Chapter 3
Text and Picture Integration in Comprehending and Memorizing Spatial Descriptions

Francesca Pazzaglia

Abstract Spatial processes are involved both in subjects' perception and motion in real-life environments, and in the comprehension of texts with spatial contents. Two experiments examined the effectiveness of pictures in supporting the comprehension of spatial text. Undergraduate students listened to the description of a town while either viewing an integrated picture, partial pictures of some landmarks, or no picture at all. Integrated pictures resulted in better comprehension especially in participants with a lower spatial working memory capacity. It is concluded that integrated pictures help low spatial ability individuals comprehend spatial relations that are explicit in the text. Integrated pictures also help any listener draw inferences about other, non specified spatial relations.

Keywords Integrated pictures · Spatial text · Working memory

3.1 Introduction

Spatial language is very common in every-day life. People often describe the shape and the position of certain objects, different spatial configurations and both indoor and outdoor environments. Particularly frequent is the task in which a person describes a particular environment thus allowing others to create a mental model, which can help them to move successfully therein. These spatial descriptions can be given from different perspectives, for example in relation to the cognitive style of the descriptor or the spatial features of the environment. Several studies (as reviewed by Devlin, 2001) have demonstrated that, in describing their environment, women generally prefer to adopt a route perspective based on egocentric terms of reference, and particularly centered on the description of salient landmarks. By contrast, men prefer to adopt a survey

F. Pazzaglia
Universitá Degli Studi di Padova, Dipartimento di Psicologia Generale "Vittorio Benussi", Via Venezia 8, 35131 Padova, Italy
e-mail: francesca.pazzaglia@unipd.it

J.-F. Rouet et al. (eds.), *Understanding Multimedia Documents*,
DOI: 10.1007/978-0-387-73337-1_3,

perspective, based on exocentric terms of reference, such as compass directions. It has also been demonstrated that description-perspectives can change according to environmental complexity. Modern cities, with straight roads intersecting at 90° angles, are more frequently described from a survey perspective than are historic centers of European cities characterised by very narrow, non-linear routes (Pazzaglia, 2000). These results demonstrate that the use of spatial language is quite complex, requiring the speaker to choose among several possibilities as regards to perspective (Taylor & Tversky, 1992; Tversky, 2003), linearisation (Levelt, 1989) and use of landmarks for their description (Daniel & Denis, 1998; Tversky & Lee, 1999a,b). Despite this complexity, people can quite efficiently describe their environment, so as to allow others, unfamiliar with it, to successfully navigate through it following verbal instructions alone. Guidebooks with verbal descriptions of places to see are a typical example. Pictures and maps often accompany and integrate verbal information. Substantial research has also addressed the interesting question regarding which sort of mental representations derive from processing spatial descriptions. In an attempt to answer this question, many studies, in the last twenty years, have focused on the concept of "spatial mental models" (as reviewed by Tversky, 1991).

3.2 Theoretical Framework

3.2.1 Spatial Mental Models and the Role of Imagery in the Comprehension of Spatial Texts

In Johnson-Laird's (1983) theory, the final mental representation emerging from the processing of a text is a mental model representing the situation described therein. In the case of spatial texts, mental models are supposed to have spatial properties isomorphic to those of the environments represented (Mani & Johnson-Laird, 1982). In their analysis of how spatial texts are understood and comprehended, Perrig and Kintsch (1985) drew a distinction between a text-based representation, which maintains the verbal characteristics of the message, and a situation model, spatial in nature. The nature and the features of mental models derived from spatial descriptions have been intensively studied during the last two decades (Tversky, 2003). Amongst the numerous questions addressed by these empirical studies, some remain particularly relevant and continue to be debated: e.g. the nature, verbal or spatial, of the mental model (de Vega, Cocude, Denis, Rodrigo, & Zimmer, 2001); its dependency on the perspective assumed in the description (Bosco, Filomena, Sardone, Scalisi, & Longoni, 1996; Pazzaglia, Cornoldi, & Longoni, 1994; Perrig & Kintsch, 1985; Taylor & Tversky, 1992); the dimension accessibility in mental frameworks (Franklin & Tversky, 1990; Franklin, Tversky, & Coon, 1992; Maki & Marek, 1997; de Vega et al., 2001).

The studies on the comprehension of spatial descriptions assume that the mental models derived from spatial descriptions are themselves spatial in nature.

Supposedly, they tend to maintain the characteristics of the original spatial configuration: people build mental models to represent significant aspects of their physical world and manipulate them when thinking and planning (Bower & Morrow, 1990). Further support to this is provided by Morrow, Bower and Greenspan (1989).

Indeed, they showed that, when people memorise a building layout and then read narratives that describe a protagonist moving around in the building, they focus on information that is relevant to the protagonist. Evidence to that supplied by the fact that objects from the room where the character was located were most accessible (Morrow, Greenspan, & Bower, 1987).

