Science Arts & Métiers (SAM)
is an open access repository that collects the work of Arts et Métiers Institute of Technology researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: https://sam.ensam.eu
Handle ID: .http://hdl.handle.net/10985/8780

To cite this version :

Jean-Marie LE YAOUANC, Christophe CLARAMUNT, Eric SAUX - A semantic and language-based representation of an environmental scene - GeoInformatica - Vol. 14, n°3, p.333-352 - 2010

Any correspondence concerning this service should be sent to the repository Administrator : scienceouverte@ensam.eu

Geoinformatica manuscript No.
(will be inserted by the editor)

A semantic and language-based representation of an environmental scene

Jean-Marie Le Yaouanc · Éric Saux · Christophe Claramunt

Received: date / Accepted: date

Abstract The modeling of a landscape environment is a cognitive activity that requires appropriate spatial representations. The research presented in this paper introduces a structural and semantic categorization of a landscape view based on panoramic photographs that act as a substitute of a given natural environment. Verbal descriptions of a landscape scene provide the modeling input of our approach. This structure-based model identifies the spatial, relational and semantic constructs that emerge from these descriptions. Concepts in the environment are qualified according to a semantic classification, their proximity and direction to the observer, and the spatial relations that qualify them. The resulting model is represented in a way that constitutes a modeling support for the study of environmental scenes, and a contribution for further research oriented to the mapping of a verbal description onto a geographical information system-based representation.

Keywords Spatial cognition · Landscape perception · Scene descriptions

1 Introduction

Early models of Geographical Information Systems (GIS) have been deeply influenced by quantitative models and geometrical representations of space [18]. Despite the interest of these approaches for cartographical applications, they do not completely reflect the way a human being perceives and describes his/her environment since he/she preferably stores and processes qualitative information [20,10,36]. Qualitative terms arise from *common sense*, *i.e.* intuitive concepts we daily manipulate to interact with our environment. Over the past few years, *naive geography* has been established as the field of study that combines modeling concepts inspired from common sense and perception of space [8]. Naive geography addresses the modeling of space using concepts derived from daily experience and human knowledge. In particular, naive geography provides a conceptual framework for the study of verbal descriptions derived from the perception of natural landscapes. The perception of

Jean-Marie Le Yaouanc
French Naval Academy Research Institute, BCRM Brest, CC 600 Lanvéoc, 29240 Brest Cedex 9, France
E-mail: leyaouanc@ecole-navale.fr

Éric Saux · Christophe Claramunt
E-mail: {saux, claramunt@ecole-navale.fr}

space encompasses cognitive principles that favor memorization of the properties of an environment, and communication of its salient properties to an external referent using natural language [14,44].

This paper extends an initial work in which the objective was to provide a linguistic analysis of a verbal description [23]. Our approach addresses the modeling and the representation of the verbal description of a scenery by an observer perceiving it. It combines a conceptual analysis of the sentences and terms, and a structure-based study of the spatial and rhythmic properties that emerge from such descriptions. We consider the case of an observer located in an unknown natural environment, perceiving his/her 360° surroundings, and who is asked to provide a description of his/her environment to an off-site addressee. The verbal description of a natural environment generally underlines salient entities, as well as the spatial relations binding them, and the overall structure of the environment. We introduce a conceptual modeling of a landscape description, supported by a semantic model that provides a bridge between the perception of a landscape and a qualitative representation. The aim of our research is to provide a representation resulting from the perception of a natural environment, that favors the search for a mapping between descriptions of sceneries and a GIS representation.

The paper is organized as follows. Section 2 briefly discusses related work on the cognitive representations of natural landscapes. Section 3 presents the experimental study, where a panel of observers are given the task of producing verbal descriptions of a natural environment. Section 4 provides a conceptualization of these verbal descriptions and introduces a conceptual model. Section 5 introduces the spatial views that reflect an environmental scene, while section 6 illustrates the potential of the approach. Finally, section 7 draws the conclusions and outlines further work.

2 Spatial perception and mental representations of space

An observer derives an egocentric perception of an environment that reflects his/her own body of knowledge and experience, and his/her interactions with the proximate environment [32]. The resulting mental representation is not limited to a combination of spatial concepts, and the perception based on common sense not an exact replication of the real world [20]. This directly reflects the close relationship between common sense, cognitive and sensorial processes [42]. This entails several issues such as how spatial knowledge is acquired by human beings, what the range of linguistic constructs manipulated to define and characterize such an environment is, and to which degree the language people manipulate affects their ability to interact effectively with spatial information.

