

Article

Matching Visual and Acoustic Mirror Forms

Ivana Bianchi ^{1,*}, Roberto Burro ², Roberta Pezzola ³ and Ugo Savardi ²

¹ Department of Humanities, University of Macerata, Macerata 62100, Italy

² Department of Human Sciences, University of Verona, Verona 37129, Italy; roberto.burro@univr.it; ugo.savardi@univr.it

³ Guitar Maestro, Fermo 63900, Italy; r.pezzola@libero.it

* Correspondence: ivana.bianchi@unimc.it; Tel.: +39-0733-258-4320

Academic Editor: Marco Bertamini

Received: 24 January 2017; Accepted: 02 March 2017; Published: 10 March 2017

Abstract: This paper presents a comparative analysis of the ability to recognize three mirror forms in visual and acoustic tasks: inversion (reflection on a horizontal axis), retrograde (reflection on a vertical axis) and retrograde inversion (reflection on both horizontal and vertical axes). Dynamic patterns consisting of five tones in succession in the acoustic condition and five square dots in succession in the visual condition were presented to 180 non-musically expert participants. In a yes/no task, they were asked to ascertain whether a comparison stimulus represented the “target” transformation (i.e., inversion, retrograde or retrograde inversion). Three main results emerged. Firstly, the fact that symmetry pertaining to a vertical axis is the most easily perceived does not only apply to static visual configurations (as found in previous literature) but also applies to dynamic visual configurations and acoustic stimuli where it is in fact even more marked. Secondly, however, differences emerged between the facility with which the three mirror forms were recognized in the acoustic and visual tasks. Thirdly, when the five elements in the stimulus were not of the same duration and therefore a rhythmic structure emerged, performance improved not only in the acoustic but also (even more significantly) in the visual task.

Keywords: visual symmetry; acoustic symmetry; mirror form detection; inversion; retrograde; retrograde inversion; dynamic stimuli

1. Introduction

Over the years psychologists have devoted a lot of attention to the perception of visual symmetry ([1]; for a review, see [2,3]). Extensive experimental research has demonstrated convincingly that mirror symmetry on a vertical axis is easier to identify than mirror symmetry on a horizontal or oblique axis [4–16], even at very short exposure times [4,9,17–21]. There is also evidence that infants are already able at the age of 4 months to distinguish symmetry on a vertical axis from other forms of visual symmetry [22–27]. This has led scholars to theorize that a hardwired mechanism underlies the human perception of mirror reflections, in particular on a vertical axis.

In contrast, only a few studies, carried out some time ago, have investigated the recognition of acoustic symmetry (known as “mirror forms”) in non-expert listeners [28–33]. Furthermore, to the best of our knowledge, there is no previous study on whether it is possible to create a systematic parallel between the ability to recognize the same mirror forms in two sense modalities, in this case visual and acoustic. The experiment described in the present paper addresses this issue and to this end the participants carried out both an acoustic recognition task and a visual recognition task. In the acoustic task, short melodies consisting of a sequence of five notes were presented to the participants, immediately followed by another melody which was one of the three possible mirror transformations of the initial melody (Figure 1a): retrograde (i.e., the notes are played in reversed order) which correspond to a reflection around a vertical axis; inversion (i.e., the initial melody is played with the

intervals inverted, so that for example a rising minor third becomes a falling minor third) which corresponds to a reflection around a horizontal axis or retrograde inversion (i.e., the inverted notes are played in retrograde) which corresponds to a double reflection around both horizontal and vertical axes. Corresponding stimuli were presented in the visual tasks (see Appendix A). In this case, the initial stimulus consisted of a sequence of five dots, appearing one after the other, and the subsequent comparison stimulus consisted of another sequence of five dots which represented one of the three mirror transformations of the initial pattern (Figure 1b).

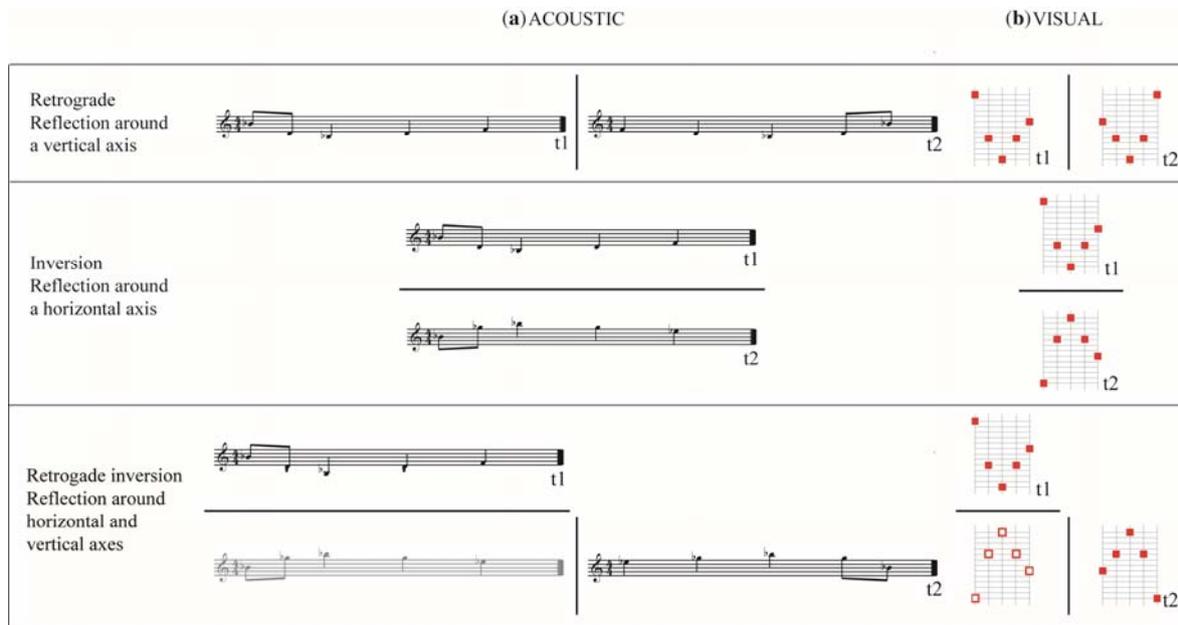


Figure 1. A representation of the three types of transformation (mirror forms) studied in: (a) acoustic task; and (b) visual task. The standard stimulus was presented first (t1), then the comparison stimulus (t2) followed after a 1 s Inter Stimulus Interval (ISI).

Visual and acoustic perception are two different sense modalities with different underlying physiological mechanisms and therefore differences in people's ability to detect the three symmetrical transformations in the two sense modalities are to be expected. However, cross modal correspondence is a well-known phenomenon in psychology [33–36]. It is implied in one of the traditional methods used in psychophysics, i.e., cross modal matching, in which people are asked to match the apparent intensity of various stimuli across two sensory modalities (e.g., adjusting the brightness of a light to match the loudness of a sound). Experimental tasks in which people are asked to identify an object which they have been shown by means of touch have also been used (for instance, in Sperry's famous split brain experiment [37]) as well as a shape-sound matching task which was first described in the "takete-maluma" study [38,39]. Over the years, cross modal correspondence has been consistently found with reference to a variety of sensory modalities but the phenomenon is very well documented for visual and acoustic modalities. The perception of an invariant relationship between audio and visual stimulation emerges at very precocious ages. At 4–5 months infants are able to detect synchrony between visual and acoustic stimuli [40,41], to make associations between sounds and changes in the direction of a movement [42,43], to match faces and voices (e.g., [44,45]) and to recognize similarities between the common rhythmic structures, tempo, and duration of auditory and visual events ([46–48]; see [49], for a review). Within this cross-modal context it is also well known that humans spontaneously describe auditory pitch spatially (e.g., [50–52]). The correspondence between symmetrical acoustic and visual patterns addressed in the present study is based precisely on this spatial correspondence: the vertical distance between the dots in the visual stimuli corresponds to the distance between the notes in the acoustic stimuli which indicates the pitch (see Appendix A).

The three transformations studied in the present paper (retrograde transformation, inversion and retrograde inversion) were widely used in theories of compositional techniques originating in the 20th century which were known as dodecaphony. Dowling [30] was the first to investigate and experiment in order to discover whether listeners with no musical experience are able to recognize these three mirror forms. He used short melodies and a short-term recognition memory task in which the participants were presented with musical sequences consisting of five tones of equal duration followed by the corresponding mirror transformation (inversion, retrograde, and retrograde inversion). After being given information about the three kinds of mirror forms, listeners were asked to judge whether the two melodies presented corresponded to one of them. Blocks of trials were run separately for each of the three forms. Performance was above chance for all three forms, although the task turned out to be difficult. At presentation times similar to those used in the present study (i.e., at five tones presented in 2.5 s), the retrograde mirror form was easier to recognize than the inversion form, and both were more easily recognized than the retrograde inversion form. However, in a faster presentation condition (i.e., five tones per second, which means that the overall duration of the stimulus was 1 s), the inversion form turned out to be the easiest with the retrograde intermediate and the retrograde inversion forms proving the most difficult. There were no differences in the ease with which the three mirror forms were recognized in another study [31] with a different experimental task and longer melodies. Participants first listened to some extracts from the *Wind Quintet* (Op. 26) and the *String Quartet* (No. 4, Op. 37) by Arnold Schoenberg and became familiar with them. Probe stimuli were then presented, consisting either of the prime form or an inversion of the prime form (beginning on the same tone), the retrograde of the prime form (without transposition) and the retrograde inversion of the prime form (the non-transposed retrograde of the inversion). Participants in the experiment were asked to recognize whether the probe stimulus was a mirror form of the *Wind Quintet* or the *String Quartet* (respectively Melodies 1 and 2). Their accuracy was above a chance level performance, but no differences in their performance emerged between the three mirror forms and this was interpreted as an indication that the listeners found it equally easy to recognize the inversion, retrograde and retrograde inversion mirror forms. However, these findings can also be interpreted simply as showing that people have an ability to recognize a certain relatedness between the comparison stimulus and one of the two standard melodies, independently of whether or not they were able to recognize the specific mirror form. Individual differences were found, with accuracy correlated to the level to which the participants had previously studied music. We will go back to the differences in outcomes relating to the experimental conditions in the final discussion.

