COGNITION

ELSEVIER

Cognition 75 (2000) 209–235

www.elsevier.com/locate/cognit

Linguistic and non-linguistic spatial categorization

L. Elizabeth Crawford*, Terry Regier, Janellen Huttenlocher

Department of Psychology, University of Chicago, 5848 S. University Avenue, Chicago, IL 60615, USA

Received 7 July 1999; received in revised form 23 July 1999; accepted 26 January 2000

Abstract

Three experiments examine the relation between linguistic and non-linguistic categorization of spatial relations. We compare linguistic and non-linguistic responses to the same spatial stimuli. Contrary to earlier claims in the literature (Hayward, W. G. & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition*, *55*, 39–84), we find that linguistic and non-linguistic spatial categories do not correspond. Rather, they appear to have an inverse relation such that the *prototypes* of linguistic categories, such as 'above', are *boundaries* in non-linguistic spatial categorization. Evidence for this inverse relation comes from linguistic acceptability judgments and the pattern of bias in participants' reproductions of location. Our findings suggest that while linguistic and non-linguistic spatial organization rely on a common underlying structure, that structure may play different roles in the two organizational systems.

Keywords: Linguistic; Non-linguistic; Spatial categorization

1. Introduction

The relation between language and perception has generally been viewed in one of two opposing manners. On the one hand, the Sapir–Whorf hypothesis (Whorf, 1956) holds that language shapes perception. By this account, the habitual use of a specific language will direct the speaker's attention to those aspects of the non-linguistic world that are captured in the linguistic categories of that language (Hoffman, Lau & Johnson, 1986; Levinson, 1996; Lucy, 1992). On the other hand, the perceptual determinism account suggests that rather than language shaping percep-

* Corresponding author. Fax: + 1-773-7026898.
E-mail address: crawford@ccp.uchicago.edu (L.E. Crawford)

PII: S0010-0277(00)00064-0

tion, it is perception that shapes language. By this account, there are relatively inflexible and universally shared perceptual categories that constrain the range of possible linguistic categories (Kay & McDaniel, 1978; Mandler, 1992).

Although these two general accounts are clearly opposed, they ultimately make a common prediction. For whether it is language that imposes its categories on perception, or perception that imposes its categories on language, the structure of one should be mirrored in the other. Thus, both accounts predict a correspondence between linguistic and non-linguistic categories.

This correspondence has been investigated in the domain of spatial categorization by Hayward and Tarr (1995). They examined linguistic and non-linguistic spatial categories and reported evidence for a correspondence between the two, supporting the shared prediction. Some of their conclusions are consistent with earlier research and some are not. Their conclusions about the structure of basic linguistic spatial categories such as 'above', 'below', 'left', and 'right' are consistent with those of others who have studied linguistic organization of space (Carlson-Radvansky & Logan, 1997; Logan & Sadler, 1996). However, their conclusions about the structure of non-linguistic spatial categories are inconsistent with earlier work that examined non-linguistic spatial categorization (Huttenlocher, Hedges & Duncan, 1991). Because it is not clear how to characterize the structure of non-linguistic spatial categories, it is not clear how they relate to linguistic spatial terms.

In this paper we explore the relation between basic linguistic spatial terms and non-linguistic spatial categories to determine whether they correspond, and thus whether the shared prediction of the Whorfian and perceptual determinism accounts is met. We begin by discussing the work of Hayward and Tarr (1995) and that of Huttenlocher et al. (1991). We suggest that these studies may have reached different conclusions about non-linguistic categorization of space because they examined different aspects of non-linguistic responses. We then argue that the categorization literature supports the measure of Huttenlocher et al. (1991) over that of Hayward and Tarr (1995). We present three experiments, motivated by this discussion, that further examine the relation between linguistic and non-linguistic spatial cognition. To preview our results, we find that non-linguistic spatial categories do *not* map isomorphically onto basic linguistic spatial categories. Instead, we find an inverse relation between linguistic and non-linguistic spatial categories; that is, the *prototypes* of linguistic spatial categories correspond to the *boundaries* between non-linguistic spatial categories. These findings indicate that, while a common underlying structure may influence both linguistic and non-linguistic categorizations of space, this structure plays different roles in these two types of categorization.

1.1. A correspondence between linguistic and non-linguistic categories?

Hayward and Tarr (1995) reported evidence of a partial correspondence between basic spatial prepositions, such as 'above', and non-linguistic spatial categories. Specifically, they suggested that the prototypes of linguistic categories are also prototypes of non-linguistic categories, and they proposed that this correspondence exists because linguistic organization of space builds on non-linguistic perceptual organization.

To determine the relation between linguistic spatial categories and non-linguistic spatial categories, Hayward and Tarr (1995) examined the two kinds of categories separately, obtaining both linguistic and non-linguistic responses to the same spatial stimuli. To examine linguistic organization, they asked participants to describe the spatial relations of a target object relative to a fixed object, or to rate how well different spatial prepositions described the relation between objects. To examine the non-linguistic organization of space, they asked participants to discriminate between different target locations, or to reproduce the locations of targets from memory. They found a striking commonality in responses on these linguistic and non-linguistic tasks. The linguistic judgments showed that linguistic spatial categories such as 'above', 'below', 'left', and 'right' have graded structure with identifiable prototypes at the vertical and horizontal axes of symmetry (see also Carlson-Radvansky & Logan, 1997; Logan & Sadler, 1996). Correspondingly, the non-linguistic responses were more accurate for stimuli located on the vertical or horizontal axis than for other stimulus locations (for related findings, see also Appelle, 1972; Vogels & Orban, 1986). Based on these results, Hayward and Tarr (1995) concluded that the cardinal axes also serve as prototypes of non-linguistic spatial categories. They hypothesized that non-linguistic spatial representations and linguistic spatial prepositions encode space in relation to the same axial prototypes, suggesting an underlying link between non-linguistic spatial representation and spatial language.

This hypothesis appears to be inconsistent with some previous research on non-linguistic spatial categorization. In an earlier study, Huttenlocher et al. (1991) found that non-linguistic spatial categories affected people's memories of the locations of dots presented inside a large circle. Specifically, they found that stimulus locations were reproduced as farther from the cardinal axes, and nearer to diagonal angular locations, than they actually were. This pattern of bias is depicted schematically in Fig. 1. They also found that this bias increased when they introduced an experimental manipulation that interfered with memory for the stimulus dot. They explained this pattern of bias as resulting from categorization. According to Huttenlocher et al. (1991), when people reproduce object locations, memory of a particular location is combined with information about the spatial category in which that object appeared, causing reproductions to be biased toward the prototype, or central value, of the category and away from the category boundaries. In addition, the more inexact the memory of the object's location, the more responses will be affected by category information. According to this explanation, the categories that participants used in the Huttenlocher et al. (1991) study had boundaries on the horizontal and vertical axes and prototypes on the diagonal locations. Thus, while Hayward and Tarr (1995) reported that the non-linguistic category prototypes lie on the cardinal axes, Huttenlocher et al. (1991) reported that the non-linguistic category prototypes lie at diagonal angular locations.

There are a number of differences between the studies of Hayward and Tarr (1995) and Huttenlocher et al. (1991) that may have led to different conclusions about non-linguistic spatial categories, and consequently, about their relation to

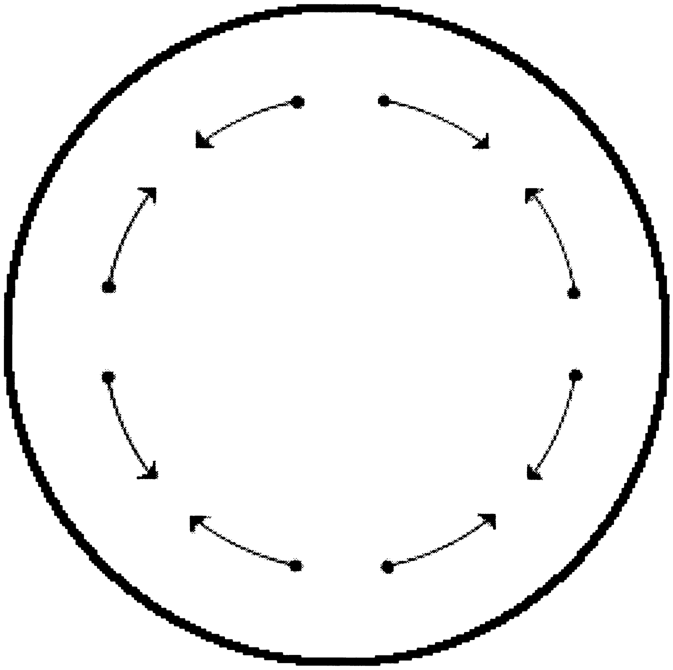

Fig. 1. Schematic representation of the bias pattern found by Huttenlocher et al. (1991). Black dots represent stimulus locations and arrowheads point in the direction of stimulus reproductions.

linguistic spatial terms. The most salient difference lies in the manner in which people's responses are analyzed to determine the location and extent of non-linguistic categories. Hayward and Tarr (1995) use accuracy in reproductions and same–different judgments of location, on the assumption that these will be greatest at category prototypes. Huttenlocher et al. (1991), in contrast, use bias in reproductions, on the assumption that bias will be directed away from category boundaries and toward prototypes. In resolving this tension, we next consider these assumptions in light of previous findings of category effects on responses in other domains.

