ISSN : 2350-0743

www.ijramr.com

International Journal of Recent Advances in Multidisciplinary Research
Vol. 06, Issue 08, pp.5143-5149, August, 2019

RESEARCH ARTICLE

THE SEMANTIC CONGRUENCY EFFECT INTO BISTABLE VISUAL PERCEPTION: A STUDY BASED ON TONES OF VOICE AS TOP-DOWN MODULATING STIMULI

***Guillermo Rodríguez**

Ph.D. in Psychology, Lecturer Assistant at School of Advertising, Universidad Jorge Tadeo Lozano, Calle 25 # 4A - 49 / Building 20, 2nd. Floor, Bogotá D.C., Colombia

ARTICLE INFO

Article History:
Received 27th may, 2019
Received in revised form
10th June, 2019
Accepted 14th July, 2019
Published online 31st August, 2019

Keywords:

Semantic congruency, Bistable Perception, Visual Consciousness, Top-down Processes

ABSTRACT

Background: Bistable images (or ambiguous images) have the chance to be interpreted in two different ways. Owing to its physical features, this type of images offers two independentvisual perceptswhich are associated with top-down and bottom-up modulating perceiving processes. It has been also stated that tones of voice can operate as semantic modulators in such a way that the voice itself can exert a top-down influence on the way an observer perceives a bistable visual stimulus, regardless of the understanding of the speech (words meaning) that operates as a modulating factor. **Objective:** The study that is outlined here aimed at establishing the modulating effect of tones of voice on the interpretation of a bistable image. The speech understanding variable was taken into consideration as a covariate. **Method:**45 participants viewed twice the image *My girlfriend or my mother-in-law* whilst listening separately to a voice of a young woman and a voice of an old woman. The first time each participant observed the image on a screen of a fixed eye-tracker device of 60 Hz. (for 20 seconds). They listened to the audio of an old woman speaking while viewing the bistable image. The second time, the image was the same, but the auditory stimulus was a young woman speaking. The auditory monologues of both the old woman and the young woman were taken from French videos. These audios were randomly presented exchanging their order so as to have a counterbalance which could control the presentation order variable. This study was an experimental-intrasubject one. **Results:** There were significant differences between the duration of the congruent visual understandings and the duration of the incongruous visual interpretations. The results show a statistically significant difference in favor of the modulated perception, that is, that a differential result was found in favor of the image relating to the semantically-congruent percept(congruent with the audio)in relation to the semantically-incongruous image. In addition, the understanding of the modulating voices (meanings of words) did not imply an effect on the visual interpretations that were congruent with the auditory modulator. **Conclusions:** Auditory stimulation can provide congruent semantic contexts with every possible interpretation of a bistable image. The use of voice tones as semantic modulators may have the possibility of exerting an effect on the decoding of bistable images, regardless of the understanding of the content of the pronounced words.

INTRODUCTION

Bistable perception is a perceptual phenomenon whereby an observer interprets the same stimulus in two different ways (Baker, Karapanagiotidis, Coggan, Wailes-Newson, & Smallwood, 2015; Brascamp, Sterzer, Blake, & Knapen, 2018; Borisyuk, Chik, & Kazanovich, 2009;Grossmann & Dobbins, 2006; Leopold,& Logothetis, 1999; Pressnitzer,& Hupé, 2006; Weilnhammer, Ludwig, Sterzer, & Hesselmann, 2014). The stimulus remains unchanged while the observer changes from one interpretation to another because the image itself offers two different possibilities of interpretation (Long,& Topino, 2004; Schauer, Kanai, & Brascamp, 2016; Rock, Hall, & Davis, 1994; Yamamoto,& Yamamoto, 2006). The two possible percepts cannot be perceived simultaneously (Rodríguez, & Castillo, 2018a).

***Corresponding author:* Guillermo Rodríguez,**
Ph.D. in Psychology, Lecturer Assistant at School of Advertising, Universidad Jorge Tadeo Lozano, Calle 25 # 4A - 49 / Building 20, 2nd. Floor, Bogotá D.C., Colombia.

Given the fact that bistable visual stimulihave two possible percepts, they can also be called ambiguous images (Gijs,& van Ee, 2006; Okazaki, Kaneko, Rodríguez, & Castillo, 2018a, Yumoto, & Arima, 2008). Likewise, the phenomenon of visual perceptual bistability can also be named visual bistability (Intaité, Kovisto, & Castelo-Branco, 2014); the jump or change from one interpretation to another is called perceptual reversal or perceptual reversibility (Clément,& Demel, 2012; Rodríguez,& Castillo, 2018a). The study of bistable perception has led to categorize various types of bistable images (Bialystok,& Shapero, 2005; Long,& Toppino, 1981, Rodríguez, & Castillo, 2018a). Three different types of these images have been identified (see in figure 1): 1. *In figure-ground reversals*, recognizable because one possible interpretation is in the background of the image, while the other has salience relative to the background (take as an example the image known as *the vase-face illusion*); 2. *In perspective reversal images*, a kind of imageswhose perception changes both orientation and perspective (e.g. Necker's cube); and, 3. Bistable images*In meaning-content reversals*, understood

as images that alternate two interpretations at the same level of salience, where each of them is different in terms of shape and meaning, such as the image of Boring (1930) *My girlfriend or my mother-in-law*.

