Shifts in the judgement of distance to a sound source behind a sonic crystal slab

Ignacio Spiousas, Pablo E. Etchemendy, Esteban R. Calcagno, and Manuel C. Eguia

Laboratorio de Acústica y Percepción Sonora,
Universidad Nacional de Quilmes
R. S. Peña 352 Bernal,
B1876BXD Buenos aires,
Argentina.

(Dated:)

Abstract

The ability of subjects to estimate the distance to a sound source in a room relies on the integration of a number of different cues: sound intensity level, direct-to-reverberant energy ratio, spectral content and binaural cues, among others. This work examines how the perception of auditory distance is modified for a particular sound field: the transmitted field of an acoustic source through a sonic crystal slab in a semi-reverberant room. A series of experiments were performed comparing the egocentric distance to a sound source passing and not passing through the sonic crystal, using an acoustical virtual environment whereby some of the auditory distance cues could be manipulated. The results obtained show that the presence of the sonic crystal introduces significant shifts on the auditory distance perception. These shifts are correlated with the spatial and spectral variation of the acoustical properties of the sonic crystal. Also, it was possible to determine the relative influence of the manipulated cues (intensity, binaural and reverberation).

PACS numbers:

I. Introduction

The human auditory system is adapted for the efficient localization of sound sources, even in the presence of reflections and diffraction of the original sound wave[1]. When a broadband acoustical source is turned on inside a room a listener is able estimate the egocentric angular localization and distance of the source, even when the energy coming directly from it is minimal compared to the energy coming from the reflections on the walls. Humans learnt to localize sound sources in these 'natural' acoustical environments formed by walls, flat surfaces and rigid obstacles[2]. In this work we are interested on how the perception of distance of an acoustical source can be modified in a 'non-natural' acoustical environment generated by a two dimensional sonic crystal (SC).

A two-dimensional SC is an acoustic metamaterial, consisting of an array of rigid cylinders in air that displays a large variation of its acoustical properties by changing its geometrical configuration[3]. SCs have been extensively studied due to their singular transmission and reflection properties. These properties vary strongly with frequency and extend from acoustic band gaps[4] to negative refraction[5], negative bi-refraction[6] and sound focusing[7]. Frequency filtering effect has been noticed when a SC slab is positioned in front of and around a sound source, giving sharp resonances and strongly frequency-dependent impulse responses[8]. Nonetheless, the effect of the changes in the sound field due to a SC slab on the perception of the spatial localization of a sound source (both angle and distance) has not been studied yet.

Several perceptual cues are known to be important for auditory distance perception (ADP) of sound sources. Intensity is one of the main and most studied cues, corresponding smaller intensities to farther sound sources. The intensity of a sound source is a relative cue of ADP because an intensity reference is needed to give an absolute evaluation of the perceived distance. Another important cue for ADP is the direct-to-reverberant energy ratio (DR). The DR is defined as the energy ratio between the intensities of the direct and reverberant field.

In reverberant environments, when moving away from sound sources the intensity of the direct sound decreases, while the energy in the reverberant field remains relatively constant, giving higher values of DR for closer sound sources. Other cues involved in ADP are: spectral cues, familiarity of the source and binaural and dynamical cues (see[9] for a recent review on ADP).

This works aims to quantify the effect of the placement of a SC slab between a sound source and a human listener on the ADP. For this purpose, first a series of binaural recordings are performed in a semi-reverberant room with and without the SC slab between the sound source and the listener. Three objective measures (that are known to be ADP cues) are derived from the recordings: intensity, direct-to-reverberant ratio and interaural cross-correlation. Next, a psychophysical experiment is performed using the personalized binaural recordings, altering the balance of the three cues. This procedure is described in the next section. Then, we present the results, aiming the attention to the cases where a shift of the ADP is induced by the presence of the SC slab. The relative influence of the different cues on the ADP and a correlation with the acoustical measurements is presented in the Discussion section. Finally, the last section concludes.

II. Methods

A. Experimental Set-up

Binaural recordings and psychophysical experiments were conducted in a rectangular room, approximately 12 × 7 × 3 m (length × width × height) with walls covered by sound-absorbing panels (pyramid polyurethane acoustical foam, 50 mm), the floor by carpet, and the ceiling by fibre glass acoustic roof panels. The average reverberation time of the room T30 (A-weighed, measured using an exponential sweep) was 0.45 s and the background noise level was 19 dBA (measured with a RION NL-32 class I level meter). The choice of a room with a small but noticeable reverberation was made because the direct-to-reverberant will be used as a cue for ADP during the experiments.

