Binaural Speech Intelligibility and Interaural Cross-Correlation Under Disturbing Noise and Reverberation

A. L. Padilla-Ortíz*, F. Orduña-Bustamante

Centro de Ciencias Aplicadas y Desarrollo Tecnológico,
Universidad Nacional Autónoma de México, Circuito Exterior S/N,
Ciudad Universitaria, A.P. 70-186, Delegación Coyoacán,
C.P. 04510, México D.F., México.
*laura.padilla@ccadet.unam.mx

ABSTRACT

Subjective tests were carried out in order to investigate speech intelligibility, and the possible relative improvements that can be obtained in practical applications to acoustic communication systems, for different forms of presentation through headphones: monaural, monophonic, binaural at 0º (in front of the listener) and binaural at ±30º (right or left, relative to the listener), played back undisturbed, and also with the addition of extreme levels of disturbing noise and reverberation, with a signal to noise ratio of $SNR = -10\ dB$, and a reverberation time of $T_{60} = 10\ s$. The influence of interaural cross-correlation (IACC) of the disturbance on speech intelligibility was also studied. Phonetically balanced words in Spanish, uttered by a female speaker, were used as speech material, which were contaminated with interaurally correlated and uncorrelated noise and reverberation. Results indicate advantages of binaural speech intelligibility under adverse listening conditions; with slight improvements observed when listening under interaurally correlated noise at an angle off-center, for azimuth angles of ±30°, relative to listening at 0º. Additionally, results show that speech intelligibility improves slightly when the disturbing reverberation has a low interaural cross-correlation (IACC).

Keywords: Speech intelligibility, binaural sound, noise, reverberation, interaural cross-correlation.

RESUMEN

Se llevaron a cabo pruebas subjetivas de inteligibilidad de la voz, y las ventajas que pueden obtenerse en aplicaciones prácticas en sistemas de comunicación acústica, utilizando diferentes formas de presentación a través de audífonos: monoaural, monofónica, biaural a 0º (frente al oyente) y biaural a ±30º (a la derecha o izquierda respecto al oyente), las señales se reprodujeron sin modificar, y también contaminadas con niveles extremos de ruido y reverberación, con una relación señal a ruido de $SNR = -10\ dB$, y un tiempo de reverberación de $T_{60} = 10\ s$. Se estudió también la influencia de la correlación interaural (IACC) en las señales de voz contaminadas con ruido y reverberación. El material de voz empleado consistió en palabras en español fonéticamente balanceadas, emitidas por una mujer; las palabras se contaminaron con señales correlacionadas y no correlacionadas de ruido y reverberación. Los resultados muestran una ventaja en la escucha biaural para la inteligibilidad de la voz ante condiciones acústicas adversas, observándose también una ligera mejoría en presencia de ruido con alta correlación interaural, cuando el escucha no está directamente frente al orador, para ángulos acimutales de ±30°, comparado con la presentación frontal a 0º. Adicionalmente, los resultados muestran que la inteligibilidad mejora cuando la reverberación tiene baja correlación interaural.

1. Introduction

Binaural rendering systems synthesize virtual 3D sound sources, providing the listener, via headphones, with the sensation of one or more sound sources in space. In addition to this, binaural rendering can also lead to significant improvement of perceived acoustic performance, especially in speech communication under adverse listening conditions involving multiple concurrent sound sources, high noise levels and reverberation. At present, interesting applications can be found for this in mobile acoustic communication systems; also, it can lead to benefits in many areas, like aviation, in order to improve the understanding of speech signals under noisy conditions received by pilots [1], or navigation system for the visually impaired people

[2], among others. One focus of our research is on optimizing binaural rendering with respect to the perceived acoustic performance; particularly, in terms of speech intelligibility. In order to validate the hypothesis that acoustic performance of mobile devices can be enhanced by using binaural rendering through headphones, and to quantify this benefit, new test procedures and criteria for acoustic evaluation specific to binaural sound systems need to be developed. Criteria for improved acoustic performance that are relevant for binaural technologies include: speech intelligibility, immunity to noise and reverberation, among others.

The present article presents methods and results that contribute towards the development of test procedures and criteria for acoustic communication performance evaluation, specific to binaural sound systems, considering aspects related to speech intelligibility, perceived acoustic performance, immunity to noise and reverberation, etc. One of the main challenges is to devise metrics that quantify the perceived acoustic communication performance of binaural systems such as speech intelligibility so that comparisons can be made with conventional systems, and to define test methodologies that can be followed to measure acoustic communication performance in the presence of multiple concurrent speech, noise sources and reverberation.

It is well known that binaural hearing improves speech intelligibility within a noisy environment [3, 4], even with negative signal-to-noise ratios [5]. It is also well known that listeners have the ability to focus attention on a single talker among multiple conversations in the presence of background noise, this is the so-called *cocktail party effect* [6, 7]. In general terms, listening with both ears allows subjects to locate sources in space and may have a significant effect on speech intelligibility [8, 9].