Bryant (1997) claims that people possess a spatial representation system that constructs spatial mental models on the basis of perceptual and linguistic information. This issue has been discussed also by de Vega et al. (2001) and Baguley and Payne (2000). Nieding and Ohler (1999) have demonstrated that six year-old children can construct spatial situation models from narratives and that these models differ from a text-based representation (van Dijk & Kintsch, 1983).

Literature on spatial descriptions shows that the description of an environment can assume two different main route and survey perspectives (Tversky, 1991). The route descriptions assume the point of view of a person who is moving along the environment. They are characterized by the use of an intrinsic frame of reference and egocentric terms, such as right, left, front and back, and have a linear organisation, given by the order in which landmarks appear along the route itself. The survey descriptions provide an overview of the spatial layout, sometimes with a strong hierarchical organization (Taylor and Tversky, 1992). An extrinsic frame of reference and canonical terms such as north, south, east and west are used. The question of whether the mental model derived from spatial descriptions is perspective dependent was investigated by some studies with different results (Bosco et al., 1996; Pazzaglia et al., 1994; Perrig & Kintsch, 1985; Taylor & Tversky, 1992). Bosco et al. (1996) found that representations of repeatedly experienced descriptions were shown to be perspective independent. Lee and Tversky (2005), even if demonstrated that switching perspective plays a significant role in comprehension of spatial texts, found that the relevance of this role, in turn, diminishes with repeated retrieval.

Studies on mental models have also considered what kind of temporary memory functions are involved in their construction. However less attention has been devoted to studying which cognitive functions are involved in the construction of spatial mental models. For example the involvement of visuo-spatial working memory has only recently been studied.

3.2.2 Visuo-Spatial Working Memory in Comprehending Spatial Descriptions

Working memory is generally defined as the dynamic control and co-ordination of processing and storage that takes place during the performance of complex cognitive tasks, such as language processing and visuo-spatial thinking (Miyake & Shah, 1999). In Baddeley's model (Baddeley, 1986; Baddeley & Hitch, 1974; Cornoldi & Vecchi, 2003; Logie, 1995) working memory is thought of as a temporary storage and processing system with a central executive and two slave sub-components: verbal working memory (VWM) and visuo-spatial working memory (VSWM).

Visuo-spatial working memory maintains and processes spatial and visual information, thus ensuring the formation and manipulation of mental images. Several studies (see below) have recently demonstrated that VSWM has a role in processing spatial texts.

Indirect evidence of the involvement of VSWM in the comprehension of spatial descriptions has emerged from data on individual differences in spatial abilities, where visuo-spatial working memory ability is related to the comprehension of spatial texts (Conte, Cornoldi, Pazzaglia, & Sanavio, 1995; de Vega, 1994; Pazzaglia & Cornoldi, 1999). Pazzaglia and Cornoldi (1999, Exp. 1) selected two groups of participants presenting no differences as regards performance on the digit span test (which measures verbal abilities) and, respectively, high and low performances on Corsi' s block task (which measures visuo-spatial abilities). Group participants were asked to listen to the description of a city and subsequently recall the spatial text. As expected, the high visuo-spatial ability group performed the memory task significantly better than the other group.

Having controlled for differences in verbal abilities, they demonstrated that the comparatively poorer performance of the low visuo-spatial ability group in the comprehension of the spatial description was effectively due to differences in spatial ability.

This experiment contributes to support the idea that the differentiation of intelligence into different components, among which spatial and verbal, allows to explain individual diversity in many every-day cognitive tasks, critical dissociations, differences between groups (Cornoldi & Vecchi, 2003). Further support in this direction is provided by the fact that even popular tests aimed at measuring intelligence are often based on the distinction between verbal and spatial (performance) intelligence. Examples are the Primary Mental Abilities (PMA) test (Thurstone & Thurstone, 1947) and Wechsler scale (Wechsler, 1981).

Other studies have shown more direct evidence of the involvement of VSWM in spatial texts processing by using a dual-task paradigm. In the dual task methodology participants have to perform a primary and a secondary task concurrently. The rationale is that performance on the primary task should be less efficient when a secondary task is presented concurrently than in the single task condition, because in the former condition the two tasks compete for the same limited resources of working memory. Many studies have explored the effects of various secondary tasks on performance during diverse cognitive activities, and it is now generally agreed that visuo-spatial tasks such as spatial tapping (continuous tapping of a series of keys or buttons) compete for maintenance of spatial information in VSWM (Farmer, Berman, & Fletcher, 1986).

Vandierendonck and De Vooght (1997) studied the comprehension of temporal and spatial relations in four-term series problems. Participants had to perform, concurrently to the problem-solving activity, an articulatory suppression task, a tapping task and a random interval repetition task. Results showed that all three secondary tasks interfered with reasoning accuracy, but that the tapping task was particularly interfering when it was performed concurrently with processing the premises of the spatial problems. This result supports the idea that VSWM is involved in constructing a mental representation of the initial data of the given spatial problem.

