



HAL
open science

Using Automatic Speech Recognition to Optimize Hearing-Aid Time Constants

Lionel Fontan, Libio Gonçalves Braz, Julien Pinquier, Michael Stone,
Christian Füllgrabe

► **To cite this version:**

Lionel Fontan, Libio Gonçalves Braz, Julien Pinquier, Michael Stone, Christian Füllgrabe. Using Automatic Speech Recognition to Optimize Hearing-Aid Time Constants. *Frontiers in Neuroscience*, 2022, 16, pp.779062. 10.3389/fnins.2022.779062 . hal-03627441

HAL Id: hal-03627441

<https://hal.science/hal-03627441>

Submitted on 3 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Using Automatic Speech Recognition to Optimize Hearing-Aid Time Constants

Lionel Fontan^{1*†}, Libio Gonçalves Braz^{2*†}, Julien Pinquier², Michael A. Stone³ and Christian Füllgrabe⁴

¹ Archean LABS, Montauban, France, ² IRIT, CNRS, Université Paul Sabatier, Toulouse, France, ³ Manchester Centre for Audiology and Deafness, School of Health Sciences, University of Manchester, Manchester, United Kingdom, ⁴ School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, United Kingdom

OPEN ACCESS

Edited by:

Qinglin Meng,
South China University of Technology,
China

Reviewed by:

Jonathan Matthew Vaisberg,
The University of Western Ontario,
Canada
Zhenyu Guo,
South China University of Technology,
China

*Correspondence:

Lionel Fontan
lfontan@archean.tech
Libio Gonçalves Braz
libio.goncalves@irit.fr

† These authors share first authorship

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 17 September 2021

Accepted: 14 February 2022

Published: 17 March 2022

Citation:

Fontan L, Gonçalves Braz L,
Pinquier J, Stone MA and Füllgrabe C
(2022) Using Automatic Speech
Recognition to Optimize Hearing-Aid
Time Constants.
Front. Neurosci. 16:779062.
doi: 10.3389/fnins.2022.779062

Automatic speech recognition (ASR), when combined with hearing-aid (HA) and hearing-loss (HL) simulations, can predict aided speech-identification performances of persons with age-related hearing loss. ASR can thus be used to evaluate different HA configurations, such as combinations of insertion-gain functions and compression thresholds, in order to optimize HA fitting for a given person. The present study investigated whether, after fixing compression thresholds and insertion gains, a random-search algorithm could be used to optimize time constants (i.e., attack and release times) for 12 audiometric profiles. The insertion gains were either those recommended by the CAM2 prescription rule or those optimized using ASR, while compression thresholds were always optimized using ASR. For each audiometric profile, the random-search algorithm was used to vary time constants with the aim to maximize ASR performance. A HA simulator and a HL simulator were used, respectively, to amplify and to degrade speech stimuli according to the input audiogram. The resulting speech signals were fed to an ASR system for recognition. For each audiogram, 1,000 iterations of the random-search algorithm were used to find the time-constant configuration yielding the highest ASR score. To assess the reproducibility of the results, the random search algorithm was run twice. Optimizing the time constants significantly improved the ASR scores when CAM2 insertion gains were used, but not when using ASR-based gains. Repeating the random search yielded similar ASR scores, but different time-constant configurations.

Keywords: random search, automatic speech recognition, hearing aids, age-related hearing loss, compression speed, attack time, release time

INTRODUCTION

Recent studies have shown that automatic speech recognition (ASR) can be used, in combination with signal-processing algorithms mimicking the effects of hearing loss (HL), to predict the speech-identification performances of older hearing-impaired (OHI) listeners [see Schädler et al. (2015) and Fontan et al. (2020a) for a discussion of the advantages of ASR-based metrics by comparison to other objective measures of speech intelligibility]. This was demonstrated for

unaided (Kollmeier et al., 2016; Schädler et al., 2018; Fontan et al., 2020a) and aided (using simulated or real hearing aids; Fontan et al., 2020b; Schädler et al., 2020) speech perception.

Based on these findings, it has been speculated that ASR-based prediction systems could also be used to assess speech-intelligibility benefits resulting from various hearing-aid (HA) configurations. Recently, Fontan et al. (2020c) used an ASR system to evaluate and to improve the insertion gains recommended by the CAM2 HA fitting rule (Moore et al., 2010b). For each of their hearing-impaired (HI) participants, 625 gain functions (corresponding to systematic variations of CAM2 gains by 0, ± 3 , or ± 6 dB) were assessed. Each gain function was applied to speech stimuli using an HA simulator. The amplified speech material was then degraded using the HL simulator developed by Nejime and Moore (1997). Based on each participant's audiogram, both the elevation of hearing thresholds and loudness recruitment were mimicked. Spectral smearing, which is also implemented in the original HL simulator to mimic the loss of frequency selectivity, was not used by Fontan et al. (2020c), since its simulation resulted in weaker correlations between ASR scores and human speech intelligibility (Fontan et al., 2020a). Finally, the amplified and degraded stimuli were fed to the ASR system for computing recognition scores. Fontan et al. (2020c) compared the benefits associated with the insertion-gain function yielding the highest ASR scores (the "optimized" gains yielding a mean improvement of 13 percentage points) to those obtained with CAM2 gains in a group of OHI participants. Significantly higher human speech-identification scores were observed for speech amplified with optimized gains than for speech amplified according to the gains recommended by CAM2. These significant improvements were observed both for word and sentence materials.