In one of the conditions studied by Dowling [30], before performing an acoustic task, the participants were presented with an analogous visual recognition task. The standard stimulus (a pattern consisting of five dots which was similar to the notational representations of the melodies in the acoustic experiment) appeared on the left-hand side of the page, while four comparison stimuli were presented on the other side of the page. One of these four configurations was the target mirror form which the participants were required to identify. Only 10 stimuli were used in this preliminary visual phase. Dowling predicted that doing the visual task first would lead to an improvement in performance in the acoustic task since it would provide more information relating to the structure of the configurations to be identified in the subsequent task. Contrary to his prediction, no significant improvements emerged in the condition including the preliminary visual task, and there was in effect only a slight tendency in that direction (i.e., 65% correct responses when there was a preliminary visual task vs. 59% in the acoustic task without the preliminary visual task). In the present study we go beyond the idea of a visual task as a preliminary training task and address a systematic direct comparison between the participants' performance with corresponding sets of visual and acoustic stimuli (using 96 stimuli for the visual condition and 96 stimuli for the acoustic condition). Moreover, our study provided a bi-directional perspective on the possible facilitator effect of performing the task across modalities as the two groups of participants performed the two tasks in a different order (i.e., first the visual task and then the acoustic task or vice versa).

There were four main specific aims in this research. The first was to verify whether naïve subjects (i.e., non-expert listeners) are able to recognize mirror forms in acoustic stimuli. The findings in

previous literature regarding this issue are controversial. Secondly, we aimed to assess whether there was a difference in performance in the two sense modalities, a comparison made possible by the fact that the participants performed both the visual and acoustic task with corresponding stimuli. This represents the main innovation in the study. The third aim, related to the second, was to explore whether there are cross-modal facilitator effects influencing the ease with which mirror forms are recognized in one sense modality after doing the same task in another sense modality (i.e., when the visual task follows the acoustic task or vice-versa). This might provide insights regarding whether the detection of symmetry in the sequential conditions analyzed in this study are specific to each sensory domain or whether cross modal facilitation effects emerge. The fourth aim was to assess the role of Rhythm. There are contrasting results from previous studies regarding the role of the rhythm of a melody (determined by the varying duration of the notes) on the ability to recognize the three mirror forms. According to some (e.g., [32,53]), non-isochronism makes recognizing the mirror forms easier. In other studies [31,54], it has been found that it is more difficult to recognize melodic patterns with non-isochronous stimuli as compared to isochronous stimuli. It has to be noted, however, that in the latter studies, the isochronous and non-isochronous stimuli also differed in other ways related to their melodic structure and therefore there is no clear evidence that isochronism was the critical factor. All of the stimuli used in the present study were made up of five elements (five tones for the acoustic stimuli and five square dots for the visual stimuli) which appeared one after the other. In one condition (isochronism), all of the elements were of the same duration, i.e., 600 ms, while in another condition (non-isochronism), they were of two different durations, i.e., 400 ms and 800 ms. The study not only made it possible to test the effect of this characteristic (isochronism vs. non-isochronism) in the acoustic condition, but also provided first indications of whether it has a similar role in the recognition of visual symmetry in short dynamic sequences.

2. Materials and Methods

2.1. Participants

180 undergraduate students and adults with no musical expertise participated in the study. All had normal or corrected to normal vision and reported normal hearing. All participants performed both the visual (V) and acoustic task (A), and all responded to both the isochronous and non-isochronous stimuli. Order and Target mirror form were the only two variables which were studied between subjects. Ninety of participants performed the visual task before the acoustic task (order VA), while the other 90 were exposed to the two tasks in the opposite order (AV). The inversion mirror form was the target for 30 participants, the retrograde mirror form was the target for 30 other participants, and the retrograde inversion mirror form was the target for the remaining 30 participants. We opted for a single target detection task rather than asking the participants to classify each stimuli according to which of the three mirror forms it represented, due to the fact that in previous studies it had been found that this type of acoustic task is difficult for non-experts and that it is easier for people to recognize one single target.

The study was approved by the Ethics Committee of the University of Verona as the local ethics committee responsible and was conducted in accordance with the Declaration of Helsinki (revised 2008). All participants gave their written informed consent in accordance with the local ethics committee requirements.

2.2. Procedure

The experiment started with a training phase during which participants were familiarized with all of the three mirror forms (inversion, retrograde, retrograde inversion) by means of three visual representations and three acoustic examples such as those shown in Figure 1, one for each of the three mirror forms. The idea behind exposing the participants to all three forms was that this would help them to understand the mechanism by contrasting or differentiating between the various different forms. The participants were allowed to hear and see examples of each of the three mirror forms as many times as they wanted until they felt they were familiar with them. The instructions followed.

They were shown on a computer screen and read out by the experimenter. In the instructions, participants were told that four series of 48 stimuli would be presented (192 stimuli in total) and that each stimulus would consist of a pair of short melodies or a pair of visual patterns which were related to each other in terms of one of the three mirror forms. One of these forms was randomly assigned to each participant as his/her “target” and it was explained to them that the task consisted of recognizing whether the pair of melodies or the pair of visual patterns in each stimulus were related in terms of their target mirror form (yes/no task). An interval of 1 s separated the presentation of the standard stimulus and the comparison stimulus. Participants had 3 s to respond (in a response sheet) before the next pair was presented.

Two of the four series of stimuli involved visual stimuli (V), one series of 48 isochronous stimuli and another series of 48 non-isochronous stimuli. The other two series involved the acoustic version of the same stimuli (A), one series of 48 isochronous stimuli and another series of 48 non-isochronous stimuli. Participants in the AV condition were first exposed to the two acoustic series and then to the two visual series; participants in the VA condition were exposed first to the visual series and then to the acoustic series. All of the participants responded to the four series of stimuli, with short pauses of four minutes between one series and the next. The overall duration of the experiment was 36 min.

A stand alone software (programming language: Actionscript 3 for Adobe AIR runtime environments) was used for the presentation of the visual and acoustic stimuli. A Dell P2210 56 cm (22 In.) screen (Dell, Round Rock, TX, USA) with a resolution of 1680 × 1050 pixels (475 × 300 mm, equivalent to approximately 43.2 × 28.1 degrees of visual angle at the recommended viewing distance 650 mm) and a refresh rate of 60 Hz was used for the experiment. The visual stimuli were made up of 1680 × 1050 pixels and thus filled the entire screen. The visual stimuli consisted of five small square red dots presented in succession, with each dot lasting 600 ms in the isochronous condition, and 400 ms or 800 ms in the non-isochronous condition, and disappearing when the next dot appeared on the screen (i.e., the inter-dot interval was null). Each stimulus started from left to right and was centered with respect to the screen. The acoustic stimuli were created using MakeMusic Finale 2011 (MakeMusic Corporate, Boulder, CO, USA) at a sampling rate of 44,100 Hz–32 bit (and presented with Creative Sound Blaster Audigy FX PCIe 5.1 (Creative Technology, Singapore, Asia) using high quality loudspeakers Audioengine A2+ (Audioengine, Austin, TX, USA). Auditory stimuli were equalized for overall sound pressure using Audition CC 2015 (Adobe Systems Software, San Jose, CA, USA). The stimuli consisted of five tones (Timbre: Guitar; attack: 10 ms; decay: 31 ms). Each tone had an overall duration of 600 ms in the isochronous condition, and 400 ms or 800 ms in the non-isochronous condition. The inter-tone spacing was null.

Both visual and acoustic stimuli lasted a total of 3 s in the isochronous condition (600 ms × 5 elements) and 3.2 s in the non-isochronous condition (800 ms × 3 elements + 400 ms × 2 elements).

2.3. Stimuli

The number of elements in each acoustic stimulus (i.e., five notes), the overall duration of the stimuli (3–3.2 s), and the duration of the interval between the standard stimulus and the comparison stimulus (i.e., 1 s) were defined based on the previous literature on the same subject (in particular [30]) for the purposes of comparison. Due to the fact that a general finding in previous studies was that it is not easy for untrained listeners to recognize acoustic mirror forms, we followed the criterion of choosing conditions—not only in terms of duration but also of the types of intervals and the range of tones (octaves)—which facilitated the recognition of the three mirror forms [29,30,32,53–55]. For instance, all of the acoustic stimuli were in a major or minor key, with each stimulus starting with an initial note (either E3 or E4) followed by other notes chosen from among the first, third and fifth notes of the key and an additional non-chord passage note from the same key (that is, either the second, fourth, sixth or seventh note). Moreover, in all of the stimuli used in the acoustic experiment, the five pitches which made up each stimulus varied within a range of one octave, which according to Pedersen [55] is associated with better performance in same/different tasks using melodic patterns. In the present study, it was also necessary to use a narrow range of tones in order to guarantee that the corresponding visual stimuli would be perceived as unified. In fact, according to the proximity

law of grouping [56], if there was too great a distance between the five elements in each visual stimulus, then there would have been a risk that the participants would perceive two or more separate groups of elements and not one unified pattern. Lastly, we also controlled other features such as the contour and length of the final interval which, according to previous literature, could influence the ability to recognize acoustic melodies in mirror forms. The complete set of variables used in the creation of the stimuli was as follows:

