Effects of personalizing hearing-aid parameter settings using a real-time machine-learning approach

Niels Søgaard JENSEN¹; Laura Winther BALLING²; Jens Brehm Bagger NIELSEN³

¹,²,³ Widex A/S, Denmark

ABSTRACT

In most hearing-aid fittings, amplification is prescribed by a fitting rationale that uses the audiogram as the main input. This approach may fail in situations where the user’s listening intention deviates from that assumed by the rationale. This shortcoming motivated a new commercially available method to self-adjust hearing-aid parameters while in a specific situation. The method is based on machine-learning algorithms that estimate the setting that optimizes user satisfaction based on user preferences in paired comparisons of parameter settings. We present results from a lab study where 20 participants with hearing loss used the method to adjust hearing-aid gain in 12 different sound scenarios with respect to three different sound attributes, and subsequently, in a double-blind assessment, compared the adjusted settings with the prescribed settings. The results showed a benefit of the method on basic audio quality. A large spread in the gain adjustments was observed, suggesting the need for more personalized settings of hearing aids. We also present anonymous user data gathered during real-life use of the method, which indicate when and why the method is used. We compare these data to general investigations of listeners’ auditory reality and suggest clinical and rehabilitative implications of using the method.

Keywords: Machine learning, hearing aids, personalization

1. INTRODUCTION

The non-linear amplification provided by modern hearing aids is typically prescribed by a fitting rationale, often with an audiogram as the only input. In addition to this basic fitting, many hearing aids offer a wide variety of sophisticated adaptive signal processing features that allow them to adjust to specific sound environments to accommodate the listening needs of the user. Features such as adaptive directional microphones, noise reduction and sound-classification systems (switching between different sets of processing parameters for different types of sound environments) are available in many hearing aids. However, the amplification scheme is typically determined based on the listening intention and needs of an average user in a given generic type of sound environment. When the listening intention and needs of a specific user deviate from those assumed by the underlying algorithms, that user’s listening experience may be less than optimal.

The common way to handle an inadequate user listening experience has traditionally been fine-tuning of the hearing aid performed by a hearing care professional (HCP). For the fine-tuning to be successful, it requires that the user is able to explain a perceived problem, and that the HCP is able to interpret the explanation and provide an adjustment of processing parameters that improves the listening experience (1). This is not always possible, which is part of the inspiration for various methods for self-adjustment of hearing aids, without the need for an HCP to be actively involved. A volume control is perhaps the simplest of such methods, but more advanced self-adjustment methods that affect multiple hearing-aid parameters have also been suggested, e.g. (2-4). These tools require a certain level of awareness of the user about the functionality of the available adjustment handles in order to be effective. Other methods have used paired comparisons performed by the user to adjust the hearing aid in an iterative manner (5). These methods only require the user to indicate a preference...
between two settings of the hearing aid. While some of these methods have shown promising results, e.g. the modified simplex method (6), they have not been used in the fitting or fine-tuning of commercially available hearing aids. A reason for this may be the substantial amount of time needed to reach the (global) optimal setting of parameters.

The increased processing power of hearing aids and, not least, the option to connect hearing aids to a smartphone to integrate its processing power in the entire hearing solution have allowed for more advanced and computationally demanding technologies like machine learning to be applied in self-adjustment of hearing aids. Making the paired-comparison approach sufficiently fast to allow hearing-aid users to use it to adjust their own hearing aids while being in daily-life situations was one of the main goals for the Interactive Hearing Aid Personalization System (IHAPS) (7). The IHAPS uses machine learning to drive a sequence of paired comparisons. In each comparison, the user assesses the degree of preference between two different settings of hearing-aid parameters. With the goal of finding an optimal setting in as few comparisons as possible, IHAPS iteratively determines the two settings for the next paired comparison based on what was learned from the paired comparisons already made by the user. This is done by continuously learning and updating a non-parametric machine-learning model of a hypothetical (unobserved) internal response function (IRF) that describes the user’s preference as a function of the hearing-aid parameter values. The final suggestion for the optimal setting, i.e. the parameter values that provide the maximum preference, is determined as the maximum of the estimated IRF. More details about this machine-learning approach can be found in (7).

