The effect of the sound level on the apparent source width and the listener envelopment

Pedro Lladó
Aalto University
Master's Programme CCIS / AAT

pedro.llado@aalto.fi

Abstract

The spatial perception of the sound is an important factor to consider when studying concert halls' perceived quality. It is in music dynamic situations when the spatial impression gains special importance. Multiple studies have tried to map the acoustic measured characteristics that correspond to apparent source width and listener envelopment. A literature review of the current knowledge about the factors that affect these spatial attributes is presented, which often focuses on the interaural cross-correlation. In parallel, the psychoacoustics factors that contribute to spatial perception, which mainly focus on interaural differences, is reviewed. While these two approaches are closely related, adding a physiologically accurate approach may help to understand how the sound level affects the spatial perception.

1 Introduction

Concert halls are venues designed so the audience enjoys the music performances at their best. Even though this enjoyment depends on multiple aspects, such as the orchestra or the piece that is being performed, the venue plays an important role on the perceived quality of the sound. Certain concert halls have gained popularity due to the subjective evaluation of listeners over time.

Multiple studies have been conducted in an attempt of matching the physical characteristics of these venues with the perceived quality. Understanding the objective features that influence the listeners' preference helps develop knowledge that would further improve the quality of the future concert halls' acoustics. The clarity of the music, the bassiness of the sound, the loudness or the spatial impression, among others, are perceived attributes that listeners often refer to when describing concert halls.

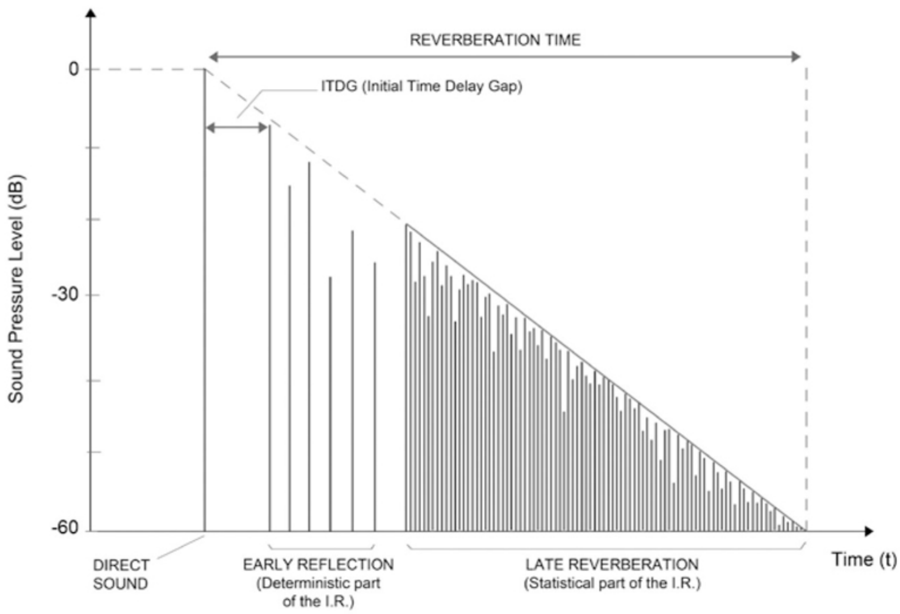


Figure 1: Initial time delay gap (ITDG) example in the context of a generic impulse response (IR). Image from D'Orazio and Garai (2017)

The spatial component of sound has been gaining importance to the point that nowadays it is one of the most studied attributes in the field. For many listeners there is a connection between concert hall preference and the spatial impression that the venue conveys. In this manuscript, a literature review about spatial sound characteristics and their perceptual fundamentals are reviewed. In particular, the focus is put on how the sound level affects the perceived spatial impression is analyzed. At the same time, fundamental literature on psychoacoustics is investigated to understand how the sound level has an effect in spatial perception.

L. Beranek (1996) proposed a list of terms to describe the subjective attributes perceived in concert halls. Since the focus of this review article is put on the spatial aspects of the sound, only the ones related to these spatial characteristics will be overviewed.

1.1 Intimacy

The first attribute found in L. Beranek (1996) related to spatial attributes is intimacy, also called presence. It is the attribute that defines the size of the space the listeners perceive they are in or whether the listener feels acoustically involved or detached from the music. Initially, it was considered to be described by the time difference between the direct sound and the first reflected sound, so called initial-time-delay gap (ITDG, or t_I) (see Figure 1).

However, intimacy has been a controversial term, since its meaning refers to a feeling that depends also on visual attributes (J. R. Hyde, 2002). L. L. Beranek (2004) stated later that maybe the ITDG was not necessarily descriptive of the concept of intimacy, and that intimacy relates to "loudness of the overall sound, since the listener assumes that performance sounds louder in a small room than in a large one". The measured ITDG was separated from the concept of intimacy, since it was considered to be an unfortunate term to name the acoustic phenomenon (J. R. Hyde, 2019). The concept of intimacy seems to be better correlated to the total sound level, even though the ITDG ignores it (Barron, 1986; Barron, 1988; Lokki & Pätynen, 2020).