- (i) *Sense modality: acoustic (A), visual (V)*. The stimuli consisted of a short sequence of five tones in the acoustic task and of five square dots appearing in sequence in the visual task. The latter “corresponded” to a visual representation of the acoustic stimuli (see Appendix A).
- (ii) *Mirror Form: inversion (INV), retrograde (RET), retrograde inversion (RETINV)*. Three comparison stimuli were obtained for each standard stimulus by means of applying one of the three mirror form transformations: inversion, retrograde or retrograde inversion.
- (iii) *Rhythm: isochronous (ISO), non-isochronous (N-ISO)*. The stimuli presented in the isochronous condition were the same as those presented in the non-isochronous condition in terms of the shape of the configuration. What varied was the duration of each of the five tones (A) or dots (V) in the configuration. The duration was fixed at 600 ms in the isochronous condition while in the non-isochronous condition, three elements had a duration of 800 ms (half-notes) and two elements had a duration of 400 ms (quarter-notes). In the non-isochronism condition a grouping effect emerged which meant that the participants perceived a kind of “rhythm” in the succession of the tones or dots. The overall duration of the stimulus (3–3.2 s) made the condition comparable to Dowling’s best performance condition (2.5 s).
- (iv) *Contour*. The shape of the contour was determined by the number of inversion points, as in [30,31], i.e., the points where increments in pitch height (for the acoustic stimuli) or spatial height (for the visual stimuli) are followed by a decrement; or, vice versa, decrements in pitch height (for the acoustic stimuli) or spatial height (for the visual stimuli) are followed by an increment. Given that each stimulus was formed of five notes (or dots), the maximum number of points where there could be an inversion was three, giving a maximum of 14 different possible contours (see Table 1). In order to contain the levels of this variable, we selected only eight out of the 14 contours (those which are not in parentheses in Table 1). The eight types of contour used in the experiment thus contained 0, 1, 2 or 3 inversion points. In the statistical analyses, contour was a random effect.

Table 1. Possible types of contour in the five element configurations, defined in terms of increment (+) or decrement (−) of each note/dot as compared to the previous note/dot in terms of pitch height (for the acoustic stimuli, i.e., notes) or spatial height (for the visual stimuli, i.e., dots).

Number of Points of Inversion	Contour
0	++++
0	----
1	+++−
1	(−−+)
1	(+−−)
1	(−+++)
1	(++−)
1	−++
2	(−+−)
2	+−+
2	(++−)
2	−+−
3	+−+
3	−+−

Note: The eight patterns used in the study are those which are not in parentheses.

- (v) *Final interval: long or short.* In [31], the length of the final interval impacted on the participants' ability to recognize the acoustic mirror forms. Despite the fact that in their experiment the stimuli used were different to those used in the present study as they were longer and characterized by wider octave extensions and a greater tonal and temporal complexity, a decision was made to take this variable into account in the creation of the stimuli with the result that there were two versions of the final interval for each of the eight contours: long (L), i.e., between seven and twelve semitones, and short (S), i.e., less than four semitones. This variable was considered as a random effect in our statistical analyses.

In total, 192 pairs of stimuli were presented to each participant in the experiment (2 Sense modalities \times 8 Contours \times 2 Final intervals \times 2 Rhythms \times 3 Mirror forms).

The Target (inversion, retrograde or retrograde inversion mirror forms) and the Order of the two sense modalities (visual-acoustic; acoustic-visual) were studied between subjects. All the other variables—Sense modality (acoustic, visual), Rhythm (isochronism, non-isochronism), Contour (8 levels) and Final Interval (long, short)—varied within subjects.

2.4. Statistics and Data Analysis

Data were analyzed in terms of the Signal Detection Theory (SDT) [57,58]. Responses were classified as either Hit (H), Correct Rejection (CR), False Alarms (FA) or Missing (M). The a priori proportion of “signals” (target stimulus) and noise (non-target stimulus) was 1/3 since all the stimuli presented to participants consisted of three different mirror forms and each participant was requested to target only one of these, meaning that there was one “yes” response and two “no” responses. This was taken into account when calculating the Hit Rate, i.e., the probability of responding “yes” on signal/target trials, and False Alarm Rate, i.e., the probability of responding “yes” on noise/non-target trials. In SDT, the binary answers (i.e., yes/no) of a set of participants are modeled as being influenced by two distinct factors, a perceptual sensitivity component and a response bias. In other words, two people with similar perceptual sensitivity capabilities may have different inclinations to answer “yes” or “no”, or they might modify their inclination (i.e., response bias) in relation to the costs/benefits associated with each of the responses. To calculate sensitivity and response bias, we used two non-parametric measures, i.e., A' and B'' , respectively [59]. We used these non-parametric measures instead of the traditional measure of sensitivity d' (d-prime) since, according to SDT, d' is unaffected by response bias (i.e., it is a pure measure of sensitivity) only if two assumptions are satisfied regarding the decision variable: (1) the signal and noise distributions are both normal, and (2) the signal and noise distributions have the same standard deviation. Since these two assumptions cannot actually be tested in yes/no tasks, non-parametric measures of sensitivity were advisable in this case. Several non-parametric measures of sensitivity and response bias have been proposed (e.g., [60,61]) but the most popular are A' and B'' . These were devised by Pollack and Norman [62]; a complete history is provided by [63]. The formulas for computing A' and B'' are:

If Hit (H) > False Alarm (FA):

$$A' = \frac{1}{2} + \frac{(H - FA) \times (1 + H - FA)}{4 \times H \times (1 - FA)} \quad (1)$$

If (FA) > (H):

$$A' = \frac{1}{2} - \frac{(FA - H) \times (1 + FA - H)}{4 \times FA \times (1 - H)} \quad (2)$$

and:

$$B'' = \frac{(1 - H) \times (1 - FA) - H \times FA}{(1 - H) \times (1 - FA) + H \times FA} \quad (3)$$

Our main focus was to test whether there were different degrees of sensitivity (A') to the three mirror forms in one or both of the sense modalities. When differences in sensitivity were found, we also studied whether there were differences in response bias (i.e., B''). If there was greater sensitivity relating to one of the mirror forms as compared to the others and if this correlated with a response

bias in the same direction, this would mean that the participants found it easier to detect that specific mirror form and that this was associated with a tendency to be more conservative in their responses (i.e., they tended to respond “no” more frequently than “yes”). If greater sensitivity was associated with no difference in response bias, then this could be taken as an indication that the participants found it easier to detect that specific mirror form but without this being associated with a specific response bias. And if greater sensitivity was associated with a response bias in the opposite direction, this could be interpreted as an indication that the participants found it easier to detect that specific mirror form as compared to the other mirror forms, and that this coexisted with a more liberal response bias (i.e., a bias towards “yes”). The A' and B'' values for each individual participant were recalculated for every interaction between the fixed effects that we were interested in studying. A series of Generalized Mixed effect Models (GLMMs) were then performed on these A' and B'' values. Based on an initial exploration of the data (see Section 3.1), we decided not to collapse the data by Subject and always entered this variable in the GLMMs models as a random effect. In all the following GLMMs, Sense Modality, Mirror Forms and Rhythm were always studied as fixed effects, while Contour, Final Interval and Subjects always entered the models as random effects.

When the range of the dependent variable is between 0 and 1 (as in the case of A'), the most suitable type of GLMM is the binomial family and logit link function [64,65]. In order to use the same type of analysis for B'' (whose values ranged between -1 and 1 , with positive numbers representing a conservative bias, i.e., a tendency to answer “no”, while negative numbers representing a liberal bias, i.e., a tendency to answer “yes”, and with 0 representing no bias), a linear transformation was preliminarily applied to the original B'' values in order to rescale them within the interval $0-1$ ($(B'' + 1)/2$).

All analyses were carried out using the statistical software program R 3.3.1, with the “lme4” [66], “car” [67], “lsmeans” [68], and “effects” [69] packages. We performed Mixed Model ANOVA Tables (Type 3 tests) via likelihood ratio tests [70–72] implemented in the “afex” package [73]. Bonferroni corrections were applied to post-hoc comparisons.

3. Results

3.1. Ease of Recognition

The scatter plots of A' values in Figure 2 provide a first overview of the participants' average sensitivity. The degree of sensitivity was assessed according to the ease with which the participant recognized the target mirror forms in the acoustic (diagrams on the left) and visual (diagrams on the right) tasks. The A' values typically range from 0 to 1 , with 0.5 indicating that the signal cannot be distinguished from noise, and 1 which corresponds to perfect performance. Values less than 0.5 may arise from sampling error or indicate that the participant was somewhat confused when responding. We know from previous literature on the recognition of acoustic mirror forms that the task is difficult for non-expert subjects, whereas no former data are available concerning how easy/difficult it is to recognize corresponding dynamic visual configurations. A straightforward comparison between the A' values obtained in our study and previous findings is not possible since in previous literature the analysis was usually based on the frequency of correct responses and did not compare the two correct responses (H, CR) or the two types of error (M, FA) as Signal Detection Theory presupposes. What is usually reported is that the number of correct responses was significantly different from chance (0.50) but that the task turned out to be difficult [30].