With a few modifications, the IHAPS procedure has been implemented in a commercially available hearing aid, the Widex EVOKE, under the name SoundSense Learn (SSL). The SSL procedure affects static gain in three frequency bands, covering the entire frequency range of the hearing aid (0.1-10 kHz). The SSL procedure is activated by the user via a hearing-aid app installed on a smartphone. When the user has completed a series of paired comparisons as described above, a final adjusted setting is determined, and the user has the option to save this as an additional program that can be activated later via the app.

In this paper we will summarize the main results from a lab study (8) that investigated the perceptual effects of using SSL to make gain parameter adjustments in different sound scenarios in order to accommodate different listening intentions. Furthermore, we will present anonymous user data, gathered during real-life use of the SSL feature that indicate when and how the feature is used in real life. These data reflect the auditory reality – i.e. the variety of experienced acoustic environments (9) – of the users, and accordingly, we will compare the data to general investigations of listener’s auditory reality. The data offer some suggestions to how the feature may contribute to the rehabilitation of the user in a clinical context.

2. LABORATORY STUDY

This section summarizes a lab study conducted at SenseLab, FORCE Technology, Hørsholm, Denmark. More details about the study can be found in (8).

2.1 Methods

Participants. Twenty participants (8 females) with sensorineural hearing loss were recruited for the study. The average age was 72 years (range: 54-83 years), and the pure tone average (PTA) hearing loss, averaged across 0.5, 1, 2 and 4 kHz, was 44 dB HL (range: 26-59 dB HL). They were all experienced hearing-aid users and were fitted bilaterally with hearing aids of various brands and models.

Hearing aids. Widex EVOKE F2 440 RIC (Receiver-In-Canal) hearing aids were used as test hearing aids. They were mounted on a KEMAR acoustic manikin and fitted to the ear canal using closed silicone ear moulds. During SSL adjustments, the participants listened to the output of the hearing aids via headphones, with the headphone signal routed through an equalizer to compensate for the acoustic effects of ear canal and headphone response. During the later comparison of settings, the participants listened (via headphones) to recordings of the processed sound made in the KEMAR ear canals.

Test protocol. The study involved two visits to the lab for each participant. At the first visit, the SSL adjustment was completed in 12 different sound scenarios in a listening room with a 22.2 multichannel loudspeaker layout. The parameter setting prescribed by the fitting software, the Universal (UNI) setting, was always used as baseline for the adjustment. The scenarios were divided into three groups of four scenarios, each of which was associated with one of three different sound attributes: Basic
Audio Quality, Listening Comfort and Speech Clarity. For each scenario, the task of the participant (in the SSL adjustment during the first visit) was to optimize the perception of the attribute associated with that specific scenario. After this first visit, all 12 scenarios were recorded in the aided ears of KEMAR. The hearing-aid settings included the individual participant’s SSL setting for each scenario and the UNI setting with all adaptive processing and the sound-classification system being active. A third setting (with the sound classification turned off) was also included. However, the results obtained with this setting will be disregarded in this paper, which focuses on the effect of SSL, but the results are available in (8); the conclusions of the analyses remain the same whether this setting is included or not. At the second visit, a direct comparison of the settings was made in each of the 12 scenarios, using a double-blind approach. Participants listened to the recordings via headphones and could switch between the settings as much as needed. The participant’s task was to rate the sound attribute associated with the given scenario for each of the settings on a continuous scale from 0 to 100. The rating of each combination of scenario and setting was done twice.

