PROCEEDINGS of the
International Symposium on Room Acoustics

15 to 17 September 2019 in Amsterdam, Netherlands

Beyond 80ms: The Subjective Effects of Sound Energy Arriving Shortly After the "Early" Sound Period

Evan GREEN[1]; Eckhard KAHLE[1]; Vincent BERRIER[1]; Emilie CARAYOL[1]

[1] Kahle Acoustics, Brussels, BELGIUM

ABSTRACT

Sound energy arriving in the first 80ms after the direct sound is the subject of much room acoustics research and design work, a time interval where reflected energy is expected to be subjectively associated or fused with the direct sound. The perceptual effects of sound arriving shortly after 80ms (usually bundled into a "late energy integral") are however less clear: energy arriving in this time period may be masked by the early sound, but may also have perceptual effects on both source and room presence.

This time period also overlaps with the calculation interval for the EDT (300ms for RT=2s), a parameter correlated with perceived running reverberation. Listener envelopment has also been shown to correlate strongly with lateral sound level after 80ms. Are there acoustical attributes in the time period after 80ms that are necessary to enhance subjective acoustical effects such as running reverberation and listener envelopment?

To establish the subjective effects of sound energy arriving between 80ms and 300ms – termed "cluster reflections" – listening tests having been conducted in binaurally reproduced 3D virtual acoustic environments with varying reflection level, delay and direction of arrival. These experiments indicate that there are distinct time and spatial regions that are responsible for generating the subjective impression of envelopment and running reverberation. Only for a particular combination of delay times between 150ms and 300ms and directions of arrival between 60° and 150° azimuth were simultaneous running reverberation and listener envelopment generated. In addition, complex masking processes between the direct, early and two time regions of the cluster reflections have been found.

Keywords: Envelopment, Running Reverberation

1. INTRODUCTION

Many of the earliest discoveries in psychoacoustics and room acoustics were achieved through the use of highly simplified synthesized sound fields generated using loudspeakers in anechoic environments. Experiments with loudspeakers to simulate a frontal direct sound plus one or two reflections led to, amongst many others, the following important discoveries in room acoustics: the Precedence Effect and the finding that early reflections before 50ms are fully integrated with the direct sound leading to an increased subjective loudness, both by Haas (1) and that a broadening of the sound source (initially called "spatial impression", now called "Apparent Source Width") is generated by early lateral reflections – Barron and Marshall (2,3). Many similar experiments were reported between the 1950s and the 1980s; the results form the foundation of our understanding of the subjective effects of early reflections.

Experiments with highly simplified sound fields lend themselves very well to the investigation of the "early" time period (up to around 50-60ms). In this time period reflections are generally fused with the source and disturbing colouration and/or echoes are only generated in extreme situations. Perhaps because the subjective effect of individually audible "late" reflections is associated with a "disturbance" of the direct sound, investigations into the effects of late sound have generally used sound fields with continuous, evenly decaying reverberation.

Such investigations have revealed that the predominant subjective acoustical effects associated with reflected sound arriving after the early time period are reverberance and listener envelopment (LEV). It has been further found that the subjective effect of reverberation should be separated into

[1] egreen@kahle.be

terminal reverberation (audible after the music stops) and *running reverberation* (audible while the music is playing).

The EDT parameter has been positively correlated with the sense of running reverberation: for a reverberation time of 2 seconds, the decay rate for the first 10dB of decay is measured over approximately 300ms. Subsequent work, by for example Griesinger (4), has identified that reflected energy arriving in the time period 150ms to 350ms is necessary to generate running reverberation.

The subjective effects of listener envelopment have been researched in great detail by Bradley and Soulodre (5), as well as Morimoto (6). One of the key results from Bradley's group is that the degree of listener envelopment can be well predicted by the late lateral strength $G_{lateral}(80ms\text{-}\infty)$ i.e. the strength integrated after 80ms measured with a figure-of-eight microphone.