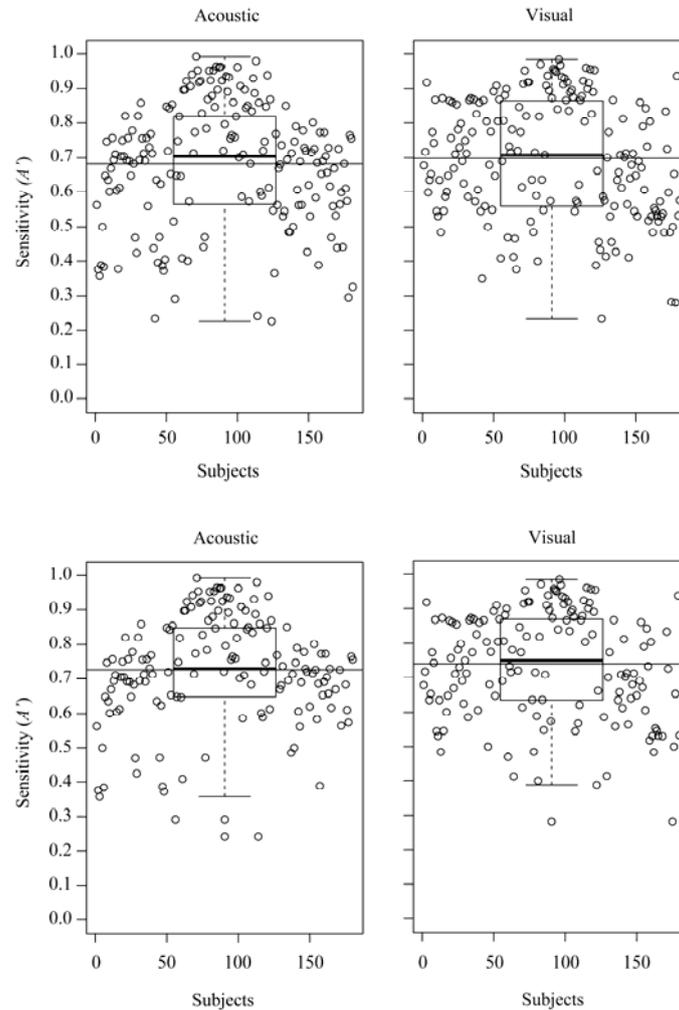


Figure 2. Scatter plots of the participants' average sensitivity (A') and corresponding box plots around the median in the Acoustic task and Visual task (the box is between the upper and lower quartile; whiskers represent the maximum and minimum values excluding outliers; the long thin horizontal line displayed within each plot represents the mean). The two top scatter plots refer to all of the sample data. The two bottom scatter plots refer to the sample data after elimination of participants with low sensitivity (i.e., $A' < 0.6$) in both tasks.

In the top two diagrams in Figure 2, the original sample of participants is displayed. They show that the average sensitivity (represented by the thin horizontal line) was $M_{A'} = 0.682$ (*Standard Deviation* (SD) = 0.177) in the Acoustic task and $M_{A'} = 0.699$ ($SD = 0.172$) in the Visual Task. The medians (displayed in the box plots) are slightly above these values. The scatter plots also reveal that while most of the participants performed overall better than chance ($A' > 0.6$: 71% in the acoustic task and 67% in the visual task), and that some of them performed very well ($A' > 0.8$: 27% in the acoustic task and 36% in the visual task), the task was in effect not very easy. Indeed a percentage of participants that is certainly not negligible had sensitivity values around 0.5 ($0.4 < A' < 0.6$: 20% in the acoustic task and 29% in the visual task), which indicates that they were not able to distinguish signal (i.e., the target mirror forms) from noise (i.e., the other two mirror forms). Moreover, 9% of participants in the acoustic task and 4% in the visual task had sensitivity values between 0.2 and 0.4. This suggests that they confused the target form which they were supposed to be looking for with one of the other two mirror forms. This is not as strange as it may seem given that the three mirror forms all consisted of a similar type of transformation in that they were all mirror reflections.

Participants with an overall sensitivity less than 0.6 in both tasks were eliminated from the dataset (i.e., around 16% of the initial 180 participants) before the subsequent statistical analyses were

done (the new dataset is provided as Supplementary Material: file “database_sym”). Since the aim of the study was to test whether different sensitivity to the “same” mirror forms emerged in the two sense modalities, participants with a sensitivity >0.6 in *one* of the two tasks but not in the other were considered meaningful cases and for this reason were included in the dataset. The new values relating to average sensitivity in the two tasks (represented by the thin horizontal line in the two bottom diagrams in Figure 2) are $M_{A'} = 0.726$ ($SD = 0.154$) in the Acoustic task and $M_{A'} = 0.740$ ($SD = 0.153$) in the Visual Task.

3.2. Ease of Recognition of the Three Mirror Forms in the Two Sense Modalities

In order to study the participants’ sensitivity to the three different Mirror Forms in the two Sense Modalities, we conducted a first GLMM on A' with Sense Modality and Mirror Forms as fixed effects. Both of these main effects turned out to be significant (Mirror Forms: $\text{ChiSq} = 53.489$, $df = 2$, $p < 0.0001$; Sense Modality: $\text{ChiSq} = 91.212$, $df = 1$, $p < 0.0001$), as was their interaction ($\text{ChiSq} = 2202.414$, $df = 2$, $p < 0.0001$). As shown in Figure 3 and as confirmed by post-hoc tests and the descriptive statistics for the raw data (i.e., Mean and SD —as reported in Table 2), the participants were more sensitive to the retrograde mirror form, i.e., a reflection on a vertical axis, in both the visual and acoustic tasks (RET_A vs. INV_A: $\text{EST} = 0.261$, $\text{SE} = 0.027$, $z\text{-ratio} = 9.352$, $p < 0.0001$; RET_A vs. RETINV_A: $\text{EST} = 0.198$, $\text{SE} = 0.029$, $z\text{-ratio} = 6.680$, $p < 0.0001$; RET_V vs. RETINV_V: $\text{EST} = 0.228$, $\text{SE} = 0.029$, $z\text{-ratio} = 7.682$, $p < 0.0001$; RET_V vs. INV_V: $\text{EST} = 0.096$, $\text{SE} = 0.027$, $z\text{-ratio} = 3.451$, $p = 0.008$).

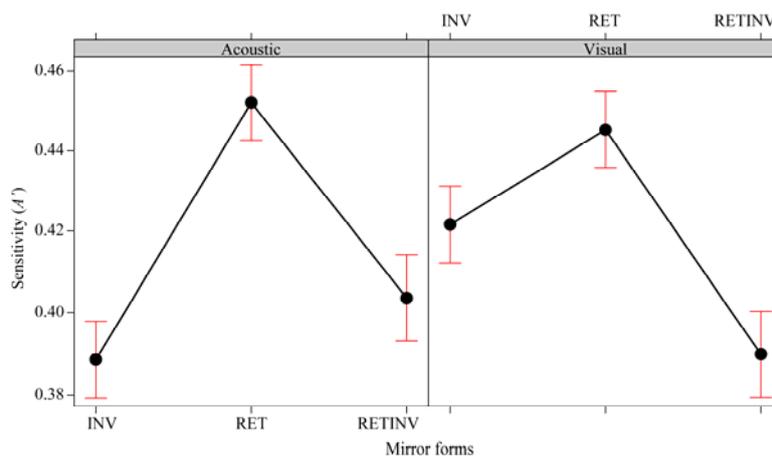


Figure 3. Effect plots of the interaction between Mirror form (Inversion (INV), Retrograde (RET), Retrograde Inversion (RETINV)) and Sense Modality (acoustic, vision) on participants’ sensitivity (A'). Results are given on the logit scale used in the Generalized Mixed effect Models (GLMM) analysis. Error bars represent a 95% confidence interval.

Table 2. Mean sensitivity (A') and mean response bias (B'') relating to the three mirror forms (Inversion (INV), Retrograde (RET), Retrograde Inversion (RETINV)) in the two sense modalities, with corresponding standard deviations (SD).

MIRROR FORM_Sense Modality	A'	B''
RET_acoustic	$M = 0.833$ ($SD = 0.139$)	$M = 0.154$ ($SD = 0.464$)
INV_acoustic	$M = 0.645$ ($SD = 0.148$)	$M = 0.167$ ($SD = 0.288$)
RETINV_acoustic	$M = 0.681$ ($SD = 0.096$)	$M = 0.381$ ($SD = 0.286$)
RET_visual	$M = 0.811$ ($SD = 0.154$)	$M = -0.124$ ($SD = 0.454$)
INV_visual	$M = 0.736$ ($SD = 0.122$)	$M = 0.019$ ($SD = 0.311$)
RETINV_visual	$M = 0.646$ ($SD = 0.136$)	$M = 0.228$ ($SD = 0.356$)

An analyses of the response biases (see Table 2, B'') revealed that, in general, greater sensitivity in the detection of the retrograde transformation was also associated with a more conservative bias

in the other two mirror forms (i.e., a tendency to respond “no”). This was found for all comparison, except one (B'' : RET_V vs. RETINV_V: EST = -0.392, SE = 0.071, z-ratio = -5.457, $p < 0.0001$; RET_V vs. INV_V: EST = -0.210, SE = 0.067, z-ratio = -3.109, $p = 0.028$; RET_A vs. RETINV_A: EST = -0.230, SE = 0.071, z-ratio = -3.209, $p = 0.02$; RET_A vs. INV_A: EST = -0.064, SE = 0.067, z-ratio = -0.954, $p = 1.000$).

A comparison between the sensitivity relating to the visual task and that relating to the acoustic task revealed that participants were more sensitive to the inversion mirror form (i.e., a reflection on a horizontal axis) in the visual task than in the acoustic task (INV_V vs. INV_A: EST = 0.137, SE = 0.003, z-ratio = 44.424, $p < 0.0001$). In terms of the percentage of correct responses (Hit + Correct Rejection), participants correctly responded to 66% of the total number of inversion mirror forms presented in the visual task and to 60% of the total number of inversion mirror forms presented in the acoustic task. As the analysis of B'' revealed (Table 2), this difference in sensitivity was also associated with more conservative responses (i.e., with a bias towards responding “no”) in the acoustic task as compared to the visual task (INV_V vs. INV_A: EST = -0.141, SE = 0.003, z-ratio = -42.347, $p < 0.0001$).

Sensitivity to the other two mirror forms also differed between the two sense modalities, with greater sensitivity to both the retrograde and the retrograde inversion mirror forms in the acoustic task (RET_A vs. RET_V: EST = 0.027, SE = 0.003, z-ratio = 9.558, $p < 0.0001$; RETINV_A vs. RETINV_V: EST = 0.057, SE = 0.003, z-ratio = 16.392, $p < 0.0001$). The percentage of correct responses (Hit + Correct Rejection) for the retrograde mirror form was 76% in the acoustic task and 73% in the visual task and for the retrograde inversion mirror form, the percentage of correct responses was 61% in the acoustic task and 58% in the visual task. The analyses revealed that participants were not only better at detecting these two mirror forms (i.e., retrograde and retrograde inversion) in the acoustic as compared to the visual task, but also that their responses for both mirror forms were more conservative (i.e., with a bias towards responding “no”) in the visual task as compared to the acoustic task (B'' : RET_V vs. RET_A: EST = 0.287, SE = 0.003, z-ratio = 84.057, $p < 0.0001$; RETINV_V vs. RETINV_A: EST = 0.125, SE = 0.003, z-ratio = 35.473; $p < 0.0001$).