Pazzaglia and Cornoldi (1999, Exp. 2) investigated the involvement of verbal and visuo-spatial WM during memorisation of short abstract and spatial texts. The spatial texts consisted of instructions that required the filling-in of cells in an imagined 4 x 4 matrix, in order to follow a route within it (Brooks, 1967). Participants had to listen to the instructions while concurrently performing either a verbal or a spatial task. Results showed an interference effect of the concurrent spatial task on the spatial sentences: Average recall of spatial sentences under the concurrent spatial condition was lower than under the concurrent verbal condition.

More recently, De Beni, Pazzaglia, Gyselinck, and Meneghetti (2005), and Pazzaglia, De Beni, and Meneghetti (2006), studied the involvement of the verbal and visuo-spatial components of working memory in the memorisation and retrieval of spatial descriptions from a route perspective. In several experiments recall and recognition of spatial and non-spatial texts were compared under different conditions of concurrent spatial and verbal tasks. In accordance with their hypothesis, both memorisation and retrieval of the spatial texts was impaired by the concurrent spatial task. By contrast, the performance in the non-spatial texts was mainly affected by the concurrent verbal task.

Studies on the involvement of VSWM in the construction of spatial mental models are relevant because, on one hand, they support the models' spatial nature, and on the other because in these studies the focus has shifted from the models' characteristics to the cognitive functions and abilities required for their construction. As a consequence, they offer the theoretical basis for developing tools and training methodologies to improve comprehension and memorisation of spatial texts. Given that spatial representations are isomorphic to spatial configurations, and that they require VSWM, we can assume that pictures, being a sort of external representation of the mental model, can help the comprehension of spatial texts. However, it remains to be established if certain kinds of pictures are more effective in supporting the implementation of the spatial model than others.

3.2.3 Discourse-Picture Integration in Spatial Descriptions: An Empirical Study

Although it has been well documented that pictures improve comprehension and memorisation of texts (Levie & Lentz, 1982; Glenberg & Langston, 1992;

Hannus & Hyona, 1999), the cognitive mechanisms of this phenomenon are not entirely known. Some authors (see for example Gyselinck & Tardieu, 1999) have suggested that the "power" of pictures consists in enhancing the construction of mental models of the texts. For this reason they affirm that only certain types of pictures are effective i.e. those which help the reader obtain relevant information so as to organise and represent it mentally. Scientific texts accompanied by graphics and illustrations are generally used as materials in research on text/picture integration, rather less attention has been devoted to the role of pictures in improving comprehension and memorisation of spatial texts, such as environments' description from either survey or route perspectives. However, the hypothesis that effective pictures are faithful external representations of an internal model can be advanced also for illustrated spatial texts. More specifically we would expect the most effective pictures to be those, which not only represent all the landmarks described in the text, but also clearly mark their reciprocal spatial relations.

To date, only a few studies have analysed the role of pictures in spatial text comprehension (e.g. Ferguson & Hegarty, 1994; Tversky & Lee, 1999a), intending spatial text as description of routes and environments. The present study intends to be an empirical investigation of this topic. More specifically, we address three questions: 1. whether pictures really improve comprehension of spatial texts; 2. which kind of pictures is particularly effective (for example by comparing pictures representing the position of single landmarks with others also representing the spatial relation between landmarks); and finally 3. whether pictures in spatial texts differently affect individuals having respectively high versus low VSWM.

In the following two experiments these questions were addressed by presenting spatial texts, either with or without pictures, to different samples of undergraduate students. The aim of the first experiment was to investigate the role of pictures in spatial text comprehension, by comparing a 'no-picture' condition with two other picture conditions: (a) the single-picture condition, in which each sentence was accompanied by a picture representing the exact location of the landmark described within and (b) the map-picture condition, in which the same pictures were inserted in the perimeter of the described environment as framework.

The second experiment used the same materials and procedure, but introduced the theme of individual differences by creating two participant groups with a high versus low spatial abilities score as measured by the Mental Rotations Test (MRT; Vanderberg & Kuse, 1978).

The hypotheses made as regards to questions 1 and 2 were that pictures would be effective in improving comprehension, but only when the relative position of the landmarks were made explicit, thereby aiding the construction of a spatial mental model. Thus, we expected the best comprehension to occur in the map-picture condition when compared to the single-picture and no-picture conditions.

As for the third question, we tested the following hypotheses: given that VSWM is involved in the construction of spatial mental models, individuals with high spatial abilities were expected to perform better than the low spatial abilities group in the no-picture condition because of their superior ability in spontaneously constructing a good spatial mental model and using the VSWM to do so. However, in the picture condition we expected a smaller difference between groups, since the low spatial abilities group was expected to use the pictures as aids more than the high spatial abilities group and hence show greater improvement.

3.3 Experiment 1

3.3.1 Method

3.3.1.1 Participants

A total of 28 (7 male and 21 female) undergraduate students from the Faculty of Psychology of the University of Padua (Italy), participated in the experiment.

3.3.1.2 Materials

- Texts

Three spatial descriptions (from Taylor & Tversky, 1992; Pazzaglia, Cornoldi, & Longoni, 1994), which adopt a survey perspective and describe three fictitious environments: a zoo, a park and a farm were used. The descriptions were all ten sentences long, consisted of 120 words and had the same number of landmarks (7). Part of the description entitled “The zoo” is shown in Table 3.1.

