The influence of different BRIR modification techniques on externalization and sound quality

Peter Maximilian Giller, Florian Wendt, Robert Höldrich

THE INFLUENCE OF DIFFERENT BRIR MODIFICATION TECHNIQUES ON EXTERNALIZATION AND SOUND QUALITY

Peter Maximilian Giller **Florian Wendt** **Robert Höldrich**
Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz
giller@student.tugraz.at, {wendt|hoeldrich}@iem.at

ABSTRACT

In the context of binaural audio, externalization refers to the sensation of virtual sound sources being located outside of the listener's head. Binaural reproduction using anechoic head-related impulse responses is known to suffer from poor externalization. The degree of externalization can be increased by reverberation, as contained in binaural room impulse responses. However, the presence of reverberation is not always desired since the original sound of a recording should usually be preserved. This study concerns the dilemma of creating well-externalized dry-sounding signals. We investigated the manipulation of either the impulse response length, the reverberation time, or the direct-to-reverberant energy ratio regarding externalization and attributes of sound quality. As expected, each condition is a compromise between externalization and sound quality. While externalization increases with increasing amount of reverberation for all methods in a similar way, our findings show that the differences between them lie in sound color and perceived naturalness.

1. INTRODUCTION

Binaural synthesis aims to produce well externalized sound images, i.e., auditory images are perceived to be located outside the listener's head, 'compact and correctly located in space' [1]. The perceptual and technical aspects of the phenomenon were the subject of earlier research [1–6]. Studies showed that the presence of reverberation can increase the degree of externalization [2, 3], but also introduce sound colorations [4]. While, for binaural rendering, it is usually desired to preserve the room impression and sound color of the original recording, this may lead to a conflict between the synthesized and the listening room known as the room divergence effect [5].

We distinguish head-related impulse responses (HRIRs), which capture the influence of reflections at the pinnae, head, and torso from the direct sound impinging at the ear canals from a certain direction, and binaural room impulse responses (BRIRs). The latter consist of the direct part, which is identical to the HRIR, followed by early reflections and diffuse reverberation from all directions, convolved with the HRIRs for the respective directions.

The goal is to add as little reverberation as necessary to HRIRs in order to increase externalization, or to reduce the reverberation of BRIRs as much as possible in order to reduce differences in sound. The influence of BRIR truncation, the simplest method to reduce reverberation, on externalization was investigated in several studies [2–4]. It was found that a minimum BRIR length of 80-100 ms is sufficient to yield externalized sound images [2, 3]. Since there is no equivalent physical phenomenon leading to truncated BRIRs, this method yields a rather artificial sound. In contrast, natural de-reverberation can either be achieved by increasing the absorption area in a room, yielding a shorter reverberation time, or by reducing the source distance, leading to a higher direct-to-reverberant energy ratio (DRR).

In this study, we investigate the influence of modifications of the reverberant part systematically in order to establish a connection between the achieved externalization and sound quality. Three basic approaches are studied, which are: (*i*) truncation of the impulse response length (temporal modification), (*ii*) manipulation of the DRR by weighting the reverberant part with a constant factor (modification of level), and (*iii*) alteration of the reverberation time by weighting the reverberant part with an exponential decay function (temporal modification of level). Section 2 describes the above mentioned modification methods. We have compared the three approaches in a listening experiment introduced in Section 3. The participants compared the degree of externalization as well as sound quality by rating naturalness and similarity to an anechoic signal. Section 4 discusses the results of the experiment.