1.2. Category effects

In research on categorization, one finding that holds across a variety of domains is that objects that are members of the same category appear to be more similar to each other than objects belonging to different categories. This greater within-category similarity can be revealed by a variety of measures, including both discrimination and estimates (or reproductions) of stimuli.

Many studies have examined how categories affect the discriminability of items. In general, the literature on 'categorical perception' shows that the discriminability of two equally distant points is greatest at the category boundary and falls off as distance from the boundary increases (Harnad, 1987). For example, Liberman, Harris, Hoffman and Griffith (1957) presented phonemes that varied continuously on the dimension of Voice Onset Time (VOT) and found that phonemes on either side of the category boundary were more discriminable than phonemes from within the same categories, even though the actual difference in VOT between the phonemes in each pair was the same. Similarly, enhanced discrimination has been found at the boundaries of color categories (Wright, 1947). In addition, studies by Goldstone (1994) have shown that learning to categorize novel objects results in increased discriminability at category boundaries. Thus, in general, discrimination between equally distant items appears to be best at the boundaries between categories.

Studies that have examined reproductions or estimates of items have shown that categorization can result in systematic biases in responses because people use information about categories to estimate object features (Bartlett, 1932; Brewer & Nakamura, 1984). For example, Nelson, Biernat and Manis (1990) asked participants to guess the height (in inches) of men and women who appeared in photographs. They found that women were judged to be shorter than men, even when their heights were comparable, suggesting that although people's estimates of male and female targets reflected the target's actual height to some extent, they also reflected the typical height of the target's gender category. In addition, when the targets were seated, and thus their heights more difficult to guess, judgments reflected gender categories more than when targets were standing. Nelson et al. (1990) concluded that participants were using the gender category to infer the targets' heights, and that this category effect was stronger when people were more uncertain about their estimates.

A parallel category effect is found in people's estimates of the dates of events. Huttenlocher, Hedges and Prohaska (1988) asked university students when certain films had been shown on campus, and found that the responses were systematically biased away from the temporal boundaries separating academic quarters. The effect was stronger for events more distant in time, which were presumably more difficult to remember. This finding is consistent with the general notion that when people are uncertain about a particular item, they use their knowledge of the category to which that item belongs in order to adjust their estimates. This category influence is stronger the more uncertain subjects are.

Experiments in the domain of geographic categories show a similar effect. Friedman and Brown (2000) asked people to estimate the latitude and longitude of cities around the world and found that their estimates were influenced by both the location of the particular city and by that city's geographic region. Cities that were less well known by people showed a stronger category influence than cities that were more well known, again suggesting an increased category effect when uncertainty was greater.

If non-linguistic spatial categories operate like other, non-spatial categories, one would expect them also to be characterized by enhanced discrimination at category boundaries, and by estimation bias away from boundaries and toward prototypes, and by an increased category effect under conditions of higher uncertainty. This may help to resolve the conflicting conclusions of Hayward and Tarr (1995) and Huttenlocher et al. (1991). Hayward and Tarr (1995) found that people were best at discriminating target locations from other, nearby locations when those target locations were presented on the cardinal axes. Discriminability was worse for items farther from the cardinal axes. In light of the above discussion, this finding suggests that category boundaries are located on or near the cardinal axes. In addition, Huttenlocher et al. (1991) found that estimates were biased away from the axial locations, and that this estimation bias increased when participants were less certain of actual stimulus locations. These findings further suggest that the axes are boundaries between non-linguistic spatial categories. Because the axes appear to be the *prototypes* of linguistic spatial categories (Carlson-Radvansky & Logan, 1997; Hayward & Tarr, 1995; Logan & Sadler, 1996), these findings taken as a whole suggest that there is no direct correspondence between linguistic and non-linguistic organizations of space.

Such a conclusion may be premature, however. The Huttenlocher et al. (1991) and Hayward and Tarr (1995) studies used different stimuli and different spatial arrays and therefore cannot be directly compared. Huttenlocher et al. (1991) displayed target stimuli within a circle, while Hayward and Tarr (1995) displayed target stimuli at locations outside a reference object. This may have caused participants in the two studies to use different spatial categories. Thus, it is possible that these studies reached different conclusions because their participants actually organized space differently, and not because of their different dependent measures. In the present experiments, we collect both linguistic judgments and non-linguistic estimation biases for the same spatial stimuli. These stimuli are similar in nature to those used by Hayward and Tarr (1995).

2. Experiment 1

In our first experiment, we present a reference object (a drawing of a television) and a target stimulus (a dot). Participants first rate how well the word ‘above’ describes the relation of the dot to the television, and then reproduce the dot’s location. We expect to find that the region for which ‘above’ is most acceptable will be on the vertical axis in the region over the TV. Our main interest, however, is what will happen in the non-linguistic estimates of location. If participants encode the stimulus location in terms of a non-linguistic spatial category that corresponds to the linguistic category, then its prototype should be at the vertical axis. In this case, we predict that estimates will be biased *toward* that axis, as in Fig. 2a. This pattern of results would constitute strong support for the hypothesis that linguistic spatial terms correspond to a non-linguistic organization of space. However, if the axes are category boundaries rather than prototypes, we predict that estimates will be biased *away* from the vertical axis, as in Fig. 2b. This pattern of responses would suggest that non-linguistic spatial categories do not map isomorphically onto linguistic categories for this array.

On each trial, the target is first rated according to how good an example of the ‘above’ relation it is, and then is reproduced from memory. Thus, it is possible that by having participants first make a linguistic response to each stimulus, we may be contaminating responses on the non-linguistic task. That is, this procedure may

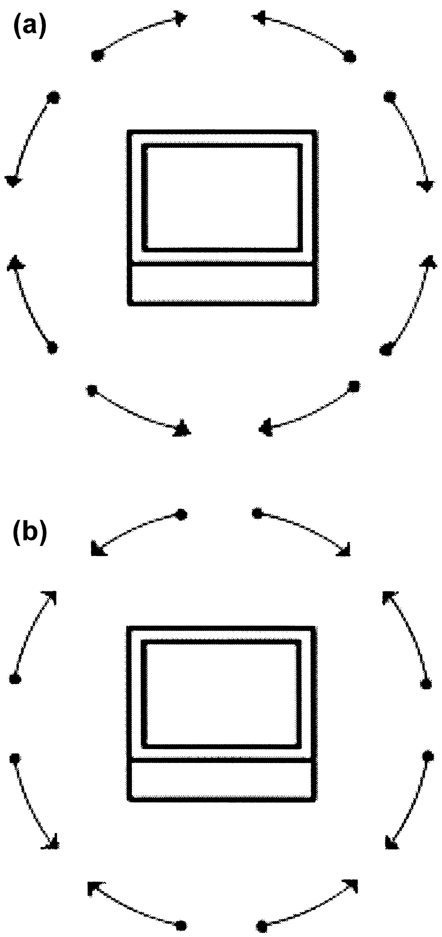

Fig. 2. (a) Schematic bias pattern indicating category prototypes on the horizontal and vertical axes. (b) Schematic bias pattern indicating category boundaries on the horizontal and vertical axes.

encourage participants to use categories that correspond to linguistic terms although they might not have done so if the location estimation task had been presented alone. If the linguistic task has any influence on the location estimates, it should increase the likelihood that reproductions reflect the linguistic organization. Thus, if we find that memories are biased *toward* the axes, we cannot tell whether this reflects participants' non-linguistic spatial categories, or the forced application of their linguistic categories. However, if we find that memories are biased *away* from the axes, we can conclude that the non-linguistic spatial categories that participants use do not map onto basic linguistic categories, and that the non-linguistic organization persists in the face of possible priming of linguistic categories.

2.1. Method

2.1.1. Participants

Ten University of Chicago students participated in this experiment. All were recruited by flyers posted around the campus and were paid $2 for their time.