Figure 1. The three types of bistable images. On the left, the Necker cube (In perspective reversal).In the center, *The vase-face illusion* (In figure-ground reversals). On the right, *My girlfriend or my mother-in-law* (In meaning-content reversals). Source: Rodríguez, & Castillo (2018a)

The factors that allow this type of images to offer more than one possible interpretation have been widely studied. First of all, it has been stated that the physical characteristics of bistable imagesbring about ambiguity, so that the resulting interpretation depends on the way in which the visual stimulus is observed as well as the areas of the image in which eyes make fixations. Likewise, the visual path that the observer makes during the corresponding observation exerts an influence on the final perception (García-Pérez, 1989; García-Pérez, 1992; Hsiao, Chen, Spence, & Yeh, 2012). Moreover, as has been stated, physical aspects of the image that have an impact on its interpretation (such as shared contours for the two possible interpretations, shapes, opposite and complementary tones, luminance, etc.), impliesa modulation of visual perception of the bottom-up type (Meng,& Tong, 2004). Secondly, it has been put forward the argument that the interpretation of a bistable image also depends on the processing of information isolated from the physical characteristics of the visual stimulus, that is, from concepts and ideas that are integrated into the perceptual process, in such a way that the perceptual outcome is being modulated in a top-down way.

In other words, the interpretation of the bistable image is determined by knowledge previously stored in memory (Intaité, Noreika, Šoliūnas, & Falter, 2013; Kornmeier, Hein, & Bach, 2009), or also by information that additionally enters into the perceptual system, generating an interpretive echo at the time of defining the meaning of the visual stimulus (Chen, & Spence, 2010).In fact, the interpretation of multistable images can be understood as a dynamic function of the brain (Sterzer, Kleinschmidt, & Ress, 2009), which is modulated both by the basic sensory processes (bottom–up) and by references outside the distal stimulus that affect its perception (top-down type processing). That is to say that perceptual reversibility is linked to both bottom-up and top-down processes (Rodríguez, & Castillo, 2018b).As regards, the alternation between one and another interpretation is often involuntary, but it is also voluntary, especially when the observers can draw their attention so as to direct their perception (Aydin, Strang, & Manahilov, 2013; Gijs & van Ee, 2006; Intaité, Koivisto, Rukšėnas, & Revonsuo, 2010). On the other hand, ithas been stated that perceptual reversals suppose the chance of being explained from the theoretical model of high and low levels.

As found in Sterzer & Rees (2009), both for the bottom-up modulation of perception, and for the one based on top-down type processing, there is a correspondence with high-level theory and low level theory, respectively. Thus, there are two ways in which perceptual reversibility arises: firstly, the low level explanatory theory, which suggests that spontaneous alternations take place in the visual cortex. This explanation is the foundation of the so-called bottom-up processing. The explanatory bottom–up model of perceptual bistability is based on the fact that a perceptual reversal occurs through an adaptation of the sensory mechanisms, in such a way that certain perceptual processing supports a specific perceptual configuration until, due to fatigue, the perceptive competitor emerges, finding sustenance in a different mechanism (Kogo, Hermans, Stuer, van Ee, & Wagemans, 2015). Secondly, it has been put forward the high level theory, which states thatthere is a dependency on central information processingin order for perceptual reversibilities to occur (Rodríguez, 2016; Sterzer, & Rees, 2009).

Modulation of visual perception: Apart from the physical characteristics of the images, such as borders, protruding elements, shared contours, among others, the way in which the observer sees them has an impact on their interpretation. Gale & Findlay (1983) stated that there are critical areas within a bistable image that favor the interpretation of each of its possible percepts. Having carried out a detailed analysis of critical lines constituting the bistable image *My girlfriend or my mother-in-law* (see figure 2), they stated that certain lines of the image allow the observer to perceive an image mostly (young woman or old woman) over the other. Performing a graphical synthesis on the original image of Boring (1930), their study referred to four specific areas, each with useful visual information for each possible interpretation. On the basis of this study, points of attentional fixation were inferredto be modulating factors so that it was possible determine which of them favored one interpretation more than the other one (Hsiao et al., 2012). In other words, certain areas and specific constitutive aspects of the bistable image cause a certain perception to be configured. Besides, when an observer fixes his/her gaze on these areas and features, the corresponding interpretation arises (Brower,& van Ee, 2006; García-Pérez, 1989). This type of perceptual modulation is understood as a bottom-up modulating factor because the features of the stimulus happen to fall in the fovea and benefit the identification of one possible percept (Hsiao et al., 2012). This modulating factor implies a sensory processing without providing information that may imply a semantic priming effect (Meng, & Tong, 2004; Rodríguez, & Castillo, 2018a).