- Verification test

Sixteen assertions, half true and half false, were formulated for each text. Half the assertions were paraphrased, half inferential (examples of assertions are reported in Table 3.2).

Table 3.1 The first five sentences of the spatial text “The zoo”. The entire text is composed by ten sentences

The Zoo
The zoo extends over a large rectangular area.
It has only one entrance in the middle of the south side of the whole frame.
In front of the entrance there is a bar.
The bar is exactly in the centre of the zoo.
In the south-west corner of the zoo there is the amusement park.
. . .

Table 3.2 Examples of paraphrased (P) and inferential (I) assertions referred to the texts "The zoo"

The zoo (assertions)
1. (P) In the middle of the south side of the whole fence there is the only zoo entrance.
2. (P) Right in the middle of the zoo there is the service bar.
3. (P) The zoo extends on a large square area.
4. (I) The amusement park is to the south-east of the bar.
5. (I) The elephants are to the north-west of the penguins.

- Pictures

Two different types of pictures were created for each descriptive sentence. The first type, the "single-picture", depicted the local spatial information contained in each sentence, i.e. one landmark and its position. The second type of pictures, i.e. the "map-pictures" was identical to the single-pictures, but assembled so as to form a map of the environment described in the text (example of single- and map-pictures, and of the sentences they referred to, are shown in Fig. 3.1).

3.3.1.3 Procedure

Each participant was tested individually for approximately 40 minutes. They were informed that the experiment required them to listen to and memorise three descriptions in order to answer a questionnaire and that two of these were accompanied by some pictures, one for each sentence. Descriptions were tape-recorded. Each participant listened to three descriptions, each in a different condition: 1. listening, 2. single-picture, 3. map-picture. The order of text and picture type presentation was counterbalanced across participants. During the single-picture and the map-picture conditions, each picture was presented on an A4-format paper, for the duration of each sentence description. Limited to the single-picture, the participants could inspect in any time all the old pictures. Immediately after having listened to each description, participants were asked to respond to the true/false assertions presented on a computer screen in random order.

3.3.2 Results and Discussion

A 3 X 2 Analysis of Variance (ANOVA) with presentation condition (listening, single-picture, map-picture) and assertion type (paraphrased vs. inferential) as factors was performed on the total correct responses given. The analysis revealed the expected best performance in the map-picture condition (M = 6.86, SE = 0.19), compared to single-picture (M = 6.14, SE = 0.20) and listening (M = 6.05, SE = 0.24), $F(2, 54) = 6.49$, MSE = 1.67, $p < 0.005$. Paraphrased

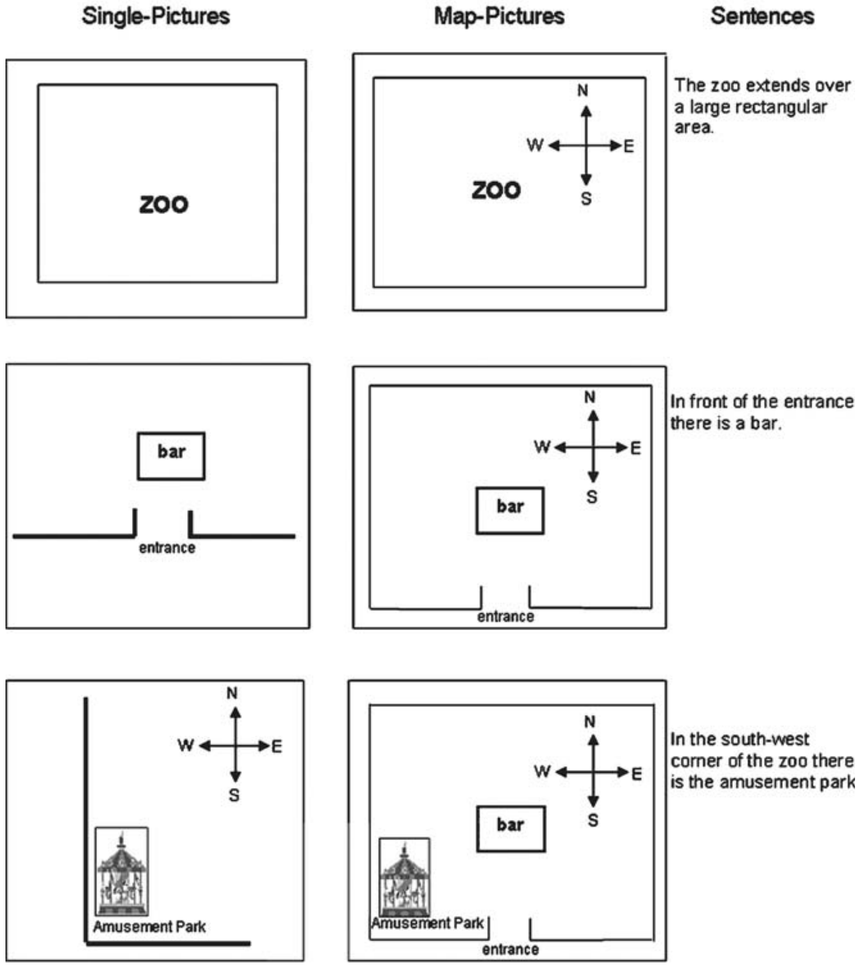

Fig. 3.1 Example of single-pictures and map-pictures used in the text "The zoo" and relative sentences

assertions were simpler than inferential ones (respectively: M = 6.85, SE = 0.17; M = 5.84, SE = 0.18), $F(1, 27) = 50$, MSE = 0.84, $p < 0.001$.

