WIDEXPRESS

MAY 2019 · ISSUE NO. 42

MACHINE LEARNING IN **WIDEX EVOKE**: PERCEPTUAL BENEFITS OF **SOUNDSENSE LEARN**

NIELS SØGAARD JENSEN, M.SC., SENIOR EVIDENCE AND RESEARCH SPECIALIST
OLIVER TOWNEND, B.SC., HEAD OF AUDIOLOGY COMMUNICATION

SUMMARY

The perceptual effects of SoundSense Learn – the option to use real-time machine learning to optimise the parameter setting of Widex EVOKE according to an individual listening intention – were explored in two studies. In both studies, the participants used SoundSense Learn in different types of sound scenarios to meet different given listening intentions. In a double-blind setup, they then compared their personal settings with two prescribed settings of EVOKE. The results from both studies showed that the use of SoundSense Learn improved both sound quality and listening comfort. Furthermore, the studies showed that the parameter adjustments were highly individual and impossible to predict, indicating the need for efficient tools to personalise the setting of hearing aids. The studies suggest that SoundSense Learn is indeed such a tool. Sharing of anonymous data generated during the use of SoundSense Learn is vital for the continuous improvement of hearing experiences for all users.

INTRODUCTION

The introduction of Widex EVOKE included an entirely new paradigm in hearing aid technology: real-time machine learning to optimise the setting of hearing aid parameters according to the individual listener's context, circumstances, intentions and emotions. The SoundSense Learn (SSL) feature allows its users to optimise their listening experience according to their individual listening intention and personal preference in any given situation. This adds another dimension to the already powerful functionality of EVOKE.

The highly advanced adaptive signal processing in EVOKE is designed to automatically accommodate the user's listening intention in the variety of different situations they encounter in everyday life. Features like Fluid Sound Technology, Variable Speed Compressor and HD Locator enable the hearing aid to adapt automatically to the acoustic environment in order to meet the listening intention in the best possible way. The automatic functionality is based on available knowledge about the Acoustic Scene, Auditory Processing and Auditory Cognition. These are three of the main elements in the Real-Life Hearing framework (Jensen, 2018a; Jensen, 2018b) which Widex uses to conceptualise and explain the entire process of hearing (see Figure 1). When it comes to the fourth Real-Life Hearing element, Evaluation & Behaviour, the adaptive functionality makes assumptions about what the hearing aid user's listening intention is in the given situation. But what happens when the assumed listening intention deviates from the actual listening intention of the user? It has often been the case that faced with this problem of automation not meeting listening intention, the user would return to the hearing care professional (HCP). Going to see the HCP did not always result in a solution, and this was no fault of the HCP. Often a solution was hard to find, because what happens in real life is difficult to transfer to the clinic. The user's description of the sound, environment or problem is often vague, and the interpretation of user descriptions into meaningful HCP changes is somewhat challenging.

It is exactly this situation that SSL addresses. When the user in a specific situation wants something different from what is provided automatically by the hearing aid, SSL offers a way to address this need in the moment, without having to consult the HCP to make an adjustment to the hearing aid during a conventional fine-tuning session and without compromising the original fitting set by the HCP. In this way, SSL may be perceived as an added bonus to the fitting.