2. MANIPULATION OF IMPULSE RESPONSES

Fig. 1 gives an illustrative example of the studied modification techniques. The dashed black line represents the envelope of a BRIR $h(t)$ on a logarithmic scale. The vertical dashed line marks the boundary between the direct part $h_{\mathrm{dir}}(t)$, which we define as the direct sound followed by all reflections from pinnae, head, and torso (the HRIR), and the reverberant part $h_{\mathrm{rev}}(t)$, containing early and diffuse reflections from the surroundings, where $h = h_{\mathrm{dir}} + h_{\mathrm{rev}}$. Each of the modifications can be understood as a time-variant weighting of each impulse response $h(t)$, as illustrated in

doi:10.25836/sasp.2019.20

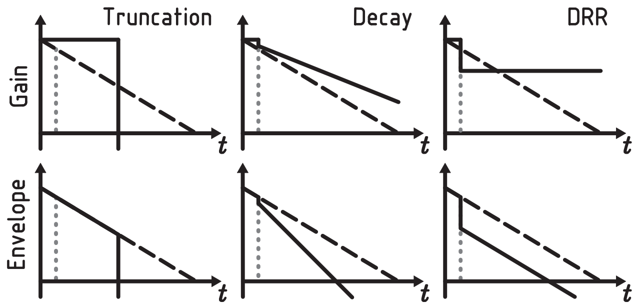

Figure 1. Illustration of the modification methods. The black and gray dashed lines show the envelope of the BRIR and the boundary between direct and reverberant part. The upper row shows the equivalent time-dependent gain, whereas the lower row shows the resulting envelope.

the upper row. The lower row of Fig. 1 illustrates the resulting envelopes.

As shown in the left column, the truncation method applies a window of length L with unit gain to the BRIR, with a short fade in the transition region where needed.

In contrast, by multiplication with an exponential decay curve the reverberation time is modified, as shown in the center column. The decay curve begins at the position of the direct sound, but, in order to leave the HRIR unaffected, is only applied to the reverberant part of the BRIR,

$$\tilde{h}_{\mathrm{rev}}(t) = h_{\mathrm{rev}}(t) \cdot 10^{-\frac{60}{20}(\tilde{T}_{30}^{-1} - T_{30}^{-1})t}, \quad (1)$$

where T_{30} is the reverberation time estimated from the original BRIR, and $\tilde{T}_{30}$ is the target reverberation time. To realize the weighting, the BRIR is split into its direct and reverberant parts, using a short overlapping window in the transition region, which are then individually weighted.

The right column of Fig. 1 exemplifies the manipulation of the DRR, which is defined as

$$DRR = 10 \lg \left(\frac{\int h_{\mathrm{dir}}^2(t) \mathrm{d}t}{\int h_{\mathrm{rev}}^2(t) \mathrm{d}t} \right). \quad (2)$$

It can be manipulated if the reverberant part of the BRIR, split analogously to the decay method, is multiplied with a constant factor.

3. LISTENING EXPERIMENT

A listening experiment was conducted to evaluate the impact of the three modifications truncation, decay, and DRR on externalization and sound quality. Externalization is known to be particularly fragile for the frontal direction which is why we investigated speech sources simulated using non-individual BRIRs direct in front of the listeners at $0°$, played back over headphones. We did not carry out individual measurements because the experiment concerns rather the relative differences to the original, unmodified BRIR than the absolute differences to a physical sound source.

The experiment consisted of three consecutive parts: The first part (part *N*) was an exploratory evaluation of the perceived naturalness of a subset of the created conditions. In the second part, participants had to rate the externalization (part *E*), and in the last part sound quality was evaluated by rating the similarity to an anechoic signal regarding different attributes of sound (part *S*).

Method		Truncation	Decay	DRR
Parameter		L (s)	T_{30} (s)	DRR (dB)
Condition i	0 (BRIR)	1.0	0.70	2.3
	1	0.350	0.60	5.3
	2	0.193	0.51	8.3
	3	0.106	0.42	11.3
	4	0.059	0.33	14.3
	5	0.032	0.23	17.3
	6	0.018	0.14	20.3
	7	0.009	0.05	23.3
	8 (HRIR)	0.003	-	∞

Table 1. Conditions created within each method of modification.