The two-dimensional SC slab used for the recordings was built with 59 PVC cylinders of diameter 16 cm arranged in a triangular lattice configuration, with a lattice constant of 22 cm. This SC was designed to have noticeable focusing effects for a frequency range close to the maximum sensitivity of the human ear. The first partial band gap (for normal incidence) extends approximately from 0.57 to 1 kHz, while the first negative-refraction focusing band is in the range between 1 and 1.5 kHz.

B. Sample Recordings

The sound samples used for the recordings consisted of thirteen one-third-octave noise bands with central frequency between 0.5 to 2 Khz and a one-sixth-octave step. Samples were prepared digitally using Matlab (Mathworks Inc.) at 96 kHz sampling frequency and 16-bit sample depth and reproduced by a two-way loudspeaker Genelec 8030 unit connected to a MOTU 896mk3 digital audio interface. The duration of the sample was 2 seconds with onset and offset rounded by a Hann window of 50 ms. Binaural recordings were made using a custom designed binaural dummy head equipped with SP-TFB-2 inter aural microphones connected to a Tascam DR40 hand-held recorder, using 96 kHz sampling frequency and 16-bit sample depth.

The recordings were made using a fixed source and setting dummy-head at perpendicular distances to the slab from 0 to 4 m with steps of 0.5 m (crosses on 1) giving a total of 117 recorded samples (13 frequency bands x 9 positions). All recordings were performed with and without the SC slab. A binaural impulse response was also obtained for all the recording points. A schematic representation of the sonic crystal slab and the source and receiver positions is shown on fig. 1. The elevation of the midpoint between ears of the dummy-head and the loudspeaker were both of 1.2 m.

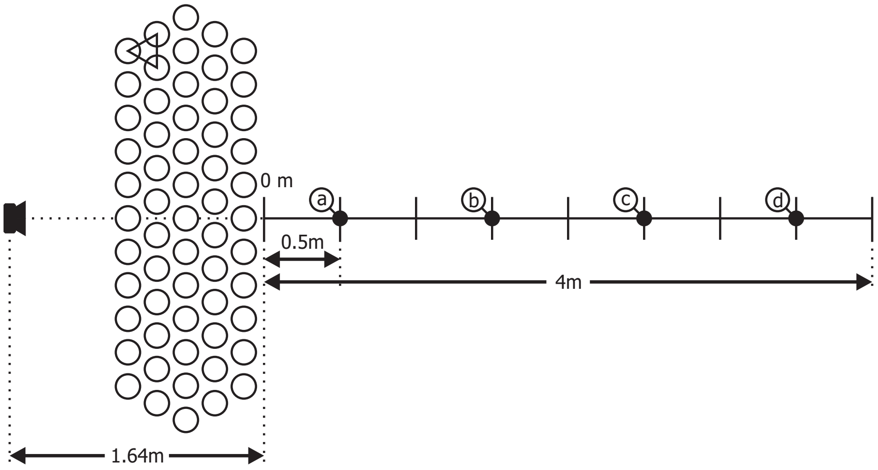

FIG. 1. - Schematic representation of the experimental configuration used for the binaural recordings. The sound source position is shown by a loudspeaker-like icon on the left of the figure. For the acoustical recordings, the dummy head was placed on each cross of the central line. The black dots show the position of the human subjects during the binaural recordings used for obtaining the psychoacoustic experiment stimuli.

C. Acoustical parameter measuremts

The most significant cue for auditory distance perception is the perceived loudness of a sound source. The Intensity is a physical magnitude strongly related to the perceived loudness that will help us outline the effect of the SC slab on this percept. We calculate a binaural intensity (BI) from the recordings adding linearly the intensity of the signal reaching both ears as follows:

$$I_{l,r} = \frac{1}{L}\int_0^L p_{l,r}^2(t)dt \tag{1}$$

$$BI = 10log_{10}[(I_l I_r)^{1/2}/I_{ref}] \tag{2}$$

where $I_{l,r}$ stands respectively for the left and right ear individual intensity, $p_{l,r}(t)$ for the A-weighted pressure field on each ear and L for the length of the recorded signal, and $I_{ref} = 10^{-12}\,\mathrm{W/m^2}$ is the reference intensity.