Gonçalves Braz et al. (2022) extended this work and combined ASR with several random-search (RS) algorithms to optimize not only insertion gains but also compression thresholds. This approach is referred to as OPRA-RS, which stands for "Objective Prescription Rule based on ASR and Random Search." Using slow time constants for the compressor of the simulated HA, optimized insertion gains and compression thresholds were determined for 12 audiometric profiles corresponding to different levels of HL severity. ASR scores yielded by the optimized parameters were significantly higher than those obtained with CAM2 (mean improvements ranged from 2 to 10 percentage points for the different RS algorithms). Significant differences were observed between RS algorithms in terms of ASR score and convergence speed.

A limitation of Gonçalves Braz et al.'s (2022) study is that only one set of time constants was used. However, aided speech intelligibility depends on the attack and release times of the HA compressor (Moore et al., 2011; Hopkins et al., 2012). Small time constants (i.e., "fast" compression) help perceiving rapid changes in loudness, such as those occurring when a weak speech sound (e.g., a consonant) precedes or follows a speech sound with higher energy (e.g., a vowel; Souza, 2002; Hopkins et al., 2012). At the same time, when fast compression is implemented in a multi-channel HA that processes each frequency channel independently, it tends to reduce spectral contrasts (i.e., by

flattening the speech spectrum) and may thus have a deleterious effect on the perception of speech formants, which are crucial for the identification of vowels (Bor et al., 2008). By causing rapid variations of the signal amplitude at the onset and offset of speech sounds, fast compression speeds can also distort the signal envelope and therefore negatively impact speech intelligibility (Stone and Moore, 1992, 2008; Stone et al., 2009). These distortions are more likely to happen when high compression ratios are used (Verschuure et al., 1996). Despite their impact on speech intelligibility, there is currently no consensus as to the best time constants that should be used: time constants used clinically and commercially in hearing aids vary broadly, with attack and release times ranging from 0.5 to 2,000 ms and 10 to 5,000 ms, respectively (Moore and Søk, 2016).

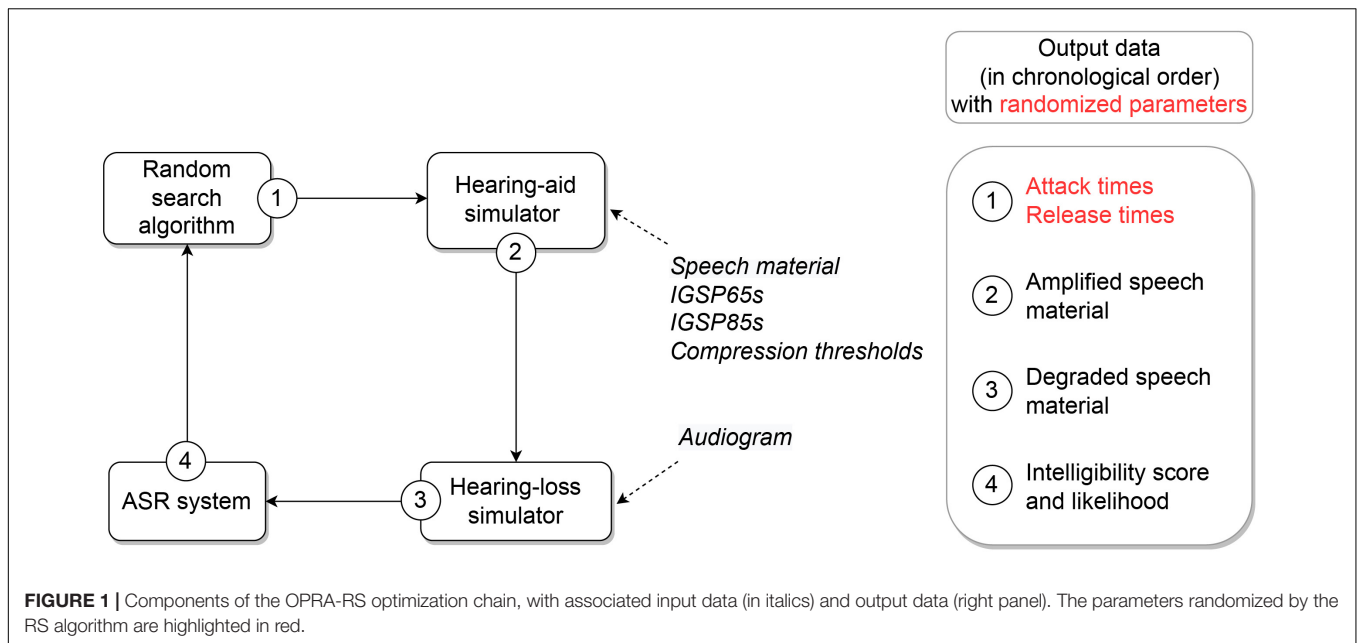
The present study extends the work of Gonçalves Braz et al. (2022) by investigating whether OPRA-RS can also be used to optimize time constants. Attack and release times were optimized for HA configurations that corresponded to the compression thresholds and/or the insertion gains recommended either by OPRA-RS or by CAM2 for the same 12 audiometric profiles as used in Gonçalves Braz et al. (2022). As these HA configurations sometimes involved high compression ratios (>3), and that in such cases, fast compression can distort the signal envelope and thus affect speech intelligibility (Souza, 2002), only "slow" compression speeds were used. To assess ASR performance, speech stimuli were first amplified using an HA simulator and then degraded to mimic the perceptual consequences of the elevation of hearing thresholds and loudness recruitment. The resulting speech signals were eventually fed to an ASR system for recognition. The optimization of compression speed was carried out twice in order to assess the reproducibility of the outcomes in terms of ASR scores and optimized time constants.