As expected, the results of Experiment 1 demonstrate that pictures presented concurrently with spatial texts improve text comprehension. Pictures are effective not only in helping the memorisation of single pieces of information, but also in aiding the formation of a spatial mental model of the text. In fact both responses to paraphrased and inferential assertions improved with picture presentation. However, this positive effect was restricted to the map-pictures condition. Only the presentation of a map-picture together with the spatial text was effective in enhancing comprehension of the spatial description. This is in accordance with our hypotheses that the map-picture condition would aid the construction of a mental model of the text by making the relation between elements of the description explicit.

This experiment contributes to demonstrate that comprehension of spatial texts is enhanced by the presence of accompanying pictures. The paradigm adopted in the experiment, i.e. comparing two different kinds of pictures, contributes also to explain the cognitive mechanisms underlining this phenomenon and results so far suggest that this is not entirely due to a dual-code effect (Paivio, 1978), but to the construction of mental models of the text.

3.4 Experiment 2

3.4.1 Objectives

Experiment 2 aimed to investigate the involvement of spatial abilities in processing spatial texts and to verify whether participants with high and low spatial abilities were differently influenced by illustrations. Existing research on spatial abilities provides evidence that they are not as a single component but that they are articulated to varying extent (Cornoldi & Vecchi, 2003; Devlin, 2001). Linn and Petersen (1985), conducting a meta-analysis of 172 studies, argued for the existence of three spatial factors: spatial perception, determining spatial relations with respect to one's own body; mental rotation, a gestalt-like analogue process; spatial visualisation, multistep manipulation of spatially presented information. Spatial abilities have been examined either through pencil-and-paper psychometric tests, such as mental rotation tests, or more real-world tasks, such as distance judgments, way-finding, pointing in the direction of unseen locations, map learning (Kirasic, 2000).

In our study, individual differences in spatial abilities were measured using the Mental Rotations Test (Vanderberg & Kuse, 1978). This choice was due to the fact that the MRT can be solved only using a global spatial ability, it is correlated with survey spatial representation (Pazzaglia & De Beni, 2001) and derives from studies on imagery. All these factors were important in our experimental procedure, where a description from a survey perspective was used and where imagery abilities could be considered important.

We expected to replicate the results found in the previous experiment on the different efficacy of single and map pictures in improving text comprehension. Regarding individual differences, we hypothesized that participants with high spatial abilities would perform better than those with low spatial abilities, in considering the spatial features of the texts. However, we expected that this effect, dramatic in the no-picture condition, would be reduced in the map-picture condition because, for individuals in the low spatial ability group, the presentation of pictures would constitute an external spatial representation i.e. a useful aid in integrating their poor internal representation.

3.4.2 Method

3.4.2.1 Participants

Participants were 36 undergraduate students (7 males, 29 females) divided in two groups with high and low spatial abilities, each comprising 18 participants. They were selected from a sample of 174 students by administering the Mental Rotations Test (MRT) (Vanderberg & Kuse, 1978). Participants with scores in the MRT lower than, or equal to the 25th percentile of the entire sample were considered as having low spatial abilities and participants with scores higher than or equal to the 75th were considered as having high spatial abilities. Two sub-groups, composed of 12 low – and 9 high-spatial individuals respectively were administered with a standardised comprehension test (Cornoldi, Rizzo, & Pra Baldi, 1991) in order to verify that they did not differ in comprehension ability. The mean scores were 7.75 and 8.11 respectively for low and high spatial ability groups, and difference was not significant, $t(19) = -0.46$, $p = 0.65$.

3.4.2.2 Materials and Procedure

The Materials and Procedure were the same as in Experiment 1, except that there were 24 assertions instead of 16.

3.4.3 Results and Discussion

A 3 X 2 X 2 ANOVA with presentation condition (listening, single-picture, map-picture), assertion types (paraphrased, inferential) and group (high and low in spatial abilities) as factors, was performed on the percentage of correct answers to the true/false assertions. As expected the analysis revealed best performance in the map condition (M = 84.81, SE = 2.22), compared to picture (M = 77.17, SE = 2.44) and listening (M = 73.26, SE = 2.63), $F(2, 68) = 9.99$, MSE = 248, $p < 0.001$. Paraphrased assertions resulted simpler than inferential ones (respectively: M = 83.10, SE = 1.90; M = 73.72, SE = 2.16), $F(1, 34) = 39.21$, MSE = 121, $p < 0.001$.

