Effects of auditory noise on the psychophysical detection of visual signals: Cross-modal stochastic resonance

Elias Manjarrez a,*, Ignacio Mendez a, Lourdes Martinez a, Amira Flores a, Claudio R. Mirasso b

a Instituto de Fisiología, Benemérita Universidad Autónoma de Puebla, 14 Sur 6301, Col. San Manuel, CP 72570, Puebla, Pue., Mexico
b Departament de Física, Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain

Received 21 September 2006; received in revised form 20 December 2006; accepted 9 January 2007

Abstract

Harper [D.W. Harper, Signal detection analysis of effect of white noise intensity on sensitivity to visual flicker, Percept. Mot. Skills 48 (1979) 791–798] demonstrated that the visual flicker sensitivity was an inverted U-like function of the intensity of different levels of auditory noise from 50 to 90 dB (SPL), without concomitant changes in the response bias. The aim of the present study was to extend these observations in the context of the stochastic resonance, a counterintuitive phenomenon in which a particular level of noise enhances the response of a nonlinear system to a weak input signal. We show psychophysical evidence in a yes–no paradigm for the existence of a stochastic resonance-like phenomenon in the auditory–visual interactions. We show that the detection of a weak visual signal was an inverted U-like function of the intensity of different levels of auditory noise. Nevertheless, for a strong visual signal the auditory noise acts in detriment of the ability of visual detection. Our results suggest that auditory noise could be employed in vision rehabilitation interventions in order to improve the detection of weak visual signals.

Keywords: Noise; Stochastic resonance; Psychophysics; Cross-modal; Multisensory

The stochastic resonance (SR) is a phenomenon in which an intermediate level of noise enhances the response of a nonlinear system to a sub-threshold (weak) input signal [33,19,7,11]. Typically, the amplitude of the nonlinear response versus the input noise is an inverted U-like function characterized by maximal enhancement of the response at a specific noise amplitude value. SR type effects have been demonstrated in physical and biological systems [18]. SR has been studied in diverse sensory systems in which the signal and noise applied were of the same sensory modality [11–14,26,17,6,34,35,25]; however, there are no studies related with the influence of noise of one sensory modality on the responses elicited by periodic stimuli of another modality. The purpose of the present study was to analyze changes in the ability of humans to detect sub-threshold (or supra-threshold) visual stimuli when a continuous auditory noise was applied. We assume that within the central nervous system there are multisensory neurons [2,27–32,9,4,5] participating in the auditory–visual cross-modal integration of the signal and noise inputs.

We performed a series of psychophysical experiments in humans following the same experimental protocol employed by Collins et al. [6], except that now we have applied auditory noise and a sub-threshold (or supra-threshold) visual stimulus. The present experiments extend previous studies about the effects of auditory noise in the multisensory integration of visual signals [8,23]. Specifically, our study contributes for the understanding of the visual–auditory information processing via SR within the central nervous system. We provide support to the recent observation that inverse effectiveness does not strictly apply to the auditory–visual multisensory enhancements during audiovisural speech perception [23]. We show that the ability of a subject to detect sub-threshold visual stimulus (brief flash) can be significantly enhanced by a particular intermediate level of auditory noise. In this context, the fact that an intermediate level of auditory noise improves visual function is attractive, because it suggests that auditory noise could be employed in vision rehabilitation interventions in order to improve detection of weak visual signals.

We performed psychophysical experiments in a sample of 15 healthy young subjects (18–28 years of age). All the procedures conformed to the principles set forth in the Declaration of Helsinki (1964) established by the World Medical Associa-
The stimulation protocol and the number of trials were similar to the protocol employed by Collins et al. [6] in the study of SR in psychophysical responses to tactile stimuli. We applied either a sub-threshold visual stimulus plus auditory noise or no stimulus plus noise, and we instructed the subjects to indicate to us when they detected a visual stimulus (Fig. 1A). The visual stimuli consisted of brief light flashes (50 ms). The stimulus was presented binocularly by means of one pair of white light-emitting diodes (LEDs) mounted inside of a pair of goggles. A pulse stimulator (MASTER-8, AMPI (Jerusalem)) provided input to the LEDs. The sub-threshold stimulus intensity was about $68 \times 10^{-8} \text{cd/m}^2$ ($0.1 \times 10^{-8} \text{W/m}^2$) and the supra-threshold stimulus about $546 \times 10^{-8} \text{cd/m}^2$ ($0.8 \times 10^{-8} \text{W/m}^2$). The auditory stimulus was white noise (Fig. 1A and E) presented binaurally by means of a pair of headphones. The output of a function generator (WaveTek (San Diego, CA) 132) provided input to the headphones and supplied the input noise. The intensity range of the noise generator was calibrated to supply white noise with a standard deviation of 65 to 75 dB (SPL).