Data analysis. The subjective ratings were analyzed using a linear mixed model ANOVA, with rating as the dependent variable, participant as a random effect, and setting, scenario and repetition as fixed effects. Separate analyses were done for the three attributes. The gain adjustments (in the three frequency bands, Low, Mid and High) in the SSL settings were registered and compared to the perceptual SSL benefit, calculated as the difference between ratings of SSL and UNI.

2.2 Results and discussion

Boxplots showing the distribution of the ratings of each combination of sound attribute and setting are shown in Figure 1. Each of the six boxplots in the figure is based on 20 (participants) x 4 (scenarios) x 2 (repetitions) = 160 data points. There were no missing data, and all data points (incl. outliers) are represented in the plots.

Figure 1 – Boxplots of ratings of the two hearing-aid settings, UNI and SSL, on each of the three attributes. The boxes indicate quartiles, while whiskers indicate minimum/maximum of non-outliers. Mean values are shown as squares, while outliers (outlier coefficient 1.5) are marked as circles.

For the Basic Audio Quality ratings, the mean difference between the SSL and UNI ratings, i.e. the mean SSL benefit, was 6.5 scale points. The statistical analysis showed a significant main effect of setting ($p<0.00001$), and a Tukey HSD post-hoc test showed a significant difference between SSL and UNI ($p<0.05$). Thus, the participants were able to improve the perceived general sound quality by use of the SSL procedure.

For Listening Comfort, the mean difference between SSL and UNI was 4.4 scale points. In this case, the main effect of setting was not significant, and the pairwise differences were therefore not tested. However, the interaction between scenario and setting was significant ($p<0.01$), indicating that the effect of setting varied across the scenarios. A post-hoc test showed that the difference between SSL and UNI was significant for one of the scenarios (traffic noise), but not for the other three scenarios. One reason for this may be that the three other Listening Comfort scenarios, as opposed to the traffic noise, included speech babble as part of the environment. This may have been a disturbing factor,
making some participants focus on speech intelligibility/clarity rather than comfort during the SSL adjustments and/or the following assessments. For the traffic noise, there may have been a higher compliance between the dictated listening intention (defined by the Listening Comfort attribute) and the intention that participants would intuitively have in that scenario.

For Speech Clarity, the mean difference between SSL and UNI was only 0.9 scale points. The main effect of setting was not significant, and no significant interaction with scenario was seen in this case. The fact that no SSL benefit on speech clarity was observed (on a group level) was not very surprising, given the nature of the gain adjustments (in three rather broad frequency bands) offered by the SSL procedure, and the fact that the UNI setting is prescribed with optimization of speech intelligibility in mind. It could be argued that adjustment of other hearing-aid parameters (e.g. time constants in the compression system) would be needed to obtain improvements on this attribute. However, it should be noted that on an individual level there were participants who obtained an SSL benefit in the Speech Clarity domain via adjustment of the gain parameters.

Across the attributes, a general trend in Figure 1 is the large variation in ratings across participants. The ratings basically span the entire rating scale. The low ratings of the UNI setting indicate that in some cases, the prescribed setting was not satisfactory. In a clinical context, fine-tuning of the hearing-aid setting would probably have been recommended, but that option was not part of the protocol. The low ratings of SSL obviously indicate that the SSL procedure did not provide a satisfactory result in all cases. This aspect is further discussed below.

For each scenario, the SSL benefit was compared to the gain adjustments (in the three bands) made during the SSL adjustments. Figure 2 shows the individual gain adjustments made in three selected scenarios (one for each attribute). In each panel in the figure, the gain adjustments in each band (left y-axis) and the corresponding SSL benefit (right y-axis) are plotted for each participant. Please note that gain adjustments between +6 and +12 dB were only possible in the Low band. In the Mid and High bands, the maximum gain increase was +6 dB to reduce the risk of introducing acoustic feedback.