One of the challenges however with using long energy integrals or decay rates estimated by linear regression is that details are lost in the statistical analysis: when integrating over a long interval of 80ms to ∞ or using linear regression, it is no longer possible to determine which shorter time periods within the longer interval are most responsible for particular subjective effects. For this reason, and even though the result is not particularly musical, it can be very instructive to listen to music in the presence of single late reflections to elucidate their subjective effects.

2. LISTENING EXPERIMENTS WITH LATE REFLECTIONS

In the following, experiments are described with single reflections and groups of reflections in the range 70ms to 300ms. The term "cluster reflections" was developed by IRCAM to distinguish between two time periods of reflection and the later reverberation – this term will be used here (7).

All sound fields were synthesized using a tool developed in-house with MaxMSP and SPAT (8) software, whereby synthesised 3D sound fields including direct sound, up to 64 reflections and reverberation can be presented over headphones. The system uses individually measured HRTFs to create a 3D binaural presentation and head-tracking was used for all experiments. To provide some consistency with experiments into LEV by Bradley's group, an excerpt of Handel's Watermusic from the Denon anechoic orchestral recordings was used (9). The overall listening level for all experiments was in the range 70dB(A) with some differences due to number and level of reflections used – the sound level was measured under the headphones using a class 1 sound level meter.

2.1 Perceptual Associations of Cluster Reflections

In a first experiment, listeners were presented with direct sound and a single reflection with delay times of either 70ms, 80ms, 100ms or 150ms. All sound was presented in the horizontal plane and the azimuth of the reflection was adjustable by the participant. Two reflection levels were tested: -10dB and -15dB with respect to the level of the direct sound. Although such reflections would typically be characterized as "disturbing", since the reflections were low in level, it was found that listeners could make valuable judgments about the envelopment created by such reflections. Only for the 150ms reflection was the echo-effect of the -10dB reflection too strong to be able to make useful judgments. The results presented are the average of three participants, all experienced with such listening tests. Initial tests indicated that a single reflection, with appropriate direction-of-arrival and level, could create the sensation of almost complete envelopment on one side of the head i.e. that reflected sound was coming from one hemisphere.

Participants were asked to judge the degree of envelopment as a percentage of being "fully surrounded" for reflections arriving horizontally. A value of 100% would indicate that the sound was perceived as filling one hemisphere. The azimuth of the reflection was altered in 10° steps. The results of this experiment are illustrated in Figure 1.

For all reflections in this time range there was a range of azimuths where a strong sense of envelopment was generated, with the stronger reflections also generating a greater sense of envelopment. The perceived envelopment was strongest for reflections arriving from 60-150° with sharp drops in the subjective envelopment for angles smaller and larger than this range.

In contrast with early reflections where subjective spatial impression (Apparent Source Width) peaks for reflections arriving from ±90° and varies smoothly according to a cosine rule (see Figure 2 reproduced from Reference 3), for cluster reflections, the degree of listener envelopment was relatively constant for reflections arriving in the range 60-150°.

Reflections in the range 0-40° were subjectively frontally associated and did not generate strong listener envelopment, with subjective envelopment rapidly increasing in the range 40-60°. In the range 150-180°, subjective envelopment dropped off again – this corresponds with findings from Morimoto (6) that sound only from the rear contributes little to the feeling of envelopment. Lateral and especially "rear lateral" sound is necessary to generate the sensation of "complete" envelopment.