3.3. Cross Modality Facilitation Effects

A second GLMM was conducted (with Order, Sense Modality and Mirror Forms as fixed effects) in order to test whether sensitivity increased as a consequence of the order in which the two tasks were performed, i.e., acoustic followed by visual or visual followed by acoustic. No main effect of Order emerged (ChiSq = 0.717; df = 1; $p = 0.396$) which indicates that sensitivity was not improved *in general* by presenting the visual and acoustic tasks in a specific order. However, as the significant interaction between Order, Sense Modality and Mirror form indicated (ChiSq = 1086.764; df = 2; $p < 0.0001$), and as post-hoc tests clarified, facilitation effects across modalities did emerge for one mirror form, i.e., the retrograde mirror form (Figure 4).

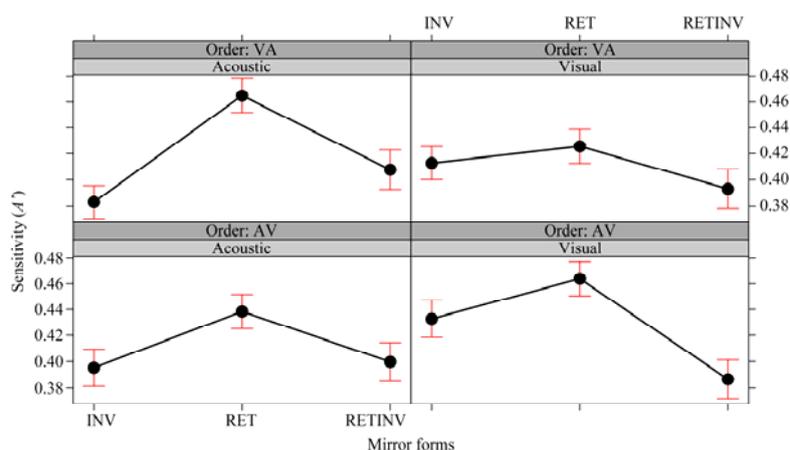


Figure 4. Effect plots of the interaction between Order (VA: Visual followed by Acoustic, AV: Acoustic followed by Visual), Sense Modality (Acoustic, Visual) and Mirror form (Inversion (INV), Retrograde

(RET), Retrograde Inversion (RETINV)) on sensitivity (A'). Results are given on the logit scale used in the Generalized Mixed effect Models (GLMM) analysis. Error bars represent a 95% confidence interval.

Sensitivity to the retrograde mirror form in the visual task was increased after participants had carried out the acoustic task. In other words, the participants' sensitivity when they were asked to detect the retrograde pattern visually (i.e., a reflection around a vertical axis) increased when the visual task followed the acoustic task (AV) as compared to when the visual task was performed first (VA) (RET_V_AV: $M = 0.871$, $SD = 0.111$; RET_V_VA: $M = 0.752$, $SD = 0.168$; RET_V_AV vs. RET_V_VA: EST = 0.155, SE = 0.038, z -ratio = 4.006, $p = 0.004$). No difference due to the order of the visual and acoustic tasks in terms of the participants' sensitivity was found for the other two mirror forms. As the analysis of B'' revealed, the difference in sensitivity found for the retrograde transformation was also associated with a change in the response bias: participants' responses turned out to be more liberal (i.e., with a bias towards "yes") when the visual task was performed after the acoustic task, while there was a slightly more conservative bias (i.e., a bias towards "no") when the visual task was performed first (RET_V_AV: $M = -0.286$, $SD = 0.493$; RET_V_VA: $M = 0.039$, $SD = 0.341$; RET_V_AV vs. RET_V_VA: EST = -0.410, SE = 0.093, z -ratio = -4.394, $p < 0.0001$).

3.4. The Effects of the Isochronism versus Non-Isochronism (Rhythm)

A third GLMM with Sense Modality and Rhythm as fixed effects was conducted to test whether sensitivity to the three mirror forms varied when the elements forming the pattern had different (non-isochronism (N-ISO)) or identical (isochronism (ISO)) durations. A main effect of Rhythm emerged (ChiSq = 329.85, $df = 1$, $p < 0.0001$) and there was a significant interaction with Sense Modality (ChiSq = 168.97, $df = 1$, $p < 0.0001$). As shown in Figure 5 and confirmed by post-hoc tests (see also the mean values reported in Table 3), participants had greater sensitivity in the non-isochronous condition in both sense modalities (ISO_A vs. N-ISO_A : EST = -0.011, SE = 0.002, z -ratio = -4.213, $p < 0.001$; ISO_V vs. N-ISO_V : EST = -0.058, SE = 0.002, z -ratio = -22.161, $p < 0.0001$). However, the effect was stronger in the visual task. In the isochronous condition, participants were equally sensitive in both the visual and acoustic tasks (ISO_A vs. ISO_V : EST = 0.005, SE = 0.002, z -ratio = 2.007, $p = 0.268$), while in the non-isochronous condition, sensitivity was greater in the visual task (N-ISO_A vs. N-ISO_V: EST = -0.041, SE = 0.002, z -ratio = -16.745, $p < 0.0001$). The analysis of B'' revealed that in the acoustic task participants were in general more conservative in their responses than in the visual task (ISO_A vs. ISO_V : EST = 0.178, SE = 0.003, z -ratio = 55.752, $p < 0.0001$; N-ISO_A vs. N-ISO_V : EST = 0.205, SE = 0.003, z -ratio = 72.260, $p < 0.0001$). While in the acoustic task there was no difference in the response bias between the two rhythm conditions (ISO_A vs. N-ISO_A : EST = -0.007, SE = 0.003, z -ratio = -2.463, $p = 0.10$), a difference was found in the visual task where participants were more conservative in the isochronous condition (ISO_V vs. N-ISO_V : EST = 0.066, SE = 0.003, z -ratio = 19.849, $p < 0.0001$)—despite the fact that the mean values indicate a very small response bias.

Table 3. Mean sensitivity (A') and mean response bias (B''), with corresponding standard deviations, relating to the three mirror forms (inversion, retrograde, retrograde inversion) in the two sense modalities and in the isochronism versus non-isochronism conditions.

MIRROR FORM_Sense Modality	A'	B''
ISO_acoustic	$M = 0.726$ ($SD = 0.165$)	$M = 0.213$ ($SD = 0.415$)
N-ISO_acoustic	$M = 0.722$ ($SD = 0.172$)	$M = 0.219$ ($SD = 0.421$)
ISO_visual	$M = 0.722$ ($SD = 0.171$)	$M = 0.010$ ($SD = 0.430$)
N-ISO visual	$M = 0.752$ ($SD = 0.168$)	$M = -0.026$ ($SD = 0.467$)

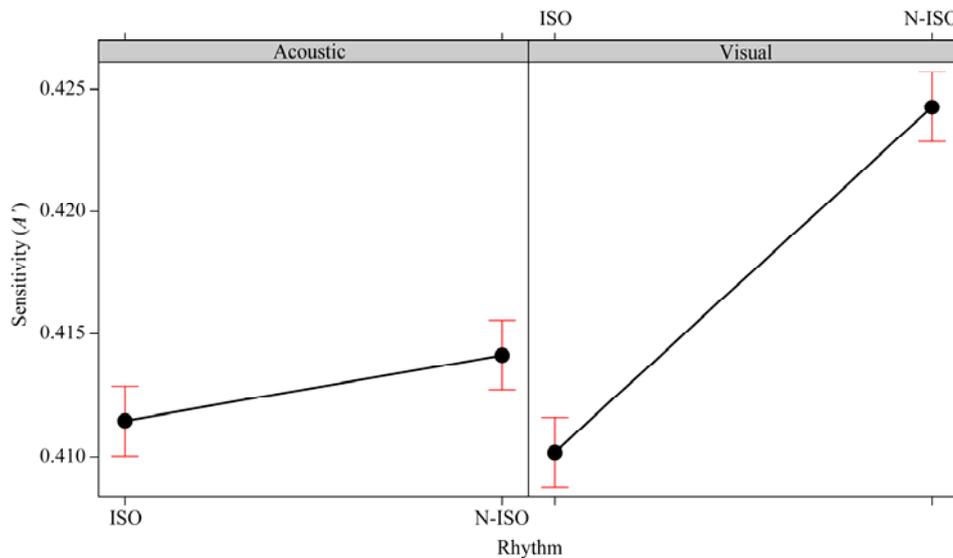


Figure 5. Effect plot of the interaction between Rhythm and Sense Modality (Acoustic, Visual) on sensitivity (A'). Results are given on the logit scale used in the Generalized Mixed effect Models (GLMM) analysis. Error bars represent a 95% confidence interval.

4. Discussion and Conclusions

One of the basic problems affecting the study of visual perception is to explain how the observer recognizes stimuli in his/her environment which undergo continuous and various transformations of shape and size (the constancy problem). In acoustic perception too the issue of whether people are able to perceive two melodies as being “the same” even if all the notes have been changed is an old problem dating back to von Ehrenfels [74] who used this example to establish his definition of Gestalt qualities. We know that melodies may be transposed, speeded up or slowed down (within certain limits) without compromising the listener’s ability to recognize them as “the same” tune. The transformations considered in this paper (i.e., mirror forms) consist of “reflections” of the melodic pattern as a whole which may take the form of an upside down transformation (inversion), a back to front transformation (retrograde) or an upside down and back to front transformation (retrograde inversion). The ability to recognize these mirror forms has been addressed in this paper from a new perspective as both the visual and acoustic recognition of corresponding symmetrical patterns was studied for the purposes of comparison. Moreover, in addition to providing new information regarding the acoustic perception of the three mirror forms to be added to the few (discordant) findings in previous literature, the study also aimed to produce new data on the visual detection of mirror symmetry when sequential patterns and a sequential detection task are considered.