Independently from the perceived intimacy, the ITDG could be, potentially, a useful metric to characterize a concert hall. However, there are two clear points of criticism to ITDG from the current state of the art point of view. The first one is that this ITDG was initially defined at one specific point in the concert hall, at a location near the center of the main floor. It has been shown that if measurement are taken in different positions in the hall, the ITDG varies greater within seats in the same concert hall than among concert halls in the same established position. The second one is that ITDG does not inform about the spatial direction from the reflections, and this has been found to be an important aspect to assess the quality of a concert hall (Lokki & Pätynen, 2020). Thus, due to the criticism to the ITDG and the new paradigm that relates intimacy to overall loudness directly, it is not further studied in this review the relation between the sound level and the intimacy attribute.

1.2 Spaciousness: apparent source width and listener envelopment

The lateral reflections give a listener the subjective spatial impression of the sound (Barron & Marshall, 1981; A. Marshall & Hyde, 1979). They were also considered by L. Beranek (1996) under the attribute spaciousness. Two attributes were defined in this context, which have some important aspects in common but differ conceptually: the apparent source width (ASW) and the listener envelopment (LEV). The ASW is defined as the size of the sound source that a listener associates with the sound. The LEV is representative of the diffusion of the reverberant sound field.

Both ASW and LEV are highly related to the ratio between the energy of the lateral reflections and the total energy. Two different measures have been extensively used to quantify this aspect: (1) The lateral fraction (LF), ratio of a figure of eight microphone compared to that arriving from all directions in the first 80 ms after the direct sound; and (2) the interaural cross-correlation (IACC), measured from microphones either in a person's or a dummy head's ears. Hidaka et al. (1995) showed that IACC was more accurate than LF, since the measured values of LF overlapped between concert halls and could not separate their perceived spatial impression.

Often, the $IACC_A$ is used to refer to the IACC coming from all angles and for the whole duration of the reverberation time (Ando, 2012). $IACC_E$ (early) and $IACC_L$ (late) are usually defined separately, since they have shown to correlate to different psychoacoustic attributes. In a first attempt, de Keet (1968) tried to relate the (1 - $IACC_{50ms}$) to the ASW. Hidaka et al. (1995) tried to correlate the ASW with different $IACC_t$, being t the time of the early reflections after the direct sound. They concluded that 80 ms correlated better across different concert halls and locations.

Due to its wavelength, IACC becomes very small below 300 Hz. Also Okano et al. (1994) found that the low frequency content influences the ASW a bit differently than the rest. The experiment results showed that ASW widens about 2 degrees for each decibel change in frequencies below 355 Hz, while maintaining the whole level constant. When the level at high frequencies was varied, holding the low-freq level constant, the change was 0.6 degrees for each decibel.

The subjective spatial impression grows with the sound level (A. Marshall & Hyde, 1979). Okano et al. (1998) studied the relationship between $IACC_{E3}$, LF_{E3} and ASW ($E3$ is used to refer to the early reflections in the three octave bands of 500 Hz, 1 kHz and 2 kHz). They found that the $IACC_{E3}$ combined with the strength factor below 250 Hz (G_{low}) correlated with the order of subjective ASW of the studied concert halls. Both ASW and LEV seem to be affected by G at mid frequencies (500 Hz - 1000 Hz) (L. Beranek, 2011). The G factor is at the same time closely related to loudness.

In this review article, the main interest will be focus in the relationship between the loudness and the spatial attributes ASW and LEV. In Section 2, the theory regarding the psychophysical aspects that cause the spatial perception of sound is studied. The literature regarding how the relationship between loudness and the spatial attributes have been investigated is reviewed in Section 3. The psychophysical evidences and theories are compared to the conducted studies in Section 4.

2 Theory and hypotheses of spatial perception and its relation to the sound level

The humans' ability to perceive the size of a sound source is, in general, not very accurate. The ability to perceive the spatial distribution of a spatially extended sound source increases when the sounds are impulsive or noise-like (Pulkki & Karjalainen, 2015). On the other hand, the perception tends to be very narrow if several sources emit sinusoids of the same frequency and wider if it is white noise. To understand this phenomena, it is important to analyze the sound from a binaural point of view.

Binaural cues are defined as the timing and level differences between the two ears, the so-called interaural time differences (ITD) and interaural level differences (ILD), respectively. These binaural cues are combined into a single perceived direction in later stages of the auditory processing. In particular, the superior olivary complex is known to be sensitive to ITD and ILD and appear to have a key role in azimuth localization (Grothe, 2003; Tollin, 2003). These cues are frequency dependent, and they can provide conflicting information when they are decoded. This can lead to a worsening of precision of the listener's ability to localize the sound source.

Griesinger (1997, 1999) presented an hypothesis which related spaciousness to the level and time behavior of the sound field. They wanted to study the spatial attributes used in concert hall acoustics from a psychoacoustics point of view. That is, exploring the relationship between the perception of spaciousness and the binaural cues. For that, it was crucial to understand how lateral reflections cause ITD and ILD fluctuations when they interfere with the direct sound.

Blauert (1970) and Griesinger (1992a, 1992b) studied the motion of continuous sound. They agreed that the sound was perceived to move if the motion was at 3 Hz or below for random movements (can be a bit higher for a single direction movement). Also, when the sound moves faster than 20 Hz, it is perceived as a shift in pitch. When the motion is in this range between 3 and 20 Hz, a stationary sound is perceived in presence of a surround. This phenomenon was called localization lag.