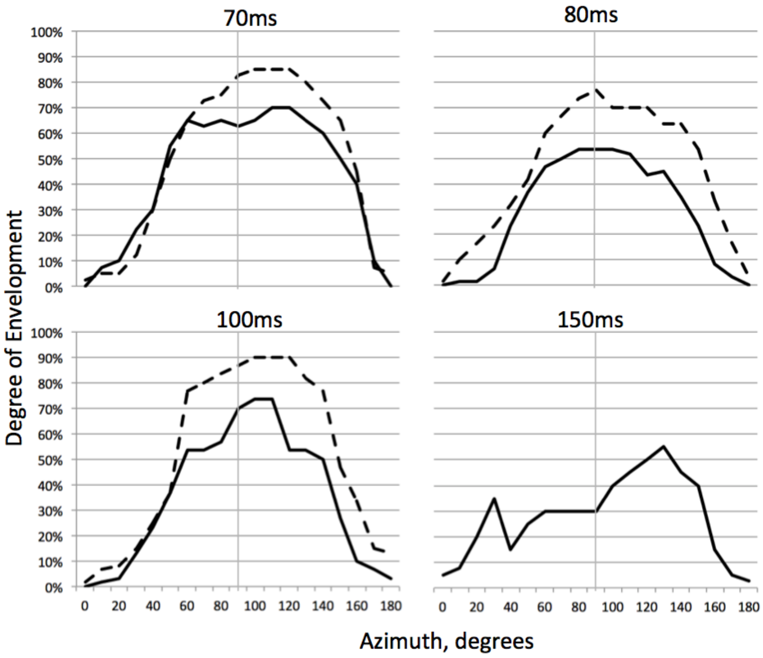

Figure 1 – Degree of envelopment generated by a single reflection as a function of azimuth, delay time and level. Solid line = -15dB reflection, dashed line = -10dB reflection. Vertical grey line indicates 90° azimuth.

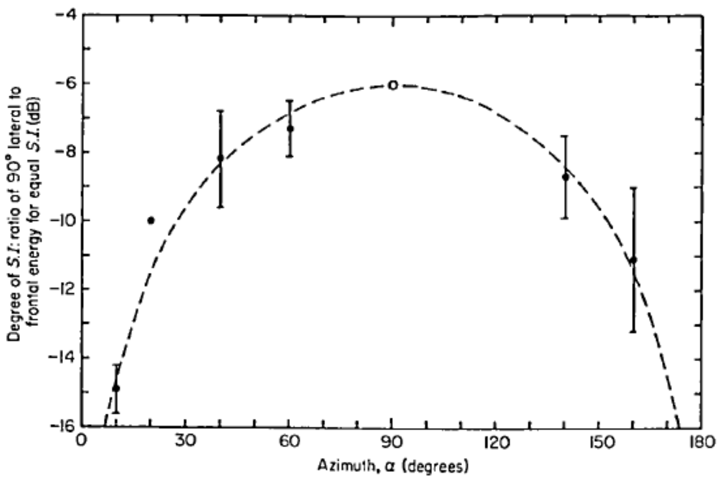

Figure 2 - Degree of spatial impression (Apparent Source Width) for a single reflection with delay 40ms (after Figure 7 in (3))

The levels used in this first experiment were all stronger than those of a specular reflection having undergone geometrical spreading i.e. level calculated from the delay time and associated path length. When the levels of the reflections were set to a level determined by geometrical spreading, even when combining four reflections with delays 70ms, 80ms, 100ms and 150ms the total reflected level was only -24dB relative to the direct sound. With four reflections presented with azimuths in the range 60-150°, the audibility of the reflections and sense of envelopment were almost absent. Only by increasing the combined level of the reflections to around -10dB relative to the direct sound level was the sense of envelopment strongly present and satisfying for music listening.

It is also important to note the quality of the sound and the quality of the envelopment generated by reflections with delays in the range 70-150ms:

- these reflections created the clear sense of the sound source being present inside a room, but with little to no impression of the size of the room;
- with a binaural headphone presentation, these reflections externalized the sound, something which reflections under 70ms generally did not do;
- for most azimuths, the 70ms reflection and 80ms reflection were not fused with the direct sound and contributed positively to the sensation of envelopment;
- the envelopment effect can be compared to being in a small room with reflective surfaces;
- reflections in this time range did not generate a perception of running reverberation.