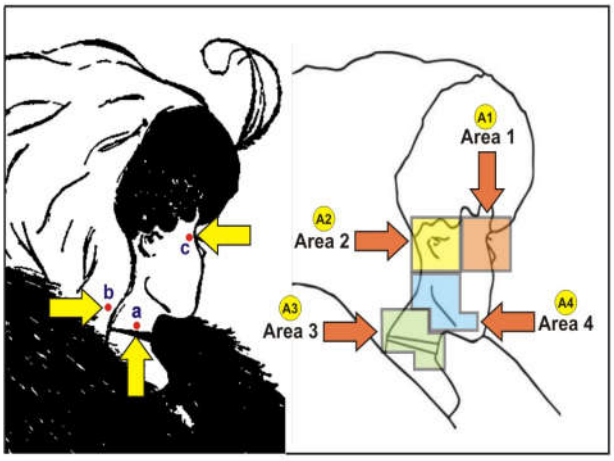


Figure 2. In the image on the left (adapted from Hsiao et al., 2012), there are fixation points, namely, a (neutral point that does not favor any interpretation), b (point that favors the interpretation of the old woman), and, c (point that favors the perception of the young woman). On the right, on the simplified online model used by Gale & Findlay (1983), there are four areas of interest considered to be critical for the perception of the bistable image *My girlfriend or my mother-in-law*. Area A1 contains defining lines of the eye and nose of the young woman; the A2 one predominantly defines both the ear of the young woman and the eye of the old woman; the area A3 refers to the mouth of the old woman; the area A4 contains a line that defines the nose of the old woman and, in turn, an outline of the jaw of the young woman.

When an observer is viewing a bistable image of the In meaning-content reversals type, his/her interpretation (perception) may be influenced if he or she is simultaneously listening to an audio that refers to a particular meaning that matches with the semantic meaning of one of the posible percepts (Yeh, Hsiao, Chen, & Spence, 2011). In this case, there is a crossmodal stimulation model where there is simultaneous stimulation in two different sensory modalities, visual and auditory. The effect of audiovisual semantic congruency is marked by the relation (in terms of semantic load) between the content of the auditory stimulus and what the image depicts (Feist,& Gentner, 2007; Goolkasian,& Woodberry, 2010; Hsiao et al., 2012; Smith, Grabowecky, & Susuki, 2007). Regarding this, it is possible to use a short story (auditory stimulation) so that its contentcan modulate the final interpretation of a bistable image (Hsiao et al., 2012). As expected, the semantic load of the story will have to be related to the semantic content of one of the possible interpretations of the bistable image in order to modulate the perception (Balcetis,& Dale, 2007). As mentioned above, this type of modulation implies a crossmodal semantic congruence, where the phenomenon of modulation arises due to the association between the information provided from the different sensory modalities, so that a unification of the visual and auditory information is assumed by the observer. This presumption that there is a perceptual unit is known as unity assumption (Vatakis,& Spence, 2007).

Likewise, when an observer faces at the same time the reception of information from the environment in different sensory modalities, a crossmodal perception is present (Lalanne,& Lorenceau, 2004). Taking into account the unity assumption phenomenon, it has been stated that an auditory stimulus can operate as a modulator of perception (Munhall, Ten Hove, Brammer, & Paré, 2009; Pressnitzer,& Hupé, 2006). As Hsiao et al. (2012) and Smith et al. (2007) suggested, it is possible to observe the conveyance of semantic congruence from the use of tones of voice. Regardless of the speech understanding, mere tones of voice are able to modulate visual perception if the semantic content of the tone (instead of semantic content of words) is related to the meaning of the visual stimulus observed (Smith et al., 2007). This implies that when an observer is observing a bistable figure such as *My girlfriend or my mother-in-law*, he/she can perceive the young woman whether, for instance, an auditory modulator appears simultaneously, which could be a female youth voice; on the other hand, if the auditory modulator is an old woman's voice, the observer will be able to notice the presence of the old woman by means of semantic congruency (Hsiao et al., 2012; Yeh et al., 2011). As regards, semantic memory emerges as a guiding principle that encourages the perceptual phenomenon of semantic congruency. The process that is involved vindicates visual processing of top-down type (Yeh et al., 2011).

MATERIALS AND METHODS

Forty-five (45) volunteers participated in this study (44.44% = women; 55.55% = men; age range between 18 and 27; average age = 23.17; SD = 6.63). None of them had physical vision or hearing problems. Experimental tests were carried out in a laboratory of psychology, with luminance and temperature control. Each participant had to observe twice, in a fixed 60 Hz. eye-tracker device, the bistable image *My girlfriend or my mother-in-law* (at a viewing distance of 60 cms.) while listening to the voice of a woman by using headphones. The auditory stimuli were presented at approximately 52 dB SPL. The image was displayed in black and white on a gray background (RGB = [127, 127, 127]). A red fixation point placed in an area of the image that did not favor any possible interpretation of it was exposed for 250 ms. before the presentation of the image. The first time each participant observed the image (for 20 seconds), he/she listened to the audio of an old woman speaking. The second time, the image was the same, but the auditory stimulus was a young woman speaking. The auditory monologues of both the old woman and the young woman were taken from French videos (just as Hsiao et al. (2012) did). These audios were randomly presented exchanging their order so as to have a counterbalance and control the presentation order variable.