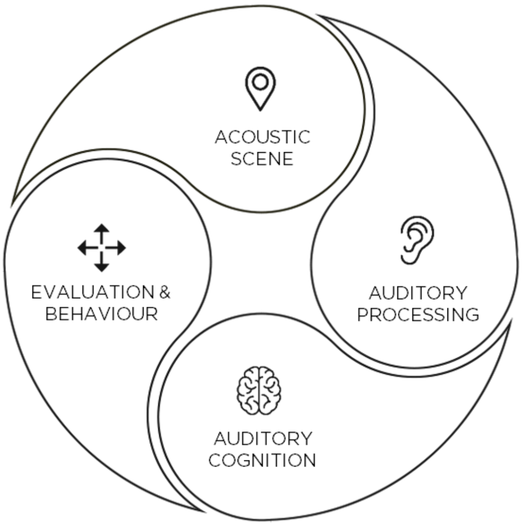


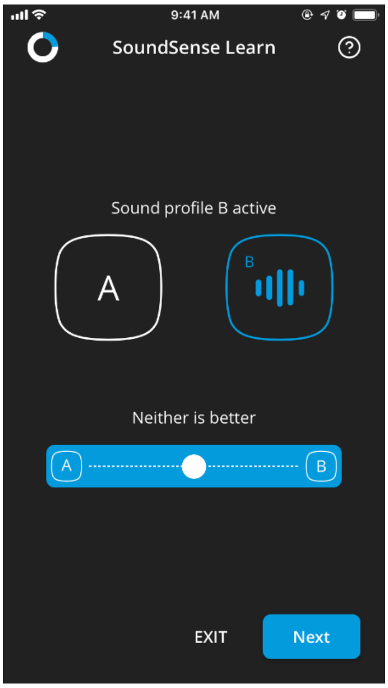
Figure 1. The Real-Life Hearing framework, showing the hearing process as a continuous looped flow involving four main elements.

In this *WidexPress*, we will summarise the results from two laboratory studies that investigated the perceptual effects of SSL (Townend et al., 2018b; Jensen et al., 2019). The two studies were based on the same experimental approach. Most importantly, they were both conducted as double-blind studies, meaning that neither the participants nor the experimenters knew what the participants listened to during the perceptual assessment.

SOUNDSENSE LEARN AT A GLANCE

In brief, the basic concept of SSL is to let the user make a series of paired comparisons of different settings of hearing aid parameters. In each comparison, the user indicates a degree of preference between the two settings. The machine-learning algorithm uses the user's response to determine the next two parameter settings to be compared, and it uses all the input provided by the user to estimate the user satisfaction as a function of the parameter setting. After just a few comparisons, the algorithm is able to provide a good prediction of the setting of parameters that maximises user satisfaction. Adding more comparisons will improve the prediction, but the optimal setting should be found with fewer than 20 comparisons. The final optimal setting can then be saved as an additional program and recalled when the user later enters the same (or the same type of) environment.

The user controls the SSL procedure via the EVOKE app, which provides an intuitive and user-friendly interface (see Figure 2). The user's task is simply to use the two buttons to switch between the two settings, A and B, and then use the slider to indicate the degree of preference between the two settings.


9:41 AM

SoundSense Learn

Sound profile B active

Neither is better

Figure 2. The SoundSense Learn user interface in the EVOKE app.

In the first version of SSL, which was tested in the two studies described in the following, the hearing aid parameters affected by the procedure include gain in three frequency bands, Low, Mid and High. These bands cover the entire frequency range of the Widex EVOKE hearing aid.

A more detailed description of the approach used by the SSL feature is provided by Townend et al. (2018b), and an in-depth description of the underlying advanced machine-learning algorithms can be found in Nielsen et al. (2015).

METHODS

The two studies were conducted at two different sites, one internal site (Site 1: Widex headquarters) and one independent external site (Site 2: SenseLab, FORCE Technology). Thus, two different test environments, two different groups of test participants and two different groups of test leaders were used. However, the basic test approach was the same in the two studies.

There were 19 participants (6 female/13 male, average age 68 years) at Site 1, and 20 participants at Site 2 (8 female/12 male, average age 72 years). All participants had a sensorineural hearing loss within the fitting range of the test hearing aid (Widex EVOKE F2 440 RIC).

The test protocol for both studies included two visits to the lab for each participant. At the first visit, the SSL adjustment was completed in several different sound scenarios, where the participant was asked to indicate their degree of preference with respect to a specific sound attribute that could be associated with a given listening intention. In this paper, we will focus on the attributes “general sound quality” and “listening comfort”. Table 1 shows a list of the sound scenarios used for the SSL adjustment for each attribute at each of the two study sites. The two sites did not use the same recordings of scenarios, but in some cases they used the same type of scenario.