We maintained the intensity of the brief flashes (50 ms) constant for each trial. Each trial consisted of a sequence of 30 (or 60) randomized presentations (Collins et al. [6] used 20 presentations), randomly distributed between “visual stimulus” and “absence of visual stimulus”. We maintained the intensity of the continuous auditory input noise constant for each trial and varied it between trials. The inter-presentation interval was 5 s (or 2 s). We applied five or six different auditory noise intensity levels in the protocol and performed one trial for each auditory noise level including the control in randomized order. This randomized order of the trials assured that the observed effects are not simply due to a modulation in attention/arousal. That means that we can exclude the possibility that subjects were simply responding more frequently when auditory noise was present.

We employed the measure of “%correct” that quantifies the percentage of trials for which a subject correctly identified the presentation of “visual stimulus” or “absence of visual stimulus” during different levels of continuous auditory noise. Subjects responded to the stimulus by drawing a small line with a pencil (Fig. 1A). The frequency of the subject’s incorrect response (response bias) was determined by an observer. The “%correct” values for each auditory noise intensity level were obtained from the formula:

$$\%\text{correct} = \left( \frac{N_{\text{correct}}}{N_{\text{total}}} \right) \times 100,$$

where $N_{\text{correct}}$ is the number of correct responses and $N_{\text{total}}$ is the number of presentations of “visual stimulus” or “absence of visual stimulus”.

Furthermore, before each protocol, we obtained psychophysical input–output curves without auditory noise to select an intensity of the visual stimulus for which the “%correct” was about 50%. Visual stimuli corresponding with a “%correct” of about 50 were considered as sub-threshold (for an equal number of “visual stimulus” and “absence of visual stimulus” presentations). We considered supra-threshold stimulus any stimulus that was close to 100.

Fig. 1B and C illustrates an example of the psychophysical response (i.e., six times the subject’s response was “yes”; Fig. 1C) when a sequence of 21 randomized presentations of “visual stimulus” or “absence of visual stimulus” were applied (arrows in Fig. 1B). Fig. 1B and C also shows that although 12 visual flashes were applied (vertical bars) the subject gave a
successful response only 6 times (i.e., 50%). Fig. 1D illustrates that in some cases (2 in this example) the subjects can produce a response bias: a false alarm.

All 12 subjects we examined (in the sub-threshold condition) exhibited a cross-modal SR-like phenomenon when sub-threshold visual stimulus was employed. Fig. 2A shows the results for these subjects. We observed that as the noise amplitude of auditory modality increases, “%correct” values of visual modality increases as well. A maximum value is reached for a particular auditory noise-amplitude level. Note that there is a particular level of auditory noise for which the detection (%correct) of the visual stimuli was improved. We observed that the graphs of the “%correct” versus the auditory input noise exhibited different profiles between subjects, showing diversity among the different individuals. From Fig. 2 we obtained that the mean maximal %correct when an intermediate level of noise was applied was 73.8 ± 15.5, and this value was significantly different from the mean %correct in control conditions 39 ± 7.8 (12 subjects; p < 0.01, Student’s t-test). This result shows that a particular level of auditory noise had an effect on the capability of detection of a sub-threshold visual stimulus.

However, when a supra-threshold visual stimulus was applied we obtained distinct results. Fig. 2B illustrates the graphs obtained from three other subjects in which the visual stimulus was supra-threshold. Note that in this case the auditory noise reduces the capability to detect the visual stimuli. The performance of the supra-threshold detection task declines with increasing auditory noise levels.