![Figure 2](image-url)

**Figure 2** – Individual SSL gain adjustments away from the UNI setting (left y-axis) in the Low, Mid and High bands for three sound scenarios representing the three attributes (left, middle and right panel). In each panel, the participants (on the x-axis) have been rank-ordered according to the magnitude of their SSL benefit. The SSL benefit (right y-axis) is indicated by black open circles. Linear regression lines have been added to indicate overall trends in gain adjustment as functions of SSL-benefit ranking. The vertical dotted line in each panel separates participants with positive and negative SSL benefits.

A general observation across the scenarios (also those not shown in Figure 2) is that the gain adjustments span the entire range from -12 dB to +6/12 dB, and that participants make very different gain adjustments. To investigate the relationships between gain adjustments and SSL benefit (visualized with the regression lines in the figure), Spearman rank correlation coefficients were calculated for all three bands in all 12 scenarios. After correction for multiple correlations, there was only one scenario with a significant correlation between the gain adjustments in one of the bands (Low) and the SSL benefit. Thus, the overall observation is that the SSL benefit cannot be predicted based on the gain adjustments, which underlines the individual nature of the adjustments made by the participants.

Across scenarios and participants, the general trend in the gain adjustments was an increase of gain
in the Low band, a decrease of gain in the High band, with the mean gain adjustment the Mid band being somewhere in between (close to 0 dB). This is also the trend seen in the examples in Figure 2 where the midpoint of each trend line indicates the mean adjustment across participants. The most noticeable adjustment away from the prescribed setting is observed on the Low band where the mean adjustment across scenarios was approximately +5 dB. While this could be seen as an indication of too little gain provided by the fitting rationale, the test procedure may have played a significant role, e.g. by not including the participants’ own voice. Had own voice been included, it may be assumed that less low-frequency gain would have been preferred. When the large individual variation is taken into account, the results do not suggest that systematic changes in the gain prescription would have been enough to meet the participants’ individual needs.

The individual SSL benefits plotted in Figure 2 show that while several participants obtained an SSL benefit, there were other participants who obtained a rather substantial disadvantage. For example, participant 10 had large SSL disadvantages in two of the three scenarios included in the figure. The most likely explanation for this is inconsistency in the preference criteria used during the SSL adjustments. Since a certain level of consistency is required for the procedure to find a true preference maximum, the consequence is that some users may not be able to complete the SSL adjustment in a way that provides a perceptual benefit.

From a clinical perspective, the results of the study indicate that the SSL procedure has the potential to provide benefits for some (but not necessarily all) users, most likely in the Basic Audio Quality domain. The results also indicate that different users have different amplification preferences – within the same sound scenario and for the same given listening intention – that are impossible to address with the gain prescribed by a single overall fitting rationale based on the audiogram. This suggests the need for more personalized hearing solutions. The observed variation in gain adjustments correspond well with the variation in adjustments made during real-life use of the procedure. This will be further explored in the following section.

3. REAL-LIFE USE OF MACHINE-LEARNING APPROACH

The laboratory results address several central aspects of the effects of personalizing hearing-aid settings using the SSL machine-learning approach. However, an additional key question is how this approach functions in hearing-aid users’ real lives. This question can to some extent be addressed by investigating the anonymous user data that are collected whenever listeners save and use SSL programs. These data include information on the gain settings and usage of saved SSL programs, as well as the users’ indication of the situation and listening intention for which the program was created.

Figure 3 – SSL adjustments in the three frequency bands in a sample of 9690 personal programs.

Looking first at the gain adjustments, Figure 3 shows 9690 personal programs (a subsample of the sample discussed in more detail below, the reduction is necessary in order to see the individual programs). Each dot represents a saved personal program, with the adjustment in gain in the Low
(bass), Mid (middle) and High (treble) frequency bands. The figure shows a wide distribution of gain
adjustments, with no discernable patterns. This is in line with the adjustments seen in the laboratory
study reported above (see Figure 2) and indicates that the gain adjustments are specific to the
individual user and the individual situation.