Speech intelligibility improves when speech and noise come from different directions. Kock [10] reported that under binaural conditions, a listener has the ability to focus on the desired signal coming from a specific direction; both noise and reverberation are eliminated effectively. Dirks and Wilson [11] demonstrated that a separation of 10° between speech and noise sources were sufficient in order to obtain a measurable change in

intelligibility. MacKeith and Coles [12] carried out experiments with two loudspeakers near the head of each subject who listened to a speech signal played from one loudspeaker and noise from the other loudspeaker. The results show that the maximum overall binaural advantage of 18 dB was observed when speech and noise signals were directed straight from the opposite side of the head (noise at +90° and speech at +180°). Several studies focus on the effects of reverberation in speech intelligibility; these have demonstrated that binaural gain reduces as the reverberation time increases [13, 14, 15]. Subjects with impaired hearing can also benefit from binaural hearing [16, 17, 18]; although these cases are not addressed in the present investigation.

Objective models for binaural speech intelligibility have been proposed. A commonly accepted model is the binaural equalization – cancellation (EC) mechanism proposed by Durlach [19], which takes advantages of the fact that signals coming from different directions cause different interaural time and level differences. The equalization step is supposed to attenuate and delay the signal from one ear with respect to each other, so that left- and right-ear signals are first matched in amplitude and phase, while in the cancellation step, the signal in one ear is then subtracted from the other ear, aiming at maximizing the speech-to-noise SNR.

In the following sections of this paper, we present experiments that extend previous research on speech intelligibility, especially with respect to binaural rendering, considering aspects such as form of presentation of the speech signal, influence of noise and reverberation, angle of presentation, and the influence of interaural cross-correlation. In the literature there are no studies of speech intelligibility under extreme listening conditions of noise and reverberation or different forms of presentation (monaural, monophonic, binaural and binaural ±30°) as studied in this work and which provide original contributions to this field.

2. Speech intelligibility tests

2.1 Experiment 1: Speech intelligibility tests with uncorrelated noise and reverberation

Speech intelligibility was investigated for different forms of presentation: monaural (speech signal in

one ear), monophonic (speech signal in both ears) and binaural (left and right ear signals recorded with the speech source in front at 0º, and at ±30º left or right of the listener). The speech signals were presented undisturbed, or disturbed with additive noise or reverberation. In the monophonic and binaural forms of presentation, involving both ears, disturbances were interaurally uncorrelated.

2.1.2 Subjects

Thirty subjects (10 female, 20 male) took part in these listening tests. Before the test, subjects were asked about their hearing health condition. All of them reported normal hearing and that they had not been exposed to notoriously loud noises in the period before the tests. A random sample of 8 subjects was audiometrically screened, all of them presenting normal hearing to within 15 dB HL. The age range was from 20 to 36 years old, with an average of 25.8. All of them were university students and Mexican Spanish native speakers. None of the participants were previously familiar with the lists of words used in the study. Listeners had no prior experience in any psychoacoustic experiments.

2.1.3 Test material

Bisyllabic words with meaning in Spanish were used in this study; prosodically, all of the words are of paroxytone type (with an accent on the next to last syllable of the word), representing the most common type of bisyllable words in Spanish [20]. Here it is important to point out that development of intelligibility test material in Spanish has so far been very limited; however, there have been some efforts on this regard [21 - 25]. Some of this research has pointed out the difficulty of making lists of monosyllabic words in Spanish because of the lack of a sufficient number of meaningful monosyllables [21, 22].

Speech material consisted of 4 different lists of words, with 50 bisyllable phonetically balanced (PB) Spanish words, commonly used in everyday conversation [26, 27] (See Appendix). Speech was produced by a female speaker born in Mexico City. In the recording, words were preceded by different carrier sentences in Spanish, which in translation are similar to: "The next word is…".

2.1.4 Recording of speech material

Binaural recording of the speech material was done in a semi-anechoic chamber using two acoustic manikins (artificial ear, mouth, head, and torso simulators). One acoustic manikin (Head Acoustics, Head Measurement System model HMS II.3) was used to play back the speech signal, while the recording was done using a second manikin (Head Acoustics, model HMS III). The use of a manikin for the reproduction of the speech signal is to avoid variations in the speech level of a human speaker in different recorded versions of the same list of words that were required in this study. The manikins were placed facing each other and the distance between them was approximately 1 m. The speech stimuli were recorded at 44.1 kHz sampling rate. These recordings were carried out at three different azimuth angles: 0º (front), +30º and -30º, where a negative sign refers to the left-hand and a positive sign to the right-hand side.

2.1.5 Signal processing of speech material

Speech material was presented undisturbed, and also under severe disturbing acoustic conditions in order to ensure a significant loss of speech intelligibility under simulated adverse listening conditions. In order to obtain the same severe degradation of the signal, considering noise or reverberation, a modulation reduction factor of $m = 0.1$ was proposed, corresponding to a Speech Transmission Index, $STI = 10\%$, which is qualitatively associated with bad intelligibility [28]. The modulation reduction factor can be calculated as follows:

$$m(F) = \frac{1}{1+\left[2\pi F\frac{T}{13.8}\right]^2} \times \frac{1}{1+10^{(-SNR)/10}} \qquad (1)$$

where F is the modulation frequency in Hz, T the reverberation time in seconds, and SNR the signal-to-noise ratio in dB. Assuming a modulation frequency of $F = 2\,Hz$, consistent with the pace of speech production present in the recordings of the test material that was used (which amounts, on average, to approximately two syllables pronounced per second), the proposed modulation reduction factor of $m = 0.1$ is then obtained, according to Equation (1), either with a

reverberation time of $T = 10\,s$, or with a signal-to-noise ratio of $SNR = -10dB$. These values of reverberation time and signal-to-noise ratio were used to generate artificially disturbed speech signals as described below. Figure 1 illustrates the way binaural recordings were processed to add noise and reverberation.