2.1.2. Materials

The experiment was run on a Macintosh IIci computer with a large (17 × 11 inches) monitor. The stimuli were pairs of objects presented on a computer screen: a (2 × 2 cm) square object resembling a TV and a black dot. The TV was always presented in the center of the computer screen and the dot was presented in 28 evenly spaced angular positions (approximately every 13°) and two predetermined radial distances (10 and 5 cm) from the center of the TV, as shown in Fig. 3. Thus, the complete distribution of dot locations formed two concentric circles of evenly spaced dots surrounding the TV.

A small piece of paper was attached to the upper left corner of the computer monitor. On it the words ‘Very Good’, ‘Good’, ‘Fair’, ‘Poor’, and ‘Very Poor’ were printed to remind participants of their verbal response options throughout the experiment.

2.1.3. Procedure

Participants faced the computer screen with the computer mouse situated in front of them. On each trial, a picture of the TV appeared in the center of the screen and remained there while one of the stimulus dots was presented. The dot appeared for 1 s and then was removed. Subjects were then asked to state ‘if the word ‘above’ is a *very good*, *good*, *fair*, *poor*, or *very poor* description of the dot’s location relative to the TV’. Each verbal response was coded on a scale ranging from 1 to 5 (1 for Very Good and 5 for Very Poor) by the experimenter who then entered the coded response using the computer keyboard. After this verbal response, the dot re-appeared at the center of the screen (inside the TV), and participants used the computer mouse to move the dot to the location where they remembered seeing it and then clicked on that location.

Each subject was given two randomly located practice trials that were immediately followed by the 56 experimental trials. The experimental trials presented the dot once in each of the possible positions and the order of presentation was randomly generated for each subject.

2.2. Results

Two responses were deleted from the data set because the mouse button was accidentally pressed before the mouse was moved.

The mean linguistic acceptability for each location was computed by taking the average acceptability rating at each of the 56 stimulus locations. These means are presented in Fig. 4. As expected, we found that the two stimuli that were located on the upper portion of the vertical axis were the best rated examples of ‘above’ and

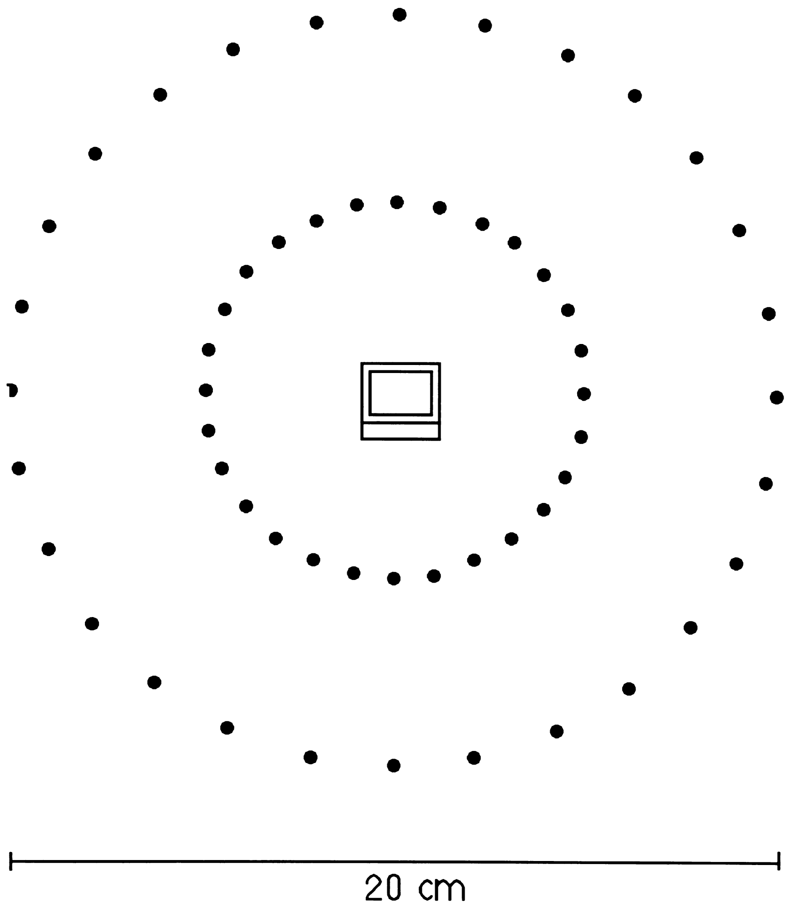

Fig. 3. Distribution of stimulus locations used in Experiment 1.

received the lowest average rating. A planned contrast showed that the rating for the stimuli presented on the top part of the vertical axis was significantly lower than for the other angular locations ($F(1, 530) = 419$, $P < 0.001$). Fig. 4 also shows that the acceptability of the other locations falls off gradually as the positions become further from the vertical axis. These results are consistent with previous research that has examined the acceptability of spatial terms (Carlson-Radvansky & Logan, 1997; Hayward & Tarr, 1995; Logan & Sadler, 1996).

Like Hayward and Tarr (1995), we also measured the accuracy of stimulus reproductions in both the horizontal and vertical dimensions. We measured the horizontal error of each response as the horizontal distance between the response location and the given stimulus location and found that on average, horizontal error was less for points on the vertical axis than for points at other stimulus locations ($F(1, 530) = 25$, $P < 0.001$). Likewise, we measured the vertical error of each response as the verti-

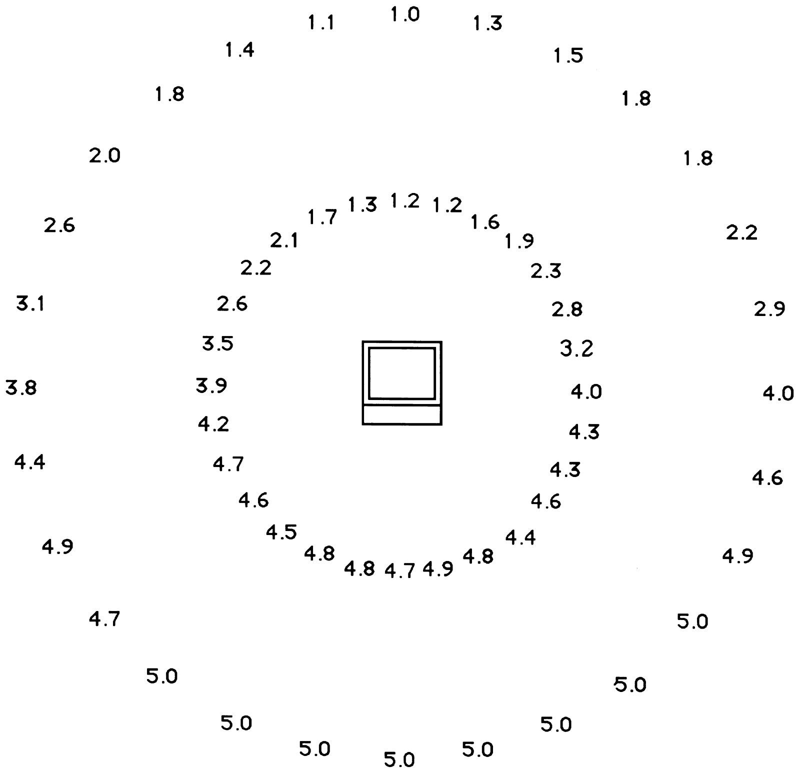

Fig. 4. Mean acceptability of *above* for each stimulus location in Experiment 1.

cal distance between the response location and the given stimulus location and found that the average vertical error was smaller for points on the horizontal axis than for points at the other angular locations ($F(1,530) = 22$, $P < 0.001$).

Like Huttenlocher et al. (1991), we also examined the *direction* of errors, or bias, in reproductions. We computed a mean response to each stimulus by averaging the X-coordinate and Y-coordinate of responses to each stimulus. These mean response locations are presented in Fig. 5. Boxes are drawn around the mean response to display the variability of responses. The width of each box represents two horizontal standard deviations and the height of each box represents two vertical standard deviations. In addition, an arrow is drawn from each stimulus location to the mean response to that stimulus to illustrate the direction of errors in responses. Fig. 5 displays the decreased horizontal variability of responses to stimuli presented on the vertical axis and decreased vertical variability of responses to stimuli presented on the horizontal axis.