The baseline setting (the starting point to which gain changes were made) was always the Universal setting prescribed by the Widex Fitting Rationale, based on the individual participant’s audiogram. During the SSL adjustment, the EVOKE hearing aids were mounted on a KEMAR acoustic manikin placed in a multi-loudspeaker setup that reproduced the sound scenarios. The participant was seated in the neighbouring control room and listened to the output from the hearing aids via headphones. The approach of placing the hearing aids on a KEMAR, and not on the participant’s ears, was chosen in order to allow for a seamless and double-blind comparison in the subsequent perceptual assessment procedure.

Site	Sound Attribute	Scenario name	Description
1	Sound Quality	Contemp1	Rock music
		Contemp2	Jazz trumpet
		Classical1	Chamber music
	Listening Comfort	Canteen1	Canteen noise w. babble
		Dinner	Café noise w. babble
		Road	Road noise
2	Sound Quality	Contemp3	R&B music
		Contemp4	Pop music
		Classical2	Debussy
		Classical3	Mozart
	Listening Comfort	Traffic	Traffic noise
		Café	Café noise w. babble
		Canteen2	Canteen noise w. babble
		Outdoors	Outdoor setting w. babble

Table 1. Sound scenarios used for SSL adjustment and the following assessment of hearing aid settings. The sound attributes were used to prescribe the listening intention during adjustment and assessment.

After the SSL adjustments were completed in all sound scenarios, and with the hearing aids still mounted on KEMAR, the different sound scenarios were then recorded with the hearing aids programmed in three different settings: 1) “No classifier”: The Universal program with the Fluid Sound Technology system turned off; 2) “Active Classifier”: The Universal program with the Fluid Sound Technology turned on; and 3) “SSL”: The SSL setting determined by the participant, using the “Active Classifier” setting as baseline.

At the second visit to the lab, the participants blindly compared the three settings by listening to the recordings via headphones. The comparison was done for each of the sound scenarios shown in Table 1, where the participant had the option to switch freely between the three settings. Each setting was rated on a continuous rating scale with respect to the relevant sound attribute. For example, in the canteen scenarios, the three settings were rated with respect to the attribute Listening Comfort (see Table 1). If one setting was perceived as better than another, it received a higher rating.

More details about the test methodology can be found in Townend et al. (2018a) and Jensen et al. (2019).

RESULTS

In the following, results from the rating experiments are presented. They are grouped according to the sound attribute that was rated, starting with sound quality.

Sound Quality

Both studies evaluated the general sound quality (also known as basic audio quality) of each of the three hearing aid settings. The sound samples used for the evaluation included both contemporary music and classical music (see Table 1). In this way, both of the new music sound classes available in EVOKE were activated in the "Active Classifier" and "SSL" settings.

The mean ratings of the three hearing aid settings for each scenario are shown separately for each site in the two upper panels in Figure 3. The bottom panel in the figure shows the mean ratings for the three settings across scenarios at each of the two sites.

For all sound scenarios, the plots show that activation of the sound classifier provided a – typically quite substantial – improvement in general sound quality. The main explanation for this observation is that the noise reduction that was activated in the "No Classifier" setting provides a reduction in gain which is detrimental to the sound quality of music. Accordingly, the noise reduction is deactivated in the two music sound classes, which explains the observed improvement. For all scenarios, adding the SSL adjustment provided a further improvement in the sound-quality rating. There was some variation in the benefit between scenarios, but the overall trend is clearly seen.

The sound-quality rating data were analysed statistically by running separate mixed-model analyses of variance (ANOVAs) for the two sites, and by using Tukey post-hoc tests to assess pairwise differences between settings. The analyses of data from both sites showed that the benefit of turning on the classifier (i.e., the difference between the mean ratings of "Active Classifier" and "No Classifier") was statistically significant. On top of this, there was a significant further benefit of performing the SSL, i.e., the difference between the mean ratings of "SSL" and "Active Classifier" (all $p < 0.05$).