2.1.5.1 Speech disturbed with reverberation

In order to obtain a reverberated speech signal, recorded speech was convolved with an artificial reverberant impulse response, generated as follows:

$$h(t) = h_0 u(t) exp\left(\frac{-6.9t}{T}\right) \qquad (2)$$

Where T is the reverberation time, $u(t)$ is one instance of a random signal with a uniform distribution of probability, zero mean, and unit variance, and h_0 is a scale factor. The reverberation time was set to $T = 10\,s$. The reverberated speech signals were obtained by convolution, as follows:

$$S_{LT} = S_L(t) * h_L(t) \qquad (3)$$

$$S_{RT} = S_R(t) * h_R(t) \qquad (4)$$

Where $S_L(t)$, $S_R(t)$ are the original (clean, undisturbed) binaural signals recorded at the left and right ears of the acoustic manikin; $h_L(t)$, $h_R(t)$ are the reverberant impulse responses, which can be the same or independent at the left and right ears. In monaural presentation, only Equations (3) or (4) were employed. A monaural mix was presented at the two ears in monophonic presentation.

The use of Equation (2) to generate artificial room impulse responses can be justified as follows. There are a number of computational acoustic methods that can generate room impulse responses which are still artificial but more realistic in their basic properties. However, we are interested in impulse responses with acoustic characteristics which are not particular to any given room, but which can be generalized to a broad class of rooms, whose main characteristic is reverberation time only. In this sense, Equation (2) randomly generates room impulse responses with no particular early reflection pattern, while always ensuring an exponential amplitude decay with the prescribed reverberation time.
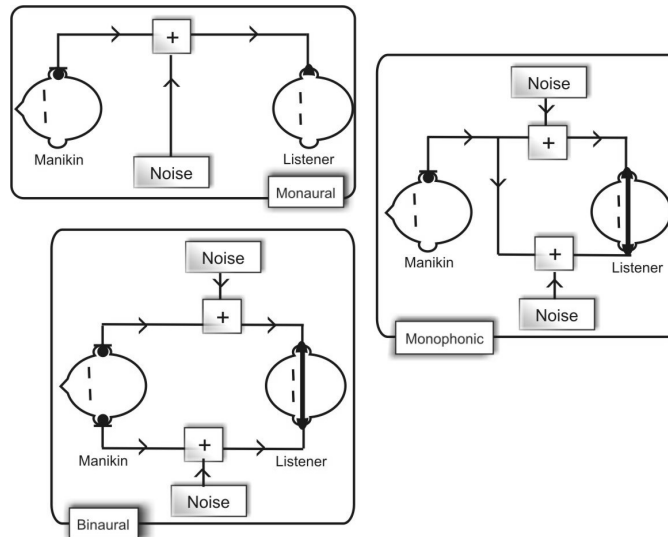

Figure 1. Preparation of the speech samples. Monaural: speech signal and disturbance presented only on right or left ear. Monophonic: same speech signal at the two ears, but independent noise and reverberation. Binaural: different signals at the two ears, binaural recording, with independent noise and reverberation.

2.1.5.2 Speech disturbed with noise

In order to contaminate the speech signals, white noise was added in the two channels of the clean speech recording. For monophonic and binaural signals, noise was different (statistically independent) for the two ears. The signal-to-noise ratio was set to $SNR = 10\ dB$.

$$S_{LN} = S_L(t) + N_L(t) \qquad (5)$$

$$S_{RN} = S_R(t) + N_R(t) \qquad (6)$$

Where $S_L(t)$, $S_R(t)$ are the binaural signals recorded at the left and right ears; $N_L(t)$, $N_R(t)$ are random noise signals with uniform distribution, zero mean, and scaled to the specified signal-to-noise ratio.

2.1.6 Presentation of the speech samples

The signals recorded with the manikin were presented to listeners in three different ways: monaural, monophonic and binaural. Table 1 shows the experimental conditions and the signal presentation.

- **Monaural (A):** The two channels of the binaural recording were averaged and the mix was played back at only one ear (right or left, letting the subjects to choose).
- **Monophonic (M):** The same monaural mix was played back at the two ears.
- **Binaural (B):** The binaural recording was played back at the two ears.

Experimental Condition	Label	Signal presentation
Monaural		Just one ear (left or right)
Clean	**A**	
Noise	**AN**	
	AT	
Reverberation		
Monophonic		Same speech signal at both ears, but statistically independent noise and reverberation
Clean	**M**	
Noise	**MN**	
	MT	
Reverberation		
Binaural		Binaural speech signal, noise and reverberation statistically independent
Clean	**B**	
Noise	**BN**	
	BT	
Reverberation		
Binaural ±30º		
Clean	**B±30º**	
Noise	**BN±30º**	
	BT±30º	
Reverberation		

Table 1. Summary of experimental listening conditions for Experiment 1.