High-spatial (M = 84.12, SE = 2.69), performed better than low-spatial group (M = 72.68, SE = 2.69), $F(1, 34) = 8.02$, MSE = 785, $p < 0.01$.

There was a significant interaction effect between presentation condition, assertion type and group, $F(2, 68) = 4.55$, MSE = 126, $p < 0.05$. A post-hoc Newman-Keul analysis (c.d. = 9.35) revealed, as shown in Fig. 3.2, different patterns of results for high and low spatial individuals in relation to pictures and assertions.

In answering paraphrased assertions high spatial abilities individuals had the same performance for both listening and pictures conditions. This is due to their very high performance (more than 85%) in all conditions. By contrast, the low spatial abilities group had a significantly poorer performance in the listening

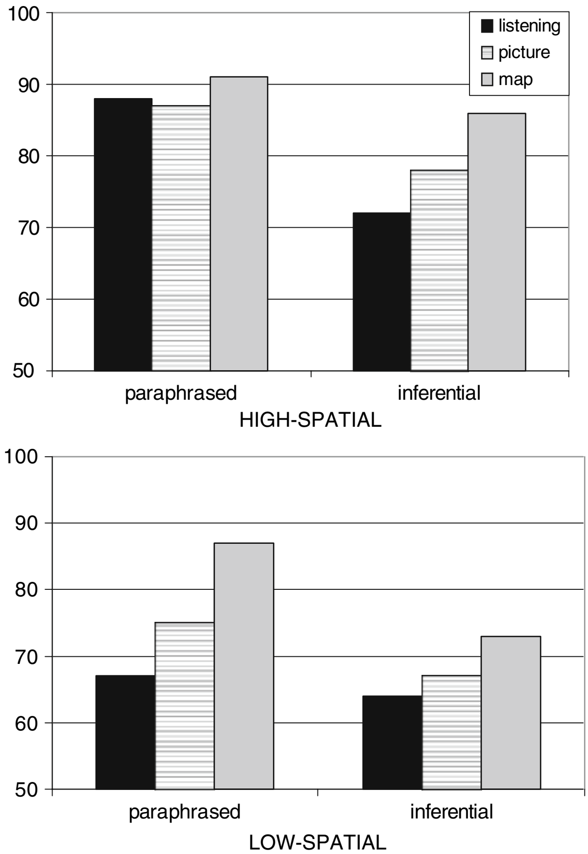

Fig. 3.2 Percentage of questions correctly responded in Experiment 2: Interaction group (high and low spatial abilities) by picture (no picture, single-picture, map-picture) by assertions (paraphrased and inferential)

and single-picture condition (with no significant difference between them), as compared to the map-picture condition. Compared to the high spatial abilities individuals, they performed worse both in the listening and single-picture conditions, but performance of the two groups was equal in the map-picture condition. Taken together, these results support the hypothesis that pictures are effective because they help the construction of an internal spatial mental model. In fact, as observed in the first experiment, only map-pictures, which make explicit the spatial relations between landmarks, are effective in improving memory performance.

It is also interesting to note that this beneficial effect is limited to individuals with low spatial abilities. A possible interpretation is that participants in the high spatial abilities group can spontaneously activate an internal spatial model, using their superior abilities in maintaining and processing visuo-spatial materials. The external representation is in this case superfluous, because their internal representation is good enough to answer the paraphrased assertions presented (see also Schnotz, Chapter 2). By contrast, low spatial individuals do not have sufficient cognitive resources to construct a good mental model. However, given the external aid, i.e. the map-picture, their performance improves sufficiently to equal that of the high spatial abilities group.

A reverse pattern of results was found for inferential assertions. In this case the low spatial abilities group showed the same, low, performance in all three conditions, although it should be noted that the difference between the listening and the map-picture conditions was close to significance. The high spatial abilities group, however, performed significantly better in the map-picture than in the listening condition. Furthermore, the two groups did not differ in the listening condition, but did differ in the predicted direction in both single- and map-picture conditions. This is consistent with our hypotheses in that when questions require a more complex and stable spatial representation individuals with good spatial abilities take advantage from inspecting the map-picture during text presentation.

Results of Experiment 2 confirm that illustrations enhance text comprehension when they make explicit the relationship between units of information contained in texts and that spatial abilities are indeed involved in processing verbal information when the content of the text is spatial.

3.5 Conclusions

Describing spatial configurations and comprehending spatial texts are tasks that involve both spatial and verbal cognitive systems. They require a conversion from an internal spatial representation to a linguistic expression, or the creation of a mental spatial representation from a verbal description. In the latter case, the presence of pictures can enhance the construction of the correspondent spatial mental model. The results of the present study suggest that of the two types of pictures used (single-picture versus map-picture), the most effective in helping to memorize a spatial text are those which describe the relation amongst landmarks present in the text. This is thought to be because this type of picture presentation aids individuals in the creation of an internal mental map, matching that of the text. Furthermore both people with high and low spatial abilities can take advantage, even if at different levels, of the use of pictures accompanying spatial texts.