This study, which was an experimental-intrasubject one, aimed to establish the modulating effect of the audios on the interpretation of the bistable image. It was also wanted to determine whether the tones of voice were able to provide crossmodal contexts for the “old woman” or “young woman” interpretations, taking into account the understanding of the content of the modulating monologues. As has been said, one audio corresponded semantically with a possible interpretation of the image (old woman), while the other did it with the other possible interpretation (young woman). The understanding of the monologues was taken into account. For this purpose, all participants were previously assessed in terms of their level of French language proficiency, using a self-assessment scale from 0 to 9, where zero meant no language proficiency, and nineimplied total mastery. This scale was designed by means of a previous test by which the value 4 was considered to be a reasonable level of listening comprehension according to five qualified bilingual English native speakers who worked as teachers in an institute of foreign languages. As for the way participants expressed their reports, they all clicked on a computer mouse to report the percept they were recognizing during each audiovisual stimulation (for each interpretation a different button was corresponding). The design of the experimental task is illustrated in figure 3:

RESULTS

The duration of each visual perceptsemantically congruent with the audio was the measure that was taken into account to make the corresponding statistical analyzes. In other words, time was taken from the moment of the first report until a report of the incongruous interpretation was expressed. Then, duration concerned to the next lapse taken from the perceptual reversal to the next perceptual reversibility was measured, and so on. In this way, the durations of the congruent and incongruous interpretations with the modulating audios were obtained. The hypothesis that was tackled was stated in the way that the semantically-congruent percept (with the soundtrack) should be dominant for a larger proportion of time that the other one during the test lapse. The results show a statistically significant difference in favor of the modulated perception, that is, that a differential result can be seen in favor of the image relating to the semantically-congruent percept(with the audio)in relation to the semantically-incongruous image (see figure 4): duration of the “old woman” percept congruent with the modulator audio: t(0.574); F = 1,931; p = 0.0425; M = 6679.036; duration of the “young woman” percept congruent with the modulator audio: t(1,861); F = 3.04; p = 0.0465; M = 9348,485).

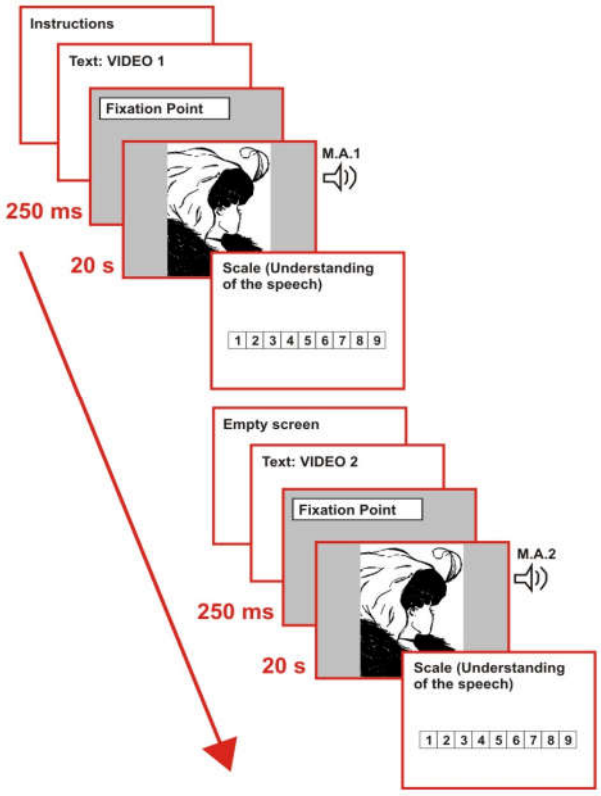
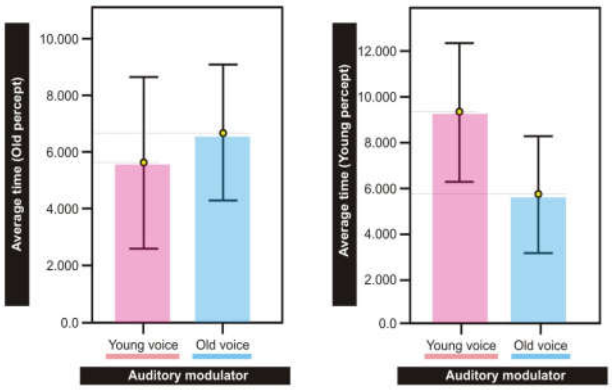
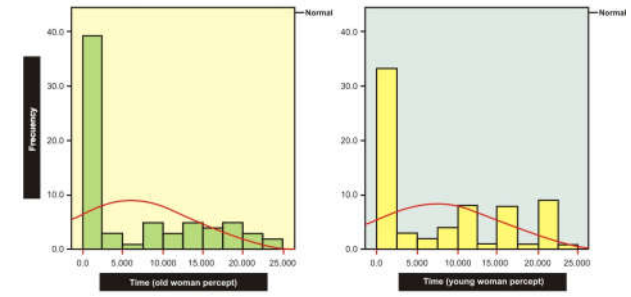