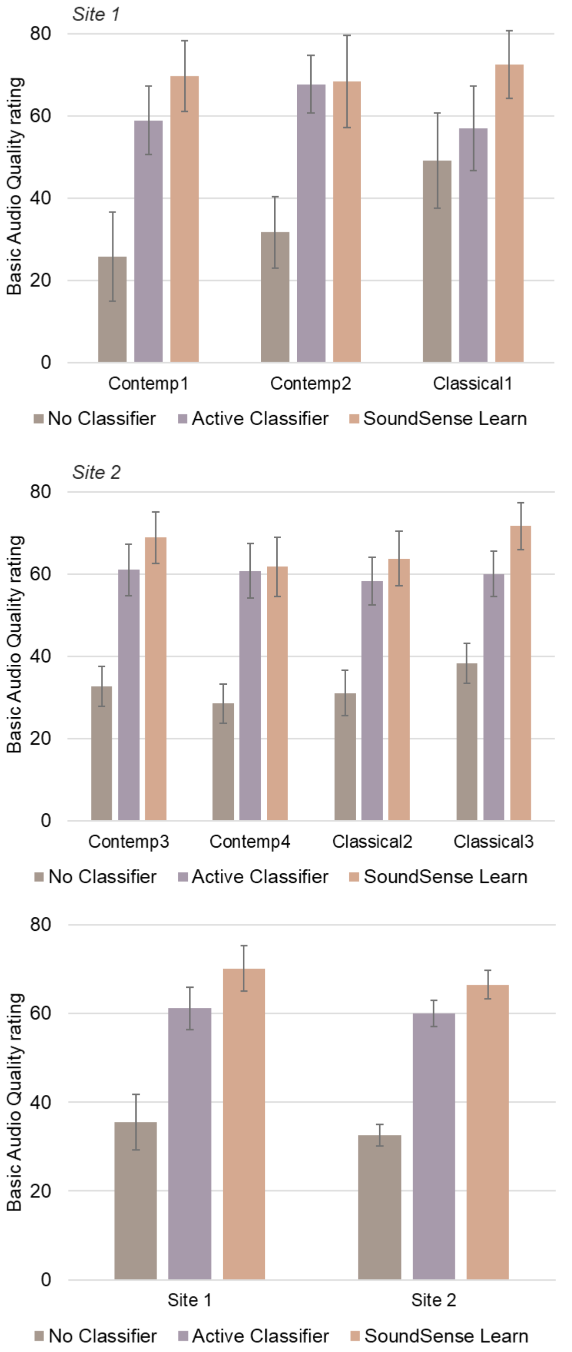


Figure 3. Mean ratings of general sound quality (basic audio quality) for each sound scenario in each of the three hearing aid settings at Site 1 (upper panel) and Site 2 (middle panel). The mean ratings across scenarios are shown in the bottom panel for both sites. Error bars indicate 95% confidence intervals.

Besides looking at the effects of SSL on the mean ratings, we also looked at the number of participants who obtained a benefit with SSL, as measured by a higher rating of “SSL” than of “Active Classifier”. We know that participants may use the rating scale differently, and a difference of, for example, 5 scale points may reflect a large auditory effect for some people and a small auditory effect for others. However, since each participant compared the three settings directly, it is fair to assume that even a small rating difference of a given sound scenario represents an audible difference for that scenario. When pooling the data from the two studies, we observed that:

- **72%** of participants were able to obtain a general sound-quality benefit with SSL, based on their average ratings across the scenarios.
- **92%** of participants were able to obtain a sound-quality benefit with SSL in at least one of the scenarios.

Listening Comfort
Listening comfort was included as a sound attribute in both studies. Participants were exposed to sound scenarios recorded in different everyday sound environments (see Table 1). The variation in scenarios ensured that different EVOKE sound classes were activated in the different scenarios.

The mean ratings of the three hearing aid settings for each scenario are shown for each site in the two upper panels in Figure 4. The bottom panel in the figure shows the mean ratings across scenarios for the three settings at each of the two sites.