In addition, Fig. 5 shows that reproductions tended to be biased away from the vertical and horizontal axes and toward the diagonals. We tested whether the angular locations of the mean responses were significantly closer to the diagonals than the

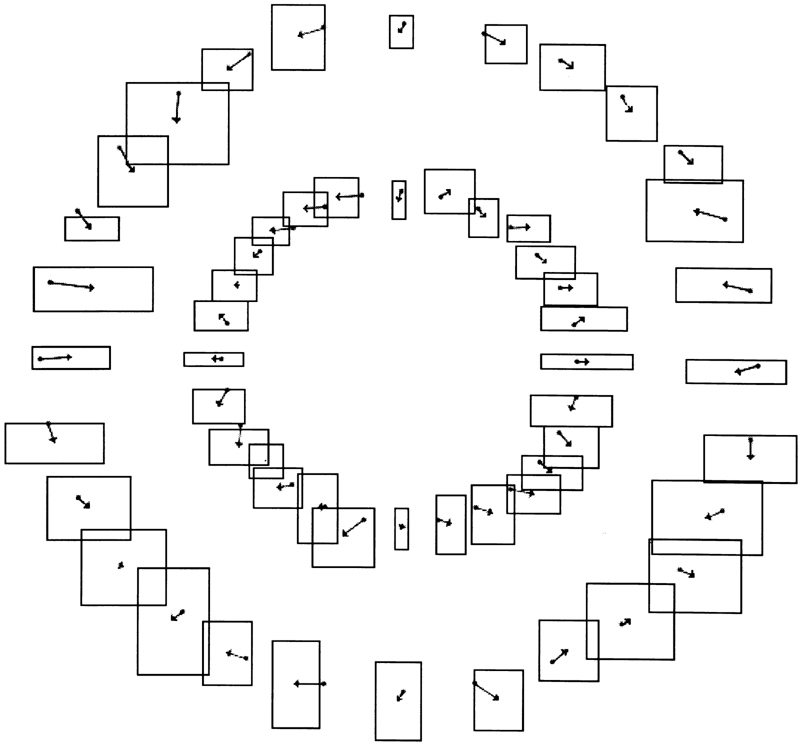

Fig. 5. Experiment 1, bias in reproductions. Black dots represent the actual stimulus location, and arrowheads represent the mean response location. Boxes represent two horizontal standard deviations and two vertical standard deviations.

angular locations of the presented stimuli. Within each quadrant, the angular difference between the stimulus and response was coded as positive if the response was shifted toward the diagonal (and away from the axes) and negative if the response was shifted away from the diagonal (and toward the axes). If there is no systematic bias in responses toward the diagonals, these differences will not differ from 0 on average. If responses are significantly biased toward the diagonals and away from the major axes, these differences will be positive on average and if responses tend to be biased toward the axes, these differences will be negative on average. The average of these differences was $+2.30°$ which is significantly different from 0 ($t(47) = 6.87$, $P < 0.001$). This analysis indicates that on average, responses are biased toward diagonal angular locations and away from the horizontal and vertical axes.

2.3. Discussion

As expected, the analysis of linguistic acceptability judgments showed that the locations for which the word 'above' is most acceptable are on the vertical axis directly above the reference object. These results are consistent with Hayward and Tarr (1995) and other studies that have examined the extension of the term 'above'. We conclude from this analysis that the vertical axis is the prototype of the linguistic category 'above'.

The analysis of reproduction accuracy shows that reproductions of stimuli that were presented on the vertical axis are more accurate in the horizontal dimension than reproductions of other locations, which is also consistent with the findings of Hayward and Tarr (1995). Likewise, reproductions of stimuli that were presented on the horizontal axis are more accurate in the vertical dimension than reproductions of other locations. These findings suggest that space is not uniformly coded and that the array's cardinal axes are coded more precisely than other locations. Although the axes appear to have a privileged status in our non-linguistic task, this finding alone does not indicate what role they play in non-linguistic spatial categorization.

We also examined the *direction* of errors in non-linguistic reproductions. Based on our review of the categorization literature, we predicted that, if participants used non-linguistic categories with prototypes on the cardinal axes, their reproductions would be biased *toward* those axes. Alternatively, if participants used non-linguistic categories for which the cardinal axes were category boundaries, then responses would be biased *away from* those axes. We found that reproductions that were near the vertical axis were consistently biased away from it. This pattern of bias indicates that the vertical axis is serving as a category boundary rather than a prototype on the non-linguistic task. Although the term 'above' denotes a region in space, that same region appears to be cut into two categories for the non-linguistic estimation task. This occurs despite the fact that participants were required to describe each dot in terms of the word 'above' before reproducing it, which might have been expected to cause them to categorize the space in terms of the linguistic category.

The bias pattern for Experiment 1 suggests that, for this array, there is no single, non-linguistic spatial category that corresponds to the linguistic term 'above'. Instead, it appears that the prototypical 'above' – the area along the vertical axis – is a boundary between two non-linguistic spatial categories, much like the quadrant categories of Huttenlocher et al. (1991).

Because we were focusing on the category 'above', our main interest was on the vertical axis in the top half of the spatial array. However, our results may be informative about other spatial regions as well. Reproductions of locations below the reference object were also systematically biased away from the vertical axis. In addition, vertical errors are smallest for locations on the horizontal axis, and location estimates near this axis tended to be biased away from it. However, this pattern of bias is not as obvious here as it is at the vertical axis and suggests that participants divide the space into four quadrant categories with boundaries on the cardinal axes.

3. Experiment 2

In our first experiment, we concluded that participants did not use a non-linguistic spatial category that corresponds to the basic linguistic category ‘above’. We showed that stimuli within the region for which ‘above’ was an acceptable description were estimated to be further away from the vertical axes than they actually were, and we suggested therefore that the vertical axis is a category boundary rather than a category prototype. However, one shortcoming of the first experiment is that the reference object was rather small, and as a result, very few stimulus dots were actually in the region directly above the reference object. Thus, our examination of the ‘above’ category was necessarily of low resolution.

This low resolution could be problematic because it might cause us to miss narrow spatial categories. Specifically, people might code location using a non-linguistic spatial category that corresponds to the area directly above the reference object, but does not extend beyond the width of the reference object. This category would share its prototype with the linguistic category ‘above’ and therefore would correspond, at least at the prototypes, with the linguistic category ‘above’. If that were the case, we would expect to find that locations presented above the reference object, but within its horizontal edges, would be reproduced as closer to the vertical axis than they actually were. However, our first experiment presented only two stimuli in the region directly above the reference object, and as Fig. 6a illustrates, these two stimuli were actually on the vertical axis. Because there were no stimuli in the regions above the reference object and between the vertical axis and the left and right edges of the object, the experiment would not have been able to detect such a narrow category. In our second experiment, we investigate this possibility by widening the reference object and including more stimuli in the region directly above it, as shown in Fig. 6b. If people use a non-linguistic spatial category centered on the vertical axis, but bounded by the edges of the reference object, this stimulus arrangement will allow us to detect it by examining bias in reproductions of stimuli within that region. If the stimuli directly above the reference object are remembered as being closer to the vertical axes than they actually were, this will suggest that there is a non-linguistic category that shares its prototype with the ‘above’ category. However, if the stimuli above the reference object are remembered as being farther from the vertical axes than they actually were, this will suggest that the vertical axis is a category boundary in non-linguistic spatial representation.

3.1. Method

3.1.1. Participants

Ten University of Chicago students participated in this experiment. They were recruited and compensated as in Experiment 1. None had participated in the earlier experiment.

3.1.2. Materials

Stimuli used in this experiment were similar to those used in Experiment 1. However, the reference object was a (8 × 1.5 cm) rectangle that was four times the width of the TV used in Experiment 1. The 56 stimulus dots were arranged in two