Another significant cue for auditory distance perception is the DR. We calculated the DR for every frequency band by filtering the impulse response using the same one-third-octave filters used for the generation of the samples. In order to separate the direct from the reverberant field we used a step window with a $2ms$ rised cosine slope centered at $t_0 = 7ms$. The 'direct' energy included also the first reflection on the floor but no lateral reflections. We used the following formula to calculate the DR for each frequency band:

$$DR_n = 10\log_{10}\left(\frac{\int h_n(t)w(t)dt}{\int h_n(t)(1-w(t))dt}\right) \tag{3}$$

where $h_n(t)$ and DR_n are the nth band filtered impulse response and direct to reverberant ratio respectively.

The DR is considered a monaural cue, but reverberation also introduces binaural cues by altering the correlation of the signal between ears[10]. The most studied binaural cue is the interaural cross-correlation (IACC). While the direct sound arrive with a strongly correlated signal (high IACC value) the reverberant field is characterized by a low correlation of the signal between ears. The IACC is the maximum value of the interaural cross correlation function $IACF(\tau)$ as a function of the time difference between ears (τ)[11]:

$$\phi_{lr}(\tau) = \frac{1}{2L}\int_{-L}^{+L} p_l'(t)p_r'(t+\tau)dt \tag{4}$$

$$IACF(\tau) = \phi_{lr}(\tau)/[I_r I_l]^{1/2} \tag{5}$$

Since we are measuring along the symmetry axis the maximum always occurs for τ values very close to zero.

In Fig. 2 we show the magnitudes of BI, DR and IACC as columns, and the conditions with-SC and without-SC as rows. Each plot displays these magnitudes as a function of the central frequency of the noise bands and the distance along the central axis. The most noticeable difference between the two conditions is an increase of these magnitudes within the negative-refraction focalization region (1-1.5 kHz) with the SC. Also, note that the position of the maximum is shifted along the axis as the frequency is changed from approximately 0.5