The plots of the listening-comfort ratings show more variation across both scenarios and test sites than the sound-quality ratings. At Site 1, the same overall trend as in the sound-quality data was observed: Activation of the classifier provided a benefit (in two out of the three scenarios), and the SSL adjustment provided a further benefit (in all three scenarios). At Site 2, the differences between settings were generally smaller. Activation of the classifier only provided a (minor) benefit in two of the four scenarios, while it provided a (minor) disadvantage in the other two scenarios. In all four Site-2 scenarios, “SSL” provided a benefit over “Active Classifier” in all four scenarios, but this benefit was only substantial for the “Traffic” scenario.

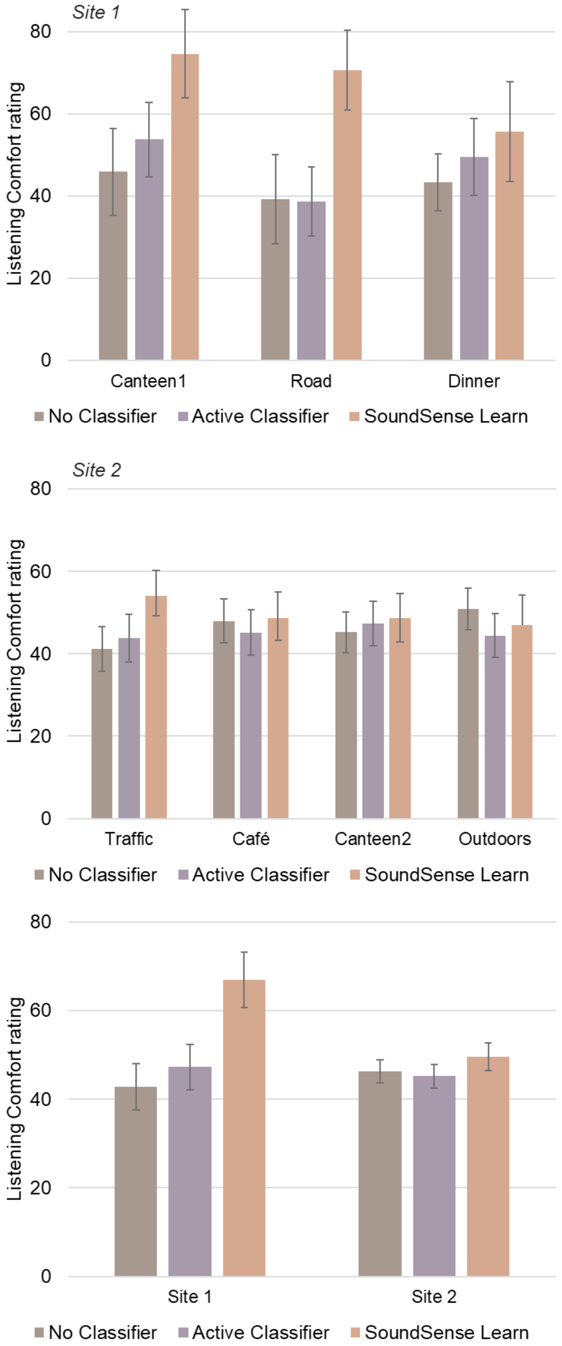


Figure 4. Mean ratings of listening comfort for each sound scenario in each of the three hearing aid settings at Site 1 (upper panel) and Site 2 (middle panel). The mean ratings across scenarios are shown in the bottom panel for both sites. Error bars indicate 95% confidence intervals.

The listening-comfort rating data were analysed statistically using the same type of mixed-model ANOVA and post-hoc tests as for the sound-quality data. The analysis of data from Site 1 showed that the benefit of SSL (across scenarios) was statistically significant ($p < 0.05$), while at site 2 it was only the SSL benefit in the “Traffic” scenario that was statistically significant.