Given the results above, the role of pictures is now clear: they constitute external representations that facilitate the creation and maintenance of an internal mental model. A possible objection is that the superiority of "listening plus picture" condition was due to the fact that pictures were available for review and participants could inspect them whenever they want. Conversely, they could not listen again to the descriptions in the "listening" condition. However, the goal of these experiments was not to demonstrate that listening plus picture was better than listening twice or more (dual-code vs. single-code), but to compare two very common and ecological conditions: listening the spatial description of an environment and listening the same description accompanied by pictures. Further, the main focus was on the comparison between the two picture conditions, identical in the procedure. In these comparisons the map-picture turned out to be superior not only respect to listening but to the single-picture condition too.

Some questions still remain unanswered. An important question is the property of the test used to assess comprehension. In our experiments we used a verification task of paraphrased and inferred information. This task allowed us to assess more directly the construction of a mental model. Free-recall is a different index commonly used to assess comprehension and memorisation of the whole content of a text, but it has the limit to be less sensitive in distinguishing between literal and inferential processes. A further index, which should be useful to our goals, is response times in the verification task. In fact, if a reader has built a mental model properly, with a lot of inferences, then these inferences should be readily and quickly available. This should result in faster verification times. It would be interesting to verify if pictures affect response times and to compare response times of participants with high and low spatial abilities.

In our experiments picture-presentation was concurrent to text-presentation, hence each picture added only the units of information contained in each sentence. Thus, the increment of knowledge offered by each picture corresponded to that given by each sentence and the construction of the mental model followed the sequential procedure typical of text processing. It would be interesting to verify the effect of pictures containing all the information of the text given either at the beginning, at the end or during the presentation.

Furthermore in the present study we compared “listening” to “listening plus picture”, finding a superiority of the latter condition, thus we can claim that pictures help the comprehension of spatial texts. Yet the question remains whether the reverse is also true, i.e. if the presence of text can help the comprehension of pictures (on this issue see also Rinck, Chapter 10). In order to answer this question it could be useful to compare a “picture” to a “listening” and to a “listening plus picture” condition. If it is true that “a picture is worth a thousand words” we would expect the “picture” condition to be more efficient than the “listening” condition but perhaps equal to the “listening plus picture” combination. We hope to answer these and other questions in future research.

Acknowledgments The author wishes to thank Sabina Scattola and Francesca Zanardi for collecting data of the two experiments.

References

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8). New York: Academic Press.

Baguley, T., & Payne, S. (2000). Long-term memory for spatial and temporal mental models includes construction processes and model structure. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53A*, 479–512.

Bosco, A., Filomena, S., Sardone, L., Scalisi, T. G., & Longoni, A. M. (1996). Spatial models derived from verbal descriptions of fictitious environments: The influence of study time and the individual differences in visuospatial ability. *Psychologische-Beitrage, 38*, 451–464.

Bower, G. H., & Morrow, D. G. (1990). Mental models in narrative comprehension. *Science, 247*, 44–48.

Brooks, L. R. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology, 19*, 289–299.

Bryant, D. J. (1997). Representing space in language and perception. *Mind and Language, 12*, 239–264.

Conte, A., Cornoldi, C., Pazzaglia, F., & Sanavio, S. (1995). Lo sviluppo della memoria di lavoro visuospaziale e il suo ruolo nella memoria spaziale./The development of the visuo-spatial working memory and its role in spatial memory. *Ricerche di Psicologia, 19*, 95–114.

Cornoldi, C., Rizzo, A., & Pra Baldi, A. (1991). *Prove avanzate di comprensione della lettura./Reading comprehension advanced tests*. Firenze: Organizzazioni Speciali.

Cornoldi, C., & Vecchi, T. (2003). *Visuo-spatial working memory and individual differences*. Philadelphia, PA: Psychology Press.

Daniel, M. P., & Denis, M. (1998). Spatial descriptions as navigational aids: A cognitive analysis of route directions. *Kognitionswisssenchaft, 7*, 45–52.

De Beni, R., Pazzaglia, F., Gyselinck, V., & Meneghetti, C. (2005). Visuo-spatial working memory and mental representation of spatial descriptions. *European Journal of Cognitive Psychology, 17*, 77–95.

Devlin, A. N. (2001). *Mind and maze: Spatial cognition and environmental behavior*. Westport, CT: Praeger Publishers/Greenwood Publishing Group, Inc.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratchpad in working memory. *The Quarterly Journal of Experimental Psychology, 38A*, 675–688.

Ferguson, E. L., & Hegarty, M. (1994). Properties of cognitive maps constructed from texts. *Memory and Cognition, 22*, 455–473.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119*, 63–76.

Franklin, N., Tversky, B., & Coon, V. (1992). Switching points of view in spatial mental models. *Memory and Cognition, 20*, 507–518.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated texts: Picture help to build models. *Journal of Memory and Language, 31*, 129–151.

Gyselinck, V., & Tardieu, H. (1999). The role of illustration in text comprehension: what, when, for whom and why? In H. van Oostendorp & S. Goldman (Eds.), *The construction of mental representations during reading* (pp. 195–218). Mahvah, NJ: Lawrence Erlbaum Associates.