Cognitive studies make a distinction between the internal level oriented to the neuropsychological processes that generate a mental representation and the external level that reproduces the structure-based organization of the observed phenomenon. Our purpose concerns the latter, and it is oriented to the modeling of an environment based on a verbal description. As shown in previous studies, an observer acting in an environment organizes the perceived space according to proximities and senses [44]. The "proximate environment" is defined as the area perceived by several senses, while the "landscape area" that extends to the horizon is perceived by sight alone [11]. The notion of *mental space* is defined as a cognitive space derived from cultural, pragmatic and linguistic concepts, and common sense-based reasoning [30]. It has been schematized by "eliminating details and simplifying features around a framework consisting of elements and the relations among them" [28]. Tversky categorizes four cognitive spaces defined by the actions achievable at different scales, the "space

of the body" that integrates our proper actions and sensations, *i.e.* the area where entities can be touched ; a "space near the body", larger than the previous one, and conceptualized as a tridimensional environment where objects are located ; "the navigational space" *i.e.* the space in which the observer has to navigate to perceive it as a whole ; and the "space of graphics" made of graphical representations of space (cf. Fig. 1). Similar categorizations of cognitive spaces have been suggested where the identified categories characterize the different spaces according to the scale of entities and their relation to the observer: "figural space" is described as being the space smaller than the body and composed of small objects, "vista space" as the region visible from a single location, "environmental space" as the region accessible via displacement or navigation, and "geographic space" the space too large to be apprehended directly [31].

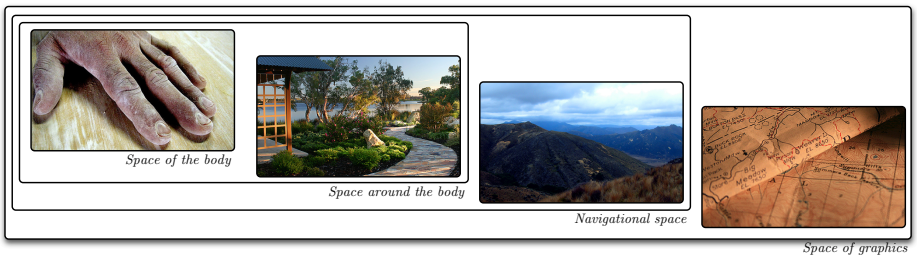

Fig. 1 Tversky's cognitive spaces

A mental representation associated to the environmental knowledge is denoted as a *cognitive map* [42], elsewhere quoted as "map in the head" [21]. Cognitive maps result from acquisitions using various modalities and perspectives, and are "fragmented, schematized, incomplete and multimodal" [43]. A "cognitive collage" aggregates sequences of cognitive maps from different scales and levels of abstraction. It is made of specific cognitive representations, including fragments with imprecise spatial information, particularly for distance or orientation metrics.

Landmarks and salient features are prominent objects used in mental representations of an environment. The role of these structural elements has been characterized by the contribution of Lynch, "The Image of the City" that refers to a cognitive representation where landmarks, nodes, paths, quarters and barriers are considered as primitive concepts of an urban environment [27]. These structural elements are typical of such an environment, but landmarks should also be considered as salient natural entities used for the orientation of an observer in a natural environment. Couclelis defines a *landmark* as a reference identifying a fixed point in the environment, and that helps a human for achieving orientating tasks [6]. A landmark is not just useful for navigation as it can also help for the location of entities located in the vicinity of salient places [1]. Landmarks are defined as spatial constructs with key characteristics that make them recognizable in the environment. Sorrows and Hirtle introduce three categories of landmarks either based on their visual properties that contrast with the other entities of the environment, physical characteristics when those play a prominent role or location in the environment, or cognitive qualities for which the meaning prevails [39]. Landmarks that play a prominent role for the qualification of a specific environment constitute key-references for the conceptualization of a natural scene.

3 Context and experimental study

The perception of natural landscapes can be directly experimented using in-situ observations or using various image-based alternatives. Photographs have been recognized as substitutes for *in-situ* landscape surveys as they facilitate the collection of verbal descriptions resulting from the perception of an environment. They have been largely used in studies oriented to the cross-comparison of similarities and differences resulting from the direct or indirect perception of landscapes [19,35,29]. It has been observed that although photographs entail some limitations for the precise identification of natural features, and some visual distortions due to the resulting planar projection of a three dimensional space, they provide a convenient alternative for the representation of visual environments [34,40]. This consequence is largely due to the high quality of digital photographs and the use of adapted lights and intensities.