METHODS

Overview of the Optimization Chain

Figure 1 describes the processing chain used to optimize time constants for a given input audiogram. At initialization, the RS algorithm randomly selects attack and release times within two ranges of possible values. These time constants, as well as the compression thresholds and the insertion gains prescribed by OPRA-RS or CAM2 for a 65- and 85-dB-SPL speech input level (IGSP65s and IGSP85s, respectively; see **Supplementary Datasheet** for more details), are transmitted to an HA simulator. The HA simulator amplifies 50 speech stimuli corresponding to five 10-word lists of the speech intelligibility test of Fournier (1951), which is the test most often used by French audiologists for speech audiometry (Rembaud et al., 2017). The amplified speech signals are then degraded by the HL simulator, according to the input audiogram. The resulting speech signals are finally processed by an ASR system developed for the French language.

A total of N iterations are used to assess N time-constant configurations. After each iteration, the ASR score and average log-likelihood of recognized words yielded by the current time constants are compared to those obtained with the best time-constant configuration found up to the current iteration. If the



current configuration yields a higher ASR score (or the same ASR score but with a higher log-likelihood) than the previous best configuration, the current configuration is used as a baseline for the next iteration. Otherwise, the best previous configuration serves as a baseline for the next iteration.

Based on the current number of iterations i , the search ranges are reduced around the baseline time constants for the next iteration, following the equation:

$$\text{search range } (i+1) = \text{search range } (i) - \frac{\text{initial search range} - (2 \times \text{stepsize})}{N} \quad (1)$$

where *stepsize* corresponds to the step (in ms) used to define possible values within the search range.

Simulation of Hearing-Aid Processing

A 5-channel HA simulator implemented in MATLAB™ (Moore et al., 2010a) was used to amplify the speech signals. The frequency ranges of the five HA channels were 0.1–0.7, 0.7–1.4, 1.4–2.8, 2.8–5.6, and 5.6–8 kHz. In each channel, the simulator used two dynamic range compressors placed in series: the wide dynamic range compression function was applied in the first compressor, while the second compressor was used as a limiter. For further details about the implementation of the HA simulator, see Fontan et al. (2020c).

Simulation of Hearing Loss

The functioning of the HL simulator, also implemented in MATLAB™, is detailed in Nejime and Moore (1997). As done in Gonçalves Braz et al. (2022), the simulator was used to mimic two of the perceptual consequences of age-related HL: Based on the input audiogram, a linear filter simulated the elevation of hearing thresholds, while loudness recruitment was simulated by raising the signal envelope (Moore and Glasberg, 1993).

Automatic Speech Recognition System

The ASR system used in the study consisted of Hidden Markov Models and Gaussian Mixture Models. It was implemented using the Julius ASR engine (Lee and Kawahara, 2009). The acoustic models were trained on approximately 100 h of French radio broadcast news. These speech recordings were not processed to mimic HA amplification or HL and did not include the 50 word recordings used in the study to evaluate time constants. The lexicon used by the ASR system only comprised the 50 target words. A more detailed description of the ASR system is given in Gonçalves Braz et al. (2022).

Test Procedure

The processing chain was used to optimize time constants for 12 audiograms, using the compression thresholds selected by OPRA-RS and the insertion gains prescribed either by OPRA-RS or by CAM2. The audiograms, shown in Figure 2, represented mean or individual audiometric thresholds falling into levels 4–7 of the Wisconsin Age-Related Hearing Impairment Classification Scale (WARHICS; Cruickshanks et al., 2020). The audiograms corresponded to mild-to-moderately severe losses, with thresholds generally increasing as a function of frequency, as is typical of age-related HL. The mean audiograms were based on the data collected by Humes (2021). Some of the hearing thresholds required by the HL simulator (corresponding to the frequencies 0.125, 0.25, 0.75, and 1.5 kHz) were not included in the mean audiograms reported by Humes (2021). Those missing thresholds were intra- or extrapolated using third-least-squares polynomial regressions. The individual audiograms corresponded to older patients (mean age: 70 years; age range: 63–78 years) with sensorineural HL. For each of the four WARHICS levels, one mean audiogram and two individual audiograms were used.