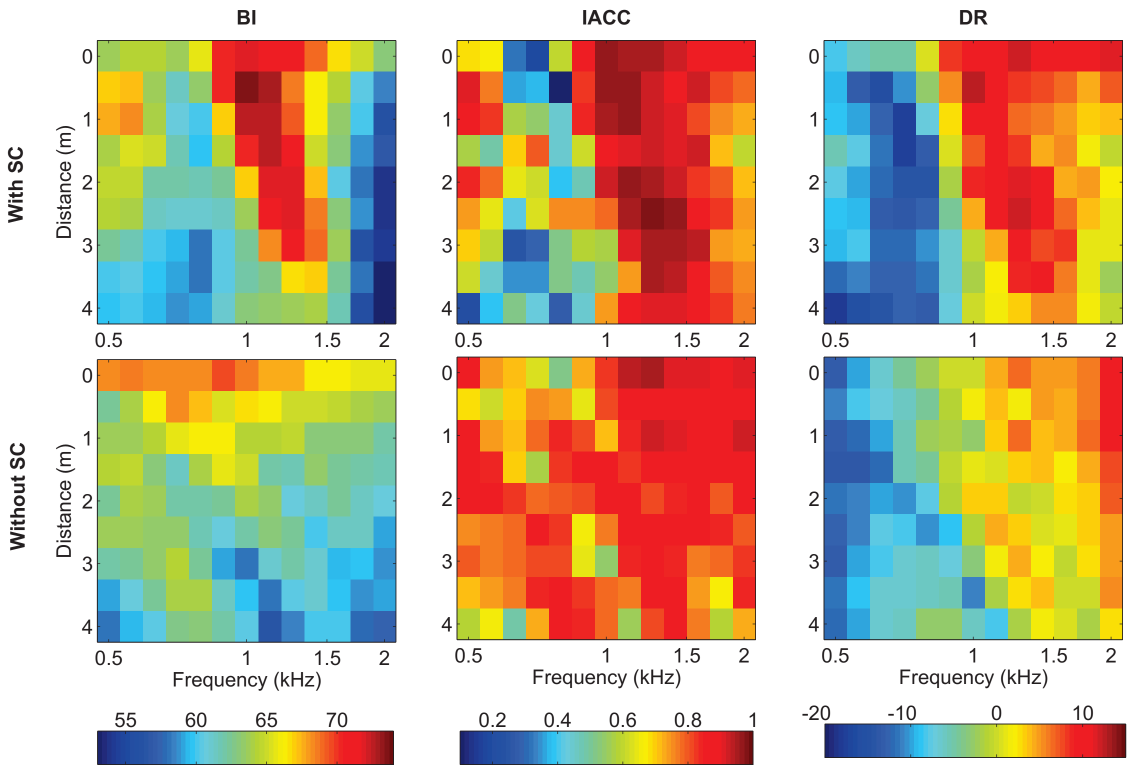

FIG. 2. Binaural intensity BI (left), interaural cross-correlation IACC (center) and direct to reverberant ratio DR (right) for one-third-octave noise bands. The top plots corresponds to the measurements made with the sonic crystal slab between the source and the receiver (condition with-SC) and the bottom plots to the room response (condition without SC). The results were obtained along the symmetry axis with a perpendicular distance to the slab of 0 to 4 m with steps of 0.5 m. The results for the BI and DR are expressed in dB, while the IACC is plotted in non-dimensional units.

m for 1 kHz to 4 m for 1.5 kHz. The focalization is clearly apparent for the BI magnitude, but a relative increase is also observed for the IACC and DR. The region of focalization is roughly the same for the three mangitudes.

As it would be expected, within the region of the partial band-gap the BI is reduced for the condition with SC. Also there is a small positive-refraction focalization around 0.5 kHz and a distance of 1 m.

D. Subjects

Four male subjects participated in the experiment (ages ranging 30-42 years). All listeners reported normal hearing and have prior experience in psychoacoustic experiments. Subjects participated in four sessions that lasted about one hour. The first two sessions were used for the recordings (with and without the SC). The following two sessions were used for the psychoacoustical experiments.

E. Stimuli

Sound samples for the stimuli were recorded using the same experimental configuration detailed above. Listeners were seated at distances: (a) 0.5, (b) 1.5, (c) 2.5, and (d) 3.5 meters (black circles in Fig.1), using the midpoint between ears as a reference , Their heads were fixed using a chin-rest support bar, and a narrow acetate sheet at the eye level for fixing gaze as a self-control. Four personalized sets of 52 recordings (13 noise bands x 4 listener positions) were obtained using the same equipment as in the dummy-head measurements, with and without the SC.

Three types of modifications were applied to the sample recordings: (a) binaural loudness normalization, the binaural loudness was calculated for all recordings (following[12]) and then the amplitudes of the recordings were modified in order to obtain the same binaural loudness for all stimuli; (b) rectangular window, all samples were convolved with a rectangular window of the same duration of the sound samples in order to eliminate the reverberation tail of the stimuli; (c) phase randomization, using STFT re-synthesis the phase information was wiped out from the sample recordings. Each of these modifications eliminates or significantly reduces the three cues for ADP that we are interested in. Modification (a) eliminates the BI cue, the modification (b) significantly reduces the DR information even when some information remains during the stimulus, and modification (c) certainly eliminates the IACC cue.

Five sets of stimuli were then obtained combining these modifications: (1) the original recordings where the three cues are present; (2) IACC-only stimuli, after modifications (a) and (b) ; (3) BI-only stimuli after modifications (b) and (c); (4) DR-only stimuli, after modifications (a) and (c); and (5) control stimuli, after modifications (a), (b) and (c). For this last type of stimuli there are still spectral differences between the two conditions due to the filtering of the SC slab.

Stimuli were modified and played using Matlab, converted through an external sound board (Focusrite Saphire LE) and presented using Sennheiser HD600 headphones, in the same environment used for the recordings.

F. Procedure

The experiment was divided into five blocks, one for each stimuli set. The first block was always the corresponding to stimulus set (1). The other blocks were performed in random order, distributed in two sessions. Each block consisted in the repeated presentation of a pair of stimuli corresponding to the same noise frequency band and position, with and without the SC, in random order (trial). The first 20 trials were considered as a training set and were discarded later. The task at each trial was to judge which stimulus was perceived *farther*. The subjects were allowed to repeat the stimuli pair, but the choice was forced. No feedback was provided. Each pair was repeated four times, giving a total of 208 test trials per block.

III. Results

Since we are interested in the shifts of the ADP, for the statistical analysis of the data we define a 'success' as a choice of the condition without-SC as the stimulus perceived farther by the subjects. This means that 'success' corresponds to the case when the sonic crystal make the source appear closer than it actually is perceived without the SC.