3.1 Conditions

Virtual sounds sources were created using a BRIR measurement of a Neumann KU 100 dummy head. This measurement constitutes the starting point for the creation of all other conditions and was obtained in the lecture room of our institute (dimensions $3\,\mathrm{m} \times 7\,\mathrm{m} \times 8.3\,\mathrm{m}$, reverberation time $T_{30} = 0.7\,\mathrm{s}$). The sound source, a Neumann KH 120 loudspeaker, was located at a distance of $2.5\,\mathrm{m}$ to the receiver, at an angle of $0°$. Both source and receiver were positioned at a height of $1.25\,\mathrm{m}$. To provide additional reflections from breast and shoulders, possibly supporting externalization, the dummy head was equipped with the torso of a Brüel & Kjær HATS. Different conditions were created by manipulating the length, reverberation time, or DRR of the measured BRIR. An anechoic $8\,\mathrm{s}$ long sequence of male speech was convolved with the resulting BRIRs corresponding to each of the created conditions.

Tab. 1 lists the parameter values of each condition. The condition with index $i = 0$ corresponds to the original, unmodified BRIR, whereas the HRIR, i.e., the BRIR truncated immediately before the arrival of the first reflection from the room, has the index $i = 8$. The parameter levels for each modification were selected heuristically based on the experience from preceding informal experiments of the authors to achieve, with regard to reverberation, a near-uniform sampling of the parameter ranges from unmodified BRIR to the HRIR, i.e., from congruence to divergence between the synthesized and the real room.

3.2 Playback and Equalization

All conditions were played back via headphones, except for the loudspeaker reference condition in the externalization part *E*. To facilitate comparative rating, participants wore open headphones throughout the whole experiment. Unfortunately, the headphone alters the sound from the loudspeaker reference somewhat as the sound has to propagate through the ear cups. The AKG K702 headphone was thus modified in order to reduce the damping of frontal sound as

doi:10.25836/sasp.2019.20

much as possible by replacing the ear cushions by self-made ones with cutouts at the front and back. Remaining differences in sound color between headphone and loudspeaker were equalized using a minimum-phase filter. Far-field responses from the loudspeaker to the dummy head, both with and without the modified headphone, were measured in an anechoic chamber. The magnitude spectra were smoothed within critical bands and the filter for the headphone signal was obtained by dividing the without-headphones magnitude frequency response by the with-headphones response.

Obviously, the modification of ear cushions distorts the frequency response of the headphone itself. To linearize the magnitude transfer function from each headphone driver to the corresponding ear canal, the headphones signals were convolved with an additional minimum-phase inverse filter of the magnitude spectrum smoothed within critical bands. This second equalization also removes the undesired contribution of the pinnae and ear canals (of the dummy head) to the headphone transfer function.

3.3 Experimental Design

Each part of the listening experiment was carried out in a MUSHRA-like procedure where participants had to rate a number of conditions of the three different modification techniques. Participants were asked to rate every condition in the presented set of stimuli with continuous sliders on a graphical user interface. They were allowed to repeat each condition at will, and audio files were played back in loop.

The configuration of the test is summarized in Tab. 2. Parts *E* and *S* each consisted of two stages. The first stage (*E*.I/*S*.I) we refer to as the *indirect* comparison. It consists of three sets presented in random order, each corresponding to one modification technique, i.e., one column of Tab. 1. While this setup enables us to draw a comparison between conditions of each modification, a cross-comparison between modifications can solely be achieved indirectly via comparison to the common reference, hidden reference, and anchor. Therefore, in the second stage (*E*.II/*S*.II), a *direct* cross-comparison of modifications was carried out – due to the large number of conditions only with a subset of conditions $i = \{2, 4, 6\}$ from each modification in the same set. These conditions were selected by informal listening to yield preferably similar ratings of externalization or similarity for all modifications in each of the three corresponding levels.

The ratings from the direct comparison x_i^{II} for conditions $i = \{0, 2, 4, 6, 8\}$ were then used to obtain complete corrected curves x_i for each listener by linear scaling and shifting of the ratings of the indirect comparison x_i^{I} for conditions $i = \{1, 3, 5, 7\}$. With $i = 1...8$, this yields a complete set of ratings

$$x_i = \begin{cases} x_i^{\mathrm{II}} & \text{even } i, \\ x_{i-1}^{\mathrm{II}} + \frac{x_{i+1}^{\mathrm{II}} - x_{i-1}^{\mathrm{II}}}{x_{i+1}^{\mathrm{I}} - x_{i-1}^{\mathrm{I}}} \left(x_i^{\mathrm{I}} - x_{i-1}^{\mathrm{I}} \right) & \text{odd } i \end{cases} \quad (3)$$

per listener for each method, allowing for a cross-comparison between the methods.