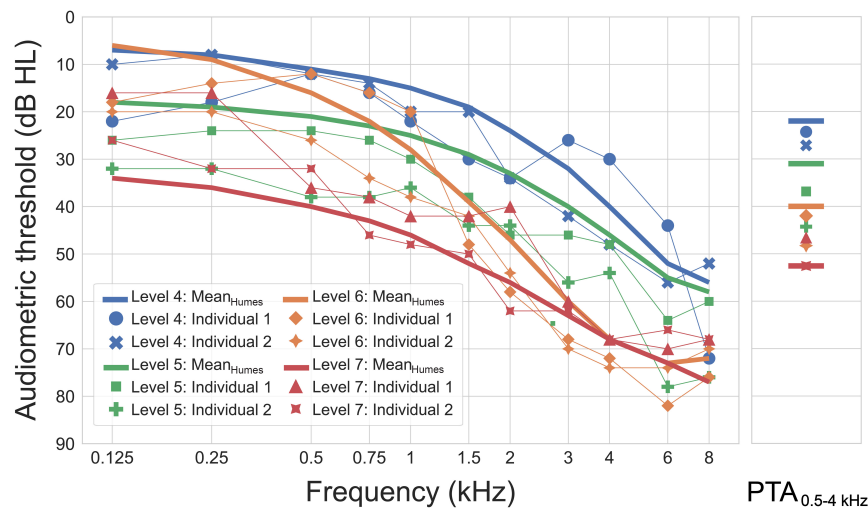


FIGURE 2 | Audiograms used as an input for the simulation of hearing loss. Corresponding pure-tone averages (PTAs) for frequencies between 0.5 and 4 kHz are shown in the right panel. Figure reproduced from Gonçalves Braz et al. (2022).

The RS algorithm that yielded the highest ASR performance in Gonçalves Braz et al. (2022) was used in the present study. This algorithm tunes all parameters (here, time constants) in all HA channels simultaneously. As in Gonçalves Braz et al. (2022), four independent RS threads were run in parallel. Each thread consisted of 1,000 iterations, during which time constants were randomly varied within predefined search ranges, using 10-ms steps. At the start of the RS, the search ranges were 10–500 ms for attack times, and 300–2,000 ms for release times. These ranges correspond to those generally associated with a slow compression system (Moore, 2008a,b; Moore et al., 2010a; Moore and Şek, 2013). For each audiogram, the final time-constant configuration yielding the highest ASR performance across the four search threads was selected. In what follows, unless explicitly mentioned, only data from the first repetition of the RS algorithm are used.

RESULTS

Figure 3 compares the ASR scores achieved with default and optimized time constants, using either the insertion gains recommended by CAM2 (left panel) or those calculated by OPRA-RS (right panel). The default time constants correspond to the fixed compression speeds used by Fontan et al. (2020c) and Gonçalves Braz et al. (2022). Those were 200, 100, 100, 100, and 100 ms for attack times, and 2,000, 1,500, 1,200, 1,000, and 1,000 ms for release times for HA channels 1–5, respectively.

With the insertion gains recommended by CAM2, it can be noticed that ASR scores tended to be higher after the optimization of time constants (median ASR score: 92%) than with default time constants (median ASR score: 88%). As Kolmogorov-Smirnov tests indicated that the ASR scores were not normally distributed ($p \leq 0.044$ in both conditions), a Wilcoxon signed-rank test was used to assess the significance of the observed difference. The results show that ASR scores

are significantly higher after the optimization of time constants ($Z = 2.8$; $p = 0.005$). In contrast with this general trend, for two out of the 12 audiograms, all time-constant configurations tested during the RS yielded lower ASR scores than those obtained with the default constants. The improvements due to the optimization of time constants seem to be larger for the most severe HLs than for milder HLs. For example, for audiograms corresponding to level 7 of the WARHICS scale, the ASR score improved by 10 percentage points on average, whereas an average improvement of 1.3 percentage point is observed for audiograms corresponding to level 4 of the WARHICS scale. A Spearman correlation was computed to assess the existence of a significant association between HL severity, represented by the pure-tone average (PTA) for frequencies of 0.5, 0.75, 1, 1.5, 2, 3, and 4 kHz, and the improvement in terms of ASR score due to the optimization of time constants. The results indicate a significant positive relationship between the two variables ($\rho = 0.62$; $p = 0.03$), that is, the higher the PTA, the larger the benefit due to the optimization of time constants.

Contrary to the ASR scores obtained with CAM2 gains, no improvement was observed after the optimization of time constants when using the gains recommended by OPRA-RS. For six out of the 12 audiograms, all time-constant configurations tested during the RS yielded lower ASR scores than those obtained with default time constants.