2.1.7 Procedure

Intelligibility tests were carried out in a room with very low background noise level (L_{eq} below 40 dBA). The subject's task was to listen to the speech material via headphones (Sennheiser HD 600) and to write down the words on an answering paper sheet, in order to allow subjects enough time to write the words, a silent pause of 3 seconds was included in the recording after each word. Spelling mistakes were not considered relevant (e.g., "jestor"/"gestor"). A total of 300 words were heard by listeners in each session. The speech intelligibility score was calculated as the percentage of words which were correctly written. For monaural presentation, only one of the headphone channels (left or right) was used, letting the subjects choose which channel to use.

The test was split in 2 sessions, and each session took typically about 45 minutes for each subject. The sound pressure level of clean speech samples presented to the subjects was measured to be L_{eq} = 70.0 dBA. Without modifying the presentation gain, speech samples contaminated with noise had an increased presentation level of L_{eq} = 80.0 dBA, this is because of the -10 dB SNR. Speech samples with reverberation had a measured presentation level of L_{eq} = 69.7 dBA.

After a first analysis of the results, it was observed that some words from the lists were consistently very difficult to understand by the majority of subjects, especially under severe reverberation disturbance. The number of these words was different in each list, leading to some inconsistent results in the comparative evaluation of speech intelligibility under different conditions which were tested with different lists of words. For that reason, some words were eliminated in some of the tests, considering for further analysis only those words that were understood by at least 33% of the subjects in each of the tests; in some cases, these led to reduced lists of less than 50 words, as shown in Tables 2 and 4.

2.1.8 Results

The results show that when the signal is clean, the intelligibility score is almost 100% in all presentations (monaural, monophonic, binaural and binaural ±30°), demonstrating that the undisturbed speech material used in our tests is in fact an adequate reference, resulting in the maximum possible intelligibility score from the subjects.

Results also indicate a significant advantage of binaural rendering systems in terms of speech intelligibility under adverse listening conditions (Figure 2), consistent with the findings of several studies [6, 8, 11, 12, 13, 14]. In the present study, and under noisy conditions, speech intelligibility in monophonic presentation (90.2%) improves over monaural presentation (79.4%) by 10.8%. Similar intelligibility scores were obtained for monophonic (90.2%) and front binaural presentations (89.7%), which is to be expected, because the monophonic signal and the left and right binaural signals are nearly indistinguishable in this case. In binaural presentation (90.8%) under noisy conditions at ±30° there is an advantage of 11.4% more intelligibility over monaural presentation (79.4%).

In reverberant conditions, speech intelligibility in monophonic presentation (74.9%) improves over monaural presentation (65.8%) by 9.1%, with further improvements over monaural presentation of 14% in front binaural (79.8%), and of 17.6% when the listener is not directly facing the talker, with azimuth angles of ±30° (83.4%). Table 2 shows the number of words that were understood by at least 33% of the subjects and which were used in the analysis of the results.

2.1.8.1 Statistical analysis

Confidence intervals were calculated, based on Student's *t*-distribution [29], with a statistical significance of 95%. Statistical hypotheses were also tested, and the statistical significance was calculated as a function of the number of subjects involved in order to optimize (reduce) the number of subjects in subsequent tests (See Figure 3).

Statistical tests show, with a significance level of 95%, that under noise and reverberation conditions, binaural intelligibility will be greater than monaural presentation. According to the results shown in Figure 3, from 10 to 15 subjects are already sufficient to attain a high statistical significance; for this reason, it was decided to include only 15 subjects in the second test.
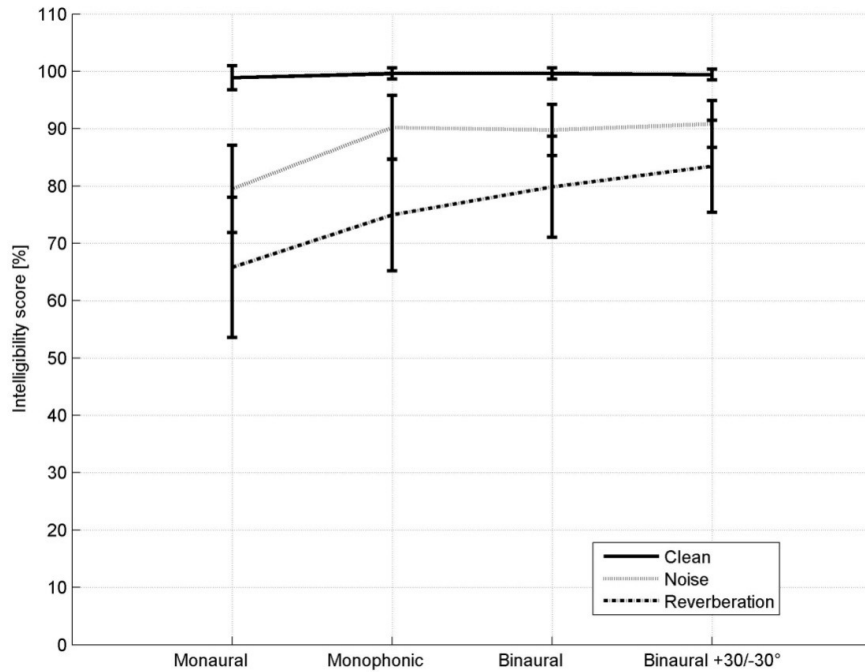
Figure 2. Percent speech intelligibility for different disturbance conditions and different forms of presentation for all 30 subjects. The error bars indicate the standard deviation.