Blauert's work showed that this motion has slightly different properties if it is originated by ITD or by ILD. Motion of ILD is perceived as a sharp source in presence of a surround, while ITD modulation is perceived as an enveloping sound without the possibility of localizing the source. This suggests that the detectors for ITD and ILD would be separated in the auditory system. After these experiments, fluctuations of ITD and ILD are thought to be responsible of spatial impression of continuous sources. This results were confirmed in a later study (Goupell and Hartmann, 2006).

This localization lag can be caused in natural conditions, such as in presence of reflections, even if the reflected energy is lower than the direct sound. Griesinger (1992a, 1992b) introduced the term pseudoangle. This is the angle that would be resolved from the ITD and ILD decoding, but that the auditory system is not fast enough to resolve due to the localization lag. In general, when a broadband stimulus is presented such as noise or music, each critical band has a different pseudoangle fluctuation. This helps to resolve the actual location of the source, and therefore, diminishes the effect of the localization lag.

The sensitivity to ITD improves when the information is consistent across different frequency bands compared to when it is presented to a single band (Buell and Hafter, 1991; Buell and Trahiotis, 1993). In other words, when the information across frequency is inconsistent, it is more difficult to perceive a single angle of the source. Thus, similarly to what happens with ITD modulation, ITD inconsistencies may increase the enveloping perception.

The ITD detection threshold decreases when the overall perceived level increases (Dietz et al., 2013; Smoski and Trahiotis, 1986; Zwislocki and Feldman, 1956). Across studies, log-ITD thresholds decreased linearly with overall level. In the case of ILD, its relation to the overall level still exists, but seems to be weaker than for ITD (Dietz et al., 2013; Rowland Jr and Tobias, 1967; Stellmack et al., 2004).

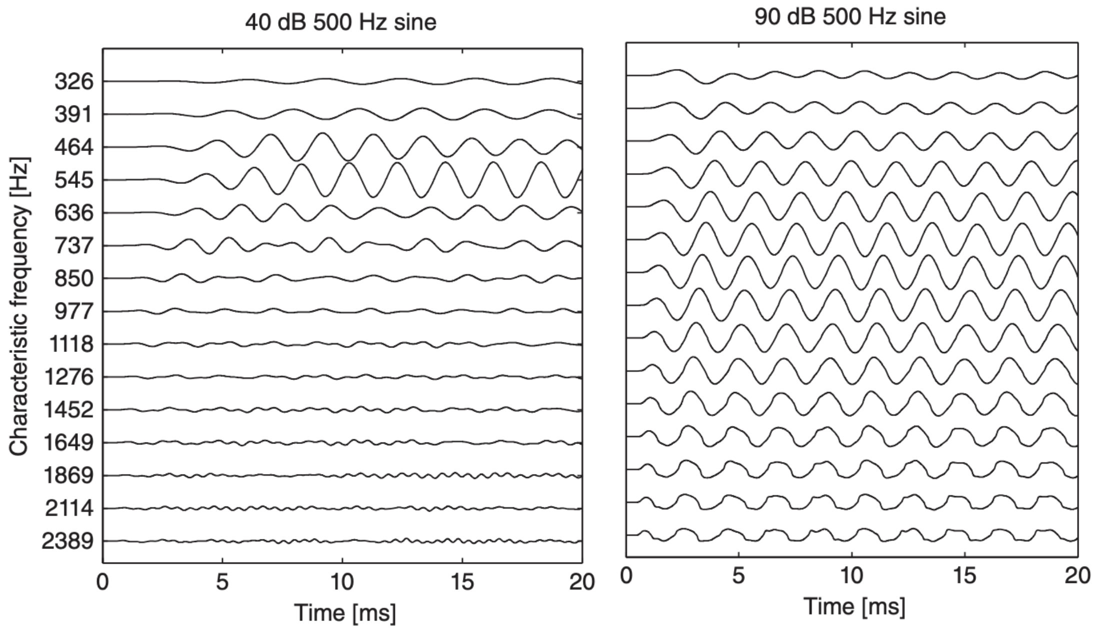


Figure 2: Normalized basilar membrane velocity response to a 500 Hz sinusoidal input at 40 dB SPL (left) and 90 dB SPL (right). The output was computed using the model from Verhulst et al. (2012). Figure 7.12 from Pulkki and Karjalainen (2015).

The mechanics of the cochlea present nonlinearities that depend on frequency, time and level. These properties have beeen characterised in animals in terms of basilar membrane input/output measurements (Rhode and Recio, 2000; Robles et al., 1986; Ruggero et al., 1997; Sellick et al., 1982). These properties can also be measured in humans by measuring the threshold of masked probe tones (Glasberg et al., 1999; Moore et al., 2002; Moore et al., 1997).

A way to visualize how the mechanics of the cochlea work is to use a computational auditory model. In this case, the model from Verhulst et al. (2012) may be good to visualize how the different frequency bands are affected by a given stimulus. This model has explained data from multiple behavioral studies, and therefore is considered to be a good representation of the functioning of the cochlea. In Figure 2, the same stimulus is presented at different levels. Each characteristic frequency in the figure represents the movement of a specific point of the basilar membrane, which presents a tonotopic distribution.