As in the analysis of the sound-quality rating data, we also looked at the number of participants who obtained a benefit with SSL on listening comfort. When counting a larger rating as a preference (or benefit) and pooling the data from the two studies, we see that:

- **69%** of participants were able to obtain a general listening-comfort benefit with SSL, based on their average ratings across all the scenarios.
- **90%** of participants were able to obtain a listening-comfort benefit with SSL in at least one of the scenarios.

Speech intelligibility/clarity

Both studies included a third sound attribute in the test protocol. For Site 1, it was the attribute “speech intelligibility”, while for Site 2 it was “speech clarity”. Statistically significant differences between the settings were not observed in any of the studies, and the results are not reported in this *WidexPress*. When listening for speech intelligibility or clarity, it is not surprising that the listening intention of the wearer matches the assumptions made by the classifier, therefore reducing the benefit of SSL. However, even though no significant differences between mean ratings appeared, it is still worth noticing that 74% of the participants in the two studies were able to obtain an SSL benefit on perceived speech intelligibility or speech clarity in at least one of the included speech scenarios. Along with the advantages for the other two attributes, this suggests that for the majority of users there are likely to be benefits from using SSL in specific scenarios.

Gain adjustments

The gain adjustments in the three frequency bands made during the SSL adjustments were registered and analysed. The 3D plot in Figure 5 shows all SSL gain adjustments made by the participants at Site 2. A similar pattern was observed in the adjustments made at Site 1.

One of the most important observations based on Figure 5 is that there are no general trends in the gain adjustments. There are no systematic changes away from 0 dB, neither on a group nor on an individual level. Such systematic changes would have indicated that a general change in the fitting rationale could offer a similar improvement, but this is clearly not the case here, nor in the much larger set of real-life data reported by Balling et al. (2019).

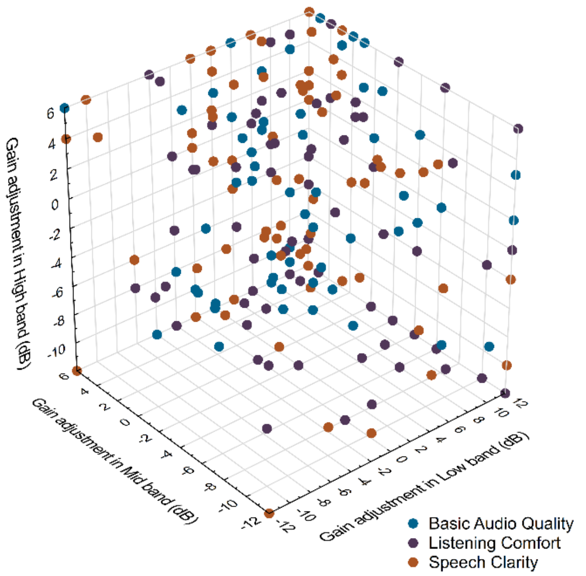


Figure 5. Individual gain adjustments made in the three frequency bands by participants at Site 2. The colour coding separates data points associated with the three different sound attributes. Each of the 20 participants contributes 12 data points. Some of the data points are coincide.

Another important observation is that the variation in the gain adjustments is substantial, spanning the entire gain range allowed by the SSL feature. Thus, participants ended up with very different gain adjustments, depending on the scenario and the listening intention. Moreover, if we analyse the gain adjustments together with the rating data, there are no connections between the adjustments made and the SSL benefits obtained.

DISCUSSION

The clearest perceptual benefits of SSL were observed in the ratings of general sound quality. These benefits were observed in all sound scenarios at both test sites. The fact that the clearest effect was observed for SSL adjustments in music scenarios is not surprising. It is quite easy for most people to relate to the sound quality of music, making it fairly easy for users of SSL to be consistent during the paired comparisons made as part of the SSL procedure. Even though the underlying machine-learning algorithms also take inconsistent user performance into account, the more consistent the user is during the SSL procedure, the higher the chances are that the algorithms will correctly identify the setting that optimises satisfaction.

The sound-quality ratings also demonstrated the substantial improvement offered by the music sound classes in EVOKE. As already explained, the effect can be attributed to the music classes' handling of the noise reduction system, which provides a much fuller sound of music.