Part		Method (abbreviation)	Reference	Hidden Reference	Conditions	Anchor
N		all	-	0	2,4,6	8
E	I	Truncation (*trc*)	LS	0	1-7	8
		Decay (*dec*)	LS	0	1-7	8
		DRR (*drr*)	LS	0	1-7	8
	II	all	LS	0	2,4,6	8
S	I	Truncation (*trc*)	8	8	1-7	0
		Decay (*dec*)	8	8	1-7	0
		DRR (*drr*)	8	8	1-7	0
	II	all	8	8	2,4,6	0
	IIa-b	all	8	8	2,4,6	0

Table 2. Experimental setup. Numbers correspond to the condition indices in Tab. 1, and *LS* refers to the loudspeaker.

Part		Task: ‘Rate the...
N		...naturalness!’
E	I-II	...externalization, compared to the reference!’
S	I-II	...similarity to the reference *in general*!’
	IIa	...similarity regarding *sound color*!’
	IIb	...similarity reg. the *amount of reverberation*!’

Table 3. Task definitions of each part of the experiment.

The first two parts, *N* and *E*, were conducted in the original room at the position of the measurement. The task definitions of each part are listed in Tab. 3. Part *N* was conducted without a reference, and, thus, the loudspeaker did not play, on which the participants were informed. The participants were asked to rate the naturalness in general, not necessarily bound to the particular room and distance to the loudspeaker. The rating had to be entered on a scale from 0 (‘very unnatural’) to 100 (‘entirely natural’). The stimuli were identical to *E*.II. In part *E*, the loudspeaker was used as a reference to remind the participants of the impression of full externalization of a distant and compact physical sound source. The signal convolved with the BRIR (cond. 0) was used as a hidden reference. The participants were asked to rate the externalization compared to the loudspeaker on a scale from 0 (‘inside head’) over 33 (‘close to the head’) to 100 (‘at the position of the loudspeaker’). They were instructed not to move their head during playback.

In part *S*.I-II, participants were asked to rate the general similarity to a reference on a scale from 0 (‘very different’) to 100 (‘identical’). In addition, they were asked to rate the similarity regarding sound color or reverberation (*S*.II.a-b) for the reduced set. The reference and hidden reference were the anechoic speech signal convolved with the HRIR (cond. 8). The conditions used were the same as in the first two parts. In order to avoid the influence of spatial attributes on the rating, all stimuli including the reference were presented monaurally by playing back the left-ear signal for both ears. Furthermore, it was conducted in an anechoic chamber to decouple the rating from the measurement room. The stimuli were presented over the modified headphones.

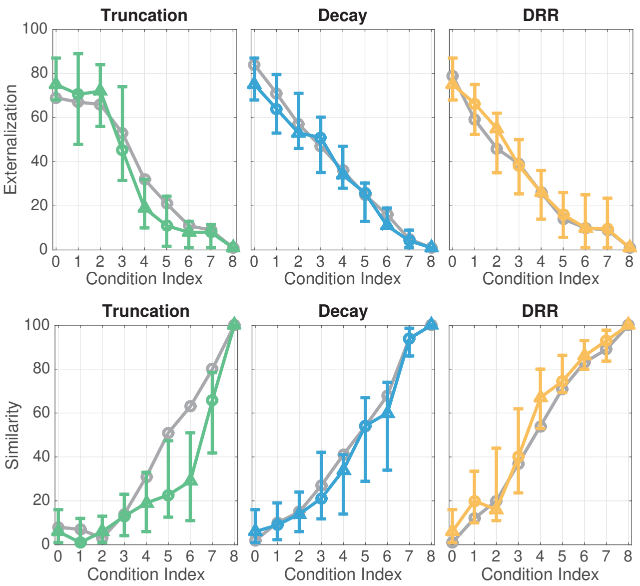

Figure 2. Median original *(gray)*, as well as the scaled ratings *(colored)* with 95 % confidence intervals. The condition indices refer to Tab. 1. Externalization is shown in the upper, and similarity in the lower row. Triangular shapes mark conditions of the direct comparison.