In a next experiment, single reflections were auditioned in the presence of direct sound at delays greater than 150ms. The echo-effect of such reflections was very strong, but valuable judgements about the subjective effect could still be made. As has been found by Griesinger (4) reflected sound arriving between 150ms and 300ms is necessary to generate a sense of running reverberation and this is supported by the experiments reported here. As noted above, when reflections in the time range 150-300ms are absent, there is a sense of a room but with almost no indication of the size or character of the space – reflections after 150ms created the sense of a large volume and, depending on the level of the reflections, of a large-room reverberation heard in parallel with the music. For delays longer than 300ms, while keeping the reflection strength constant, the strength of running reverberation began to diminish – Griesinger has indicated that this time cut-off is related to the source material and masking from subsequent musical material (4).

Increasing the level of the reflections (while keeping delay times constant) tended to increase the perceived volume of the space and of the perceived reverberation time. Very strong subjective differences in spatial quality due to the direction of arrival of reflections in the time range 150-300ms were heard:

- when reflections in the range 150-300ms arrived from an azimuth in the range 60 – 150°, the perceived effect was of simultaneous running reverberation and listener envelopment;
- reflections from more frontal directions generated a sense of running reverberation but without envelopment: the subjective feeling was that the music was in front of you and that you were looking in on the event;
- reflections from azimuthal directions greater than 150° also generated running reverberation, but again with little envelopment.

The results of the two experiments described above are summarized in Figure 3: large blue dots indicate delay times and azimuths most responsible for generating listener envelopment while small green dots indicate combinations where running reverberation is strongly perceived. Where envelopment and running reverberation are perceived simultaneously, the dots overlap.

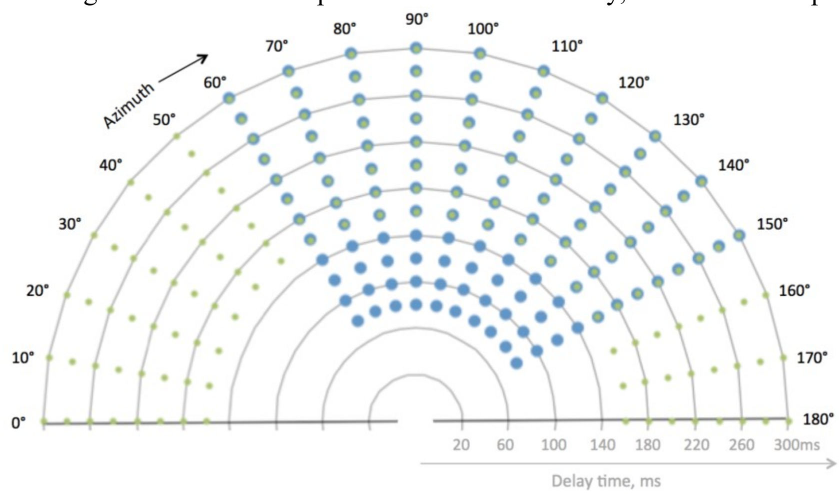

Figure 3 – Summary of reflection azimuths and delay times predominantly responsible for generating listener envelopment (blue dots) and running reverberation (green dots).

2.2 Masking of Cluster Reflections

A second set of experiments were carried out, this time with multiple cluster reflections along with early reflections. A "typical" location in a concert hall was simulated with a source-receiver distance of 20m (G_{direct} = -6dB).

Eight early reflections were generated with delays in the range 30-60ms (delay times all prime numbers to avoid simple relationships) and levels were set according to their path length (specular reflection and geometrical spreading), azimuths for the early reflections were frontal and in the range ±45° (4 each side, non-symmetrical) with the resulting $G_{early(0\text{-}80ms)}$ being 0dB. Reflections in the range 60-80ms were not included as these tended to be ambiguous with regard to attribution to the source or to the room.

Eight cluster reflections were generated with delay times in the range 90-300ms: again the delay times were prime numbers to avoid simple relationships and the level was initially set according to the reflection path length, leading to a total G_{late} for all eight reflections of -8dB. An equal number of reflections was presented on the left and right, the azimuth range selected was ±70° to ± 150°, corresponding to the range in Experiment 1 which generated the maximal subjective envelopment effect. Again, no reverberation was used in these experiments. A list of all reflections with delay times and azimuths is given in Table 1.