The ratings of listening comfort also showed clear benefits of SSL. However, the benefits varied more across sound scenarios and test sites than the sound-quality benefits. An explanation could be that it was more difficult for participants to relate to listening comfort than to sound quality, at least in some of the scenarios. This could have affected both the SSL adjustments and the following perceptual assessments. Another possible explanation for the difference between sites was a difference in the definitions of listening comfort offered at the two sites. The definition used at Site 1 focused mainly on the ability to tolerate being in a given sound scenario for a long time, whereas the definition used at Site 2 focused on the pleasantness of the sound and the balance between elements in the sound picture. The latter may have been more difficult both to obtain and assess, especially because speech (babble) was part of three of the four listening-comfort scenarios used at Site 2. Interestingly, for the one scenario ("Traffic") that did not include speech, a clear SSL benefit was observed on listening comfort, similar to the results observed for the "Road" scenario included at Site 1.

At both test sites, no mean SSL benefits were observed for perceived speech intelligibility and speech clarity. The lack of a significant group effect was not very surprising: partly because comparison was made to a prescribed setting of EVOKE that was developed to optimise speech understanding, and partly because the SSL adjustments only affected the frequency response (and not, for example, parameters of the compression system). However, it is still noteworthy that 74% of the participants were able to obtain an SSL benefit on speech intelligibility or speech clarity in at least one of the speech scenarios. This indicates that SSL, at least in some situations, can also provide an improvement when the listening intention is to understand speech.

A striking observation is that all 39 participants in the two studies were able to obtain an SSL benefit in at least one of sound scenarios they were exposed to. Within each of the two main perceptual domains, sound quality and listening effort, around 70% of the participants were able to make SSL adjustments that provided a benefit across the scenarios, while around 90% of the participants were able to obtain an SSL benefit in at least one scenario. The data are summarised in Figure 6. These findings indicate that the SSL procedure is so easy to complete that a vast majority of users will be able to obtain a benefit from it, at least in some situations and for many across the board.

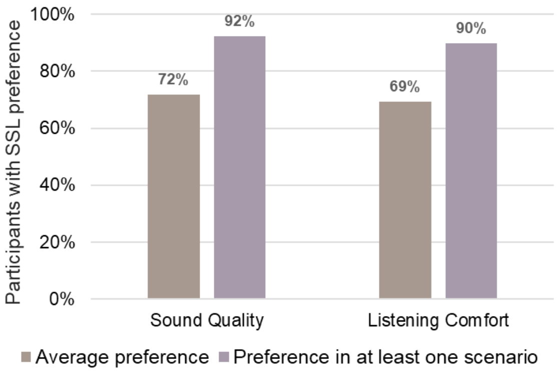


Figure 6. The share of participants in the two studies (combined N =39) showing a preference for SSL on sound quality and listening comfort, either averaged across all scenarios or in at least one of the scenarios.

Overall, the rating data show the potential of SSL to offer perceptual benefits in a variety of situations. The question is of course how well these lab benefits transfer to real life outside the lab? Actually, the results match well with the findings by Balling et al. (2019), who conducted a large-scale international survey to investigate users' experiences of Widex EVOKE. They found that a vast majority of survey respondents had been able to obtain a benefit with SSL in a real-life situation. Furthermore, they found that most of the respondents who had used SSL would recommend the feature to others.

The real-life confirmation of these results is all the more important considering the unavoidable limitations of the lab design. One such limitation is that the controlled study design used in the two lab studies to some extent contradicts the basic idea behind SSL: To allow users to make improvements in situations they decide themselves and according to their own listening intentions. In these studies, participants were ‘forced’ to make the adjustments, whether they felt a need or not, and the listening intention was dictated by the test protocol. This design – which was necessary for a controlled study – completely ignores that participants might intuitively have an entirely different listening intention in a given sound scenario than the one defined by the study. Spontaneous comments made by some participants indicated that this mismatch did indeed occasionally happen. For example, after some of the listening-comfort scenarios used at Site 2, where speech was part of the background noise, the comment was made that it felt “impolite” to listen in on the conversations of others. This type of experience may have affected both the SSL adjustments and the subsequent ratings.

