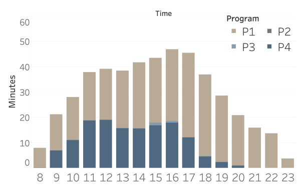
Using Subject 2 and 4 as contrasting examples in Figure 2d and 2h notice how the P4 usage pattern changes between weekdays and weekends. Subject 2 uses P4 more throughout the day (Mon - Fri), and only uses the program sparingly during evenings in the weekend. Subject 4 prefers P4 primarily during afternoons in the weekends, whereas



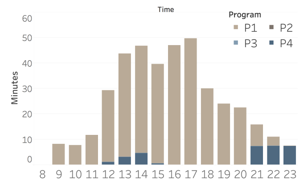
(a) Mon-Fri for subject 1



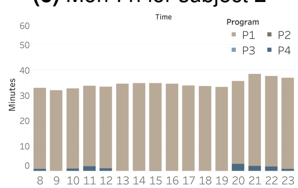
(b) Sat-Sun for subject 1



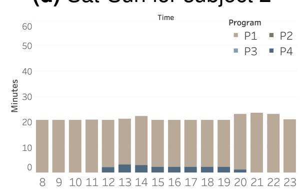
(c) Mon-Fri for subject 2



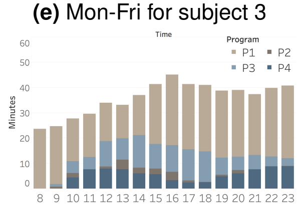
(d) Sat-Sun for subject 2



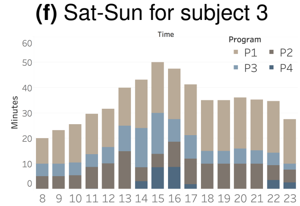
(e) Mon-Fri for subject 3



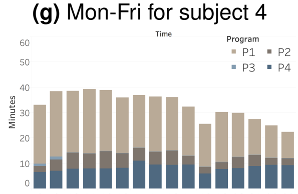
(f) Sat-Sun for subject 3



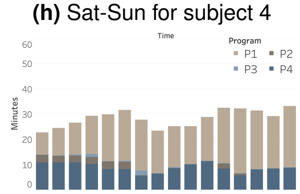
(g) Mon-Fri for subject 4



(h) Sat-Sun for subject 4



(i) Mon-Fri for subject 5



(j) Sat-Sun for subject 5

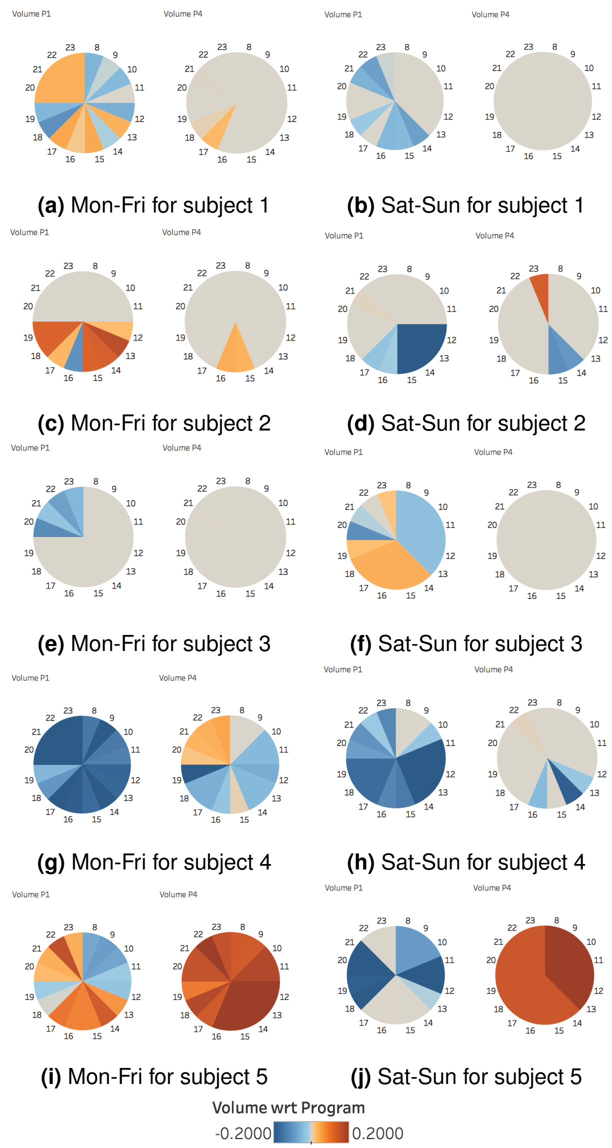
Figure 2: Program usage over time, from 8AM to 12AM. P1 is beige, P2 is brown, P3 is light blue and P4 is dark blue. The left hand columns represents usage over weekdays, and the right ones represents usages in weekends.

this usage patterns is not found during weekdays.

Volume and program interactions

An additional parameter to investigate when modeling the behavioral patterns are the volume change interactions. The volume interaction can be interpreted as a fine tuning of the desired auditory scene, by increasing or decreasing the intensity, thus zooming in or out of an auditory scene. In Figure 3 a comparison of the 5 test subjects and their usage of volume with respect to program can be observed. The light to dark blue colors reflect decreasing volume, while the yellow to orange gradients reflect an increase in gain. It can be observed that most subjects decrease the volume in P1 during the weekend. Subject 4 prefers to primarily reduce the volume, in contrast with Subject 5 which prefers to mostly increase the volume. In these cases we hypothesize that the gain settings of the devices might need to be adjusted. Subject 1 adjusts the volume both up and down from Monday through Friday, whereas the volume is only decreased during weekends.