Probability of success p was calculated as a function of frequency and position for each subject and for the results of the four subjects combined, using a binomial model.

In Fig. 3 we display the obtained p values for the pooled data of the four subjects and for the five type of stimuli. For the sake of clarity we classify the possible results into three cases: (a) the upper bound of the confidence interval of p is less than 0.5, this would correspond to the case when the source is consistently perceived closer without the SC (white color in Fig. 3); (b) the confidence interval of p includes the 0.5 value, this would correspond to 'ambiguous' cases where the source was neither consistently perceived closer or farther with the SC (gray color in Fig. 5); (c) the lower bound of the confidence interval of p is greater than 0.5, this would correspond to the 'success' case when the source is consistently perceived closer with the SC (red color in Fig. 5).

A replicated G-test of independence was performed for each frequency and position and for the five types of stimuli, in order to see if the results from the four subjects could be combined into a single data set. The heterogeneity G-value was not significant for most of the data (82 percent). The 18 percent of the data for which the null hypothesis (that all the responses come from the same distribution) cannot be accepted, corresponds predominantly to the 'ambiguous' cases (16 percent of total) where the confidence interval for p included the 0.5 value. Only two cases (both for position d and central frequency 1.2 kHz and for stimuli type 1 and 4) occur for 'success' cases where $p > 0.5$.

IV. Discussion

Form Fig. 3 it is apparent that the 'success' cases (red areas) where significant shifts of the ADP were observed occur within the same region in frequency and position. The results for the stimuli set 1 (all cues) display the larger region, including eight cases. For stimuli types 2-4 three, seven and four cases are observed respectively, all within the same region. Comparing with the first row of Fig 2. it is clear that this region correspond to the focalization of the SC. Seven of the cases correspond to the negative-refraction focalization, and one case (position a, central frequency 0.5 kHz) to the positive-refraction focalization. There are two cases, ocuppying the centre of the focalization region, for which ADP shifts