Due to the active function of the cochlea, a single narrowband noise, or a sinusoid, may excite multiple points of the basilar membrane. Thus, multiple frequency bands are excited. Since multiple frequency bands are excited, the ITD and ILD values have to be analyzed carefully across frequency bands. If the ITD and ILD values are coherent, the ITD and ILD sensitivity may increase, decreasing the detection threshold for both of them, specially for ITD. But if these interaural cues are not coherent and/or fluctuate over time, the listener envelopment and the apparent width source may increase.

3 Subjective studies of spaciousness and its relation to the sound level

From the concert hall literature point of view, the relationship between perception and measurements has been an important topic in the field. A link that connected the objective measurements to the acoustic halls preference was always considered very interesting to improve the techniques in concert halls design. In this section, the focus is put on the relationship between loudness and spatial perception.

In the 1970's, studies from Gottlob (1973) and Schroeder et al. (1974) conducted subjective tests where the listeners' preferences were asked. Binaural recordings of classical orchestral music were used, and loudness was held constant during the judgments. The initial-time-delay gap for all the studied halls was considered similar, and therefore, not an attribute that could influence their preferences. They found that the IACC correlated well with the listening preferences. It was considered one of the three orthogonal acoustical measures, together with the reverberation time and the "definition" (energy ratio between the first 50 ms and the total energy).

Ando (2012) and Ando and Raichel (1998) combined subjective test analysis, measurements and auditory models to understand the mapping between the subjective preference of acoustic halls to their physical qualities. Their perceptual tests focused on A-B preference both for binaurally recorded orchestra music and synthesized scenarios where the delay of reflections and the interaural properties were controlled. They transformed these A-B test results into linear scale value of preference that allowed their comparison and analysis. They found that IACC was one of the 4 orthogonal measures of acoustical quality of concert halls, together with RT, G, and ITDG. However, there is some correlation between IACC_E and G, because both increase by increasing the intensity of early reflections. In fact, IACC_{E3} and IACC_A are highly correlated, but IACC_{E3} is a more sensitive measure, with approximately double the spread of values among concert halls.

Potter (1993) investigated the measured acoustic properties that influenced the perceived spaciousness of the sound. His studies were found on the central spectrum theory for binaural hearing, which has a physiological basis. This model analyzes the phase-locking in the auditory neurons to analyze the fine structure of interaural level differences and interaural phase differences. From them, they derived a new metric, the central modulation coefficient (CMC).

To understand the CMC, it is important to introduce the central activity pattern (CAP). The CAP is a three dimensional representation of a binaural acoustic signal where the three dimensions are frequency, delay and the power of the signal. The delay, in this context, represents the relative delay between the left and right ears,

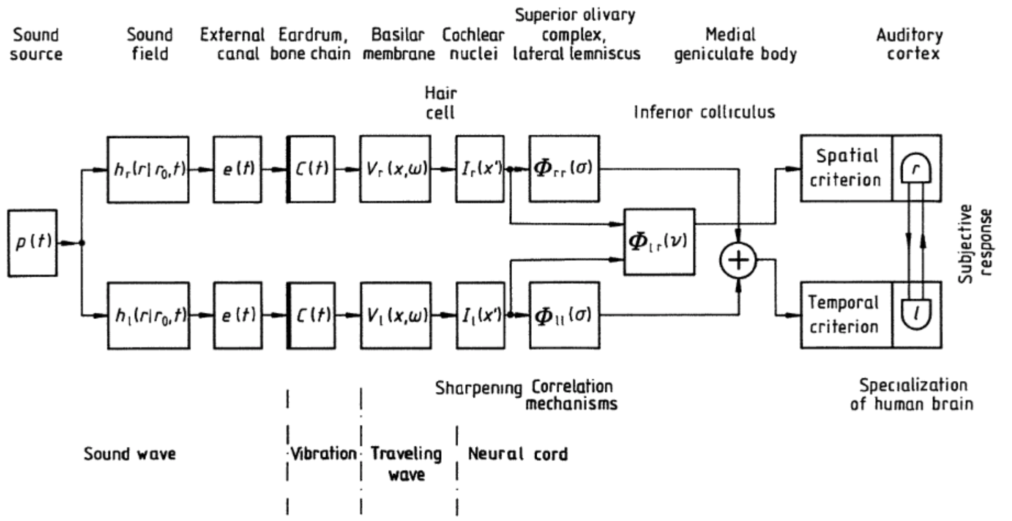


Fig. 5.1. **Model of processors in the auditory pathways for subjective preference judgments: $p(t)$: source signal in the time domain; $h(r|r_0, t)$: impulse response between the source point r_0 and the ear entrance of a listener sitting at r; $e(t)$: impulse response between ear entrance and eardrum; $c(t)$: impulse response of the bone chain between eardrum and oval window including the transformation factor into vibration motion at the eardrum; $V(x, \omega)$: wave form of the basilar membrane, where x is the position along the basilar membrane measured from the oval window: $\Phi_{ll}(\sigma)$ and $\Phi_{rr}(\sigma)$: autocorrelation processors; $\Phi_{lr}(\nu)$: interaural cross-correlation processor; $\oplus$: signals combined. Output signals of a spatial cross-correlation mechanism and temporal autocorrelation mechanisms are assumed to be interpreted by different hemispheres, labeled r and l, respectively**

Figure 3: Auditory model described in Ando (2012). This model was used to analyze the measurements and match them to the results obtained when they were presented to the subjects in the preference perceptual test.

similar to the delay-line network proposed by Jeffress (1948) to extract interaural phase differences. A lateralization function was defined by integrating the CAP over the frequency range of interest. The CMC was defined as the average modulation depth of the lateralization function. This CMC was physiologically motivated, and showed a high degree of correlation to both subjective judgments on ASW and objective measures of IACC_A. This was considered to establish a psycho-physical basis for the use of IACC_A in room acoustics.