Studies on symmetry detection in dot patterns have confirmed that the visual system is efficient at detecting symmetry, even at very brief presentation times, i.e., stimulus presentations less than 160 ms (e.g., [17,18,21]), during which it is not possible to apply intentional pair-wise comparison strategies based on eye movement or shifts in spatial attention. Processes involving pair-wise comparisons and the intentional matching of corresponding dot-pairs have been hypothesized to support symmetry detection (e.g., [75]), especially when highly demanding tasks are involved, such as the discrimination of perfectly symmetrical versus nearly symmetrical patterns [76]. The type of stimuli and the task used in the study presented in this paper are different from those used in traditional visual symmetry detection tasks (in both static and dynamic conditions) in which the pattern elements are presented simultaneously (for a review and application of various types of methods see [77]). In our experiment, the elements forming each pattern (i.e., the five square dots) appeared in succession over approximately 3 s and not simultaneously. Moreover the two “halves” forming the symmetrical configuration were also presented in succession (i.e., one after the other) divided by a one second Inter Stimulus Interval (ISI). Many researchers have shown that symmetry is easier to detect than a simple repeated pattern when the symmetry is within an object, i.e., there is

a “within-object relation”, while repetition is easier to detect when the matching parts belong to different objects, i.e., there is a ‘between-object relation’ (e.g., [78–80]). In a sense, the conditions which were used in the present study can be considered to be more similar to a between-object matching than a within-object matching and as such it is very different from the automatic rapid symmetry detection conditions reported by most researchers. The task described in this paper likely involved pairwise comparison processes and stimulated matching strategies to compare the two series of elements, which made the task more demanding and difficult.

Three main findings emerged. Firstly, our results provide further evidence (in addition to the findings in previous literature regarding static and dynamic visual configurations—see the papers cited in the introduction) that symmetry around a vertical axis is particularly easy to recognize. Our findings show that this also applies to the case of sequential visual configurations and acoustic stimuli. The preference for this type of symmetry seems to be even more evident in the case of auditory perception than it is for visual perception as indicated by the following findings from the present study: (i) acoustic sensitivity to symmetry around a vertical axis (the retrograde mirror form) was significantly greater than sensitivity to the corresponding visual stimuli (see Figure 3); and (ii) recognizing the retrograde mirror form in the visual task was easier after the acoustic task but not vice versa (Figure 4). This suggests that even though the reflection around a vertical axis (the retrograde mirror form) was one of the two mirror forms which the participants found easier to detect visually, performing the acoustic task first increased their sensitivity to the visual equivalent, whereas performance in the acoustic task remained the same with or without prior visual “training”.

Secondly, despite the fact that a preference for symmetry around a vertical axis (the retrograde mirror form) was common to both sense modalities, the participants did not in general display the same sensitivity in the two types of task. Differences between the two sense modalities emerged for all three mirror forms: reflections around a vertical axis (retrograde) and reflections around a vertical and horizontal axis (retrograde inversion) were more easily recognized in the acoustic task than in the visual task, while reflections around a horizontal axis (inversion) were more easily recognized in the visual task than in the acoustic task. These differences, in addition to the fact that we did not find a general training effect (i.e., a main effect) depending on the order in which the tasks were performed, suggests that the two tasks were relatively independent.

Thirdly, in the present study, performance in the acoustic task improved when the five elements in the stimuli were not of the same duration and a rhythmic structure therefore resulted. This outcome is in agreement with previous evidence on the role of rhythm in tasks involving the recognition of acoustic mirror forms [32,53] but it contrasts with other results indicating that better performances were associated with isochronous melodies (e.g., [31]—although in this case longer melodies and other melodic features also co-varied). The findings in our study demonstrate that non-isochronism facilitates not only the recognition of acoustic mirror forms, but also greatly enhances performance in the recognition of sequential visual patterns (Figure 5). This is an interesting new result concerning the visual perception of symmetry. The different duration of the five elements in the sequence creates grouping effects. We speculate that these local organizations, rather than the “global spatial structure” (which is three seconds long), might provide a useful hint. The spatial relationships perceived between these sub-units (which are simpler and shorter) offer anchor points for a pair-wise comparison strategy to be applied. In our experimental conditions, the overall duration of the event to be judged was around 7 s: the first pattern was presented over approximately three seconds followed by a one second ISI and then the other 3 s long pattern. Iconic memory has a much lower capacity than echoic memory (e.g., [81,82]). Therefore it is reasonable to suppose that the grouping effects provided by rhythm in the non-isochronous conditions offer simplification strategies for the pair-matching comparison which is particularly useful for visual memory.

Our main focus in this paper was to study the participants’ sensitivity to the three mirror forms. However, whenever a significant difference in sensitivity emerged, the response bias was also analyzed. A constant finding in all the analyses of B'' carried out in the paper was that a lower degree of sensitivity was associated with a conservative response bias. This means that, in this study, when

the participants found it difficult to identify the target (i.e., in the case of the mirror forms which they were less sensitive to), they tended to deny the presence of the target pattern.

One of the thought-provoking results emerging from this study regards the difficulty with which inversion mirror forms (i.e., reflections around a horizontal axis) were recognized in the acoustic task. This confirms the evidence found by Dowling [30] in similar presentation conditions (i.e., five tones in 2.5 s). When the results of the visual task were compared to those of the acoustic task, it became clear that it was more difficult to detect the inversion transformation as compared to the retrograde transformation in both sensory modalities but the participants' sensitivity to inversion was significantly greater in the visual task than in the acoustic task. In Dowling's [30] original study, the inverse transformation was more easily detected than the retrograde transformation in the faster condition (i.e., when the five tones were presented in 1 s). On the one hand, this finding indicates the sensitivity of the outcomes to the specific temporal condition considered but this is not dissimilar to the fact that in other dynamic conditions symmetry detection is easier with dynamic as compared to static presentations only at certain optimal frequencies (see [77]). On the other hand, it provides cues for speculation about the differences between the two conditions. In the faster condition studied by Dowling [30], the five tones forming each melody were not simultaneously presented but nearly so: all five tones were presented within 1 s. The overall acoustic event was quite long (the probe stimulus, 1 s, was followed by a 2 s ISI and then by the comparison stimulus, another 1 s), but the global spatial structure of the five tones could be grasped as a whole. Conversely, in the slower presentation condition (two tones/s) which was similar to that used in this study, the overall acoustic structure cannot be grasped unless it is sequential. In the latter condition, i.e., when the global structure is difficult to grasp as a whole, the identity of the elements forming the sequence is probably decisive as it permits the invariance between the two "halves" of the symmetrical configuration to be noticed. This identity is preserved in the retrograde transformation (if not transposed) but not in the inversion transformation. Indeed, in the retrograde transformation, the pitch of the notes in the initial stimuli and in the comparison stimuli remains unchanged: only the order in which they are played is reversed (i.e., the first tone in the comparison stimulus is the last in the initial stimulus; the second tone in the comparison stimulus is the fourth tone in the initial stimulus etc.—see Figure 1a). Conversely, in the inversion mirror form, only the first note remains unchanged in the standard and comparison stimuli. The other four tones change in pitch with respect to the initial tones (see Figure 1a). Evidence of the importance of identity of pitch in the retrograde transformation emerges from the studies showing that when the retrograde transformation is transposed (and therefore the pitch of the notes no longer corresponds to that of the initial melody), people's ability to detect symmetry severely deteriorates [33].

The finding that sensitivity to inversion was lower than sensitivity to the retrograde transformation in both sensory modalities is however in line with overall evidence that upside-down inversions in mirrors (and in general in perception) represent a more severe violation than left to right reversals (e.g., [83–87]).

Taken as a whole, the results of our experiment might contribute to the discussion on the role of global and local factors in the perception of symmetry in sequential versus simultaneous conditions when various sense modalities are considered (for a revision of how the importance of global factors as opposed to local elements have been incorporated into many different theoretical accounts or models of the visual detection of mirror symmetry, see [13]).

The long-standing debate on how many mechanisms underlie the detection of mirror symmetry (e.g., [7,19,20,75,79,88–90]) might also be revived by the addition of new hypotheses regarding cross-modal or modal-specific processes. The issue of the existence of independent or related processing paths underlying the perception of visual and acoustic symmetry is well worth investigating as is the question of whether there are independent or related processing paths underlying the perception of symmetry in dynamic and static visual configurations. This might be further extended to embodied experiences of mirror symmetry related to the proprioception of one's own body structure or of specific body movements as well as to haptic symmetry perception (e.g., [91–94]). We might also ask about the relationship between the proprioceptive and tactile perception

of symmetry and the acoustic processing of symmetry in congenitally blind people whose visual experiences of symmetry are lacking [60,95–100]. When assessed in terms of brain processing, these topics would add to the list of open questions on the subject revised by Bertamini and Makin [101]. We know that the detection of visual and haptic symmetry appears to rely on common brain areas such as the lateral occipital complex in sighted individuals and that in both early blind and sighted (but blindfolded) control subjects, the detection of tactile symmetry is associated with a network implicating frontal and parietal cortical areas (i.e., the medial frontal and superior parietal cortices) [95]. However, in the case of early blind individuals, a significant activation in the retinotopic (i.e., primary visual cortex) and object-selective areas (i.e., lateral occipital and fusiform cortices) was also observed. The activation observed in blind subjects in the early visual cortex during tactile discrimination is in line with previous evidence of cross-modal cortical plasticity in cases of blindness [102]. We might ask how this relates to the cortical activation which occurs during the detection of acoustic symmetry.

Supplementary Materials: The following are available online at www.mdpi.com/link, Table S1: database_sym.