Furthermore, thanks to the suggestion of an anonymous reviewer, we performed a second series of experiments in which the response bias was measured (see methods in Fig. 1). Fig. 2A and B (triangles) shows graphs of the subject’s incorrect response (response bias; i.e., the subject’s response is “yes” in the absence of visual flashes). These results indicate that the subject’s criterion is not shifting with noise level.

Since the physiological operation of the visual system is disturbed by a variety of multisensory factors that are both unavoidable and unceasing, the study of perturbed visual sig-
nals has been considered worthy of careful exploration by many psychologists and neurophysiologists since 1935. It is interesting that previously to the discovery of the “stochastic resonance phenomenon” in 1981 [3] a substantial number of publications (see references in Harper [8]) provided evidence of a similar phenomenon in the interactions between the auditory and the visual system; i.e., that the visual flicker sensitivity was an inverted U-like function of the intensity of different levels of auditory noise. In these studies the sensitivity to visual flicker was determined by the critical flicker frequency (CFF). The CFF may be considered the transition point for an intermittent light source, at which the flicker sensation disappears, to be replaced by the sensation of continuous stimulation. With the use of this estimation and the computation of the response bias, Harper [8] performed careful experiments in humans to examine the effect of auditory white noise on sensitivity to visual flicker. Harper found that the sensitivity to visual flicker was a “sawtooth”-like function of the intensity of different levels of auditory noise from 40 to 100 dB (SPL), but an inverted U-like function of the intensity of different levels of auditory noise from 50 to 90 dB (SPL), which was a curve comparable to previous descriptions, claiming that there is a particular level of auditory noise for which the flicker sensitivity can be enhanced.

The aim contribution of Harper’s study is the demonstration that the auditory–visual U-like function can be observed without concomitant changes in the response bias. The response bias measurement is important because it is possible that the subjects become more liberal regarding his/her criterion for subject’s criterion is not shifting with noise level. For this reason, and thanks to the suggestion of an anonymous reviewer, we performed a second series of experiments in which the sub-threshold condition suggest that medium levels of auditory noise could expand the dynamical range of multisensory neurons. Harper [8] that the effect of auditory noise is not always either deleterious or innocuous, but may even be beneficial to improve visual signal transmission for certain intermediate levels of auditory noise. In particular, we have employed a different psychophysical test to examine the effects of auditory noise on visual detection thresholds in a yes–no paradigm. The results for the sub-threshold condition suggest that medium levels of auditory noise increase performance on the visual detection task. To the contrary, performance on the supra-threshold detection task declines with increasing auditory noise levels. We interpreted our findings in the context of multisensory interactions [28] and stochastic resonance [18].

The possible neural substrates that could explain our results merit a discussion. The improvement of detection of a weak visual stimulus for an intermediate level of auditory noise can be explained by the facilitation mechanism associated with the convergent synaptic inputs from auditory and visual afferent pathways on multisensory neurons. This possibility is consistent with evidence that in the superior colliculus (SC) and in the cerebral cortex there are multisensory neurons exhibiting overlapping cross-modal receptive fields from visual and auditory pathways [27,30,32,29,9,4]. Furthermore, neuroanatomical and electrophysiological studies in animals have identified cortical regions where afferents from the auditory and visual cortices converge, thus exhibiting cross-modal interactions [4]. Therefore, we suggest that the cross-modal neurons located within the SC and those multisensory neurons located within the cerebral cortex produce the SR-like phenomenon that we observed in the present study.

Another possible neural substrate for the auditory–visual SR-like effect could be the impact of auditory noise on the “dynamical range” of SC multisensory neurons. The “dynamical range” of a neuron is defined as the mean difference in activity from threshold to saturation. We suggest that an intermediate level of auditory noise could expand the dynamical range of multisensory neurons, thus producing a concomitant increase in the capability of detection of a weak visual stimulus. This suggestion is supported by evidence that in the auditory–visual interactions the size of the dynamical range of multisensory neurons is larger than the size of each unisensory dynamical range [20].