Okano et al. (1994) studied the ASW of individual octave bands of symphonic music presented to the subjects on a multiple loudspeakers setup. It was found that the bands at 500 Hz, 1 kHz and 2 kHz were equally important at determining the spaciousness. These lead to the formulation of IACC_{E3} (expected to correlate with ASW) and IACC_{L3} (expected to correlate with LEV).

However, more recent studies found that the IACC relationship to ASW may be misleading. Käsbach et al. (2014) conducted a subjective test where binaural signals were presented to the subjects, which were asked to report the ASW. The ITD fluctuation was controlled by applying a different delay to each frequency band while the IACC was maintained constant. They found out that the ITD fluctuation under 1 kHz was crucial for ASW perception.

3.1 The role of sound level on spaciousness

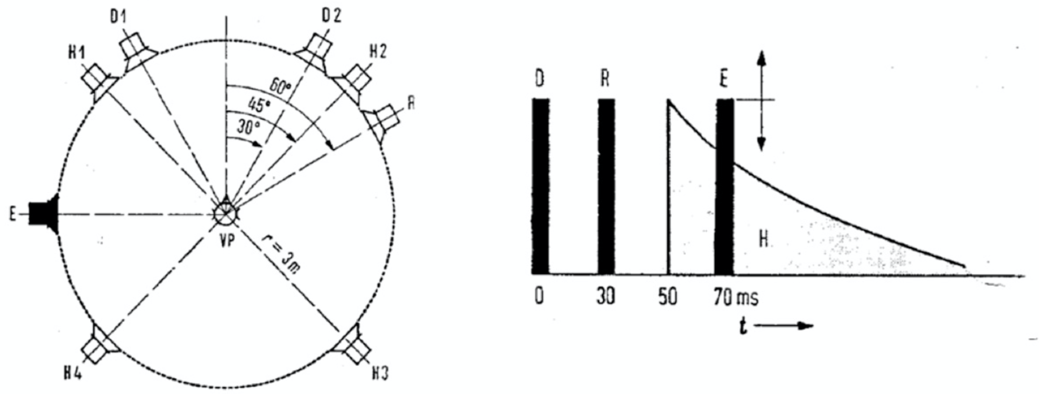
The relationship between the sound level and spatial impression is often studied on how the reflections affect the perceived loudness. In order for $IACC_{E3}$ to be large (and thus, the ASW), a number of the early reflections must arrive at the listener's ears from lateral directions and the ITDG needs to be reasonably short. Not all early reflections need to be lateral, because non-lateral reflection also increase G_{mid} (L. Beranek, 1996). Okano et al. (1994), in the same study mentioned earlier, also found that increasing the G_{low} or decreasing the $IACC_{E3}$ both resulted in a larger ASW.

Lokki, Pätynen, Tervo, et al. (2011) conducted a perceptual test on the influence of early reflections on loudness. They presented a direct sound with simulated reflections between 5 ms and 120 ms. They found that listeners often prefer concert halls with enough early reflections because they contribute to increase the binaural loudness. But it is also interesting to examine the opposite direction of influence, understanding how the level of the sound also affects the spatial impression.

There is a phenomenon related to spatial impression that occurs when the hall "wakes up". Louder sounds generate a sense of envelopment compared to more quiet sounds. This effect was already perceived in early studies (Barron, 1971; de Keet, 1968; A. H. Marshall, 1966). Kuhl (1978) mentioned that music dynamics not only influences the perceived loudness, but also the spatial impression. However, this spatial impression enhancement, is considered to be a pleasant but an unconscious effect. The number of reflections that can be perceived increase when the sound gets louder (Green & Kahle, 2019; Pätynen et al., 2014).

It is very interesting to check the findings from Wettschureck (1976) in their PhD thesis, which was nicely summarized by Green and Kahle (2019). They studied the perception threshold of reflections depending on the presentation level from the masking point of view. In an anechoic space, they presented a synthesized to direct sound followed by two discrete reflections and a 2 s reverberation starting after 50 ms. The first lateral reflection was always 30 ms after the direct sound, fixed at 50° azimuth. The second reflection was always 70 ms after the direct sound, but varying its position and the presentation level. The direct sound, first lateral reflection and reverberation were adjusted to achieve a direct-to-reverberant ratio of 0 dB. Anechoic speech was used and the level was varied between 20 dB and 80 dB in 5 dB steps to find the perception threshold of the reflection (see Figure 4).

They found that at quiet listening levels, the direct sound dominates scene, being clearly localizable but the effect of the room would be very weak. The reflections from the front and side would be then noticeable when the level is increased. When



D = direct sound as phantom source via loudspeakers D1, D2

R = fixed lateral early reflection, loudspeaker R

H = 2 second reverberation (500Hz value) starting at 50ms,
reproduced via 4 loudspeakers H1 – H4

E = 70ms reflection with adjustable level and location, loudspeaker E

Figure 4: Diagram representation of the perception threshold of reflections depending on the presentation level. Figure from Green and Kahle (2019), after Wettschureck (1976).

the sound is even louder, the rear reflections are perceived, since the threshold for the back is higher than the rest. This front and lateral reflections were expected to increase the ASW, which was later measured by de Keet (1968). The threshold to notice the reflections lowers when the overall listening level increases, which is in accordance with studies from A. H. Marshall (1966) and Kuhl (1978). Even though no formal tests were conducted with other stimuli, it was thought that results with other type of sounds would be similar.