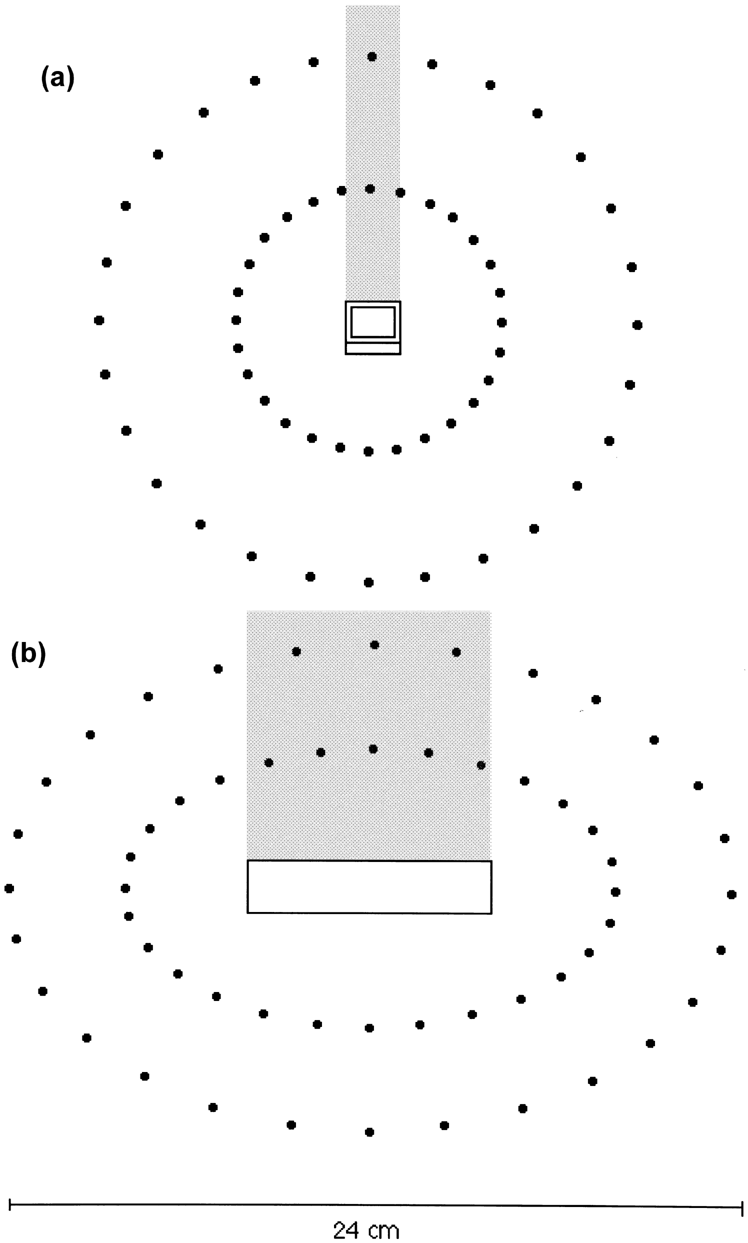

Fig. 6. (a) The shaded region represents hypothetical category bounded by the width of the reference object used in Experiment 1. (b) The shaded region represents hypothetical category bounded by the width of the reference object used in Experiment 2.

concentric ellipses, as shown in Fig. 6b. Again, the verbal responses ‘Very Good’, ‘Good’, ‘Fair’, ‘Poor’, and ‘Very Poor’ were posted on a piece of paper attached to the outside edge of the computer monitor.

3.1.3. Procedure

The procedure was identical to that of Experiment 1, except the reference object was now referred to as a box. Participants judged each dot as to how good an example of the ‘above’ relation it was, relative to the box, and then reproduced the dot’s location using the computer mouse.

3.2. Results

As in Experiment 1, judgments of Very Good, Good, Fair, Poor, or Very Poor were converted to a five-value ordinal scale. The average judgments are shown for each stimulus location in Fig. 7. As in Experiment 1, the top half of the vertical axis was the most acceptable position for ‘above’ ($F(1,532) = 280$, $P < 0.001$) and acceptability decreased gradually for locations farther from upright vertical.

Horizontal and vertical errors were analyzed as in Experiment 1. As in Experiment 1, horizontal error was less for points on the top half of the vertical axis than for points at other stimulus locations ($F(1,532) = 12.79$, $P < 0.001$). In addition, vertical error was less for points on the horizontal axis than for other stimulus locations ($F(1,532) = 13.862$, $P < 0.001$).

Fig. 8 presents the mean response locations for each stimulus. Boxes are drawn

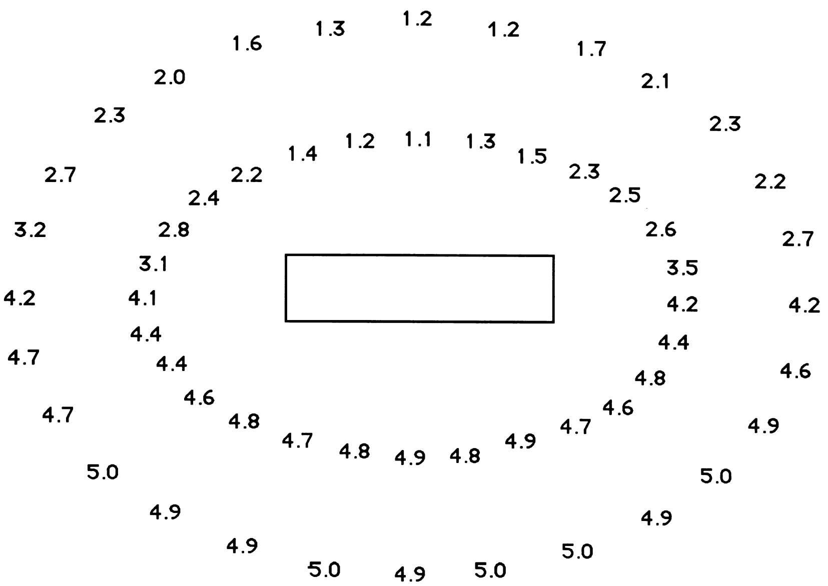

Fig. 7. Mean acceptability of *above* for each stimulus location in Experiment 2.

around each mean to display the variability of responses. The width of each box represents two horizontal standard deviations and the height represents two vertical standard deviations. An arrow is drawn from the actual stimulus location to the mean response location to display the direction of errors. As in Experiment 1, the horizontal variability of responses to stimuli presented on the vertical axis appears to be lower than the horizontal variability of responses to other stimuli. Similarly, the vertical variability of responses to stimuli presented on the horizontal axis appears to be lower than the that of responses to other stimuli.

Stimuli in the region directly above the reference object, within the left and right edges of the object, appear to be reproduced as further away from upright vertical than they actually were. Mean responses to the six stimuli in that region (excluding those directly on the vertical axis) were repelled an average of 3.03° away from the vertical axis, which is significantly different from 0 ($t(5) = 3.79$, $P < 0.02$). As in Experiment 1, we also tested whether reproductions were biased away from the cardinal axes across the entire stimulus array. Within each quadrant, the angular difference between the stimulus and mean response was coded as positive if the response was shifted away from the axes and negative if the response was shifted toward the axes. The average of these differences was +1.11°, a difference that is significantly different from 0 ($t(47) = 2.853$, $P < 0.01$). This result indicates that responses tend to be biased toward diagonal angular locations and away from the horizontal and vertical axes.

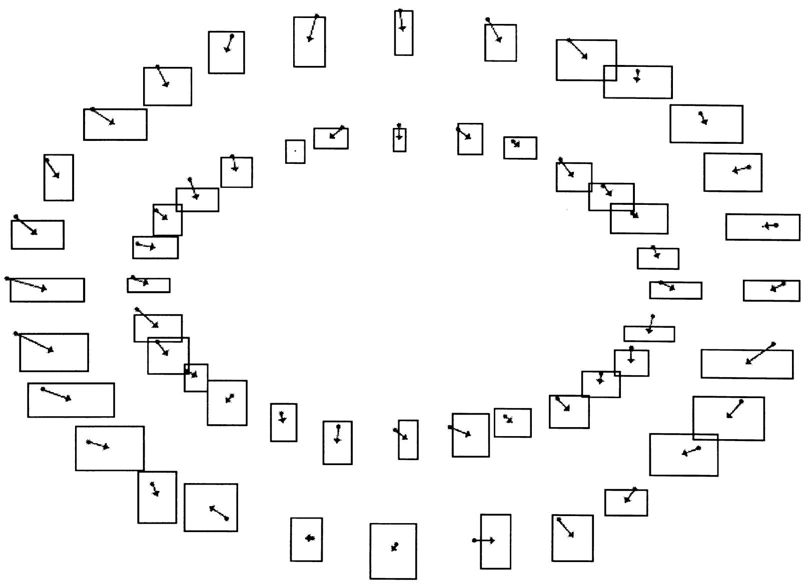

Fig. 8. Experiment 2, bias in reproductions. Black dots represent the actual stimulus location, and arrowheads represent the mean response location. Boxes represent two horizontal standard deviations and two vertical standard deviations.

3.3. Discussion

In our second experiment, we attempted to use a stimulus array that would be more likely to reveal a non-linguistic spatial category that corresponds to the basic linguistic category ‘above’. We widened the reference object so that more target locations fell in the region directly above the object. We found that, linguistically, these stimuli were judged to be good examples of ‘above’. We also found that, non-linguistically, estimates were significantly biased *away* from the vertical axes, although this pattern is less visually obvious here than in Experiment 1. These results, like those from Experiment 1, suggest that participants were using a set of non-linguistic spatial categories that were divided by the vertical axis rather than centered on it. Thus, they again suggest a non-correspondence between linguistic and non-linguistic spatial categories.

4. Experiment 3

The findings from the first two experiments suggest that there is an inverse relation between basic linguistic spatial terms, such as ‘above’, and non-linguistic spatial categories. That is, the *prototypes* of linguistic spatial categories correspond to the *boundaries* between non-linguistic spatial categories. This conclusion is supported by two experiments that investigated both linguistic and non-linguistic spatial category structure.