The reproducibility of the ASR scores was assessed by comparing the outcomes of the two repetitions of the RS algorithm. For CAM2, the median ASR score achieved during the second repetition of the algorithm (91%) was very close to the score achieved during the first repetition (92%); a Wilcoxon test revealed that no significant difference existed between the ASR scores yielded by each of the repetitions ($Z = -1.7$; $p = 0.10$). For OPRA-RS, all ASR scores remained equal across repetitions of the RS algorithm.

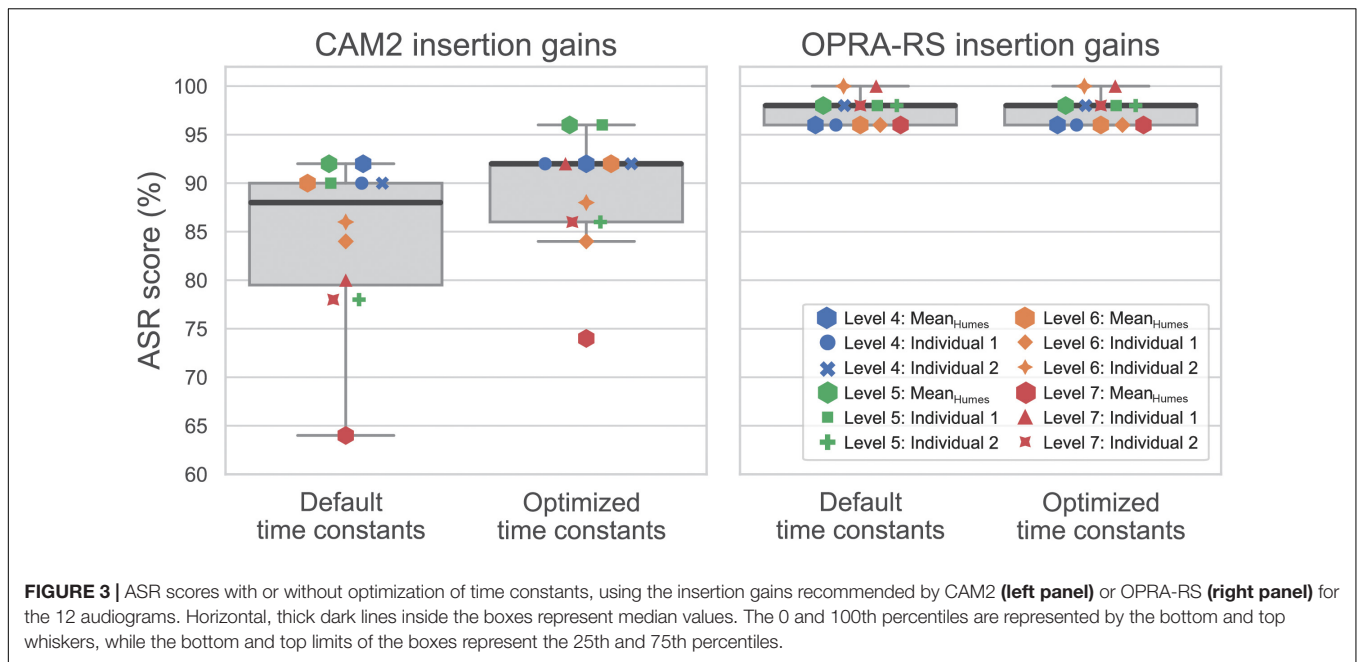


FIGURE 3 | ASR scores with or without optimization of time constants, using the insertion gains recommended by CAM2 (left panel) or OPRA-RS (right panel) for the 12 audiograms. Horizontal, thick dark lines inside the boxes represent median values. The 0 and 100th percentiles are represented by the bottom and top whiskers, while the bottom and top limits of the boxes represent the 25th and 75th percentiles.