Green and Kahle (2019) duplicated the reflection thresholds using binaural signals. They investigated the thresholds for speech and music, at different delay times and directions of arrival. The results for speech were reasonably similar to the ones measured by Wettschureck (1976). The rear reflections threshold was similar to the front one, probably caused by the binauralization issues regarding front-back confusions. For the music thresholds experiment, a solo cello anechoic recording was used, due to its even sound level in contrast to speech. Even though the delay of the reflections had an effect on the thresholds, the results showed similar trends, varying the thresholds depending on the presentation level and the direction of arrival of the reflections (see Figure 5).

4 Discussion and future work

In the concert hall acoustics literature, there is a generalized trend of relating the spatial impression to the IACC. Most studies agree on the correlation between the

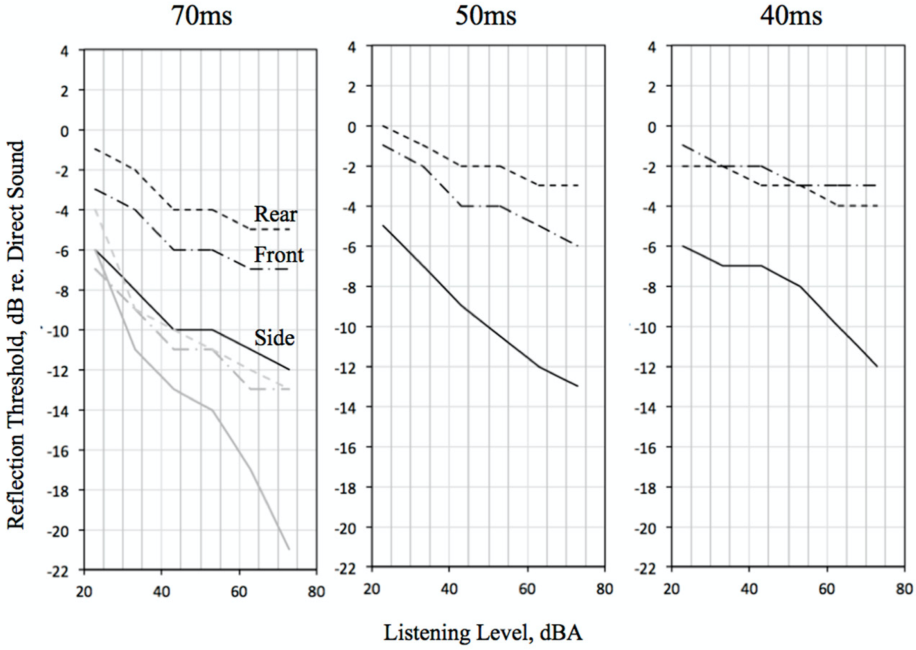


Figure 5: Reflection thresholds for music (black) and speech (gray) depending on the level, delay and direction. Figure from Green and Kahle (2019).

$IACC_E$ and the ASW, as well as the $IACC_L$ and the LEV. But in more recent studies, this relation doesn't seem to be very clear.

From the psychoacoustics community point point of view, the IACC seems to be an intermediate measure that one can compute the interaural cues from. This is probably due to that IACC is not clearly related to perception *per se*. However, it is assumed that the superior olivary complex is able to detect separately ITD and ILD values. For this reason, it would be interesting to go a bit further and relate the ASW and LEV to these binaural features directly.

After reviewing the perceptual effects caused by moving individual sources, it seems possible that the ASW and LEV could be measured in terms of ITD and ILD modulation. Some studies related the binaural cues consistency over frequency and over time to the spatial perception (Blauert, 1970; Buell and Hafter, 1991; Buell and Trahiotis, 1993; Griesinger, 1992b). To confirm this hypothesis, Goupell and Hartmann (2006) demonstrated that the IACC was not well detected if it wasn't in combination of relatively large ITD and ILD variations. However, the IACC seems to correlate well with binaural perception if the signal is broadband. This should be taken into account in the future when quantifying spaciousness.

In the concert hall acoustics literature, the G and the IACC are often thought to be orthogonal, and that combined, they convey a certain spatial impression. On the other hand, there are reasons to think that IACC is not orthogonal to perceived loudness, since both are modified by the number of early reflections, and this was concluded in some of the aforementioned studies (Lokki, Pätynen, Kuusinen, et al., 2011; Okano et al., 1994).

The studies about the relationship between sound level and the binaural cues sensitivity may provide some fundamentals to understand this relationship. By analyzing the basics of the active function of the cochlea, it seems a possible explanation to understand why the increase in presented level also increases the listeners' sensitivity to detect reflections.

Even without these reflections, loud sounds excite more areas of the basilar membrane than quiet sounds. This may cause a greater number of ITD and ILD values over frequency, which may end up generating conflicting cues over frequency and time. These inconsistencies could explain, at least partially, why louder sounds are perceived as being more enveloping, or presenting a wider perceived source, due the ITD and ILD modulation over time theory presented in 2.