However, one possible concern with this conclusion is that the non-linguistic bias we measured may actually spring from linguistic sources. For example, when participants see a dot above and to the right of the reference object, they may mentally encode this perception in a representation that is linguistic in nature: ‘above and to the right’. This internal linguistic representation may then be used to produce both a linguistic response (the utterance ‘above and to the right’), and a non-linguistic response (the reproduction of location). Should this be the case, the bias patterns we found in our first two experiments would actually reveal an underlying linguistic representation of location, rather than non-linguistic spatial categories.

Is there such an underlying linguistic mental representation giving rise to both linguistic and non-linguistic observed responses? We can determine this by examining the kinds of errors that participants make. Consider, for example, our dot presented above and to the right of the box. If a participant said ‘left’ instead of ‘right’, this would be a linguistic error. In contrast, if a participant placed the dot on the left side of the box rather than the right, this would be a non-linguistic error. Critically, there should be at least as many non-linguistic errors as linguistic errors, if both linguistic and non-linguistic responses stem from the same underlying linguistic representation. Why? Because only the non-linguistic responses involve a change in representational format, from the linguistic to the non-linguistic. We assume here that the processing route from the underlying linguistic representation to a non-linguistic response will be noisier than the route from the linguistic representation to a linguistic response.[1] Thus, if there is such a shared linguistic representation, we would expect a greater number of errors in non-linguistic responses than in linguistic responses. Alternatively, if linguistic errors are found to be more frequent than non-linguistic errors, this would be evidence against this view.

To investigate this possibility, we ran two additional experiments in which we present the same stimuli as before, and ask participants to describe the stimulus locations linguistically. In Experiment 3a, we use the same stimuli as in Experiment 2 and we compare the number of linguistic errors to the number of comparable reproduction errors from Experiment 2. In Experiment 3b, we use the same stimuli as in Experiment 1 and we ask participants to both linguistically describe and non-linguistically reproduce the location of each stimulus. We then compare errors in linguistic and non-linguistic responses to the same stimuli.

4.1. Experiment 3a

4.1.1. Method

4.1.1.1. Participants Ten students from the University of Chicago participated in the study on a volunteer basis. All were native English speakers and none had participated in the previous two experiments.

4.1.1.2. Materials Stimuli used in this experiment were identical to those used in Experiment 2, except the piece of paper that had been taped to the monitor for that study was removed.

4.1.1.3. Procedure Participants were seated in front of a large computer monitor that showed the rectangular box and a randomly selected dot. The words 'The dot is' appeared in the upper left corner of the screen and participants were asked to type in the rest of a sentence that described the location of the dot relative to the box. They were instructed to mention the box in the sentence, to keep the sentence reasonably brief, to not use metric units such as inches, and to avoid polar coordinate and clock face descriptions. The dot and box remained on the screen while participants typed their descriptions.

We coded each linguistic response for left–right and above–below errors. A response was coded as a left–right error if the dot was actually to the right of the center of the reference object and was described as being to the left, or if it was actually to the left and was described as being to the right. A response was coded as an above–below error if it was actually above the center of the reference object and was described as below or if it was actually below the center and described as above. Responses were not coded as errors if participants corrected themselves within their response (e.g. '…very high to the right of the – to the left, I mean, of the box'); thus, our measure of linguistic error rate is a conservative one. We also re-examined the non-linguistic data from Experiment 2 for left–right and above–below errors. A non-

[1] Otherwise, the specifically linguistic character of this hypothesized representation would be called into question.

linguistic response was coded as a left–right error if the target dot was actually presented to the left of the center of the reference object but was reproduced to the right of center, or vice versa. Similarly, a non-linguistic response was coded as an above–below error if the dot was actually presented above the center of the reference object and was reproduced below the center, or vice versa. For both the linguistic and non-linguistic data, stimuli that appeared on the vertical axis were not coded for left–right error and stimuli that appeared on the horizontal axis were not coded for above–below error.

4.1.2. Results and discussion

There were no cases of above–below errors in linguistic descriptions collected in this experiment or in non-linguistic reproductions from Experiment 2. There were also no left–right errors in the non-linguistic reproductions from Experiment 2. In contrast, there were 25 left–right errors in the linguistic descriptions from Experiment 3a, 4.8% of the total number of coded descriptions. Participants made between 0 and 8 linguistic left–right errors, and 6 (out of 10) participants made at least one left–right error. The average number of left–right errors per participant was 2.5, which is significantly different from 0 ($t(9) = 2.548$, $P < 0.03$), the non-linguistic error rate.

The results from Experiment 3a suggest that linguistic left–right errors occur at a higher rate than non-linguistic left–right errors. This finding argues against the notion that linguistic coding mediates non-linguistic reproduction of location. Thus, it offers support for the claim that location reproductions reflect genuine non-linguistic spatial categories, rather than covert linguistic coding.

The comparison between Experiments 3a and 2 is only suggestive because they were run at different times and the participants may have differed in a number of ways that were not controlled for. In addition, because we did not collect non-linguistic reproductions in this experiment, we were unable to determine if the left–right linguistic errors that we found could have caused participants to make left–right non-linguistic errors for the same stimuli. A more powerful test of the hypothesis that linguistic coding mediates reproduction of location would be to have the same participants produce linguistic descriptions and non-linguistic reproductions for the same stimulus locations.

4.2. Experiment 3b

In a follow-up to Experiment 3a, we ran an experiment in which participants first described (linguistically) and then reproduced (non-linguistically) the location of each stimulus within the same trial. We reasoned that if linguistic coding is used to reproduce locations, we would find at least as many non-linguistic errors as linguistic errors, and more specifically, that non-linguistic and linguistic errors would occur within the same trial. Such a finding would suggest that the reproduction task we used in Experiments 1 and 2 might actually be linguistic in nature, and that location reproductions are not an appropriate measure of non-linguistic categories.

4.2.1. Method

4.2.1.1. Participants Twenty students from the University of Chicago were paid $4 for their participation. All were native English speakers and none had participated in the previous experiments.

4.2.1.2. Materials Stimuli used in this experiment were identical to those used in Experiment 1, except the piece of paper that had been taped to the monitor for that study was removed.

4.2.1.3. Procedure Participants were seated in front of a large computer monitor that showed the rectangular box and a randomly selected dot. The words 'The dot is' appeared in the upper left corner of the screen and subjects were asked to type in the rest of a sentence that described the location of the dot relative to the box. As in Experiment 3a, participants were instructed to mention the box in the sentence, to keep the sentence reasonably brief, to not use metric units such as inches, and to avoid polar coordinate and clock face descriptions. The dot and box remained on the screen while participants typed their descriptions. Once the sentence was complete, participants pressed the return key and the sentence and the dot were removed. A response dot then appeared inside the box and participants used the mouse to move this dot to the location where the target dot was just shown and clicked the mouse to register their estimate of the stimulus location.

We coded linguistic and non-linguistic responses for left–right and above–below errors in the same manner as in Experiment 3a. For both the linguistic and non-linguistic responses, stimuli that appeared on the vertical axis were not coded for left–right error and stimuli that appeared on the horizontal axis were not coded for above–below error.

4.2.2. Results

Fig. 9 presents the total number of errors of each type (linguistic versus non-linguistic) and direction (above–below versus left–right). As the figure indicates, there were 44 linguistic left–right errors, which is 4.2% of the total number of coded descriptions. In contrast, non-linguistic left–right errors, linguistic above–below errors, and non-linguistic above–below errors occurred on 0.7, 0.2, and 0.8% of the trials, respectively. An analysis of variance, blocked by subjects, showed a significant main effect of error type ($F(1,57) = 22.27$, $P < 0.001$) and a significant effect of error direction ($F(1,57) = 9.50$, $P < 0.01$). There was also a significant interaction between error type and error direction ($F(1,57) = 18.26$, $P < 0.001$). A Tukey HSD test of pairwise comparisons revealed no significant difference between above–below and left–right errors within the non-linguistic error type ($t(57) = -0.2$, $P < 0.84$). However, within the linguistic error type, there are significantly more left–right errors than above–below errors ($t(57) = 8.35$, $P < 0.001$). In addition, within the above–below direction, there is no significant difference between the number of linguistic and non-linguistic errors ($t(57) = -1.19$,