A second interesting aspect is the situations and listening intentions that SSL is used for, which
gives us information about the real-life use of personalization and the needs that the users seek to
fulfill. This information stems from questions introduced with SSL version 2.0 where the users are
asked where they are (what we call situation in the following) and what their hearing goal is (what we
call intention), before they start the pairwise comparisons. Users must indicate one situation from a list
including 12 predefined generic situations, while they can choose none, one, or two intentions from a
list of 9 predefined generic intentions. Most programs in our sample (77%) represented one intention,
while 19% had two intentions, and 4% none. Figure 4 shows the distribution of answers in a sample of
13,813 personal programs created with SSL version 2.0 from 5,448 unique users.

First considering situations and intentions separately, we see from the left panel of Figure 4 that
‘Home’ is by far the situation where most SSL programs are created. This
dominance of the home
situation corresponds well with findings from a study that used ecological momentary assessment to
investigate the auditory reality of a group of hearing-aid users (10). It may reflect both that many users
spend most of their time at home and that many home situations are ones in which it is relatively easy
to conduct the pairwise comparisons, particularly for the ‘TV’ intention which is the most frequent
intention in home situations. Aside from home, situations are fairly evenly distributed, with a
substantial part – just under 20% - being the kind of noisy situations that tend to be challenging to
hearing-aid users (‘Restaurant’ and ‘Noisy venue’). Turning to the intentions, we see from the right
panel of Figure 4 that the top four intentions are relatively well-defined ones – conversation, TV, noise
reduction and music – that together account for 82% of all intentions (counting each intention as one
instance, whether it occurs alone or together with another intention). The patterns are very similar if
we analyze the number of uses of the programs instead of the number of programs, indicating that
program use is approximately evenly distributed for different situations and intentions.

![Figure 4 – Distribution of situations and intentions in the sample of SSL programs, excluding programs with no indication of intention. The right panel shows each category as a proportion of all intention labels.](image)

Alternatively, situations and intentions may be considered jointly, and compared to general
investigations of the auditory reality of people with and without hearing loss, e.g. (11-13). In such a
comparison, three characteristics of the SSL data are important to keep in mind: First, SSL is intended
for situations in which the sound is not entirely satisfactory, which is in line with some (e.g. (11)) but
not all (e.g. (12, 13)) more general investigations. Secondly, SSL is not intended for – and not suitable
for – all users, but relevant for those that have an interest in self-adjusting their hearing aids and are
capable of doing the comparisons systematically through a smartphone. Thirdly, creating a personal
program with SSL involves listening and responding in a way that is not suited to all situations and
intentions, which may overrepresent some situations and underrepresent others in the SSL data.
compared to listeners’ general auditory reality.

However, even with these differences, the distributions of situations and intentions in SSL are comparable to what we see in the literature. We focus here on the Common Sound Scenarios (CoSS) suggested by Wolters et al. (12) because these summarize findings from ten studies. The CoSS framework includes three overall listening intentions: speech communication, focused listening, and non-specific listening, with seven different tasks categorized into one of these three overall intentions. Each of the seven tasks is in turn exemplified by two representative scenarios, one in quiet and one in noise. Importantly, the CoSS focuses on common scenarios that are representative of hearing life across a wide range of listeners, rather than on an exhaustive list of possible listening scenarios or a catalogue of situations that are problematic for hearing-aid users.

The comparison between CoSS and SSL user data may be conducted in two ways, starting from either of the two categorizations. First, taking the CoSS categorization as a starting point, each of the 14 scenarios may be matched to an SSL situation-intention combination, with matches generally being straightforward. The prominence of speech communication, which is part of six out of 14 scenarios in CoSS, is also reflected in the SSL data, where approximately one third of all intentions are labelled ‘Conversation’. This is all the more remarkable because conversations – particularly one-on-one – are not ideal situations for conducting the pairwise comparisons that are required to construct a new personal program. This challenge is likely to be especially pronounced for conversations over the phone, which are rated as both very frequent and very important in CoSS but are relatively rare in the SSL categorizations (only just 1.8% of intentions are ‘Phone’). This makes sense when we consider the difficulty of completing the SSL procedure while also keeping a phone conversation going.