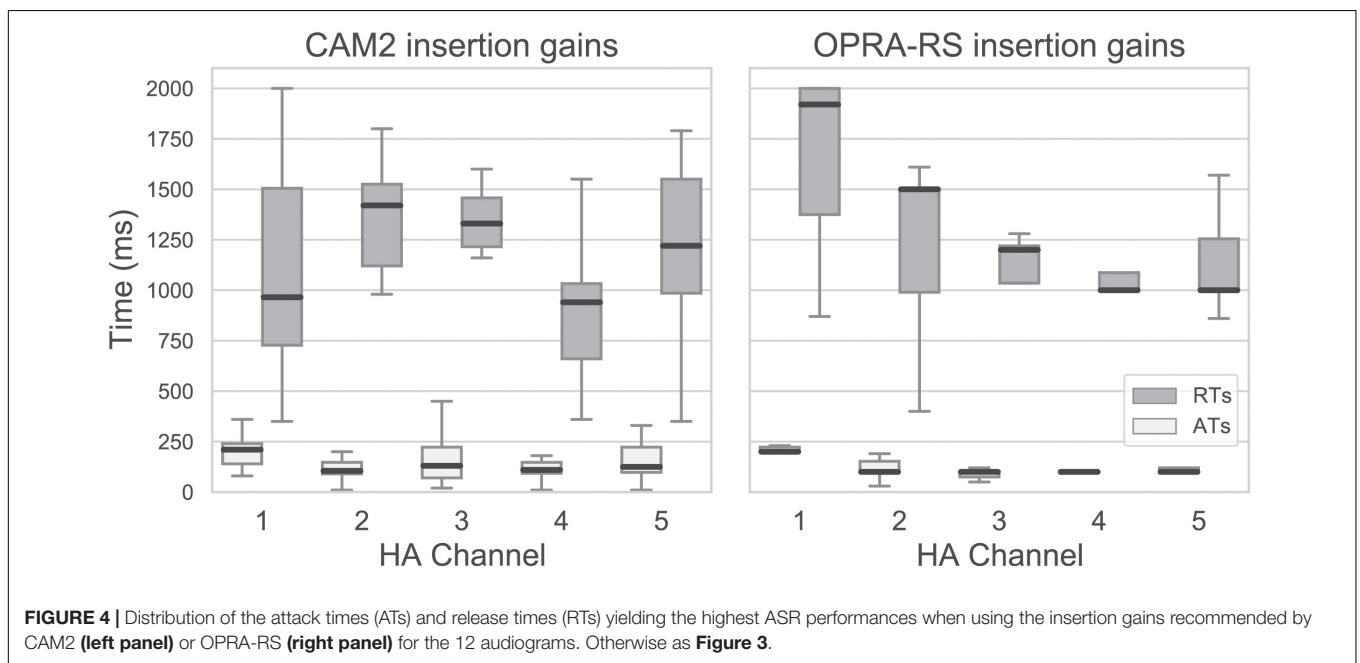


FIGURE 4 | Distribution of the attack times (ATs) and release times (RTs) yielding the highest ASR performances when using the insertion gains recommended by CAM2 (left panel) or OPRA-RS (right panel) for the 12 audiograms. Otherwise as **Figure 3**.

Figure 4 shows the distribution of attack and release times yielding the highest ASR performances for the 12 audiograms with the insertion gains recommended by CAM2 (left panel) or by OPRA-RS (right panel). In the cases for which better ASR scores were achieved with the default time constants used by Gonçalves Braz et al. (2022), these default values were retained as best configurations.

Median attack and release times across channels are 135 and 1,190 ms, respectively, for CAM2 gains, and 100 and 1,200 ms, respectively, for OPRA-RS gains. Contrary to attack

times, optimized release times span the entire possible range of values. Optimized time constants seem less variable for OPRA-RS than for CAM2. This is at least partially due to the fact that, for OPRA-RS, a larger proportion of the optimized time constants correspond to the default constants used by Gonçalves Braz et al. (2022).

Finally, the time constants obtained during the two repetitions of the RS algorithm were compared. As default time constants corresponded to fixed values, the HA configurations for which the best ASR scores were achieved with default time constants were excluded from this analysis. For the remaining HA

configurations ($N = 16$), the median absolute differences across repetitions were 85 and 355 ms for attack and release times, respectively. The minimum and maximum absolute differences were 0 and 420 ms for attack times, and 20 and 1,640 ms for release times.

DISCUSSION

This study provides proof of concept that RS can be used for the optimization of HA time constants for a given audiometric profile. This approach might prove particularly useful since there is currently no consensus as to the time constants that should be used to maximize speech intelligibility for a HI individual (Moore and Şek, 2016). It has been shown that knowledge of the HA user's cognitive abilities might help to choose slow or fast compression (Gatehouse et al., 2003; Souza and Sirow, 2014), but the results of studies addressing the relationship between hearing abilities and optimal time constants are heterogeneous (Hopkins et al., 2012; Moore and Şek, 2016). Within this context, OPRA-RS represents a novel approach that, given the audiometric profile of the HA user, can be used to systematically explore a large number of time-constant configurations and assess their impact in terms of speech intelligibility.

ASR scores for the optimized time constants were reproducible across repetitions of the RS algorithm, but were associated with different combinations of time constants. This is possibly due to an interaction between attack and release times, as the two parameters were optimized simultaneously. Future studies should optimize each parameter independently to assess their reproducibility. It might also be interesting to extend in future studies the search ranges used for attack and release times, which were limited in the present study to values generally associated with slow compression.

For the mild-to-moderately-severe HLs used in the this study, the optimization of time constants yielded significant improvements in ASR scores for CAM2, but not for OPRA-RS. In addition, the improvements observed for CAM2 were small (4 percentage points, corresponding to 2 out of the 50 words used in the study). These observations are likely due to ceiling effects in the two test conditions, even before the optimization of time constants. Indeed, it was observed that more severe HLs, yielding the lowest ASR scores with CAM2 and default time constants, were associated with higher improvements after the optimization of time constants. To limit such ceiling effects and thus to assess if clinically significant benefits can be obtained, future studies should use more challenging experimental conditions (e.g., speech materials that are shorter and/or presented in noise).