Figure 3. Procedure utilized in the present study. After the instructions were given to the participants, they had to fixate a fixation point (250 ms.) which did not favor any of the possible percepts of the bistable image. Then, the bistable image was presented on the screen while listening to the modulator audio (20s.). This procedure was repeated so as to change the modulating audio. After each participant reported the percepts that were recognized, he/she filled out the scale referring to the understanding of speech. (A.M.1 = Audio modulator 1; A.M.2 = Audio modulator 2). Source: own design.

Figure 4. Comparison between the averages of the duration of each reported interpretation considering the congruent and incongruous modulating audios. Source: own design. A subsequent analysis was performed, because it was established that the distribution of data (durations of congruent interpretations) was not normal (see figure number 5). The statistic test applied was the Kolmogorov-Smirnov test (p = 0.000, for the duration of the "old woman" interpretation congruent with the modulator audio; p = 0.000, for the duration of the "young woman" interpretation congruent with the modulator audio). The average value of the duration of each interpretation was: M = 6149.89 (SD = 7881.14) for "old woman" and, M = 7554.18 (SD = 8349.91), for "young woman".





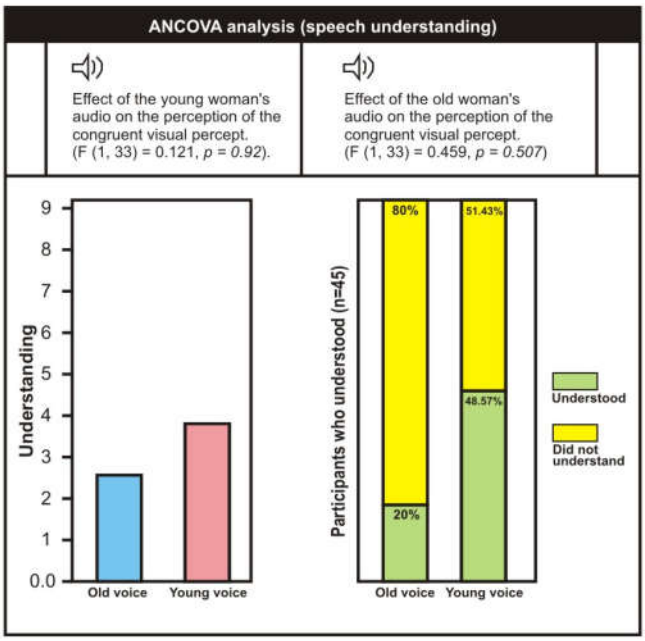


Figure 6. Speech understanding analysis. ANCOVA analysis and percentage distribution.Source: own design

The nonparametric test performed afterwards was the Mann-Whitney U test. In statistical terms, evidence was found to reject the hypothesis of equality in the distribution of the duration of the interpretation "old woman" congruent with its modulating audio (Mdn = 4923.90; U = 701.3; p = 0.0412). On the other hand, the duration of the interpretation of the percept "young woman"relating to its modulating audio was statistically significant in favor of the congruent perception (Mdn = 10602.28; U = 471.3; p = 0.0472). Besides, the covariate "speechunderstanding" was taken into consideration, for which a dichotomization of this variable was applied, so that values less than 3 implied the non-understanding of the content of the speech; higher values were referred to an understanding of the verbal content of each audio.The results show that there is no a strong relation between the "understanding of speech" and the perception of the two possible percepts of the bistable image when they were congruent with the modulating audio (see figure 6). The results presented below show the relationship between the interpretations that were congruent with the modulating audios assumed as "understood" ($\chi 2 = 3.86$): audio of old woman: $\chi 2$ (1, N = 72) = 1,553, p = 0.215 (1,553 <3.86; 0.215> 0.05); young woman audio: $\chi 2$ (1, N = 72) = 1,548, p = 0.216 (1,548 <3.86; 0.216> 0.05); all associations (both audios): $\chi 2$ (1, N = 72) = 2,309, p = 0.131 (2,309 <3.86; 0.131> 0.05). Apart from having made these analyses, a covariance analysis (ANCOVA) was made (see figure 6) to determine if the speechlistening comprehension variable had incidence on the modulating