Further studies are needed to confirm the relation of sound level to ASW and LEV by means of analyzing interaural cues fluctuation. It would be also interesting to analyze if these spaciousness attributes are remotely related to sound externalization, since interaural cues fluctuation has been also related to the perceived distance to a sound source Catic et al., 2013. This may provide an interesting line of research to understand the relationship between loudness and spatial impression.

References

Ando, Y. (2012). *Concert hall acoustics* (Vol. 17). Springer Science & Business Media.

Ando, Y., & Raichel, D. R. (1998). Architectural acoustics: Blending sound sources, sound fields, and listeners.

Barron, M. (1986). Intimacy, loudness and sound level in concert halls. *Proc. 12th ICA, Toronto, Paper E4-3*, *17*.

Barron, M. (1971). The subjective effects of first reflections in concert halls—the need for lateral reflections. *Journal of sound and vibration*, *15*(4), 475–494.

Barron, M. (1988). Subjective study of british symphony concert halls. *Acta Acustica united with Acustica*, *66*(1), 1–14.

Barron, M., & Marshall, A. H. (1981). Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure. *Journal of sound and Vibration*, *77*(2), 211–232.

Beranek, L. (1996). Concert and opera halls. *Acoust. Soc. Amer.*

Beranek, L. (2011). The sound strength parameter g and its importance in evaluating and planning the acoustics of halls for music. *The Journal of the Acoustical Society of America*, *129*(5), 3020–3026.

Beranek, L. L. (2004). Comments on "intimacy"and itdg concepts in musical performing spaces. *The Journal of the Acoustical Society of America*, *115*(5), 2403–2403.

Blauert, J. (1970). Zur tragheit des richtungshorens bei laufzeit-und intensitatsstereophonie. *Acustica*, *23*, p287–293.

Buell, T. N., & Hafter, E. R. (1991). Combination of binaural information across frequency bands. *The Journal of the Acoustical Society of America*, *90*(4), 1894–1900.

Buell, T. N., & Trahiotis, C. (1993). Interaural temporal discrimination using two sinusoidally amplitude-modulated, high-frequency tones: Conditions of summation and interference. *The Journal of the Acoustical Society of America*, *93*(1), 480–487.

Catic, J., Santurette, S., Buchholz, J. M., Gran, F., & Dau, T. (2013). The effect of interaural-level-difference fluctuations on the externalization of sound. *The Journal of the Acoustical Society of America*, *134*(2), 1232–1241.

de Keet, V. (1968). The influence of early lateral reflections on the spatial impression. *Proc. 6th Int. Cong. Acoust., Tokyo*, *2*.

Dietz, M., Bernstein, L. R., Trahiotis, C., Ewert, S. D., & Hohmann, V. (2013). The effect of overall level on sensitivity to interaural differences of time and level at high frequencies. *The Journal of the Acoustical Society of America*, *134*(1), 494–502.

D'Orazio, D., & Garai, M. (2017). The autocorrelation-based analysis as a tool of sound perception in a reverberant field. *Rivista di estetica*, (66), 133–147.

Glasberg, B. R., Moore, B. C., & Stone, M. A. (1999). Modelling changes in frequency selectivity with level. *Psychophysics, physiology and models of hearing* (pp. 143–154). World Scientific.

Gottlob, D. (1973). *Comparison of objective acoustic parameters in concert halls with results of subjective experiments*. Dissertation, Goettingen University.

Goupell, M. J., & Hartmann, W. M. (2006). Interaural fluctuations and the detection of interaural incoherence: Bandwidth effects. *The Journal of the Acoustical Society of America*, *119*(6), 3971–3986.

Green, E., & Kahle, E. (2019). Dynamic spatial responsiveness in concert halls. *Acoustics*, *1*(3), 549–560.

Griesinger, D. (1992a). Ialf-binaural measures of spatial impression and running reverberance. *Audio Engineering Society Convention 92*.

Griesinger, D. (1992b). Measures of spatial impression and reverberance based on the physiology of human hearing. *Audio Engineering Society Conference: 11th International Conference: Test & Measurement*.

Griesinger, D. (1997). The psychoacoustics of apparent source width, spaciousness and envelopment in performance spaces. *Acta Acustica united with Acustica*, *83*(4), 721–731.

Griesinger, D. (1999). Objective measures of spaciousness and envelopment. *Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction*.

Grothe, B. (2003). New roles for synaptic inhibition in sound localization. *Nature Reviews Neuroscience*, *4*(7), 540–550.

Hidaka, T., Beranek, L. L., & Okano, T. (1995). Interaural cross-correlation, lateral fraction, and low-and high-frequency sound levels as measures of acoustical