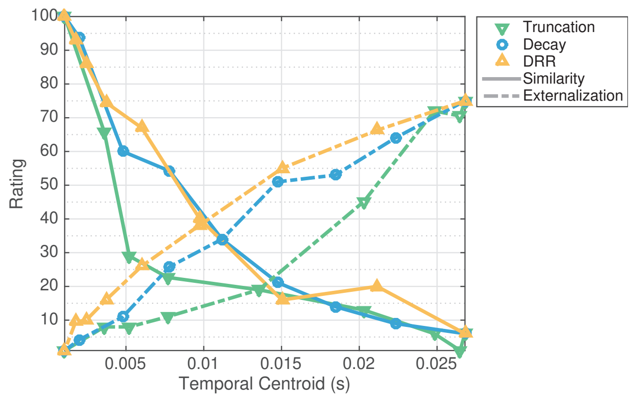

Figure 3. Median scaled ratings of externalization and similarity over the temporal centroid of the weighted BRIRs of each condition.

4. RESULTS

Twenty-one experienced listeners participated in the experiment. A Friedman test showed that the effect of the BRIR modification is significant with $p < 0.05$ within every part of the experiment. Fig. 2 shows the median ratings of experiments *E*.I-II and *S*.I-II for each of the modifications truncation, decay, and DRR. For brevity, we will refer to the methods simply as *trc*, *dec*, and *drr* in the following, with, e.g., *dec* i denoting the i-th condition of decay modifications. The externalization ratings are shown in the upper and the similarity ratings in the lower row. The results of the indirect comparisons, *E*.I and *S*.I, are plotted as gray lines in the background. Colored curves represent the ratings corrected based on the direct comparison, *E*.II and *S*.II. The conditions tested in the direct comparison are marked with triangular shapes.

As expected, the overall relation between modification depth and externalization as well as similarity is monotonic, where increasing modification depth leads to a decrease in externalization and an increase in similarity to the HRIR.

4.1 Representation on a Common Axis

The ratings of the different modifications are not directly comparable due to the different nature of the respective varied parameter. In order to relate them to a common physical measure, we computed the temporal centroid of the impulse responses of all conditions. The impulse responses were priorly weighted with the average spectrum of the anechoic speech signal order to limit the evaluation to the frequency content presented to the participants. Fig. 3 shows the scaled externalization and similarity ratings against the temporal centroid on the horizontal axis. Note that the externalization curves almost describe a straight line for *dec* and *drr*, indicating a good prediction for the methods that have a physical equivalent. Note also the downward deviation of the curve for *trc*, showing that similarly externalized conditions are associated with more late energy in the BRIR. The curves for similarity do not differ much between *dec* and *drr*, but, again, the curve for *trc* deviates towards lower ratings. The temporal centroid is primarily unrelated to fluctuations of binaural cues or timbral properties, which are known to have an effect on externalization [2, 6]. Thus, we do not claim it to be the best approach for a fair comparison. However, it appears to be a suitable measure to connect the ratings to the physical properties of the signal content and reveal differences between the methods therein.

In the following, we will only consider the ratings obtained in the direct comparison, since ratings for naturalness and the sub-attributes of similarity were recorded with the reduced stimulus set for simplicity reasons. The ratings of parts *N* and *S*.II-IIb are shown in Fig. 4. In order to compare each of these attributes to the corresponding externalization rating, *E*.II is shown in the background as a gray line. We performed paired comparisons between the conditions of each part using the Wilcoxon signed-rank test.