While the above user interaction over a 10 week period can be inferred directly from the program change and volume adjustment, we subsequently in follow-up audiological sessions with the subjects found that the behavioral patterns were aligned with the aggregated program usage history data continuously collected over 4 months by the devices. Subsequently we interviewed the test subjects to determine what defined their program and volume preferences. The P1 program was preferred in most listening scenarios because it allows the users to selectively shift their attention omnidirectionally to any sound sources. However, when encountering more challenging acoustical environments, the three alternative program settings were selected, whether the aim was to enhance speech intelligibility, attenuate ambient sounds or remove background noise. Additionally



(a) Mon-Fri for subject 1 **(b)** Sat-Sun for subject 1

(c) Mon-Fri for subject 2 **(d)** Sat-Sun for subject 2

(e) Mon-Fri for subject 3 **(f)** Sat-Sun for subject 3

(g) Mon-Fri for subject 4 **(h)** Sat-Sun for subject 4

(i) Mon-Fri for subject 5 **(j)** Sat-Sun for subject 5

Figure 3: Program usage with respect to volume gain, from 8AM to 12AM. For each column the left figure is P1 and right figure is P4. Left hand columns represents usage over weekdays, and the right are usages in weekends.

users increased or reduced the perceived loudness of these settings by continuously adjusting the volume.

Perspectives

These results indicate that the users predominantly preferred to combine volume adjustments with settings providing an open frontal focus coupled with a natural attenuation of ambient sounds in 74% of the usage time. This differ from earlier studies reporting that an omnidirectional focus was only chosen in respectively 37% [11] and 33% [1] of listening scenarios. In contrast to earlier studies using simulated sound environments [2] our findings are based on the actual acoustic environments encountered by users over several weeks of usage. It is difficult to compare these studies, as the data generated in our study represent snapshots of user intents triggered by the changing auditory context throughout daily life. When compared to earlier studies, the quality of sound enabled by recent advances in digital signal processing provided by the state of the art devices used here is also likely altering how the auditory focus is perceived subjectively. The method of continuous data collection may facilitate long term personalization of auditory interfaces not limited to hearing instruments but encompassing next generation hearables in a wider sense. We propose that our data driven approach could potentially be used to individualize settings based on continuous interaction with Internet of things connected devices. In turn providing a dynamically optimized personalization, inferred from learned behavioral patterns.

Acknowledgements

This work is supported by the Technical University of Denmark, Copenhagen Center for Health Technology (CACHET) and the Oticon Foundation. We would like to thank Eriksholm Research Centre and Oticon A/S for providing resources.

References

[1] Gabriel Aldaz, Sunil Puria, and Larry J Leifer. 2016. Smartphone-Based System for Learning and Inferring Hearing Aid Settings. *Journal of the American Academy of Audiology* 27, 9 (2016), 732–749. DOI: http://dx.doi.org/10.3766/jaaa.15099

[2] Yngve Dahl and Geir Kjetil Hanssen. 2016. Breaking the sound barrier: designing for patient participation in audiological consultations. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3079–3090. DOI: http://dx.doi.org/10.1145/2858036.2858126

[3] Adrian Davis, P Smith, M Ferguson, D Stephens, and I Gianopoulos. 2007. Acceptability, benefit and costs of early screening for hearing disability: A study of potential screening tests and models. *Health technology assessment-Southampton* 11, 42 (2007).

[4] Wouter a. Dreschler, Hans Verschuure, Carl Ludvigsen, and Søren Westermann. 2001. ICRA Noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. *International Journal of Audiology* 40, 3 (2001), 148–157. DOI: http://dx.doi.org/10.3109/00206090109073110

[5] Malene Patsche Kjeldsen and Ben Matthews. 2008. Talking about hearing: Designing from users' problematisations. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*. ACM, 3. DOI: http://dx.doi.org/10.1145/1463160.1463237

[6] Sergei Kochkin, Douglas L. Beck, Laurel A. Christensen, Cynthia Compton-Conley, BJ Fligor, Patricia B. Kricos, and RG Turner. 2010. MarkeTrak VIII: The impact of the hearing healthcare professional on hearing aid user success. *Hearing Review* 17, April (2010), 12–34.

[7] F R Lin, Luigi Ferrucci, Yang An, J O Goh, Jimit Doshi, E J Metter, Christos Davatzikos, Michael A Kraut, and Susan M Resnick. 2014. Association of hearing impairment with brain volume changes in older adults. *Neuroimage* 90 (2014), 84–92.

[8] Florian Mueller and Matthew Karau. 2002. Transparent hearing. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*. ACM, 730–731.

[9] Jonathan E Peelle, Arthur Wing, Arthur Wingfield, and Arthur Wing. 2016. The neural consequences of age-related hearing loss. *Trends in Neurosciences* 39, 7 (jul 2016), 486–497. DOI: http://dx.doi.org/10.1016/j.tins.2016.05.001

[10] Miikka Tikander. 2009. Usability issues in listening to natural sounds with an augmented reality audio headset. *Journal of the Audio Engineering Society* 57, 6 (2009), 430–441.

[11] Brian E Walden, Mary T Cord, and Ole Dyrlund. 2011. Predicting hearing aid microphone preference in everyday listening. *Journal of the American Academy of Audiology* 396, April 2003 (2011), 365–396.

[12] World Health Organization. 2011. Grades of hearing impairment. (2011). http://www.who.int/pbd/deafness/hearing