effect of the audios on the reported visual percepts. Thus, the effect of the young woman audio on the perception of the congruent visual percept did not turn out to be significant (F (1, 33) = 0.121, p = 0.92), after listening comprehension control. The old woman had a similar effect on the perception of the congruent visual percept (F (1, 33) = 0.459, p = 0.507). A further MANCOVA analysis was made so as to confirm that the understanding of the speech did not have an effect on the modulation of the visual percepts. Regarding this, the two dependent variables were both the time of the old woman percept and the time of the young woman percept; auditory modulators were the intrasubject factor, and the speechlistening comprehension was the covariate. The results confirmed that there was no incidence of the listening understanding on the duration of the congruent percepts: Λ = 0.99, F (1, 67) = 0.87, p = 0.769. On the other hand, by taking into consideration this analysis it was also confirmed that auditory stimuli did not have significant impact on congruent visual perception: Λ = 0.977, F (1, 67) = 1.592, p = 0.211.

DISCUSSION

In accordance with the results obtained, it can be concluded that the modulating audios influenced significantly the interpretation of the bistable image. A difference is evidenced in favor of interpretations that are congruent with the tone of voice provided as an auditory modulator. Preliminary studies on semantic modulation of visual perception that supported the visual task used based on a bistable images paradigm, indicated that, although it is possible to use a tone of voice as a modulator (Hsiao et al., 2012, Smith et al., 2012), the effect of semantic congruency can arise, although bottom-up modulating factors could have an impact on the final perception, causing an effect on the reading of semantic congruency phenomenon. Given the fact that the present study did not take into account ocular fixation areas observed during the verbal reports referring to the two possible percepts of the bistable image, there is a chance that the bottom-up modulating factors associated with ocular fixation points could have played a decisive role in the final effect of semantic congruency. However, in the domain of bistable perception, it has been assumed that the observer's own conditions alluding to his/herability for both learning and tracking the visual field, add up to the way in which perceptual reversibilities (perceptual reversals) emerge. Just when an auditory modulator with a semantic load is present as equivalent to the semantic contents of the possible interpretations of the ambiguous image, the observer can do a search of what is associated with the auditory perception. What is more, the observer can retain one of the possible percepts through a control of his/her own eye movements. Besides, the recognized percept can be perceived by means of the visual tracking paths that are implied in the observation. The studyconducted by Gale & Findlay (1983) show a relation between the observed areas of the bistable image used (*My girlfriend or my mother-in-law*) and the reports of interpretations, in favor of one percept or the other one. The study that is outlined here was particularly focused on reviewing the modulating effect of tones of voice, whereby each tone corresponded to each possible interpretation of the bistable image, but without comparing bottom-up modulations, specifically referring to the control of the areas of ocular fixation (although it should be clarified that it was considered a fixation point prior to the exposure of the bistable image: that point was located in specific coordinates of the image that do not favor any interpretation, according to preliminary research findings). What is indicated by means of the findings can presuppose that it is possible modulate communication (in a semantically way) based on a bistable images paradigm. In addition, it can be stated that the tone of voice itself can be enough to make the modulating effect, regardless of the full understanding of what is conveyed by the speech. As the statistical tests indicated, for both interpretations congruent with each audio (young woman and old woman), evidence was found to accept the hypothesis of incidence of the modulating audio on the perceptual configuration times congruent with eachmodulating auditory stimulus. On the other hand, the speechlistening comprehension (speech understanding as a covariate) was not a factor that exerted an influence on the perceptual performance of the participants, that is, on their reports congruent with the auditory modulating stimulus. It could suggest that the semantic congruency was conveyed by the voices themselves rather than the content of the speech. As mentioned in previous paragraphs, the cut-off point to assume that there was an understanding was value 3. This fact implies an issue that could bias the observation of the influence of speech understanding, considering that this scale was designed by taking into account a pre-test in which the value 4 was linked to a reasonable level of listening comprehension according to five qualified bilingual English native speakers of an institute of foreign languages. While this scale was assessed by teachers who had their corresponding academic qualifications, the scale itself has not been validated yet as a cognitive assessment research tool.It would implythe need to conduct future research projects to elucidate on a more precise manner the role played by the modulating auditory content in connection with the modulation effect mediated only by tones of voice.

Conclusion

Auditory stimulation can provide congruent semantic contexts with every possible interpretation of a bistable image. The use of tones of voice as semantic modulators exerts an effect on the decoding of images, beyond the understanding of the semantic content of pronounced words. The present study demonstrates that the semantic information provided by an auditory soundtrack is able to influence the perceptual interpretation of a bistable image.However, theoretically speaking, modulating factors of the bottom-up and top-down typesare involved in bistable perception, which presupposes that future research should tend to study crossmodal sensory stimulation (where each sensory modality is semantically congruent with the other) by considering the bottom-up influence. As regards, further research should also consider all aspects relating to attention factors that are associated with bottom-up modulations. These modulating attention aspects are based mainly on both physical characteristics of visual stimuli and demarcated pathways by the gaze of the observers, that is to say, the points of ocular fixation and salient characteristics of each constitutive element of bistable images.