Presentation	Number of words
Monaural	
Noise	44
Reverberation	31
Monophonic	
Noise	49
Reverberation	41
Binaural	
Noise	44
Reverberation	40
Binaural ±30º	
Noise	49
Reverberation	46

Table 2. Signal presentation under noisy and reverberant uncorrelated condition and number of words used in the analysis of the results.
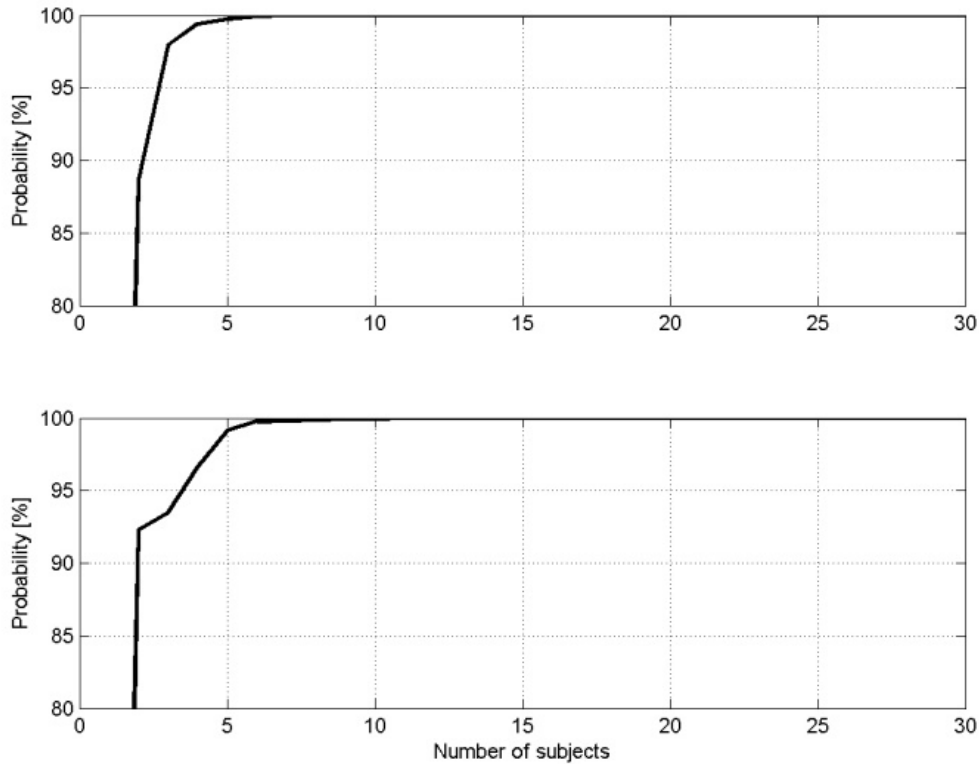
Figure 3. Statistical significance of hypothesis tests for different cases:
"binaural intelligibility greater than monaural intelligibility with noise" (top) and
"binaural intelligibility greater than monaural intelligibility with reverberation" (bottom).

2.2 Experiment 2: Speech intelligibility tests with correlated noise and reverberation

In order to investigate the effects of interaural cross-correlation (IACC) in binaural speech intelligibility, similar tests were conducted, but now with the same (fully correlated) noise and reverberation disturbance in both ears. The interaural cross-correlation evaluates the similarity of the signals at the two ears of the listener according to the following definition:

$$IACC_{t_1 t_2} = max \left| \frac{\int_{t_1}^{t_2} S_L(t) S_R(t+\tau) dt}{\sqrt{\int_{t_1}^{t_2} S_L^2(t) dt \times \int_{t_1}^{t_2} S_R^2(t) dt}} \right| \qquad (7)$$

where S_L and S_R represent the binaural signals measured at the left and right ears, respectively, t_1 and t_2 define the time limits of the binaural signals, and τ is the conventional time delay interval of -1 to +1 ms within which we search for the maximum of the correlation [30]. IACC values range from 0 to +1. A value of +1 means they are fully correlated, and 0 means they have no correlation at all. In this experiment, the same noise and reverberation (interaurally correlated) signal was used in both ears so that monaural presentation was not considered in this experiment. Only monophonic, binaural and binaural ±30° presentations were used.

2.2.1 Subjects

A total of fifteen subjects (6 female, 9 male) participated in this experiment, with the same characteristics as in the previous test. The age range of the subjects was from 21 to 38 years old, with an average of 26.9. The group of subjects was different from that in the first experiment. Some of them were audiometrically screened to ensure normal hearing.

2.2.2 Preparation of the speech samples

The binaural recordings used in this experiment were the same as in the previous experiment, but speech stimuli were contaminated with the same noise and reverberation signal in both ears (See Figure 4). Thus, disturbances are now interaurally correlated, rather than uncorrelated, as before. Binaural ± 30º presentation was processed in the same way as for the binaural 0º presentation as shown in Figure 4.