Table 1 – Reflection set #1. Grefl is the equivalent G for the single reflection.

Direct and Early Refl.	Delay, ms	Azimuth	Grefl, dB	Cluster Reflections	Delay, ms	Azimuth	Grefl, dB
D.S.	0	0°	-6.0	-	-	-	-
E1	29	16°	-9.5	C1	89	71°	-14.1
E2	31	-15°	-9.7	C2	109	-70°	-15.2
E3	37	25°	-10.3	C3	127	82°	-16.1
E4	41	-27°	-10.7	C4	139	-84°	-16.6
E5	43	32°	-10.8	C5	173	126°	-18.0
E6	47	-37°	-11.2	C6	239	-137°	-20.2
E7	53	42°	-11.7	C7	281	-131°	-21.3
E8	59	-45°	-12.1	C8	293	134°	-21.6

Using the Handel motif, the subjective effect of the early reflections alone (with no cluster reflections) on the direct sound was of strengthening and of modest source-broadening. The sound with only early reflections was not externalized in headphone listening tests.

When listening with only direct sound and cluster reflections (no early reflections, G_{early} -6dB; G_{late} -8dB), the sound was fully externalized and the sensation was of simultaneous running reverberation and envelopment. In this test, the lack of reflected energy between the direct sound and first cluster reflection at 90ms was clearly audible and the effect was disconnected and not musical.

With all reflections engaged (direct sound, early reflections and cluster reflections; G_{early} 0dB; G_{late} -8dB) the decay was much smoother, but the early reflections masked the cluster reflections which were only just audible; the sense of running reverberation and envelopment in this situation was very weak. Reducing the cluster reflections by 2dB (G_{early} 0dB; G_{late} -10dB) resulted in them being almost fully masked and all effects of running reverberation and envelopment were gone – this result corresponds with Bradley's findings that if the late sound level is 10dB less than the early level, the late sound becomes almost fully masked (10). Increasing the level of the cluster reflections above this level increased the perceived running reverberation and envelopment.

From this experiment it can be concluded that early reflections mask the cluster reflections and that, for an audible running reverberation and envelopment to be perceived, the cluster reflections should not be less than around 8dB weaker than the early sound level (direct sound plus early reflections).

The experiment above was repeated with direct sound, frontal early reflections and cluster reflections, but this time the cluster reflections were also frontal with azimuths in the range ±45° (Reflection set #2, Table 2). All other parameters were kept constant.

Table 2 – Reflection set #2. Changes to reflection set #1 highlighted.

Direct and Early Refl.	Delay, ms	Azimuth	Grefl, dB	Cluster Reflections	Delay, ms	Azimuth	Grefl, dB
D.S.	0	0°	-6.0	-	-	-	-
E1	29	16°	-9.5	C1	89	**16.0**	-14.1
E2	31	-15°	-9.7	C2	109	**-15.0**	-15.2
E3	37	25°	-10.3	C3	127	**25.0**	-16.1
E4	41	-27°	-10.7	C4	139	**-27.0**	-16.6
E5	43	32°	-10.8	C5	173	**32.0**	-18.0
E6	47	-37°	-11.2	C6	239	**-37.0**	-20.2
E7	53	42°	-11.7	C7	281	**42.0**	-21.3
E8	59	-45°	-12.1	C8	293	**-46.0**	-21.6

Reflection levels were again set according to the path length for a specular reflection. Unsurprisingly the perception of the entire sound field shifted to the front and envelopment was significantly reduced; nevertheless, the perceived " spatial spread" of the sound was around 50% in the horizontal plane, and this even though all reflections were coming from the frontal quarter-circle.