Turning next to the SSL data as a point of departure, we focused on those combinations of situation and intention that made up at least 1% of all programs, resulting in a list of 20 situation-intention combinations. The most frequent of these, TV and conversation at home, are also rated as highly frequent based on the Wolters et al. literature review, while the less frequent combinations in SSL tend to have medium or low frequency in CoSS.

The comparison also reveals an important difference between the SSL data and CoSS, namely that the categories where the situation is restaurant, noisy venue or large hall in the SSL data are not directly represented in CoSS. A substantial part of this group of intentions are likely to be cases where the background noise is speech, in other words similar to the cocktail party scenario that is frequently referred to in hearing-aid research and development. This particular scenario is not concretely represented in CoSS, though it can fit within the intention and task combinations in the framework. Wolters et al. comment on the absence of this type of scenario in CoSS, arguing that although it is certainly challenging and often important to hearing-aid users, it is not frequent in their literature review. However, its prominence in the SSL data, which derive from situations where the user is not satisfied with the hearing-aid sound, supports its continued use in hearing-aid research and development.

There are two more general implications of these analyses. Firstly, and specifically for SSL, the data show that usage is distributed across a wide range of situations and intentions, suggesting that the feature is used as intended, providing the possibility of personalization in many different real-life scenarios. Secondly, and more generally, the fact that these data derive specifically from cases where the user is in some way not satisfied with the sound of their hearing aids is a weakness if we want to investigate people’s auditory realities generally. However, it is also a strength – and, we would argue, a more important one – because it means that the data on situations, intentions, and settings are informative about exactly the situations that hearing-aid development wants to address. Here, the same caveat applies as for several of the specific scenarios discussed above, namely that the data we get from SSL usage are that subset of situations where it is possible for the user to listen to pairwise comparisons. Nevertheless, we do see patterns that may inform hearing-aid research and development in interesting ways.

In the present context, we have investigated the data with the purpose of getting a sense of the auditory reality of SSL-users considered jointly. However, this same type of data may also be considered for the individual user, with the purpose of supporting their hearing rehabilitation. This will be possible when data from the individual user’s SSL and other personal programs – with their consent – shortly become available to their HCP in the relevant fitting software. This may allow the HCP to get a better sense of the auditory reality of their individual customer and help them make any generally relevant adjustments to the hearing-aid settings based on this.
4. FINAL REMARKS

Though the laboratory and user data presented above go a long way towards understanding the effects of personalizing gain settings, several interesting research questions remain. One central issue is how the availability of self-adjustment affects hearing-aid users’ hearing life in general. The current analyses have shown that sound quality can be improved using SSL and that SSL is used in a range of diverse situations; elsewhere we have found that users are generally happy with SSL (14). One could speculate that the possibility of adjusting gain settings at any time could give hearing-aid users a sense of empowerment in relation to their hearing loss, which in turn could contribute to higher engagement with their treatment and potentially higher satisfaction. In other words, in addition to any direct effects in those situations where SSL helps the users adjust appropriately to their situation and intention, users might also experience more indirect effects on their hearing-aid satisfaction from the possibility of making such adjustments in any situation, even though they only make use of it in a subset of situations. While empowerment is not a well-defined concept, e.g. (15), there are clearly interesting questions to explore in this connection, both with a focus on user satisfaction and on user independence and self-efficacy. In addition, and related to this, it remains to be investigated systematically which user groups benefit (most) from the availability of such a functionality.

ACKNOWLEDGEMENTS

The authors would like to thank Thor Bundgaard Nielsen and Søren Vase Legarth from SenseLab, FORCE Technology, and Ole Hau from Widex A/S for their contributions to the lab study, and Lasse Lohilahti Molgaard from Widex A/S for his contributions to the analysis of the real-life data.

REFERENCES