Our experimental study has been conducted in the context of several semi-natural landscapes located in France. The experimentation was performed using a panel of 23 participants (18 males and 5 females), non-experts in GIS, and with no previous knowledge of the considered landscapes. The principles of the experiment are as follows. Several 360° panoramic images were displayed on a computer screen. Photographs were presented in an interactive manner using a Java interface, and the observers were able to explore the scenery by rotating a given view, as they would have done by rotating their body in a natural environment. Photographs were displayed in a large size in order to limit the restriction due to the difference of accuracy between an environment perceived *in situ*, and that perceived through a photograph. Participants were also able to zoom in on a specific area or entity in order to minimize depth-of-field errors. Sceneries were selected in order to favor an unbiased perception and orientation of the observer, since they were composed of entities with a high structural role (*e.g.*, a footpath crossing the scenery from left to right of the observer, a lake nearby the observer or a footpath running along the seaside that clearly makes the distinction between land and the ocean). After a quick overview of the panoramic photograph using a dynamic interface[1], the panelists were asked to perform a virtual tour of the scene, viewing an observer-centered photograph, and spontaneously describe this photograph in order to favor recognition of the site by an off-site addressee. Participants were also asked to specify when they were rotating the photograph (on the left or right) and ideally use relative relations to translate this movement as they would have done when rotating *in situ*. Verbal descriptions were recorded using a data storage device and were not to exceed 5 minutes.

Fig. 2 Experimental panorama of a semi-natural environment

An example of description resulting from the perception of the scenery presented in Fig. 2 is as follows: "*I'm on a footpath that runs along a castle that was certainly constructed during the Middle Ages, and a pond. In front of me, there is a little valley with the castle on the left of it and at the horizon, I can distinguish a mountain range. Behind me, there is the pond with a large meadow behind and a forest far away*"

[1] The experimental setup is available at `http://experimentation.yaou.org/`

The verbal descriptions generated by the panelists have led to the following results. Most of the participants describe the panoramic photograph from left to right, taking a reference point to begin their description. Most verbal descriptions are implicitly organized with a hierarchy. 60 % of the participants first describe the scene as a whole with sentences like "I am in a mountainous region". The salient concepts structuring the scene are first mentioned but not located. Next, participants describe the content of the landscape, and the way these concepts are related to the others. This shows evidence of a hierarchical perception of space where the landscape is first perceived and described as a whole, before being specifically described, which confirms previous studies and evidence of hierarchies in the perception of spatial information [15]. When observing the environment, humans perceive distinct concepts that are part of the landscape. Most descriptions contain entities directly related to the observer's cultural and ontological background, and his/her common sense. A scene description encompasses salient entities, as their role is prominent in the structure and organization of the environment. Participants identified human-made (50%), relief (30%) and vegetation (15%) entities.

These entities are qualitatively associated to others using spatial relations. The observers interpret photographs of an environmental scene by using spatial relations to qualify the location of the concepts such as "behind the mountain", "in front of the house", "in the background", "in the foreground", "in the long distance", proximity adjectives such as "near", "close to", "far from", "further", directional relations such as "to the right of" and a few constructs that generate a tridimensional representation of the scenery such as "above" or "below". Most of these terms are relative constructs associated to the location of another entity identified in the landscape. It also appears that the roles played by the entities selected by the observer are determined by their contribution to the hierarchical organization of the environment. This confirms Tversky's intuition on the organization of environment that impacts on the structure of a verbal description and precedes its linear structure [44]. The location of these concepts depends on their proximity to the observer. The proximity spaces that emerge from these descriptions, vary according to the terms used.

The experiment also highlighted the prominent role played by directional relations (50%) which are used twice as often as proximity (30%) and topological constructs (20%). This stresses the fact that directional relations are among the most appropriate for structuring a panoramic view. It is also worth noting the trivial role played by quantitative measurements in the descriptions although two participants used metric data in their description to specify a proximity. This may have been caused by the use of photographs which lack the depth-of-field of three-dimentional space to prompt the descriptions. Moreover, these quantitative measurements are widely associated to an imprecise linguistic term such as "*about* 500 meters from my position" or "*around* 400 meters from the house".