From experience in real rooms, in particular during works to improve stage acoustics, we had expected that subjective distance to the source would increase when late reflections came from frontal directions. This was certainly the case, with a distinct increase in perceived source-receiver distance when the late reflections were frontal, but what was surprising was that the running reverberation also sounded stronger and smoother when it came from the front. It is unclear why this is.

It might be due to the fact that the same number of reflections were compressed into a smaller azimuthal range and were therefore less spatially separated and less individually perceivable than in the previous experiment, leading to a stronger sense of reverberance. It is also evident that there is some sort of psychoacoustic interaction between perceived strength of reverberation, perceived clarity and subjective source-receiver distance. When the cluster reflections are frontal, clarity is reduced and subjective source receiver distance is increased. This leads to many questions, such as: which aspects of the sound drive this sensation, what are the processes in our auditory processing that lead to this sensation, and how much does expectation play a role in this sensation? This is a topic that alone deserves much more research attention.

In the final experiment, reflection set #1 was again used with cluster reflections arriving from ±70° to ± 150°. This time however the eight cluster reflections were divided into a group of four early cluster reflections with delay times in the range 90-150ms and four late cluster reflections with delay times in the range 150-300ms to investigate possible masking effects between these delay ranges – this division was based on the results of section 2.1 above i.e. those reflections which did and did not generate running reverberation. With the direct sound, early reflections and all late cluster reflections (150-300ms) initially all set at levels according to the reflection path length, the level of the early cluster reflections (90-150ms) was first reduced. The effect of reducing the 90-150ms reflections was to reduce the masking of the late cluster leading to an increase in the perceived running reverberation.

By reducing the level of the early cluster reflections, a reduction in the envelopment would be expected, but since both the early and late cluster reflections generate envelopment, this seemed to be offset by unmasking the envelopment effect of the late cluster reflections: overall the change in envelopment was not conclusive and hard to distinguish in the context of other subjective changes. In these experiments where the early cluster level was reduced by up to 4dB, envelopment tended not to reduce while perceived running reverberation did increase. More refined experiments would be required to fully understand the masking effects of early cluster reflections on late cluster reflections. Nevertheless, this simple test indicates that there is masking present between these time ranges.

3. DISCUSSION

By combining the findings in the existing literature on running reverberation, listener envelopment and masking, one could have in fact hypothesized all of the above results. Research by Griesinger (4) has already demonstrated that running reverberation is predominantly generated by reflected sound energy between 150ms and 350ms, while Bradley (5) has found that listener envelopment is primarily generated by lateral reflections arriving after 80ms. Taking these results together, one could have expected that lateral reflections arriving between 150ms and 300ms would simultaneously create running reverberation and listener envelopment. The above experiments demonstrate that this is indeed the case.

Furthermore, the results presented here indicate the following "perceptual time periods" in the impulse response:

- *direct sound*;
- *early reflections* arriving after the direct sound to 60ms. In most situations, these reflections are fused with the source (for music);
- an *ambiguous time range* between 60ms and 80ms where reflections can be attributed either to the source or to the room depending on the delay time, direction-of-arrival and level of the reflection;
- *early cluster reflections* arriving between 80ms and 150ms. These reflections are attributed to the room, and if above the masking threshold they create the perceived effect of a small room. Depending on the direction of arrival, they can generate envelopment but do not create running reverberation;
- *late cluster reflections* arriving between 150ms and 300ms. These reflections are always attributed to the room and if above the masking threshold they generate running reverberation. As the level of such reflections is increased, they create the sense of a larger reverberant space with a longer reverberation time. Depending on the direction of arrival, these reflections can also generate listener envelopment simultaneously with running reverberation.

These time periods broadly align with those found by researchers at IRCAM: the main difference being that the IRCAM group defined two early reflection periods from 20-40ms and 40ms-80ms called R1 and R2, along with one cluster period R3 from 80ms-160ms and finally a late sound period from 160ms-∞ (11).