An additional possible explanation for the cross-modal SR-like effect could be that noise alters the coherence between pairs of auditory neurons. This possibility is supported by experimental [15] and theoretical [1,2,21] evidence that in the absence of deterministic peripheral stimuli, neurons exhibit a spiking behavior completely induced by noise, through a phenomenon known as coherence resonance or internal stochastic resonance [15,1,2,21,10]. The blueprint of this phenomenon is that the coherence of the system increases as a U-like function of the noise level.

On the other hand, according to these speculations, it would be interesting to study the changes in the dynamical range (to periodic visual stimuli) of multisensory coupled neurons in the presence of different levels of auditory noise.

Stein et al. [31] showed that if auditory and visual stimuli are applied in close temporal and spatial proximity, then multisensory neurons in the SC can increase their firing rate beyond the firing rate expected by adding the impulses exhibited by each unimodal input separately. This phenomenon is known as “multisensory integration” or “inverse effectiveness”. The blueprint of inverse effectiveness is that maximal multisensory response enhancements can occur under circumstances where the constituent unisensory stimuli are minimally effective in evoking responses [16,29]. Our results are inconsistent with this mechanism exhibited by SC multisensory neurons, but consistent with the recent results obtained by Ross et al. [23]. They found a clear visual enhancement of speech comprehension in noisy environments, thus showing that there is a particular intermediate level of auditory noise for which the auditory–visual multisensory integration is enhanced. In this context, our results (for sub-threshold stimuli) and those obtained by Ross et al. [23] indicate that inverse effectiveness does not strictly apply to these multisensory enhancements. Nevertheless, the depression in “%response” for strong (supra-threshold) visual stimulus (Fig. 2B) is consistent with the principle of inverse effectiveness. It is possible that at higher intensities of the unisensory stimuli redundant information is provided by both sensory inputs, and therefore the need for multisensory integration is lessened.
The diversity of responses between subjects may be attributed to their different sensitivity to visual stimuli. Discrepancies in individual threshold can be explained by variations in processes such as expectancy, motivation, attention, interest in the consequences of the response, fatigue, etc. In our experiments, these influences may participate together with other factors, such as dissimilarities in skin thickness of the eyelid, properties of the cornea and lens, pupil size, receptor density, and afferent fiber thresholds. Additionally, irregularity of the background activity at the superior colliculus, thalamic and cortical levels, and other unknown causes may explain the different profiles observed in the psychophysical graphs obtained from individual experiments (Fig. 2).

Our results are consistent with the evidence that illumination of the caudal photoreceptor interneurons in the sixth abdominal ganglion of the crayfish enhances the efficiency of transmitting mechanosensory signals [19]. They are also consistent with other studies that suggest a behavioral role of the cross-modal interactions of two different types of sensory receptors. For example, Collins et al. [5] observed that when cutaneous and muscle spindle receptors were simultaneously activated, movement illusions usually increased more than those produced by tendon vibration alone and thus they are likely to reflect the central combination of both inputs. It might be possible that such illusory movements could also be elicited by the addition of mechanical noise on the tendon (and not only by tendon vibration) via the multisensory SR mechanism.

In conclusion the present investigation extends the Harper’s study and documents the first explicit description of the occurrence of psychophysical visual–auditory cross-modal stochastic resonance. Our results are consistent with the fact that auditory noise could be employed in vision rehabilitation interventions in order to improve detection of weak visual signals. This suggestion is supported by the evidence that unisensory visual noise can improve the signal detection in human visual perception [26,24].

Acknowledgments

We thank Robert Simpson for proof reading the English manuscript. This work was partially supported by CONACyT J36062-N (E.M), PIFI-FOMES-BAUAP and VIEP-BAUAP I3/SAL/06-1 (E.M), México and from MEC (Spain) and FEDER project CONOC2 (FIS2004-00953). E.M is grateful for the hospitality and support of the Universitat de les Illes Balears. C.R.M is grateful for the hospitality and support of the BUAP. We thank A. López and P. Linares for technical assistance in some experiments.

References