4.2 Externalization

As expected, the full BRIR (cond. 0) is externalized best, whereas the HRIR (cond. 8) is not externalized. All other conditions are rated significantly lower than the BRIR. The relation between externalization and the degree of modification, i.e., the amount of reverberation that is being removed from the BRIR, is monotonic within each method. Paired comparison showed that the differences of neighboring levels within one modification, as well as the differences of any of the levels to the full BRIR or the HRIR, are significant.

The BRIR yielded median ratings of only around 80 %. We attribute the lower ratings mainly to individual differences between the listeners' and the employed generic HRIR, as well as remaining timbral differences since equalization was carried for the dummy head. With the loudspeaker available for comparison, the setup was particularly sensitive.

doi:10.25836/sasp.2019.20

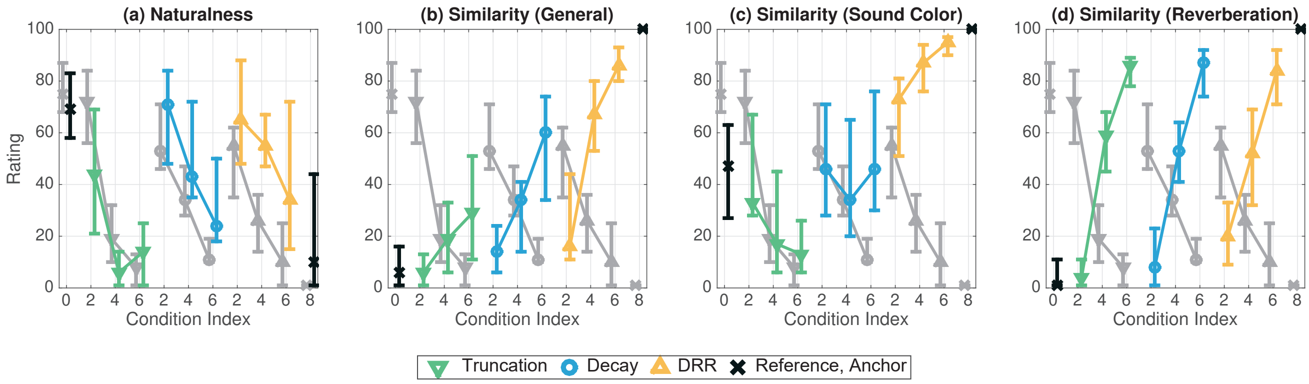

Figure 4. Median and 95 % confidence intervals of the ratings of the direct comparison (*E*.II and *S*.II-IIb). Gray lines in the background represent externalization, and colored lines in the foreground naturalness and similarity to the HRIR. The condition indices refer to Tab. 1.

4.3 Naturalness

The rating of naturalness in Fig. 4(a) similarly tends to decrease monotonically towards less reverberant conditions. Monotony appears to be interrupted between *trc* 4 and 6, but the difference is not significant. Truncation yields lower ratings than the other methods: The maximum rating is significantly lower than the maximum ratings of the other methods. All three conditions of *trc* do not significantly differ from the HRIR, whereas all conditions of the *drr* method and *dec* 2 and 4 do. Furthermore, all truncated conditions are rated significantly lower than the the BRIR, whereas for decay and DRR only cond. 6 is. The rating of the HRIR exhibits a high variation, indicating that anechoic conditions are not necessarily perceived unnatural.

4.4 Similarity

The general similarity (Fig. 4(b)) to speech convolved with the HRIR increases monotonically towards less reverberant conditions within each method. The differences between the levels of each method are significant. All conditions but *trc* 2 differ significantly from the BRIR in their rating. Cond. 2 exhibits the best externalization ratings for each of the modifications, and *trc* 2 is rated significantly better externalized than all conditions but the BRIR. It received a lower similarity rating than the other two methods, too. An increase in externalization with decreasing similarity to the dry reference was expected. However, there appear to be differences between the modifications, as *dec* 4 was rated higher than *trc* 4 in similarity, despite also being rated higher in externalization. The same is true for *drr* 4 and *trc* 6.