Another aspect ignored by the study design is the feeling of psychological ownership associated with being actively involved in the fitting and adjustment of one’s own hearing aid (Convery et al., 2011). This feeling of being empowered may amplify the satisfaction associated with a pure perceptual benefit and thereby strengthen the entire experience of using SSL in real life.

The gain adjustment made by the participants during the SSL procedure showed a substantial variation across participants, sound scenarios and listening intentions, as illustrated in Figure 5. This is very much in line with data gathered from real-life use of SSL by Widex EVOKE users all over the world (Balling et al., 2019). The variations in gain adjustments, and the fact that it is impossible to establish clear relationships between the adjustments and the sound scenarios, the listening intentions and the resulting SSL benefits, indicate that individual listening needs and preferences are highly personal and impossible to predict. The results suggest that SSL offers a way to obtain a personalised setting of Widex EVOKE that addresses the specific listening intention of the user.

FUTURE PERSPECTIVES

The introduction of SoundSense Learn was the first step on a new journey for Widex. Driven by data generated in real time by hearing aid users, Widex hypothesised that this same data could help users identify hearing aid settings for themselves to optimise their listening experience. The results shown here confirm that hypothesis, and Widex will continue on the journey with real-life data-based insights. This journey is part of the continuous development of the Widex Real-Life Hearing framework.

Looking into the near future, the positive outcomes of using data to improve the hearing experiences of all users will only be achieved with access to more shared data. The anonymous experiences of one user, recorded in data generated by use of SSL, can potentially help another user. Sharing of data is therefore vital in the journey to improve hearing experiences for everyone, so HCPs and users can help each other for the greater good.

CONCLUSIONS

The SoundSense Learn adjustments resulted in improvements of self-rated sound quality and listening comfort, indicating the perceptual advantages of using the SoundSense Learn procedure to accommodate individual listening intentions.

A high share of participants were able to obtain a perceptual benefit of SoundSense Learn in one or more of the different sound scenarios, indicating that most EVOKE users can expect to benefit from using SoundSense Learn.

Large variations and no systematic trends were observed in the SoundSense Learn gain adjustments, indicating the need for the option to personalise adjustments to accommodate individual listening intentions. This option is provided by the SoundSense Learn feature.

Sharing of anonymous data generated during the use of SoundSense Learn is vital for the continuous development of data-driven solutions that offer improved hearing experiences for all users.

REFERENCES

Balling L.W., Townend O. & Switalski W. 2019. Real-life Hearing Aid Benefit with Widex EVOKE. *Hearing Review*, 26(3), 30-36.

Convery E., Keidser G., Dillon H. & Hartley L. 2011. A self-fitting hearing aid: Need and concept. *Trends in Amplification*, 15(4), 157-166.

Jensen N.S. 2018a. Real-Life Hearing - Part 1: The theory behind. *WidexPress*, 39.

Jensen N.S. 2018b. Real-Life Hearing - Part 2: Assessment and solutions. *WidexPress*, 40.

Jensen N.S., Hau O., Nielsen J.B.B., Nielsen T.B. & Legarth S.V. 2019. Perceptual Effects of Adjusting Hearing-Aid Gain by Means of a Machine-Learning Approach Based on Individual User Preference. *Trends in Hearing*, 23, 1-23.

Nielsen J.B., Nielsen J. & Larsen J. 2015. Perception-based Personalization of Hearing Aids using Gaussian Processes and Active Learning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(1), 162-173.

Townend O., Nielsen J.B. & Balslev D. 2018a. SoundSense Learn – Listening Intention and Machine Learning. *Hearing Review*, 25(6), 28-31.

Townend O., Nielsen J.B. & Ramsgaard J. 2018b. Real-life applications of machine learning in hearing aids. *Hearing Review*, 25(4), 34-37.