Funding statement: This study was 100% funded by the University of Bogotá Jorge Tadeo Lozano. The funds used were provided by the Department of Research, Creation and Extension of this university.

REFERENCES

Aydin, S., Strang, N. C. andManahilov, V. 2013. Age-related deficits in attentional control of perceptual rivalry. *Vision Research*(77), 32-40.https://doi.org/10.1016/j.visres.2012.11.010

Baker, D. H., Karapanagiotidis, T., Coggan, D. D., Wailes-Newson, K. andSmallwood, J. 2015. Brain networks underlying bistable perception. *NeuroImage*(119), 229-234.https://doi.org/10.1016/j.neuroimage.2015.06.053

Balcetis, E. andDale, R. 2007. Conceptual set as top-down constraint on visual object identification. *Perception*(36), 581-595.https://doi.org/10.1068/p5678

Bialystok, E. andShapero, D. 2005. Ambiguous benefits: the effect of bilingualism on reversing ambiguous figures. *Developmental Science, 8*(6), 595-604.https://doi.org/10.1111/j.1467-7687.2005.00451.x

Boring, E. 1930. A new ambiguous figure. *American Journal of Psychology*(42), 444-445.http://dx.doi.org/10.2307/1415447

Borisyuk, R., Chik, D. andKazanovich, Y. 2009. Visual perception of ambiguous figures: synchronization based neural models. *Biological Cybernetics*(100), 491-504.DOI 10.1007/s00422-009-0301-1

Brascamp, J., Sterzer, P., Blake, R. andKnapen, T. 2018. Multistable perception and the role of the frontoparietal cortex in perceptual inference. *Annual review of psychology*, *69*. 77-103.DOI:10.1146/annurev-psych-010417-085944

Brouwer, G. J. andvan Ee, R. 2006. Endogenous influences on perceptual bistability depend on exogenous stimulus characteristics. *Vision Research*(46), 3393-3402. http://dx.doi.org/10.1016/j.visres.2006.03.016

Chen, Y. C. andSpence, C. 2010. When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. Cognition, 114(3), 389-404.https://doi.org/10.1016/ j.cognition .2009.10.012

Clément, G. andDemel, M. 2012. Perceptual reversal of bi-stable figures in microgravity and hypergravity during parabolic flight. *Neuroscience letters*(507), 143-146.https://doi.org/10.1016/j.neulet.2011.12.006

Feist, M. andGentner, D. 2007. Spatial language influences memory for spatial scenes. *Memory and cognition*(35), 283-296.DOI: 10.3758/BF03193449

Gale, A. andFindlay, J. 1983. Eye-movement patterns in viewing ambiguous figures. In *Eye movements and psychological functions: international views* (pp. 145-168). Hillsdale NJ: LEA.ISBN: 0-89859-281-X

García-Pérez, M. 1989. Visual inhomogeneity and eye movements in multistable perception. *Perception & Psychophysics*(46), 397-400.https://link.springer.com/content/pdf/10.3758/BF03204995.pdf

García-Pérez, M. A. 1992. Eye movements and perceptual multistability. In E. Chekaluk, & K. Llewellyn (Eds.), *The role of eye movements in perceptual processes* (pp. 73-109). Amsterdam - London - New York - Tokyo: North-Holland.https://doi.org/10.1016/S0166-4115(08)61743-4

Gijs, B. andvan Ee, R. 2006. Endogenous influences on perceptual bistability depend on exogenous stimulus characteristics. *Visual Research*(46), 3393-3402.https://doi.org/10.1016/j.visres.2006.03.016

Goolkasian, P. andWoodberry, C. 2010. Priming effects with ambiguous figures. *Attention, Perception & Psychophysics*(72), 168-178.https://link.springer.com/article/ 10.3758/APP.72.1.168

Grossmann, J. K. andDobbins, A. C. 2006. Competition in bistable vision is attribute-specific. *Vision Research*(46), 285-292.https://doi.org/10.1016/j.visres.2005.06.002

Hsiao, J., Chen, Y., Spence, C. andYeh, S. 2012. Assessing the effects of audiovisual semantic congruency on the perception of a biestable figure. *Consciousness and Cognition*(21), 775-787.https://doi.org/10.1016/j.concog.2012.02.001

Intaité, M., Koivisto, M., Rukšėnas, O. andRevonsuo, A. 2010. Reversal negativity and bistable stimuli: attention, awareness, or something else? *Brain and Cognition*(74), 24-34.https://doi.org/10.1016/j.bandc.2010.06.002

Intaité, M., Kovisto, M. andCastelo-Branco, M. 2014. Event-related potential responses to perceptual reversals are modulated by working memory load. *Neuropsychologia, 56*, 428-438.https://doi.org/10.1016/j.neuropsychologia. 2014.02.016