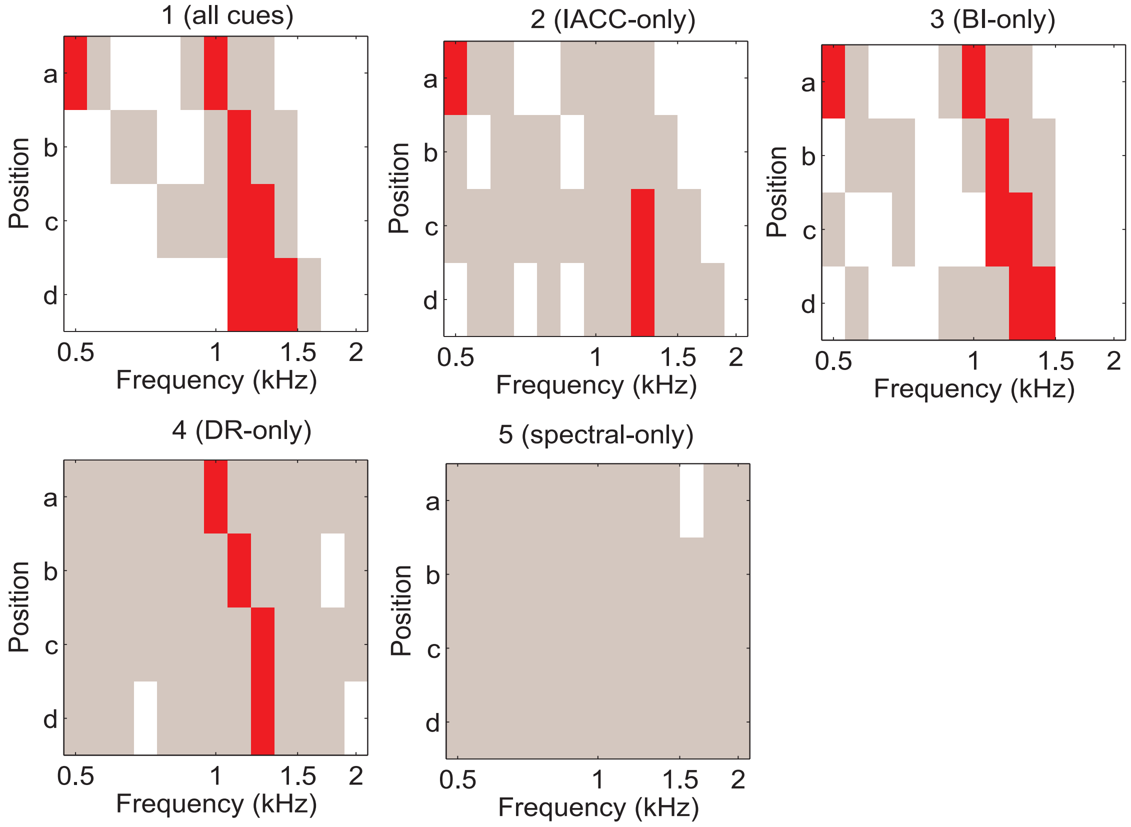

FIG. 3. Probability of 'success' (perceiving the source with-SC condition closer) for each stimuli set (1-5) as a function of frequency and position. The results were classified into three categories according to their confidence limits ($\alpha = 0.05$). Red color corresponds a probability significantly over 0.5. White color corresponds to a probability significantly lower than 0.5. Gray color corresponds to probabilities not significantly different from 0.5 (chance)

were significant for the stimuli types 1-4: frequency = 1.2 kHz and positions (c) and (d). No significant shifts were observed for stimuli type 5.

The stimuli type 1 included all the cues (BI, IACC and DR) and, as it was reasonable to expect, their results have the lower number of 'ambiguous' cases. The second result with less ambiguous cases is that of stimuli type 3 (BI cue). This indicates that the BI cue is the more relevant for the shifts in ADP. Also it replicates the region of 'success' cases of the results for stimuli type 1 in all but one case (position d, frequency 1.12 kHz).

The third result ordered from lower to higher number of 'ambiguous' cases is that corresponding to stimuli type 2 (IACC cue). Even when there are only three cases of significant ADP shifts, the two regions of focalization and the band-gap are still distinguishable. This implies that when only phase information is present and all the stimuli have the same binaural loudness there is still a significant ADP shift. This is a remarkable result since it indicates that the focalization is not an amplitude only phenomenon. It is likely that the increasing of the phase coherence between the two ears due to the convergence of the wavefronts created by the negative-refraction also makes the source appear closer than it really is.

For the results of corresponding to stimuli type 4 (DR cue) there are four cases of significant ADP shifts, all within the negative-refraction region. But almost all other positions and frequencies correspond to 'ambiguous' cases. This means that when there are neither phase nor intensity information, with the exception of the strong focalization region, it is very difficult to judge which of the stimuli (with and without the SC) is farther.

Finally, the only difference that remains between the stimuli type 5 with and without SC is spectral. Even when changes in the spectral content could also be associated to changes in the ADP, these changes are not sufficient to elicit a shift in ADP for the narrowband stimuli used here. This experiment was used as a control for the previous ones, and justify a posteriori the choice of BI, IACC and DR as the only cues affecting the ADP with the SC.

A correlation study was performed between the pooled data of the four subjects for each stimuli type and: (a) the difference of binaural loudness with and without the SC, (b) the difference of IACC with and without the SC, and (c) the difference between the square roots of the direct-to-reverberant energy ratios (before taking logarithm) with and without the SC. We used the binaural loudness and the square root of the direct-to-reverberant energy ratios since these are the psychophysical magnitudes that scale with the ADP.

The results of the Pearson correlation coefficients along with their confidence intervals are displayed in Table 1. Stimuli types 1-3 give highly significant (p¡0.001) correlations with all the cues, with all lower bounds of the confidence interval higher than 0.5. This is expected, since the magnitudes BL, IACC and DR are also correlated between them (the focalization region is the same, see Fig. 2). However, stimuli types 1 and 2 give higher correlation values with the binaural loudness cue (r¿0.85). For stimuli type 4, there is a highly significative r value (p¡0.001) of 0.47 with the DR cue only. A correlation value of 0.5 is also obtained with the BL cue but with lower significance (p=0.005). Finally no significant correlation was found for the results of stimuli type 5.