In terms of masking, the general finding from existing literature is that, with sufficient energy in a given time period, earlier time periods will mask the next later time period. The experiments described in this paper support this conclusion and indicate that there are complex masking interactions that affect the subjective strength of running reverberation and envelopment. A question for future research is the degree to which and for how long energy in certain time periods masks later time periods? The experiments reported here indicate that the relevant time periods may be relatively short, and that long integrals (e.g. 80ms - ∞ as used in G_{late} and $G_{late,lateral}$) would tend to hide the relevant details. Griesinger has proposed a parameter RR160, being equal to the ratio of energy integrated between 160ms and 320ms to the total energy arriving before 160ms, as an indicator for running reverberation (4). From the results of the experiments above, this would seem to be a useful parameter, however Griesinger has found that RR16 does not "always give a good match" with subjective impression. The question of masking, and in which time periods and spatial zones, seems key to understanding the energetic balances necessary to produce running reverberation and envelopment and more work on this front is necessay.

4. CONCLUSIONS

Experiments with single reflections have shown that reflections arriving in the time range 70-150ms generate listener envelopment, with the strongest effect created by reflections arriving from azimuths between 60 and 150°. Outside this azimuthal range, the degree of envelopment drops off quickly. Higher reflection sound levels have been shown to result in a greater perceived degree of envelopment.

Reflections in this time range do not however generate running reverberation – for this effect to be perceived, reflections in the time range 150-300ms are required. In general, all directions-of-arrival generate running reverberation, however the subjective effects of reflections in this time arriving frontally or laterally are highly significant. Frontal reflections with azimuths up to ±45° arriving after 150ms produce a subjectively stronger running reverberation and with a greater sense of source-receiver distance compared to the same reflections arriving from lateral directions. For reflections arriving between 150-300ms with azimuths between 60 and 150°, a simultaneous perception of running reverberation and listener envelopment can be generated.

These results generally support previous findings regarding the relevant time range for the generation of running reverberation and provide a refinement regarding the relevant time and spatial ranges for listener envelopment – it is postulated that if the perception of reflected sound and of running reverberation drops off after 300ms, then it would be logical that the contribution of reflected sound after 300ms to listener envelopment should also diminish.

Finally, listening tests have established that audible masking effects exist between the time range 90-150ms and 150-300ms, the latter being responsible for the generation of running reverberation. The existing of masking between these time ranges, and the possible release of masking to enhance running reverberation is clearly of relevance to concert hall design and the development of more refined acoustical parameters to predict running reverberation and envelopment.

REFERENCES

1. Haas H. Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache. Acustica. 1951;1(2).
2. Barron M. The subjective effects of first reflections in concert halls – the need for lateral reflections. Journal of Sound and Vibration. 1971;15(2).
3. Barron M, Marshall H. Spatial impression due to early lateral reflections in concert hall – the derivation of a physical measure. Journal of Sound and Vibration. 1981;77(2).
4. Griesinger D. How loud is my reverberation. Audio Engineering Society Convention. 1995;98.
5. Bradley J, Soulodre G. The influence of late arriving energy on spatial impression. J Acoust Soc Am. 1995;97(4).
6. Morimoto J, Iida K, Sakagami K. The role of reflections from behind the listener in spatial impression. Applied Acoustics. 2001;62.
7. Carpentier T, Noisternig M, Warusfel O. Twenty Years of Ircam Spat: Looking Back, Looking Forward. 41st International Computer Music Conference (ICMC), Sep 2015, Denton, TX, United States.
8. SPAT software: https://forumnet.ircam.fr/product/spat-en/
9. Denon Anechoic orchestral music recording. Audio CD, Denon Records, 1995. ASIN: B0000034M9.
10. Bradley J, Reich R, Norcross S. On the combined effects of early- and late-arriving sound on spatial impression in concert halls. J. Acoust. Soc. Am. 2000;108 (2).
11. Kahle E. Validation d'un modèle objectif de la perception de la qualité acoustique dans un ensemble de salles de concerts et d'opéras. Doctoral thesis, IRCAM, Paris. 1995.