The similarity to the HRIR regarding sound color is shown in Fig. 4(c). While the DRR ratings increase monotonically and significantly towards less externalized conditions, the truncation ratings actually decrease with a significant difference between levels *trc* 2, 4, and 6. Moreover, all conditions of *drr* are rated significantly better than all conditions of *trc*. Although no distinct trend is visible for the decay method, *dec* 6 is rated significantly higher than *dec* 4. While none of the ratings of *dec* differs significantly

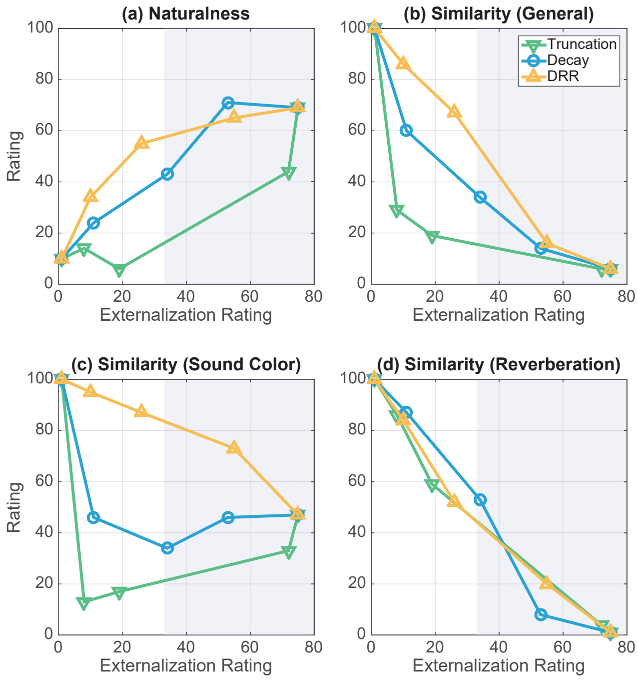

Figure 5. Median of the ratings of the direct comparison, where the naturalness and similarity ratings (parts *N* and *S*.II-IIb) are plotted against externalization (*E*.II). The highlighted area marks conditions considered externalized.

from the BRIR, *trc* 4 and 6 are rated lower than, and all of the DRR modifications are rated higher than the BRIR.

Fig. 4(d) shows the similarity regarding reverberation. At first view, cond. 4 and 6 of each method are perceived similarly reverberant. The ratings within each method increase monotonically and significantly with modification depth (towards less externalized conditions). Yet, in contrast to sound color, *trc* 2 is rated lower than cond. 2 of the other methods. All conditions are rated significantly higher than the BRIR, except for *trc* 2 and *dec* 2. While, again, *drr* 2 is rated higher than *trc* 2, the differences are less pronounced than for sound color. The similarity ratings of the truncation method with regard to sound color decrease while increasing with regard to reverberation, from which it may be inferred that the difference in similarity is due to timbral artifacts. The increase within the DRR method for both attributes may indicate that this method is less likely to produce those artifacts.

4.5 Discussion

We gain a more intuitive view on the data by plotting naturalness and similarity against externalization on the horizontal axis (Fig. 5). The highlighted area marks externalized conditions, i.e., ratings exceeding the 'close to the head' threshold. No modification provides well-externalized conditions that, at the same time, yield satisfactory ratings of similarity to the HRIR in general and with regard to reverberation. While there is little variation between the methods regarding reverberation, the differences in general similarity are highest at conditions with poor externalization, where the DRR manipulation is preferred over the other two methods. The differences in similarity regarding sound color for externalized conditions are more distinct. Again, the DRR method is preferred while truncation received the lowest ratings. This trend is similar for naturalness.

Informal listening lead us to the impression that truncated BRIRs of medium and short lengths stand out from the other methods with timbral colorations, most likely caused by comb filters due to interference of the early reflections with the direct sound. This may very well explain lower ratings regarding sound color and naturalness. In contrast to the modification of the reverberation time or the DRR, the early reflections are unaffected by truncation until very short lengths. Reducing the reverberation, however, may lead to a de-masking of the otherwise inaudible comb filters. Though, it should be noted that anechoic conditions, while sensitive to timbral artifacts, are rarely encountered in real life.