2.2.3 Procedure

The procedure for experiment 2 was the same as in the previous experiment; only that in this case, the undisturbed (clean) signal was not used. Table 3 shows the experimental conditions and the signal presentation for Experiment 2. As it was indicated, monaural presentation was not used in this test. A total of 150 words were heard by the listeners in a single 45 minute session.
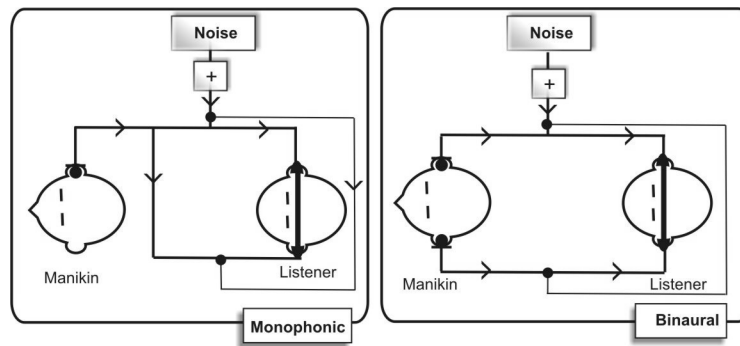

Figure 4. Preparation of the speech samples for experiment 2.
Monophonic: same speech, noise and reverberation signal at the two ears.
Binaural: binaural recording, same noise and reverberation signal.

Experimental Condition	Label	Signal presentation
Monophonic **Noise** **Reverberation**	**MN** **MT**	Same speech, noise and reverberation signal at both ears
Binaural **Noise** **Reverberation**	**BN** **BT**	Binaural speech signal, with the same noise and reverberation disturbance in both ears.
Binaural ±30º **Noise** **Reverberation**	**BN±30º** **BT±30º**	

Table 3. Summary of experimental listening conditions for Experiment 2.

2.2.4 Results

Figure 5 shows that under correlated noise conditions, speech intelligibility in binaural ±30º presentation (97.5%) improves over binaural ±30º presentation (90.8%) under uncorrelated noise by 6.7%, while on the other hand, speech intelligibility improves for uncorrelated reverberant conditions in monophonic (at the two ears) and binaural presentations compared with correlated reverberation. In reverberant conditions and monophonic presentation, speech intelligibility improves by an additional 7.7%, for uncorrelated (74.9%) compared with correlated reverberation (67.2%). These results agree with those from other study [31], which have shown that speech intelligibility tends to be poor for the monophonic presentation when the signals (speech and disturbance) are the same in both ears (IAAC = 1). Table 4 shows the number of words that were

understood by at least 33% of the subjects, and which were used in the analysis of the results. Table 5 shows the calculated values of IACC for the different presentations.

Presentation	Number of words
Monophonic	
Noise	42
Reverberation	24
Binaural	
Noise	50
Reverberation	37
Binaural ±30º	
Noise	49
Reverberation	40

Table 4. Signal presentation under noisy and reverberant correlated condition and number of words used in the analysis of the results.
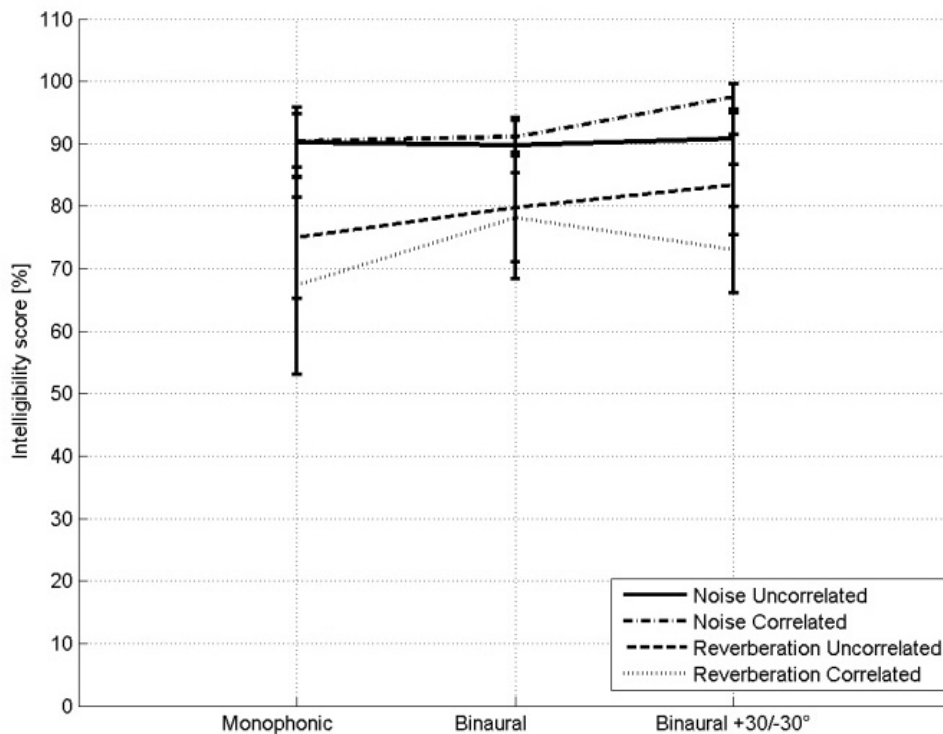

Figure 5. Percent speech intelligibility with correlated and uncorrelated noise and reverberation for different forms of presentation. The error bars indicate the standard deviation. Results for uncorrelated disturbances are from Experiment 1.