Intaité, M., Noreika, V., Šoliūnas, A. andFalter, C. M. 2013. Interaction of bottom-up and top-down processes in the perception of ambiguous figures. *Vision Research*(89), 24-31.https://doi.org/10.1016/j.visres.2013.06.011

Kogo, N., Hermans, L., Stuer, D., van Ee, R. andWagemans, J. 2015. Temporal dynamics of different cases of bi-stable figure-ground perception. *Vision Research*(106), 7-19.https://doi.org/10.1016/j.visres.2014.10.029

Kornmeier, J., Hein, C. M. andBach, M. 2009. Multistable perception: when bottom-up and top-down coincide. *Brain and Cognition*(69), 138-147.https://doi.org/10.1016/ j.bandc.2008.06.005

Lalanne, C. andLorenceau, J. 2004. Crossmodal integration for perception and action. *Journal of Physiology - Paris*(98), 265-279.https://doi.org/10.1016/j.jphysparis.2004.06.001

Leopold, D. A. andLogothetis, N. K. 1999. Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences, 3*(7), 254 - 264.https://doi.org/10.1016/S1364-6613(99)01332-7

Long, G. andToppino, T. 1981. Multiple representations of the same reversible figure: implications for cognitive decisional interpretations. *Perception*(10), 231-234.https://doi.org/10.1068/p100231

Long, G. andToppino, T. 2004. Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychol. Bull.*(130), 748–768.DOI: 10.1037/0033-2909.130.5.748

Meng, M. andTong, F. 2004. Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*(4), 539 - 551.doi: 10.1167/4.7.2

Munhall, K., Ten Hove, M., Brammer, M. andParé, M. 2009. Audiovisual integration of speech in a bistable illusion. *Current Biology*(19), 735-739.doi: 10.1016/j.cub .2009.03.019

Okazaki, M., Kaneko, Y., Yumoto, M. andArima, K. 2008. Perceptual change in response to a bistable picture increases neuromagnetic beta-band activities. *Neuroscience Research*(61), 319-328.DOI: 10.1016/j.neures.2008.03.010

Pressnitzer, D. andHupé, J.M. 2006. Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology*(16), 1351-1357.https://doi.org/10.1016/j.cub.2006.05.054

Rock, I., Hall, S. andDavis, J. 1994. Why do ambiguous figures reverse? *Acta Psychologica*(87), 33-59.https://www.ncbi.nlm.nih.gov/pubmed/7985524

Rodríguez, G. 2016. La reconfiguración perceptual de imágenes aplicada al desarrollo del pensamiento divergente en el aula de clase. *Revista Q*, *11*(21), 61-81.DOI: 10.18566/revistaq.v11n21.a05

Rodríguez, G. andCastillo, H. 2018a. Bistable perception: neural bases and usefulness in psychological research. *International Journal of Psychological Research, 11*(2), 63-76.DOI: 10.21500/20112084.3375

Rodríguez, G. andParra, H. 2018b. Tareas de búsqueda visual: modelos, bases neurológicas, utilidad y prospectiva. *Universitas Psychologica*, *17*(1).https ://doi.org/ 10.11144/Javeriana.upsy17-1.tbvm

Schauer, G., Kanai, R. andBrascamp, J. W. 2016. Parietal theta burst TMS: Functional fractionation observed during bistable perception not evident in attention tasks. *Consciousness and Cognition* (40), 105-115.https://doi.org/10.1016/j.concog.2016.01.002

Smith, E., Grabowecky, M. andSusuki, S. 2007. Auditory-visual crossmodal integration in perception of face gender. *Current Biology*(17), 1680-1685.https://doi.org/10.1016/j.cub.2007.08.043

Sterzer, P. andRees, G. 2009. Bistable Perception and Consciousness. Encyclopedia of Consciousness. 93-106.https://doi.org/10.1016/B978-012373873-8.00011-6

Sterzer, P., Kleinschmidt, A. andRees, G. 2009. The neural bases of multistable perception. *Trends in Cognitive Sciences, 13*(7), 310-318.DOI: 10.1016/j.tics.2009.04.006

Vatakis, A. andSpence, C. 2007. Crossmodal binding: Evaluating the "unity assumption" using audiovisual speech stimuli. *Perception & Psychophysics, 69*(5), 744-756.https://doi.org/10.3758/BF03193776

Weilnhammer, V., Ludwig, K., Sterzer, P. andHesselmann, G. 2014. Revisiting the Lissajous figure as a tool to study bistable perception. *Vision Research*(98), 107-112.https://doi.org/10.1016/j.visres.2014.03.013

Yamamoto, S. andYamamoto, M. 2006. Effects of the gravitational vertical on the visual perception of reversible figures. *Neuroscience Research*(55), 218-221.https://doi.org/10.1016/j.neures.2006.02.014

Yeh, S., Hsiao, J., Chen, Y. andSpence, C. 2011. Interplay of multisensory processing, attention, and consciousness as revealed by bistable figures. *i-Perception, 2*(8), 910. https://doi.org/10.1068/ic910

****