Acknowledgments: This work was supported by funds of the Department of Human Sciences at the University of Verona under Grant (ex60%MIUR). We thank two anonymous reviewers whose observations have helped us to improve the quality of the paper

Author Contributions: Ivana Bianchi, Roberta Pezzola and Ugo Savardi conceived and designed the experiment; Roberto Burro and Roberta Pezzola performed the experiment; Roberto Burro and Ivana Bianchi analyzed the data; Ivana Bianchi, Roberta Pezzola, Roberto Burro and Ugo Savardi wrote the paper and contributed to its revision.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A

The set of acoustic and corresponding visual standard stimuli used in the study are shown in Figure A1).

In figure A1, the section headed A_IS and A_NIS shows the notes played in the isochronous acoustic condition (i.e. A_IS, in which all sounds were of the same duration, in this case 600 ms) and the notes played in the non-isochronous acoustic condition (i.e. A_NIS, in which the sounds were of two different durations with the chromes lasting 400 ms and the quarter-notes lasting 800 ms). For readers who are not familiar with musical notation, please note that the flat and sharp alterations are only signaled at the first note (as by convention) but they hold until the end of the bar, which in our case means until the end of the stimulus.

The corresponding visual stimuli are shown in the section headed V: in the isochronous condition, each dot remained on the screen for 600 ms before disappearing; in the non-isochronous condition, some dots (i.e. those corresponding to the chromes in the acoustic stimuli) lasted 400 ms and the others lasted 800 ms. The grids in the background of the visual stimuli are shown here so that the correspondence between the spatial distance and the difference in pitch in the acoustic stimuli is clear, but these were not displayed in the actual stimuli shown to the participants.

Each of the 32 standard stimuli was presented in association with the three corresponding mirror transformations: retrograde, inversion and retrograde inversion. This means that 96 pairs of standard and comparison stimuli (32 × 3) were presented to the participants in the acoustic task and 96 corresponding pairs (32 × 3) were presented in the visual task.

A_IS and A_NIS		V	A_IS and A_NIS		V

Figure A1. The standard stimuli used in the experiment. The figure shows the notes played in the isochronous acoustic condition and in the non-isochronous acoustic condition (A_IS and A_NIS), and the corresponding visual stimuli (V).

References

1. Van der Helm, P.A. Symmetry perception. In *The Oxford Handbook of Perceptual Organization*; Wagemans, J., Ed.; Oxford University Press: Oxford, UK, 2014; pp. 108–128.
2. Giannouli, V. Visual symmetry perception. *Encephalos* **2013**, *50*, 31–42.
3. Treder, M.S. Behind the Looking-Glass: A Review on Human Symmetry Perception. *Symmetry* **2010**, *2*, 1510–1543.
4. Carmody, D.P.; Nodine, C.F.; Locher, P.J. Global detection of symmetry. *Percept. Mot. Skills* **1977**, *45*, 1267–1273.
5. Corballis, B.; Beale, I. *The Psychology of Left and Right*; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1976.
6. Goldmeier, E. Similarity in visually perceived forms. *Psychol. Issues* **1972**, *8*, 1–136.
7. Herbert, A.M.; Humphrey, G.K. Bilateral symmetry detection: Testing a ‘callosal’ hypothesis. *Perception* **1996**, *25*, 463–480.
8. Wagemans, J. Detection of Visual Symmetries. *Spat. Vis.* **1995**, *9*, 9–32.
9. Locher, P.; Nodine, C. The Perceptual Value of Symmetry. *Comput. Math. Appl.* **1989**, *17*, 475–484.
10. Locher, P.; Wagemans, J. The Effects of Element Type and Spatial Grouping on Symmetry Detection. *Perception* **1990**, *22*, 565–587.
11. Rock, I.; Leaman, R. An experimental analysis of visual symmetry. *Acta Psychol.* **1963**, *21*, 171–183.
12. Royer, F.L. Detection of symmetry. *J. Exp. Psychol. Hum. Percept. Perform.* **1981**, *7*, 1186–1210.
13. Wagemans, J. Characteristics and models of human symmetry detection. *Trends Cogn. Sci.* **1997**, *1*, 346–352.
14. Wenderoth, P. The salience of vertical symmetry. *Perception* **1994**, *23*, 221–236.
15. Wenderoth, P. The effects of dot pattern parameters and constraints on the relative salience of vertical bilateral symmetry. *Vis. Res.* **1996**, *36*, 2311–2320.

16. Zimmer, A.C. Foundations for the measurement of phenomenal symmetry. *Gestalt Theory* **1984**, *6*, 118–157.
17. Barlow, H.B.; Reeves, B.C. The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vis. Res.* **1979**, *19*, 783–793.
18. Julesz, B. Figure and ground perception in briefly presented isodipole textures. In *Perceptual organization*; Kubovy, M., Pomerantz, M., Eds.; Erlbaum: Hillsdale, N.J., USA, 1981; pp. 27–57.
19. Palmer, S.E.; Hemenway, K. Orientation and symmetry: Effects of multiple, rotational, and near symmetries. *J. Exp. Psychol. Hum. Percept. Perform.* **1978**, *4*, 691–692.
20. Tyler, C.W.; Hardage, L.; Miller, R.T. Multiple mechanisms for the detection of mirror symmetry. *Spat. Vis.* **1995**, *9*, 79–100.
21. Wagemans, J.; Van Gool, L.; Dydewalle, G. Detection of Symmetry in Tachistoscopically Presented Dot Patterns: Effects of Multiple Axes and Skewing. *Percept. Psychophys.* **1991**, *50*, 413–427.
22. Baylis, G.C. Visual Parsing and Object-Based Attention: A Developmental Perspective. In *Cognitive Neuroscience of Attention*; Richards, J.E., Ed.; Lawrence Erlbaum Associates, Inc.: Mahwah, NJ, USA, 1998; pp. 251–286.
23. Bornstein, M.H.; Ferdinandsen, K.; Gross, C.G. Perception of symmetry in infancy. *Dev. Psychol.* **1981**, *17*, 82–86.
24. Bornstein, M.H.; Krinsky, S.J. Perception of Symmetry in Infancy—The Salience of Vertical Symmetry and the Perception of Pattern Wholes. *J. Exp. Child Psychol.* **1985**, *39*, 1–19.
25. Fisher, C.B.; Ferdinandsen, K.; Bornstein, M.H. The role of symmetry in infant form discrimination. *Child Dev.* **1981**, *52*, 457–462.
26. Humphrey, G.K.; Humphrey, D.E. The role of structure in infant visual pattern perception. *Can. J. Psychol.* **1989**, *43*, 165–182.
27. Rhodes, G.; Geddes, K.; Jeffery, L.; Dziurawiec, S.; Clark, A. Are average and symmetric faces attractive to infants? Discrimination and looking preferences. *Perception* **2002**, *31*, 315–321.
28. Balch, W.R. The role of symmetry in the good continuation ratings of two-part tonal melodies. *Percept. Psychophys.* **1981**, *29*, 47–55.
29. Dowling, W.J. Recognition of interventions of melodies and melodic contours. *Percept. Psychophys.* **1971**, *9*, 348–349.
30. Dowling, W.J. Recognition of melodic transformations: Inversion, retrograde and retrograde inversion. *Percept. Psychophys.* **1972**, *12*, 417–421.
31. Krumhansl, C.L.; Sandell, G.J.; Sergeant, D.C. The Perception of Tone Hierarchies and Mirror Forms in Twelve-Tone Serial Music. *Music Percept.* **1987**, *5*, 31–77.
32. Tovey, D. *The Forms of Music*; Meridian Books: Cleveland, OH, USA, 1956.
33. Dowling, W.J.; Fujitani, D.S. Contour, interval, and pitch recognition in memory for melodies. *Acoust. Soc. Am.* **1971**, *49*, 524–531.
34. Marks, L. *The Unity of the Senses: Interrelations Among the Modalities*; Academic Press: New York, NY, USA, 1978.
35. Marks, L.E. Cross-modal interactions in speeded classification. In *Handbook of Multisensory Processes*; Calvert, G.A., Spence, C., Stein, B.E., Eds.; MIT Press: Cambridge, MA, USA, 2004; pp. 85–105.
36. Spence, C. Crossmodal correspondences: A tutorial review. *Atten. Percept. Psychophys.* **2011**, *73*, 971–995.
37. Sperry, R.W. Hemisphere disconnection and unity in conscious awareness. *Am. Psychol.* **1968**, *23*, 723–733.
38. Köhler, W. *Gestalt Psychology*; Liveright Publication: New York, NY, USA, 1929.
39. Köhler, W. *Gestalt Psychology: An introduction to New Concepts in Modern Psychology*; Liveright Publication: New York, NY, USA, 1947.
40. Bahrick, L.E. Infants' perception of substance and temporal synchrony in multimodal events. *Infant. Behav. Dev.* **1983**, *6*, 429–451.
41. Bahrick, L.E. Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events. *Child. Dev.* **1988**, *59*, 197–209.
42. Lewkowicz, D.J. Infants' response to temporally based intersensory equivalence: The effect of synchronous sounds on visual preferences for moving stimuli. *Infant. Behav. Dev.* **1992**, *15*, 297–323.
43. Spelke, E.S.; Born, W.S.; Chu, F. Perception of moving, sounding objects by four-month-old infants. *Perception* **1983**, *12*, 719–732.
44. Barrack, L.E.; Netto, D.; Hernandez-Reif, M. Intermodal Perception of Adult and Child Faces and Voices by Infants. *Child. Dev.* **1998**, *69*, 1263–1275.