V. Conclusions

In this work we found significant shifts in the auditory distance perception (ADP) due to the presence of a sonic crystal slab (SC). Subjects consistently reported that sound sources were perceived closer in the presence of the SC within the focalization region. These ADP shifts were mainly related to an increase of the intenisty of the signal in the focalization region of the SC, but they were also observed after that all the stimuli were normalized in loudness. This indicates that there are also phase and temporal effects that contribute to the ADP shifts in the presence of the SC. These contributions were studied through the interaural cross-correlation (IACC) and direct-to-reverberant energy ratio (DR) cues. The presence of only one of these cues (intensity IACC or DR) was enough to elicit an ADP shift around the center of the negative-refraction focalization region (1.2 kHz and 3 m from the SC).

ACKNOWLEDGMENTS

We acknowledge Ramiro Vergara for useful discussions. This work was partially funded by CONICET and UNQ.

References

[1] J. C. Middlebrooks and D. M. Green, “Sound localization by human listeners.”, Annual review of psychology **42**, 135–59 (1991).

[2] H. Colburn and A. Kulkarni, “Models of sound localization”, Sound source localization 272–316 (2005).

[3] L. Sanchis, F. Cervera, J. Sánchez-Dehesa, J. V. Sanchez-Perez, C. Rubio, and R. Martinez-Sala, “Reflectance properties of two-dimensional sonic band-gap crystals”, The Journal of the Acoustical Society of America **109**, 2598 (2001).

[4] J. V. Sanchez-Perez, D. Caballero, R. Martinez-Sala, C. Rubio, J. Sánchez-Dehesa, F. Meseguer, J. Llinares, and F. Gálvez, “Sound Attenuation by a Two-Dimensional Array of Rigid Cylinders”, Physical Review Letters **80**, 5325–5328 (1998).

[5] L. Feng, X.-P. Liu, M.-H. Lu, Y.-B. Chen, Y.-F. Chen, Y.-W. Mao, J. Zi, Y.-Y. Zhu, S.-N. Zhu, and N.-B. Ming, “Acoustic Backward-Wave Negative Refractions in the Second Band of a Sonic Crystal”, Physical Review Letters **96**, 014301 (2006).

[6] M.-H. Lu, C. Zhang, L. Feng, J. Zhao, Y.-F. Chen, Y.-W. Mao, J. Zi, Y.-Y. Zhu, S.-N. Zhu, and N.-B. Ming, “Negative birefraction of acoustic waves in a sonic crystal.”, Nature materials **6**, 744–8 (2007).

[7] B. Gupta and Z. Ye, “Theoretical analysis of the focusing of acoustic waves by two-dimensional sonic crystals”, Physical Review E **67**, 036603 (2003).

[8] I. Spiousas and M. C. Eguia, “Directional impulse response of a large cavity inside a sonic crystal.”, The Journal of the Acoustical Society of America **132**, 2842–51 (2012).

[9] P. Zahorik, D. Brungart, and A. Bronkhorst, “Auditory Distance Perception in Humans : A Summary of Past and Present Research”, Acta Acustica united with Acustica **91**, 409–420 (2005).

[10] E. Larsen, N. Iyer, C. R. Lansing, and A. S. Feng, “On the minimum audible difference in direct-to-reverberant energy ratio.”, The Journal of the Acoustical Society of America **124**, 450–61 (2008).

[11] Y. Ando, H. Sakai, and S. Sato, “Formulae Describing Subjective Attributes for Sound Fields Based on a Model of the Auditory-Brain System”, Journal of Sound and Vibration **232**, 101–127 (2000).

[12] B. C. J. Moore and B. R. Glasberg, “Modeling binaural loudness”, The Journal of the

Acoustical Society of America **121**, 1604 (2007).

	1 (all cues)		2 (IACC-only)		3 (BI-only)		4 (DR-only)		5 (spectral-only)	
	r	CI	r	CI	r	CI	r	CI	r	CI
BL	0.85	0.75/0.91	0.68	0.5/0.8	0.87	0.79/0.93	0.5	0.27/0.68	0.008	-
IACC	0.6	0.39/0.75	0.54	0.32/0.71	0.53	0.3/0.7	0.2	-	-0.03	-
DR	0.73	0.57/0.84	0.73	0.57/0.83	0.72	0.56/0.83	0.47	0.22/0.65	0.07	-

TABLE I. Pearson correlation coefficients r between the pooled data of the four subjects and three measures of the difference between condition with-SC and without-SC: BL (difference of binaural loudness), IACC (difference of IACC values), and DR (difference of the square roots of the direct-to-reverberant energy ratios). The confidence intervals (CI) are given when the r values are significantly different from chance (p¡0.05).

List of Figures