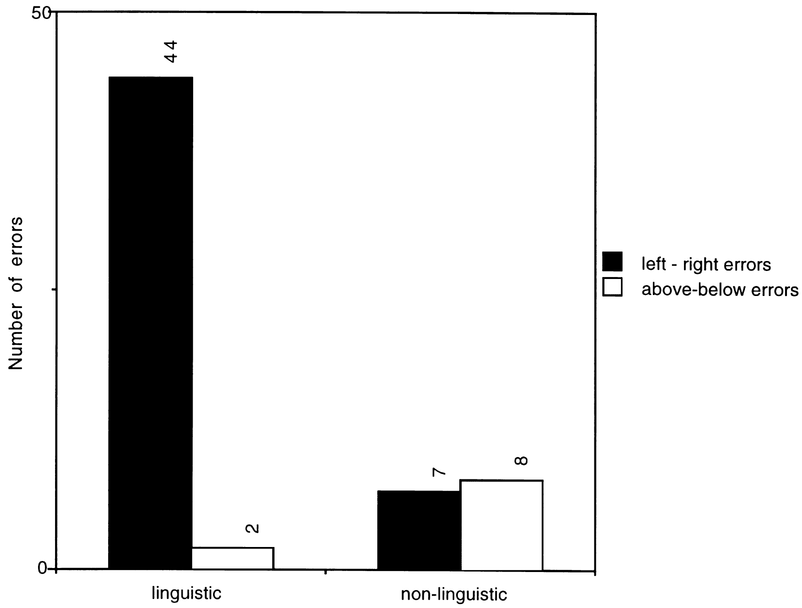

Fig. 9. Number of errors of each type (linguistic or non-linguistic) and direction (above–below or left–right).

$P < 0.24$) but within the left–right direction, there were significantly more linguistic errors than non-linguistic errors ($t(57) = 7.36$, $P < 0.001$).

A closer inspection of the errors revealed that the linguistic left–right errors and the non-linguistic left–right errors never occurred within the same trial. Furthermore, the linguistic left–right errors were not committed by only one or two participants, but were spread out among several participants. The number of linguistic left–right errors per participant ranged from 0 to 8 with an average of 2.2 per subject, which is significantly different from 0 ($t(9) = 3.50$, $P < 0.01$); 75% of participants made at least one linguistic left–right error.

Fig. 10 shows the bias pattern for reproductions. As in the first two experiments, an arrow is drawn from the actual stimulus location to the mean response location to display the direction of errors. Boxes are drawn around each mean, displaying the variability of responses. The width of each box represents two horizontal standard deviations and the height represents two vertical standard deviations. We tested whether the angular locations of the responses were significantly closer to the diagonals than were the angular locations of the presented stimuli. Within each quadrant, the angular difference between the stimulus and response mean was coded as positive if the response was shifted toward the diagonal and negative if the response was shifted away from the diagonal. The average of these differences was $+2.8°$, a difference which is significantly different from 0 ($t(47) = 9.9$, $P < 0.0001$). This result indicates that responses tend to be biased toward diagonal angular locations and away from the horizontal and vertical axes.

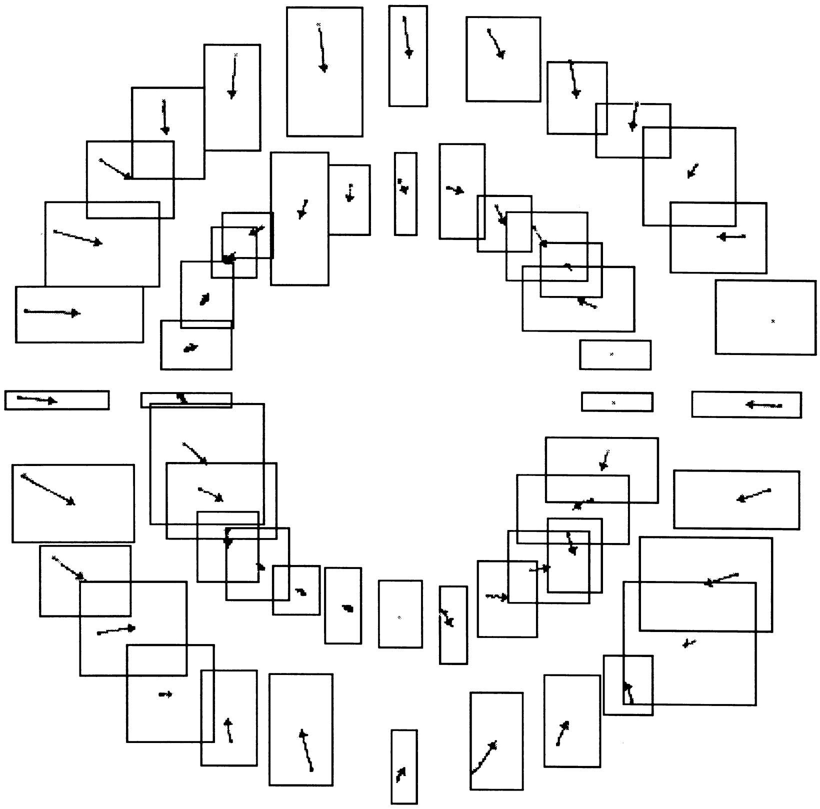

Fig. 10. Experiment 3b, bias in reproductions. Arrows are drawn from the actual stimulus location to the mean response location. Box width represents two horizontal standard deviations and box height represents two vertical standard deviations.

4.2.3. Discussion

The purpose of Experiment 3b was to determine whether linguistic coding might be mediating non-linguistic estimates of stimulus location. If that were the case, then we would have expected to find at least as many non-linguistic errors and linguistic errors, and we would have expected to find that the non-linguistic errors and the linguistic errors occurred on the same trial. Instead, we found significantly more linguistic errors than non-linguistic errors, and also found that none of the non-linguistic errors occurred on the same trial as a linguistic error. Most of the linguistic errors were left–right errors, with very few above–below errors, whereas the non-linguistic errors are almost evenly divided between left–right and above–below errors. This pattern of results suggests that the causes of linguistic and non-linguistic errors are unrelated, and therefore that linguistic coding does not mediate non-linguistic reproduction of location.

Experiment 3b conceptually replicates some of our earlier findings. The proportion of left–right linguistic errors (4.2%) was comparable to the proportion of left–right errors found in Experiment 3a (4.8%), even though different stimulus distributions were used. In addition, the pattern of bias in the non-linguistic responses is consistent with the non-linguistic findings of Experiments 1 and 2.

5. General discussion

In the experiments presented here, we examined the possibility of a correspondence between linguistic and non-linguistic structuring of space. In our first experiment, we presented a spatial array that included a central reference object and various target locations. We obtained linguistic acceptability judgments and non-linguistic reproductions of the targets in order to determine if linguistic spatial categories correspond to non-linguistic spatial categories. Instead of a correspondence, we found an inverse relation between linguistic and non-linguistic categorization of space. Specifically, the vertical axis, which serves as a category prototype in spatial language, serves as a category boundary in non-linguistic organization of space. To examine non-linguistic categories with higher resolution, we ran Experiment 2 which used a wider reference object with more target stimuli in the region above it. Here too, the vertical axis serves as a category boundary in non-linguistic spatial organization. In contrast to Hayward and Tarr (1995), we conclude that linguistic and non-linguistic categories do not have corresponding prototypes in this spatial array.

We also investigated whether the task we used to examine non-linguistic spatial categories might actually reveal covert linguistic coding, rather than non-linguistic categories. We found that participants made more left–right errors in linguistic descriptions than in their non-linguistic reproductions and that linguistic errors and non-linguistic errors did not occur within the same trials. This is not what we would expect to find if people were using language to produce non-linguistic reproductions.

The present experiments analyze bias in reproduction to determine non-linguistic categories. However, it is possible that bias could also arise from sources other than categories. For example, bias in reproductions could result from the motoric process of moving a mouse to respond. Such a motoric explanation seems unlikely for a number of reasons. First, the pattern of bias in reproductions is consistent with findings of discrimination for dot location; the regions in which we (and Huttenlocher et al., 1991) find that stimuli are estimated as closer together than they actually are correspond to the regions in which Hayward and Tarr (1995) and Munnich, Landau and Dosher (1997) find locations are more difficult to discriminate from each other. This suggests that the reproduction task reflects the same underlying representation as the discrimination tasks used in other studies. In addition, Engebretson (1994) explicitly compared reproduction and discrimination for two spatial arrays. Although those arrays differed from the ones investigated here, the studies demonstrated that comparable category effects could be found in both reproductions and same–different judgments. Finally, as noted in Section 1; Huttenlocher et al.

(1991) found that the pattern of bias increased as memory for the particular stimulus location was degraded, which is consistent with a category explanation of bias but not with a motoric explanation.