Hannus, M., & Hyona, J. (1999). Utilization of illustrations during learning of science textbook passages among low and high-ability children. *Contemporary Educational Psychology, 24*, 95–113.

Johnson-Laird, P. (1983). *Mental models*. Cambridge, MA: The MIT Press.

Kirasic, K. C. (2000). Age differences in adults' spatial abilities, learning environmental layout, and wayfinding behavior. *Spatial Cognition and Computation, 2*, 117–134.

Lee, P. U., & Tversky, B. (2005). Interplay Between Visual and Spatial: The Effect of Landmark Descriptions on Comprehension of Route/Survey Spatial Descriptions. *Spatial Cognition and Computation, 5*, 163–185.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Levie, W. H., & Lentz, R. (1982). Effects of text illustration: A review of research. *Educational Communication of Technology Journal, 30*, 195–232.

Linn, M. C., & Petersen, A. C. (1985). "Emergence and characterization of sex differences in spatial ability: A meta-analysis" *Child Development, 56*, 1479–1498.

Logie, R. H. (1995). *Visual-spatial working memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Maki, R. H., & Marek, M. N. (1997). Egocentric spatial framework effects from single and multiple points of view. *Memory and Cognition, 25*, 677–690.

Mani, K., & Johnson-Laird, P. (1982). The mental representation of spatial descriptions. *Memory and Cognition, 10*, 181–187.

Miyake, A., & Shah, P. (1999). Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.). *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY, US: Cambridge University Press (pp. 1–27).

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language, 28*, 292–312.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language, 26*, 165–187.

Nieding, G., & Ohler, P.(1999). Der Einfluss von Protagonisten Zielstrukturen auf raeumliche mentale Modelle beim narrativen Textverstehen von Kindern./The effect of protagonists' movements on spatial mental models in narrative text comprehension by children. *Sprache and Kognition, 18*, 146–158

Paivio, A. (1978). Comparisons of mental clocks. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 61–71.

Pazzaglia, F. (2000). Tipologie di descrizione di ambienti in funzione delle caratteristiche ambientali, individuali, e di familiarità con i luoghi. (Typology of environmental descriptions in function of environmental, individual, and familiarity characteristics of place). *Giornale Italiano di Psicologia, 27*, 133–159.

Pazzaglia, F., & Cornoldi, C. (1999). The role of distinct components of visuo-spatial working memory in the processing of texts. *Memory, 7*, 19–41.

Pazzaglia, F., Cornoldi, C., & Longoni, A. (1994). Limiti di memoria e specificità di rappresentazione nel ricordo di descrizioni spaziali "dall'alto" o "entro il percorso". (Limits of memory and the specificity of representations in the memory of survey and route descriptions). *Giornale Italiano di Psicologia, 21*, 267–286.

Pazzaglia, F., & De Beni, R. (2001). Strategies of processing spatial information in survey and landmark-centred individuals. *European Journal of Cognitive Psychology, 13*, 493–508.

Pazzaglia, F., De Beni , R., & Meneghetti, C. (2006). The effects of verbal and spatial interference in the encoding and retrieval of spatial and nonspatial texts, *Psychological* Research, Feb 16; [Epub ahead of print; DOI: 10.1007/s00426-006-0045-7].

Perrig, W., & Kintsch, W. (1985). Prepositional and situational representations of text. *Journal of Memory and Language*, *24*, 503–518.

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language, 31*, 261–292.

Thurstone, L. L., & Thurstone, T. G. (1947). *Primary mental abilities*. New York: Psychological Corporation.

Tversky, B. (1991). Spatial mental models. In. G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 109–145). San Diego, CA: Academic Press.

Tversky, B. (2003). Structures of mental spaces: How people think about space. *Environment and Behavior, 35*, 66–80.

Tversky, B., & Lee, P. U. (1999a). Pictorial and verbal tools for conveying routes. In C. Freksa & D. M. Mark (Eds.), *Spatial Information Theory. Cognitive and computational foundations of geographic information science: International Conference COSIT'99* (pp. 51–64). Heidelberg: Springer-Verlag.

Tversky, B. & Lee, P. U. (1999b). Why do speakers mix perspectives? *Spatial Cognition and Computation, 1*, 399–412.

Vanderberg, S. G., & Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599–604.

Vandierendonck, A., & De Vooght, G. (1997). Working memory constraints on linear reasoning with spatial and temporal contents. *The Quarterly Journal of Experimental Psychology, 50A*, 803–820.

de Vega, M. (1994). Characters and their perspectives in narratives describing spatial environments. *Psychological-Research/Psychologische-Forschung,* 56, 116–126.

de Vega, M., Cocude, M., Denis, M., Rodrigo, M. J., & Zimmer, H. D. (2001). The interface between language and visuo-spatial representation. In M. Denis, R. H. Logie, C. Cornoldi, M. de Vega, & J. Engelkamp (Eds.), *Imagery, Language and Visuo-Spatial Thinking* (pp. 109–136). Hove, East Sussex: Psychology Press Ltd.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* New York: Psychological Corporation.