Presentation	Interaural cross-correlation	
	Uncorrelated	Correlated
MN	0.2955	1
MT	0.0860	1
BN	0.2840	0.9898
BT	0.0780	0.9769
BN+30	0.2549	0.7324
BT+30	0.1049	0.9094
BN-30	0.2209	0.7334
BT-30	0.0748	0.9113

Table 5. Interaural cross-correlation for monaural, monophonic, binaural, and binaural ±30° presentations in Experiment 1 and 2.

3. Conclusions

Speech intelligibility tests were carried out under extreme conditions of artificially produced noise and reverberation. A signal-to-noise ratio of $SNR = -10\ dB$ was considered, while everberation time was $T_{60} = 10\ s$. Different forms of presentation and interaurally correlated and uncorrelated disturbances were considered. The results from these tests show that speech intelligibility is affected by noise and reverberation, and that under these adverse listening conditions, speech intelligibility can be improved by binaural rendering, as compared with monaural and monophonic presentations.

In our results, speech intelligibility with the talker at 30° is larger for correlated noise than for uncorrelated noise, which is consistent with the findings of another study [31]. This result is also consistent with, and can be explained partly on the basis of, the widely accepted equalization-cancellation model for binaural speech intelligibility mentioned in the Introduction. According to the binaural E-C model, interaurally correlated noise enables the cancellation step, while interaurally uncorrelated noise makes cancellation difficult or impossible. The other important factor is lateralization of the speech signal, which generates interaural differences that the binaural hearing mechanism exploits in order to emphasize the speech-to-noise ratio, improving speech intelligibility.

On the other hand, our results for correlated and uncorrelated reverberation show that speech intelligibility tends to be larger for uncorrelated reverberation than for correlated reverberation. These results can also be explained along the lines of the equalization-cancellation mechanism because under uncorrelated reverberation, the reverberant impulse responses are different in both ears, and the interaural phase relationship remains time invariant, and for that reason, the binaural hearing mechanism is able to equalize and cancel the unwanted signal and to improve speech intelligibility. Paradoxically, this is unlike the case of interaurally uncorrelated noise. In that case, the interaural phase relationship varies randomly with time and cannot be overcome by the binaural hearing mechanism in order to achieve the cancellation step.

This comparative study of binaural speech intelligibility under extreme conditions of noise and reverberation, different forms of presentation, and different interaural cross-correlation constitute original contributions to this field.

References

[1] Begault, D., Wenzel, E., *"Techniques and Applications for Binaural Sound Manipulation in Human-Machine Interfaces"*, J. Aviation Psyhology, 1992, pp. 1-22.

[2] Loomis, J., Golledge, R. y Klatzky, R., "*Navigation system for the blind: Auditory display modes and guidance*", Presence: Teleoperators and Virtual Environments 7, 1998, pp. 193-203.

[3] Bronkhorst, A. W., *"The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions"*, Acustica, Vol. 86, 2000, pp. 117-128.

[4] Hawley, M. L., Litovsky, R. Y., and Culling J. F., *"The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer"*, J. Acoust. Soc. Am., Vol. 115, 2004, pp. 833 – 843.

[5] Koenig, W., *"Subjective effects in binaural hearing"*, J. Acous. Soc. Am., Vol. 22, 1950, pp. 61-62.

[6] Cherry, E. C., *"Some experiments on the recognition of speech with one and two ears"*, J. Acoust. Soc. Am., Vol. 25, 1953, pp. 975 – 979.

[7] Pollack, I., and Pickett, J. M., *"Stereophonic listening and speech intelligibility against voice babbles"*, J. Acous. Soc. Am. Vol. 30, 1958, pp. 131 – 133.

[8] Hirsh, I. J., *"The relation between localization and intelligibility"*, J. Acoust. Soc. Am., Vol. 22, 1950, pp. 196-200.

[9] Hawley, M. L., Litovsky, R. Y., and Colburn, H. S., "*Speech intelligibility and localization in a multiple-source environment*", J. Acoust. Soc. Am., Vol. 105, No. 6, 1999, pp. 3436 – 3448.

[10] Kock, W. E., *"Binaural localization and masking,"* J. Acous. Soc. Am., Vol. 2, 1950, pp. 801-804.

[11] Dirks, D. D., and Wilson, R. H., *"The effect of spatially separated sound sources on speech intelligibility"*, J. Speech Hear. Res., Vol. 12, 1969, pp. 5 - 38.

[12] MacKeith, N. W., and Coles, R. R. A., *"Binaural advantages in hearing of speech"*, J. Laryngol. Otol., Vol. 85, 1971, pp. 213-232.

[13] Moncur, John P., Dirks, D., *"Binaural and monaural speech intelligibility in reverberation"*, J. Speech Hear. Res., Vol. 10, 1967, pp. 186 - 195.

[14] Nábělek, A.K., and Pickett, J. M., *"Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids"*, J. Acous. Soc. Am., Vol. 56, No. 2, 1974, pp. 628 – 638.