It is also possible that bias in reproductions could reflect the specific distribution of presented objects that were shown, rather than non-linguistic spatial categories. If that were the case, the bias pattern would change if we showed a different distribution of stimuli. This also seems unlikely because a comparable pattern of bias has been found in studies that varied the stimulus distributions (Crawford, Huttenlocher, Hedges & Engebretson, 1999). In addition, Huttenlocher et al. (1991) found the same pattern of bias whether dots were arranged in concentric circles or in a uniform, grid-like pattern similar to the one used by Hayward and Tarr (1995) and Munnich et al. (1997). Thus, the bias pattern does not seem to reflect participants' representation of the particular stimulus distribution that was shown.

In the experiments presented here, we examined two similar spatial arrays; thus, it is not clear if the relation we found between linguistic and non-linguistic organization would also hold for other arrays. However, there is some evidence that a similar relation occurs between linguistic and non-linguistic organization of three-dimensional space. Franklin, Henkel and Zangas (1995) found that when people reproduce locations of objects in surrounding space, their reproductions tend to be biased away from the canonical front location, even though this location is the prototypical location of the linguistic category, 'front'. This result suggests that our conclusions may also hold for spatial arrays other than the ones used in our experiments.

Taken as a whole, the results of these experiments indicate that there is not a direct correspondence between linguistic and non-linguistic spatial categories. However, the results also indicate that the two systems are not independent; they rely on a common underlying structure, the cardinal axes, but these axes appear to play different roles in linguistic and non-linguistic categorization of space.

A possible objection might be that composite linguistic descriptions, such as 'above and to the right' do map onto non-linguistic spatial categories, and therefore there is a correspondence between linguistic and non-linguistic categories. While it may be the case that some composite linguistic descriptions map onto non-linguistic categories, this is not directly relevant to our main question. Given the compositionality and flexibility of language, it is likely that no matter how space is categorized non-linguistically, linguistic descriptions could be composed that reflect those categories. The critical point is that descriptions such as 'above and to the right' are composites, and are not the basic linguistic terms of the English lexicon. Thus, such a correspondence would not be directly relevant to our question of the relation between basic linguistic terms and non-linguistic spatial categories.

As noted in Section 1, a correspondence between linguistic terms and non-linguistic categories is expected by both of the main schools of thought concerning the relation between categories in thought and those in language. According to the Whorfian view (Whorf, 1956), linguistic categories shape non-linguistic thought. Recent empirical support for this view comes from research on speakers of Tzeltal and Guugu Yimithirr, languages that describe space in terms of absolute directions (such as 'uphill' and 'downhill'), and from speakers of Dutch, which codes spatial relations in terms of egocentric directions (such as 'left' and 'right') (Levinson, 1996; Pedersen, 1995). This work used non-linguistic tasks to show that informants tend to encode spatial relations between objects in a manner consistent with their language. Thus, Tzeltal and Guugu Yimithirr speakers encoded absolute directions, whereas Dutch speakers encoded egocentric relations. For the spatial arrays we examined, a Whorfian argument would be the following: by labeling regions of space with linguistic terms such as 'above' and 'below', language carves space into categories. These linguistic categories then influence non-linguistic thought and responses on non-linguistic tasks. As a result, non-linguistic measures will reveal non-linguistic spatial categories that reflect the structure of linguistic spatial categories.

An alternative account of the relation between linguistic and non-linguistic spatial categories is that non-linguistic spatial cognition shapes spatial language. For example, Landau and Jackendoff (1993) suggested that language has fewer spatial prepositions than object names because the visual system codes spatial relations at a coarser level of detail than properties of objects. Others have posited that the structure of linguistic spatial categories may be grounded in specific, presumably universal, aspects of spatial representation (Munnich et al., 1997; Regier & Carlson-Radvansky, 2000).

Mandler (1992) argues that this non-linguistic structure is *categorical* in nature (see also Kay & McDaniel, 1978). Specifically, she argues that in the course of language acquisition, linguistic terms such as 'on' and 'in' are mapped to pre-existing categories or 'image schemas' (Lakoff, 1987) which have been abstracted from perceptual information. According to Mandler, ''Children do not have to consider countless variations in meaning suggested by the infinite variety of perceptual displays with which they are confronted; meaningful partitions have already taken place…What remains for children to do is to discover how their language expresses these partitions.'' (Mandler, 1992, p. 599). When applied to the spatial categories we examined, this view would hold that rather than language carving up undifferentiated spatial cognition, non-linguistic spatial schemas are formed prior to language acquisition, and only later become labeled linguistically. A consequence of this mapping of language onto pre-existing schemas is that the structure of linguistic spatial categories will reflect the structure of non-linguistic spatial categories.

The results of the present experiments would not have been predicted either by the view that linguistic categories shape perception, or by the opposing view that perceptual categories shape language. According to both views, the structure of linguistic categories should correspond to the structure of non-linguistic categories. In contrast, we find that the two kinds of categories bear an inverse relation to each other: the cardinal axes take the role of *prototypes* in linguistic spatial categories, and *boundaries* in non-linguistic spatial categories. This finding suggests another possible relation between language and perception. Linguistic and non-linguistic categorization may capitalize on the same underlying structure, but in different ways. As a result, that structure may play different roles in linguistic and non-linguistic categorization.

Acknowledgements

This research was supported by The National Science Foundation, Grant SBR9720313. We thank Dale Barr, Valerie Chase, Alex Francis, Barbara Luka and two reviewers for helpful comments on an earlier draft of this manuscript. We thank Alice Cho for her help with data collection.

References

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: the ''oblique effect'' in man and animals. *Psychological Bulletin*, *78*, 266–278.

Bartlett, F. C. (1932). *Remembering: a study in experimental and social psychology*. Cambridge: Cambridge University Press.

Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. In R. S. Wyer, & T. K. Srull, *Handbook of social cognition, vol. 1* (pp. 119–160). Hillsdale, NJ: Erlbaum.

Carlson-Radvansky, L. A., & Logan, G. D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language*, *37*, 411–437.

Crawford, L. E., Huttenlocher, J., Hedges, L. V., & Engebretson, P. E. (1999). *Spatial categories and sensetivity to spatial distributions of objects. Paper presented at The Midwestern Physchological Association Conference, Chicago, IL.*

Engebretson, P. H. (1994). *Category effects on the discriminability of spatial location*. Unpublished Doctoral Dissertation, University of Chicago.

Franklin, N., Henkel, L. A., & Zangas, T. (1995). Parsing surrounding space into regions. *Memory & Cognition*, *23*, 397–407.

Friedman, A. & Brown, N. R. (2000). Reasoning about geography, submitted.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*, 178–200.

Harnad, S. (1987). *Categorical perception*. Cambridge: Cambridge University Press.

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition*, *55*, 39–84.

Hoffman, C., Lau, I., & Johnson, D. R. (1986). The linguistic relativity of person cognition: an English-Chinese comparison. *Journal of Personality and Social Psychology*, *51*, 1097–1105.

Huttenlocher, J., Hedges, L., & Prohaska, V. (1988). Hierarchical organization in ordered domains: estimating the dates of events. *Psychological Review*, *95*, 471–484.

Huttenlocher, J., Hedges, L., & Duncan, S. (1991). Categories and particulars: prototype effects in estimating spatial location. *Psychological Review*, *98*, 352–376.

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, *54*, 610–646.

Lakoff, G. (1987). *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Landau, B., & Jackendoff, R. (1993). ''What'' and ''where'' in spatial language and spatial cognition. *Behavioral and Brain Sciences*, *16*, 217–265.

Levinson, S. C. (1996). Frames of reference and Molyneux's question: cross linguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett, *Language and space* (pp. 109–170). Cambridge, MA: MIT Press.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 359–368.

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett, *Language and space* (pp. 493–530). Cambridge, MA: MIT Press.

Lucy, J. (1992). *Language diversity and thought: a reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Munnich, E., Landau, B. & Dosher, B. A. (1997). *Universals of spatial representation in language and memory*. Poster presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia, PA.

Nelson, T. E., Biernat, M. R., & Manis, M. (1990). Everyday base rates (sex stereotypes): potent and resilient. *Journal of Personality and Social Psychology*, *59*, 664–675.

Pedersen, E. (1995). Language as context, language as means: spatial cognition and habitual language use. *Cognitive Linguistics*, *6*, 33–62.

Regier, T. & Carlson-Radvansky, L. A. (2000). Grounding spatial language in perception: an empirical and computational investigation, submitted.

Vogels, R., & Orban, G. A. (1986). Decision factors affecting line orientation judgments in the method of single stimuli. *Perception & Psychophysics*, *40*, 74–84.

Whorf, B. (1956). *Language, thought and reality*. New York: Technology Press of MIT and Wiley.

Wright, W. D. (1947). *Researches on normal and defective colour vision*. St. Louis, MO: Mosby.