quality in concert halls. *The Journal of the Acoustical Society of America*, *98*(2), 988–1007.
Hyde, J. R. (2002). Acoustical intimacy in concert halls: Does visual input affect the aural experience. *Proceedings of the Institute of Acoustics*, *24*.
Hyde, J. R. (2019). Discussion of the relation between initial time delay gap (itdg) and acoustical intimacy: Leo beranek's final thoughts on the subject, documented. *Acoustics*, *1*(3), 561–569.
Jeffress, L. A. (1948). A place theory of sound localization. *Journal of comparative and physiological psychology*, *41*(1), 35.
Käsbach, J., May, T., Oskarsdottir, G., Jeong, C.-H., & Chang, J. (2014). The effect of interaural-time-difference fluctuations on apparent source width [7th Forum Acusticum ; Conference date: 07-09-2014 Through 12-09-2014]. *Proceedings of Forum Acusticum*.
Kuhl, W. (1978). Räumlichkeit als komponente des raumeindrucks. *Acta Acustica united with Acustica*, *40*(3), 167–181.
Lokki, T., & Pätynen, J. (2020). Auditory spatial impression in concert halls. *The technology of binaural understanding* (pp. 173–202). Springer.
Lokki, T., Pätynen, J., Kuusinen, A., Vertanen, H., & Tervo, S. (2011). Concert hall acoustics assessment with individually elicited attributes. *The Journal of the Acoustical Society of America*, *130*(2), 835–849.
Lokki, T., Pätynen, J., Tervo, S., Siltanen, S., & Savioja, L. (2011). Engaging concert hall acoustics is made up of temporal envelope preserving reflections. *The Journal of the Acoustical Society of America*, *129*(6), EL223–EL228.
Marshall, A., & Hyde, J. (1979). Some preliminary acoustical considerations in the design for the proposed wellington (new zealand) town hall. *Journal of Sound and Vibration*, *63*(2), 201–211.
Marshall, A. H. (1966). A note on the importance of room cross-section in concert halls. *Journal of Sound and Vibration*, *5*(1), 100–112.
Moore, B. C., Alcántara, J. I., & Glasberg, B. R. (2002). Behavioural measurement of level-dependent shifts in the vibration pattern on the basilar membrane. *Hearing research*, *163*(1-2), 101–110.
Moore, B. C., Glasberg, B. R., & Baer, T. (1997). A Model for the Prediction of Thresholds, Loudness, and Partial Loudness. *J. Audio Eng. Soc.*, *45*(4), 224–239.
Okano, T., Beranek, L. L., & Hidaka, T. (1998). Relations among interaural cross-correlation coefficient (iacc e), lateral fraction (lf e), and apparent source width (asw) in concert halls. *The Journal of the Acoustical Society of America*, *104*(1), 255–265.
Okano, T., Hidaka, T., & Beranek, L. L. (1994). Relations between the apparent source width (asw) of the sound field in a concert hall and its sound pressure level at low frequencies (gl), and its inter-aural cross correlation coefficient (iacc). *The Journal of the Acoustical Society of America*, *96*(5), 3268–3268.
Pätynen, J., Tervo, S., Robinson, P. W., & Lokki, T. (2014). Concert halls with strong lateral reflections enhance musical dynamics. *Proceedings of the National Academy of Sciences*, *111*(12), 4409–4414.

Potter, J. M. (1993). On the binaural modelling of spaciousness in room acoustics.

Pulkki, V., & Karjalainen, M. (2015). *Communication acoustics: An introduction to speech, audio and psychoacoustics*. John Wiley & Sons.

Rhode, W. S., & Recio, A. (2000). Study of mechanical motions in the basal region of the chinchilla cochlea. *The Journal of the Acoustical Society of America*, *107*(6), 3317–3332.

Robles, L., Ruggero, M. A., & Rich, N. C. (1986). Basilar membrane mechanics at the base of the chinchilla cochlea. i. input–output functions, tuning curves, and response phases. *The Journal of the Acoustical Society of America*, *80*(5), 1364–1374.

Rowland Jr, R. C., & Tobias, J. V. (1967). Interaural intensity difference limen. *Journal of Speech and Hearing Research*, *10*(4), 745–756.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., & Robles, L. (1997). Basilar-membrane responses to tones at the base of the chinchilla cochlea. *The Journal of the Acoustical Society of America*, *101*(4), 2151–2163.

Schroeder, M. R., Gottlob, D., & Siebrasse, K. (1974). Comparative study of european concert halls: Correlation of subjective preference with geometric and acoustic parameters. *The Journal of the Acoustical Society of America*, *56*(4), 1195–1201.

Sellick, P., Patuzzi, R., & Johnstone, B. (1982). Measurement of basilar membrane motion in the guinea pig using the mössbauer technique. *The journal of the acoustical society of America*, *72*(1), 131–141.

Smoski, W. J., & Trahiotis, C. (1986). Discrimination of interaural temporal disparities by normal-hearing listeners and listeners with high-frequency sensorineural hearing loss. *The Journal of the Acoustical Society of America*, *79*(5), 1541–1547.

Stellmack, M. A., Viemeister, N. F., & Byrne, A. J. (2004). Monaural and interaural intensity discrimination: Level effects and the "binaural advantage". *The Journal of the Acoustical Society of America*, *116*(2), 1149–1159.

Tollin, D. J. (2003). The lateral superior olive: A functional role in sound source localization. *The neuroscientist*, *9*(2), 127–143.

Verhulst, S., Dau, T., & Shera, C. A. (2012). Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission. *The Journal of the Acoustical Society of America*, *132*(6), 3842–3848.

Wettschureck, R. (1976). *Über die abhängigkeit raumakustischer wahrnehmungen von der lautstärke* (Doctoral dissertation). PhD Thesis, Technical University of Berlin.

Zwislocki, J., & Feldman, R. (1956). Just noticeable differences in dichotic phase. *The Journal of the Acoustical Society of America*, *28*(5), 860–864.