[15] Libbey, B., Rogers, P., *"The effect of overlap-masking on binaural reverberant word intelligibility"*, J. Acoust. Soc. Am., Vol. 116, No. 5, 2004, pp. 3141 - 3151.

[16] Nábělek, A.K., and Mason D., *"Effect of noise and reverberation on binaural and monaural word identification by subjects with various audiograms"*, J. Speech Hear. Res., Vol. 24, 1981, pp. 375 - 383.

[17] Bronkhorst, A. W., and Plomp R., *"Binaural speech intelligibility in noise for hearing-impaired listeners"*, J. Acoust. Soc. Am., Vol. 86, No. 4, 1989, pp. 1374 – 1383.

[18] Beutelmann, R., and Brand, T., *"Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners"*, J. Acoust. Soc. Am., Vol. 120, 2006, pp. 331 – 342.

[19] Durlach, N.I., *"Equalization and cancellation theory of binaural masking-level differences"*, J. Acoust. Soc. Am., Vol. 35, No. 8, 1963, pp. 1206-1218.

[20] Zubick, H., Irizarry, L. M., Rosen, L., Feudo, P. Kelly, J. H., and Strome, M., "*Development of speech audiometry materials for native Spanish-speaking adults*", Audiology, Vol. 22, 1983, pp. 88-102.

[21] Tato, J.M., Lecciones de Audiometría. Buenos Aires, Argentina. El ateneo, 1949.

[22] Ferrer, O., *"Speech Audiometry: A discrimination test for Spanish language"*, Laryngoscope, Vol. 70, 1960, pp. 1541-1551.

[23] Cancel, C. A., "*Multiple-choice intelligibility lists for Spanish speech audiometry*", International Audiology, Vol. 4 (2), 1965, pp. 91-93.

[24] Berruecos, T., Rodriguez, J., "*Determination of the phonetic percent in the Spanish language spoken in Mexico City, and the formation of PB lists of trochaic words*", Int. Audiol., Vol. 6 (2), 1967, pp. 211-216.

[25] Benitez, L., Speaks, C., "*A test of speech intelligibility in the Spanish language*", Int. Audiol., Vol. 7 (1), 1968, pp. 16-22.

[26] Castañeda G. R., Pérez R. S., *"Análisis fonético de las listas de palabras de uso más extendido en logoaudiometría"*, Anales de la Sociedad Mexicana de Otorrinolaringología. No. 1, 1991, pp. 23-30.

[27] Salinas, G. "Validación mediante índices acústicos de material de voz para medir la inteligibilidad". Tesis de Licenciatura (2007). Universidad Nacional Autónoma de México. México, D.F.

[28] Houtgast, T., and Steeneken, H. J. M., *"A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria"*, J. Acoust. Soc. Am., Vol. 77, No. 3, 1985, pp. 1069-1077.

[29] Spiegel M., Schiller J., Srinivasan A., Probability and Statistics (C-4). Schaum's Outlines. Third edition. New York, Mc. Graw Hill, 2009, pp. 108-150.

[30] Fletcher N.H. Springer Handbook of Acoustics. 1 Ed. Springer, 2007, pp. 1182.

[31] Licklider, J., *"The influence of interaural phase relations upon masking of speech by white noise,"* J. Acoust. Soc. Am., Vol. 20, 1948, pp. 150–159.

Appendix: List of words used in the tests

	List 1	List 2	List 3	List 4
1	nube	cedros	Corea	radio
2	dejo	gose	dulce	mide
3	críos	meta	siglo	sello
4	cuales	sella	yeso	mini
5	beca	duda	taches	Irma
6	clavo	pluma	pur	algún
7	surco	fierro	vino	cano
8	pleno	lina	níquel	dieta
9	flaca	pista	monte	tigre
10	torno	duelo	Nilo	hombro
11	grasa	nombre	canción	unos
12	neta	cumbres	reto	reto
13	medios	disco	tecleo	freno
14	miden	ciega	tarde	cerca
15	reno	trenza	flanes	sudo
16	nina	manto	busto	Diego
17	botes	cebra	turco	seda
18	perros	timbre	viernes	jaque
19	calor	dije	quepa	pisen
20	duna	norte	celtas	prensa
21	ellos	perla	hacia	pili
22	sigo	cena	cama	calle
23	piano	celo	lloro	lila
24	choca	tira	luces	cardo
25	llenos	lince	premios	hambre
26	suela	jalan	damas	brazo
27	duque	nuca	pajes	sebo
28	mimo	noche	cabe	lacre
29	cita	codo	Carmen	plato
30	diosa	nena	corta	tapia
31	selva	miope	ciega	senda
32	caro	puse	libre	clame
33	cierta	Viena	deme	liso
34	crean	cero	mismo	curas
35	une	pacto	surco	sones
36	gestor	laca	tina	control
37	listo	niña	vienen	tape
38	pera	himno	regla	Carmen
39	cifra	cera	suela	lista
40	prima	alla	dardo	feria
41	simple	talco	fino	dante
42	persa	seda	cielo	dones
43	toro	conde	necio	seta
44	deme	tiro	dota	nave
45	veinte	saco	trance	nulo
46	dime	dique	padre	buque
47	lenta	lista	pardo	queso
48	celda	seso	onda	esos
49	tiendas	sigo	nadie	siete
50	nada	cura	pica	asno