We investigated our hypothesis that the general rating of similarity may be decomposed into the ratings with regard to sound color and the amount of reverberation. We used multiple linear regression in order to determine the contribution of the constrained to the general ratings, as well as the interaction thereof. With y denoting the general similarity rating, and x_{sc} and x_{rev} the ratings regarding sound color and reverberation, we compared models for every possible combination of x_{sc} and x_{rev}, the interaction term $\bar{x} = \sqrt{x_{\mathrm{sc}} x_{\mathrm{rev}}}$ (the geometric mean), and an additive constant. We used the BIC [7] and R^2 as criteria for model selection. The model that simultaneously minimizes the BIC and maximizes R^2 is the the geometric mean $y \sim c \cdot \bar{x}$, with $c = 0.9$ and $R^2 = 0.69$.

5. CONCLUSION

In this contribution, we presented an experiment comparing three modification techniques regarding externalization and sound quality. We defined sound quality as the similarity to an anechoic reference signal, as we are interested in BRIRs that yield decent externalization while preserving the original sound of a recording at the same time. We showed that, for the present experiment, it is plausible to explain overall similarity by the similarity regarding sound color and reverberation as the main contributing factors.

For each method, we saw a monotonic relationship between modification depth and a decrease in externalization, associated with an increase in similarity. Beyond that, the different methods are incommensurable. A comparison between the the methods may, however, be drawn either by relating them to a common quantity, which can be either a physical measure (in our case: the temporal centroid), or another response variable recorded for the same conditions (in our case: externalization). Since the comparison via a third quantity must be interpreted with caution, we consider the latter more meaningful.

Each of the methods has its own benefits. While truncation is obviously the right choice to yield short impulse responses, it has no physical equivalent and may thus sound unnatural. It can lead to timbral artifacts which, again, to avoid is the strength of the DRR modification. To modify the reverberation time seems to be a good compromise, since it also reduces the effective length of the BRIR and may therefore be combined with truncation.

Our findings may contribute to future research in two different ways: On the one hand, they provide a foundation for the investigation of hybrid modification methods to combine, e.g., the good timbral properties of the DRR and decay method with truncation in order to yield short impulse responses. On the other hand, an analysis of the modified BRIRs regarding the binaural cues may help to further understand the mechanisms of externalization.

6. REFERENCES

[1] W. M. Hartmann and A. Wittenberg, "On the externalization of sound images," *J. Acoust. Soc. Am.*, vol. 99, no. 6, 1996.

[2] J. Catic, S. Santurette, and T. Dau, "The role of reverberation-related binaural cues in the externalization of speech," *J. Acoust. Soc. Am.*, vol. 138, no. 2, 2015.

[3] S. Li, R. Schlieper, and J. Peissig, "The effect of variation of reverberation parameters in contralateral versus ipsilateral ear signals on perceived externalization of a lateral sound source in a listening room," *J. Acoust. Soc. Am.*, vol. 144, no. 2, 2018.

[4] R. Crawford-Emery and H. Lee, "The subjective effect of BRIR length on perceived headphone sound externalization and tonal coloration," in *Audio Eng. Soc. Conv. 136*, AES, 2014.

[5] S. Werner, F. Klein, T. Mayenfels, and K. Brandenburg, "A summary on acoustic room divergence and its effect on externalization of auditory events," in *8th Int. Conf. on Quality of Multimedia Experience (QoMEX)*, 2016.

[6] H. G. Hassager, F. Gran, and T. Dau, "The role of spectral detail in the binaural transfer function on perceived externalization in a reverberant environment," *J. Acoust. Soc. Am.*, vol. 139, no. 5, 2016.

[7] A. E. Raftery, "Bayesian model selection in social research," *Sociological methodology*, vol. 25, 1995.

 doi:10.25836/sasp.2019.20