REFERENCES

- Bor, S., Souza, P., and Wright, R. (2008). Multichannel compression: effects of reduced spectral contrast on vowel identification. *J. Speech Lang. Hear. Res.* 51, 1315–1327. doi: 10.1044/1092-4388(2008/07-0009)
- Cruickshanks, K. J., Nondahl, D. M., Fischer, M. E., Schubert, C. R., and Tweed, T. S. (2020). A novel method for classifying hearing

impairment in epidemiological studies of aging: the Wisconsin Age-Related Hearing Impairment Classification Scale. *Am. J. Audiol.* 29, 59–67. doi: 10.1044/2019_AJA-19-00021

Fontan, L., Cretin-Maitenaz, T., and Füllgrabe, C. (2020a). Predicting speech perception in older listeners with sensorineural hearing loss using automatic speech recognition. *Trends Hear.* 24:4522331216520914769. doi: 10.1177/2331216520914769

Finally, it should be determined if, as shown by Fontan et al. (2020c) for the fine-tuning of insertion gains, the benefits observed in ASR performance due to the fine-tuning of time constants translate into speech-intelligibility benefits for actual listeners with age-related HL, and if these benefits are clinically relevant. Also, in the present study, CAM2 as a baseline prescription since it was used in the previous experiments on ASR-based optimization of HA parameters (Fontan et al., 2020c; Gonçalves Braz et al., 2022). It should be determined if significant improvements are also observed for those prescription rules that are more widely used in clinical practice, such as NAL-NL2 (Keidser et al., 2011).

DATA AVAILABILITY STATEMENT

The datasets generated and analyzed for this study can be obtained from the corresponding authors for any research purpose.

AUTHOR CONTRIBUTIONS

LF initiated the idea. LG designed and implemented the random-search algorithms, under the supervision of JP. MS provided scientific advice about fitting algorithms, and the hearing-aid and hearing-loss simulations. LF, LG, MS, and CF analyzed and interpreted the data. LF and CF wrote the manuscript. All authors approved the final version of the manuscript.

ACKNOWLEDGMENTS

The experiments presented in this manuscript were carried out using the OSIRIM platform (<http://osirim.irit.fr/>) that is administered by the Institut de Recherche en Informatique de Toulouse (IRIT) and supported by the French National Center for Scientific Research (CNRS), the Occitanie Region, the French Government, and the European Regional Development Fund (ERDF). MS was funded by the NIHR Manchester Biomedical Research Centre (United Kingdom).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2022.779062/full#supplementary-material>