45. Walker-Andrews, A.S.; Bahrick, L.E.; Raglioni, S.S.; Diaz, I. Infant's bimodal perception of gender. *Ecol. Psychol.* **1991**, *3*, 55–75.
46. Allen, T.W.; Walker, K.; Symonds, L.; Marcell, M. Intrasensory and intersensory perception of temporal sequences during infancy. *Dev. Psychol.* **1977**, *3*, 225–229.
47. Mendelson, M.J.; Ferland, M.B. Auditory-visual transfer in four month old infants. *Child. Dev.* **1982**, *3*, 1022–1027.
48. Pickens, J.N.; Bahrick, L.E. Do infants perceive invariant tempo and rhythm in auditory-visual events? *Infant. Behav. Dev.* **1997**, *20*, 349–357.
49. Bahrick, L.E.; Pickens, J.N. Amodal relations: The basis for intermodal perception and learning. In *The Development of Intersensory Perception: Comparative Perspectives*; Lewkowicz, D., Ed.; Erlbaum: Hillsdale, NJ, USA, 1994; pp. 205–233.
50. Stumpf, K. *Tonpsychologie [Psychology of Tones]*; S. Hirzel: Leipzig, Germany, 1883.
51. Pratt, C.C. The spatial character of high and low tones. *J. Exp. Psychol.* **1930**, *13*, 278–285.
52. Carnevale, M.J.; Harris, L.R. Which direction is up for a high pitch? *Multisens. Res.* **2016**, *29*, 113–132.
53. White, B.W. Recognition of distorted melodies. *Am. J. Psychol.* **1960**, *73*, 100–107.
54. Frances, R. *La Perception de la Musique*, 2nd ed.; Librairie Philosophique J. Vrin: Paris, France, 1972. (In French)
55. Pedersen, P. The Perception of Octave Equivalence in Twelve-Tone Rows. *Psychol. Music* **1975**, *3*, 3–8.
56. Wertheimer, M. Untersuchungen zur Lehre von der Gestalt. II. *Psychol. Forsch.* **1923**, *4*, 301–350. (In German)
57. Green, D.M.; Swets, J.A. *Signal Detection Theory and Psychophysics*; Wiley: New York, NY, USA, 1966.
58. Stanislaw, H.; Todorov, N. Calculation of Signal Detection Theory measures. *Behav. Res. Methods Instrum. Comput.* **1999**, *31*, 137–149.
59. Donaldson, W. Measuring Recognition Memory. *J. Exp. Psychol. Gen.* **1992**, *121*, 275–277.
60. Nelson, T.O. A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychol. Bull.* **1984**, *95*, 109–133.
61. Smith, W.D. Clarification of sensitivity measure A'. *J. Math. Psychol.* **1995**, *39*, 82–89.
62. Pollack, I.; Norman, D.A. A nonparametric analysis of recognition experiments. *Psychon. Sci.* **1964**, *1*, 125–126.
63. Macmillan, N.A.; Creelman, C.D. Triangles in ROC space: History and theory of “nonparametric” measures of sensitivity and bias. *Psychon. Bull. Rev.* **1996**, *3*, 164–170.
64. Papke, L.E.; Wooldridge, J. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *J. Appl. Econ.* **1996**, *11*, 619–632.
65. Baum, C.F. Modeling proportions. *Stata J.* **2008**, *8*, 299–303.
66. Bates, D.; Machler, M.; Bolker, B.M.; Walker, S.C. Fitting linear mixed-effects models using lme4. *arXiv* **2015**, arXiv:1406.5823.
67. Fox, J.; Weisberg, S. *An R Companion to Applied Regression*, 2nd ed.; Sage: Thousand Oaks, CA, USA, 2011.
68. Lenth, R.V. Least-Squares Means: The R Package lsmeans. *J. Stat. Softw.* **2016**, *69*, doi: 10.18637/jss.v069.i01.
69. Fox, J. Effect displays in R for generalised linear models. 2003. Available online: <http://psfaculty.ucdavis.edu/bsjjones/effectdisplays.pdf> (accessed on 7 March 2017).
70. Barr, D.J. Random effects structure for testing interactions in linear mixed-effects models. *Front. Psychol.* **2013**, *4*, doi:10.3389/fpsyg.2013.00328.
71. Barr, D.J.; Levy, R.; Scheepers, C.; Tily, H.J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **2013**, *68*, 255–278.
72. Bates, D.; Kliegl, R.; Vasishth, S.; Baayen, H. Parsimonious Mixed Models. *arXiv* **2015**, arXiv:1506.04967.
73. Singmann, H.; Bolker, B.M.; Westfall, J. AFEX: Analysis of Factorial Experiments. R package version 0.16–1; R Foundation for Statistical Computing, Ed.; Vienna. Available online: <http://CRAN.R-project.org/package=afex> (accessed on 26 January 2017).
74. Von Ehrenfels, C. Über Gestaltqualitäten. *Vierteljahrsschr. Wiss. Philos.* **1890**, *14*, 249–292. (In German)
75. Wagemans, J.; Van Gool, L.; Swinnen, V.; Van Horebeek, J. Higher-order structure in regularity detection. *Vis. Res.* **1993**, *33*, 1067–1088.
76. Huang, L.; Pashler, H. Symmetry detection and visual attention: A “binary-map” hypothesis. *Vis. Res.* **2002**, *42*, 1421–1430.
77. Niimi, R.; Watanabe, K.; Yokosawa, K. The dynamic-stimulus advantage of visual symmetry perception. *Psychol. Res.* **2008**, *72*, 567–579.

78. Baylis, G.C.; Driver, J. Obligatory edge assignment in vision: The role of figure and part segmentation in symmetry detection. *J. Exp. Psychol. Hum. Percept. Perform.* **1995**, *21*, 1323–1342.
79. Bertamini, M.; Friedenber, J.D.; Kubovy, M. Detection of symmetry and perceptual organization: The way a lock-and-key process works. *Acta Psychol.* **1997**, *95*, 119–140.
80. Koning, A.; Wagemans, J. Detection of symmetry and repetition in one and two objects: Structures versus strategies. *Exp. Psychol.* **2009**, *56*, 5–17.
81. Ögmen, H.; Herzog, M.H. A New Conceptualization of Human Visual Sensory-Memory. *Front. Psychol.* **2016**, *7*, 830–845.
82. Winkler, I.; Cowan, N. From sensory to long-term memory: Evidence from auditory memory reactivation studies. *Exp. Psychol.* **2005**, *52*, 3–20.
83. Bianchi, I.; Savardi, U. *The Perception of Contraries*; Aracne: Roma, Italy, 2008.
84. Bianchi, I.; Savardi, U. The relationship perceived between the real body and the mirror image. *Perception* **2008**, *37*, 666–687.
85. Diamond, R.; Carey, S. Why Faces Are and Are Not Special. An Effect of Expertise. *J. Exp. Psychol. Gen.* **1986**, *115*, 107–117.
86. Wright, A.A.; Roberts, W.A. Monkey and human face perception: Inversion effects for human faces but not for monkey faces or scenes. *J. Cogn. Neurosci.* **1996**, *8*, 278–290.
87. Troje, F.N.; Westhoff, C. The Inversion Effect in Biological Motion Perception: Evidence for a “Life Detector”? *Curr. Biol.* **2006**, *16*, 821–824.
88. Dakin, S.C.; Watt, R.J. Detection of bilateral symmetry using spatial filters. *Spat. Vis.* **1994**, *8*, 393–313.
89. Labonte, F.; Shapira, Y.; Cohen, P.; Faubert, J. A Model for Global Symmetry Detection in Dense Images. *Spat. Vis.* **1995**, *9*, 33–55.
90. Jenkins, B. Component processes in the perception of bilaterally symmetric dot textures. *Percept. Psychophys.* **1983**, *34*, 433–440.
91. Ballesteros, S.; Manga, D.; Reales, J.M. Haptic discrimination of bilateral symmetry in 2-dimensional and 3-dimensional unfamiliar displays. *Percept. Psychophys.* **1997**, *59*, 37–50.
92. Ballesteros, S.; Millar, S.; Reales, J.M. Symmetry in haptic and in visual shape perception. *Percept. Psychophys.* **1998**, *60*, 389–404.
93. Ballesteros, S.; Reales, J.M. Visual and haptic discrimination of symmetry in unfamiliar displays extended in the z-axis. *Perception* **2004**, *33*, 315–327.
94. Bauer, C.; Yazzolino, L.; Hirsch, G.; Cattaneo, Z.; Vecchi, T.; Merabet, L.B. Neural correlates associated with superior tactile symmetry perception in the early blind. *Cortex* **2015**, *63*, 104–117.
95. Cattaneo, Z.; Bona, S.; Bauer, C.; Silvanto, J.; Herbert, A.M.; Vecchi, T.; Merabet, L.B. Symmetry Detection in Visual Impairment: Behavioral Evidence and Neural Correlates. *Symmetry* **2014**, *6*, 427–443.
96. Cattaneo, Z.; Fantino, M.; Silvanto, J.; Tinti, C.; Pascual-Leone, A.; Vecchi, T. Symmetry perception in the blind. *Acta Psychol.* **2010**, *134*, 398–402.
97. Cattaneo, Z.; Vecchi, T.; Fantino, M.; Herbert, A.M.; Merabet, L.B. The effect of vertical and horizontal symmetry on memory for tactile patterns in late blind individuals. *Atten. Percept. Psychophys.* **2013**, *75*, 375–382.
98. Driver, J.; Baylis, G.C.; Rafal, R.D. Preserved figure-ground segregation and symmetry perception in visual neglect. *Nature* **1992**, *360*, 73–75.
99. Verma, A.; Van der Haegen, L.; Brysbaert, M. Symmetry detection in typically and atypically speech lateralized individuals: A visual half-field study. *Neuropsychologia* **2013**, *51*, 2611–2619.
100. Wagemans, J. Parallel visual processes in symmetry perception: Normality and pathology. *Doc. Ophthalmol.* **1998**, *95*, 359–370.
101. Bertamini, M.; Makin, A.D.J. Brain Activity in Response to Visual Symmetry. *Symmetry* **2014**, *6*, 975–996.
102. Merabet, L.B.; Thut, G.; Murray, B.; Andrews, J.; Hsiao, S.; Pascual-Leone, A. Feeling by sight or seeing by touch? *Neuron* **2004**, *42*, 173–179.