- Fontan, L., Farinas, J., Segura, B., Stone, M. A., and Füllgrabe, C. (2020b). "Using automatic speech recognition to predict aided speech-in-noise intelligibility," in *Proceedings of the Speech-In-Noise Workshop*, (Toulouse), doi: 10.13140/RG.2.2.27785.19043
- Fontan, L., Le Coz, M., Azzopardi, C., Stone, M. A., and Füllgrabe, C. (2020c). Improving hearing-aid gains based on automatic speech recognition. *J. Acoust. Soc. Am.* 148, EL227–EL233. doi: 10.1121/10.0001866
- Fournier, J.-E. (1951). *Audiométrie Vocale : Les Épreuves D'intelligibilité Et Leurs Applications Au Diagnostic, À L'expertise Et À La Correction Prothétique Des Surdités [Speech Audiometry: Speech-Intelligibility Tests and their Application for the Diagnosis, Survey and Hearing-Aid Rehabilitation of Hearing Losses]*. Paris: Maloine.
- Gatehouse, S., Naylor, G., and Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *Int. J. Audiol.* 42, S77–S85. doi: 10.3109/14992020309074627
- Gonçalves Braz, L., Fontan, L., Pinquier, J., Stone, M. A., and Füllgrabe, C. (2022). OPRA-RS: a hearing-aid fitting method based on automatic speech recognition and random search. *Front. Neurosci.* 16:7779048. doi: 10.3389/fnins.2022.779048
- Hopkins, K., King, A., and Moore, B. C. J. (2012). The effect of compression speed on intelligibility: simulated hearing-aid processing with and without original temporal fine structure information. *J. Acoust. Soc. Am.* 132, 1592–1601. doi: 10.1121/1.4742719
- Humes, L. E. (2021). Further evaluation and application of the Wisconsin Age-Related Hearing Impairment Classification System. *Am. J. Audiol.* 30, 359–375. doi: 10.1044/2021_AJA-20-00207
- Keidser, G., Dillon, H., Flax, M., Ching, T., and Brewer, S. (2011). The NAL-NL2 prescription procedure. *Audiol. Res.* 1, 88–90. doi: 10.4081/audiores.2011.e2
- Kollmeier, B., Schädler, M. R., Warzybok, A., Meyer, B. T., and Brand, T. (2016). Sentence recognition prediction for hearing-impaired listeners in stationary and fluctuation noise with FADE: empowering the attenuation and distortion concept by Plomp with a quantitative processing model. *Trends Hear.* 20:233121651665579. doi: 10.1177/2331216516655795
- Lee, A., and Kawahara, T. (2009). "Recent development of open-source speech recognition engine Julius," in *Proceedings of the Annual Summit and Conference of the Asia-Pacific Signal and Information Processing Association (APSIPA ASC 2009)*. Sapporo 131–137.
- Moore, B. C. J. (2008a). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*. Hoboken, NJ: Wiley, doi: 10.1121/1.2982418
- Moore, B. C. J. (2008b). The choice of compression speed in hearing aids: theoretical and practical considerations and the role of individual differences. *Trends Amplif.* 12, 103–112. doi: 10.1177/1084713808317819
- Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (2010b). Development of a new method for deriving initial fittings for hearing aids with multichannel compression: CAMEQ2-HF. *Int. J. Audiol.* 49, 216–227. doi: 10.3109/1499202090329674
- Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (2010a). Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task. *J. Acoust. Soc. Am.* 128, 360–371. doi: 10.1121/1.3436533
- Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (2011). Determination of preferred parameters for multichannel compression using individually fitted simulated hearing aids and paired comparisons. *Ear Hear.* 32, 556–568. doi: 10.1097/AUD.0b013e31820b5f4c
- Moore, B. C. J., and Glasberg, B. R. (1993). Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech. *J. Acoust. Soc. Am.* 94, 2050–2062. doi: 10.1121/1.40747
- Moore, B. C. J., and Şek, A. (2013). Comparison of the CAM2 and NAL-NL2 hearing aid fitting methods. *Ear Hear.* 34, 83–95. doi: 10.1097/AUD.0b013e3182650adf
- Moore, B. C. J., and Şek, A. (2016). Preferred compression speed for speech and music and its relationship to sensitivity to temporal fine structure. *Trends Hear.* 20:2331216516640486. doi: 10.1177/2331216516640486
- Nejime, Y., and Moore, B. C. J. (1997). Simulation of the effect of threshold elevation and loudness recruitment combined with reduced frequency selectivity on the intelligibility of speech in noise. *J. Acoust. Soc. Am.* 102, 603–615. doi: 10.1121/1.41973
- Rembaud, F., Fontan, L., and Füllgrabe, C. (2017). L'audiométrie vocale en France: état des lieux [Speech audiometry in France: current clinical practices]. *Cah. Audition* 6, 22–25.
- Schädler, M. R., Hülsmeier, D., Warzybok, A., and Kollmeier, B. (2020). Individual aided speech-recognition performance and predictions of benefit for listeners with impaired hearing employing FADE. *Trends Hear.* 24:2331216520938929. doi: 10.1177/2331216520938929
- Schädler, M. R., Warzybok, A., Hochmuth, S., and Kollmeier, B. (2015). Matrix sentence intelligibility prediction using an automatic speech recognition system. *Int. J. Audiol.* 54, 100–107. doi: 10.3109/14992027.2015.1061708
- Schädler, M. R., Warzybok, A., and Kollmeier, B. (2018). Objective prediction of hearing aid benefit across listener groups using machine learning: speech recognition performance with binaural noise-reduction algorithms. *Trends Hear.* 22:233121651876895. doi: 10.1177/2331216518768954
- Souza, P. E. (2002). Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends Amplif.* 6, 131–165. doi: 10.1177/10847138020060402
- Souza, P. E., and Sirow, L. (2014). Relating working memory to compression parameters in clinically fit hearing aids. *Am. J. Audiol.* 23, 394–401. doi: 10.1044/2014_AJA-14-0006
- Stone, M. A., and Moore, B. C. J. (1992). Syllabic compression: effective compression ratios for signals modulated at different rates. *Br. J. Audiol.* 26, 351–361. doi: 10.3109/03005369209076659
- Stone, M. A., and Moore, B. C. J. (2008). Effects of spectro-temporal modulation changes produced by multi-channel compression on intelligibility in a competing-speech task. *J. Acoust. Soc. Am.* 123, 1063–1076. doi: 10.1121/1.2821969
- Stone, M. A., Moore, B. C. J., Füllgrabe, C., and Hinton, A. C. (2009). Multichannel fast-acting dynamic range compression hinders performance by young, normal-hearing listeners in a two-talker separation task. *J. Audio Eng. Soc.* 57, 532–546.
- Verschuure, J., Maas, A. J., Stikvoort, E., de Jong, R. M., Goedegebure, A., and Dreschler, W. A. (1996). Compression and its effect on the speech signal. *Ear Hear.* 17, 162–175. doi: 10.1097/53500003446-199604000-00008

Conflict of Interest: This study is part of the development of a future product/service by Archean LABS intended for hearing-aid audiologists. CF acted as a scientific consultant.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Fontan, Gonçalves Braz, Pinquier, Stone and Füllgrabe. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.