In order to describe their environment, the observers took a perspective that depended on the frame of reference used [41]. Levinson distinguishes three frames of reference that are intrinsic, absolute or relative depending on the point of view of the observer and the described entities [25]. The intrinsic frame of reference is a binary relation in which the location of an entity is defined in relation to another one. The absolute frame of reference is also a binary relation where the location of a given entity is defined by a fixed support (*e.g.*, cardinal directions). Last but not least, the relative frame of reference is a ternary system since the location of an entity is given both from the point of view of the observer, and the location of another entity of the environmental scene. Experiments show that observers mainly combine two frames of reference, the relative and intrinsic ones, depending on the location of the entities. However, participants are most likely to use a relative frame of

reference to locate salient entities. This confirms previous work on the prominent role played by relative frames of reference [41].

4 Conceptual model of an environmental scene

We define an *environmental scene* as the 360° environment, perceived and described by an observer from a static point of view. It is associated to an anthropocentric description, *i.e.*, what an observer perceives from his/her location. As humans tend to perceive their environment using different levels of spatial perception [11], an environmental scene is qualitatively structured using the concept of *proximity spaces* respectively defined according to the actions an observer is able to achieve in them. They are defined as follows :

- The *space of the body* as introduced by Tversky [43]. It corresponds to the space that integrates our own actions and sensations. It contains easily accessible and recognizable objects.
- The *experienced space* that can be easily apprehended by moving around the space of the body.
- The *distant space* is the environment located between the experienced space and the space at the horizon, and is hardly accessible without a significant displacement.
- The *space at the horizon* is made of silhouettes that constitute the boundaries of the forms of relief, *i.e.* the boundary between land and sky [5].

Cognitive spaces reflect different levels of interaction that vary with the scale of the entities composing them [31]. They are influenced by the field of vision of the observer, *i.e.*, their distance from the observer [11], and the actions the observer is able to perform in them [43]. Similarly, proximity spaces are determined by the actions the observer is able to perform in them, and their distance from him/her. The way the boundaries of proximity spaces are conceptualized by an observer depends on the landscapes studied. They might be materialized by *fiat* regions of space, *i.e.*, with non-well defined boundaries as there is an uncertainty on the location of their boundaries. Alternatively, they can be revealed by qualitative discontinuities that correspond to a physical reality, and are then considered as *bona fide* boundaries [38].

The description of an environmental scene can be considered as a one-dimensional semantic time-line [24]. This timeline is rhythmed by the succession of sentences, forms, landmarks and relations that structure the environment. This reflects the fact that space is structured through the use of periodic, recognizable signs and rhythmed by similarities and observable changes [7].

From a conceptual point of view, two definitions of space coexist. On the one hand, Newton argues that the existence of space is independent from the existence of entities composing it. On the other hand, Leibniz argues that space is purely relative, never empty and that it is defined by the entities composing it. Our position is close to the latter as we argue that an environmental scene cannot be defined without specification of the location of entities composing the environmental scene, and the relationships that relate them.

Observable entities are materialized by geographical lexemes, *i.e.*, abstract units of morphological analysis in linguistics that correspond to a set of forms taken by a single word [3]. These units include vegetal entities such as meadows or forests, structural and geomorphological features such as mountains or valleys, water bodies such as lakes or marshlands, and human-made features such as roads or buildings. These geographical entities are closely related to common-sense concepts. Geographical entities that are commonly used in everyday

discourse are essential for the apprehension of our daily natural environment. They often appear in cartographic symbols, but geographers or cartographers do not always provide a formal definition of them. Since they are mainly used for the description of an application domain, their physical existence cannot be questioned [37]. However, the formalization of these concepts requires an ontology in order to clearly reflect their semantics.

Ontology is the branch of metaphysics that addresses the nature of spatial characteristics of beings and things that exist [45]. Conceptual models applied to information systems provide a more pragmatic meaning defining ontologies by their main functions as "a specification of a conceptualization" [12]. One of the interests of these approaches is to infer facts that were previously unknown by the system [13]. The main objective of these approaches is to precisely describe the concepts of the application, and to formalize them through a theoretical framework, in order to convey the semantics and syntax appropriate for a specific domain or application. The design of an application ontology consists in a systematic description of the features that characterize this application, and the relations between them and the features specific to the domain.

Since GIS databases rest on ontological commitments [9] and the descriptions of landscapes need to be semantically well-structured, we associate a semantic meaning to the terms of the description using an ontology. The construction of our application ontology is based on the knowledge resulting from a taxonomy derived from a topographic database provided by the French Institut Géographique National. This vector database covers the French territory, at a scale of 1:10 000. One third of the collected descriptions are used for the design of the application ontology. A top - down approach is applied for the modeling process, since the entities identified by the description, and referencing geographical entities of the environment are associated to semantic categories, either anthropomorphic or natural [17]. The application ontology is also composed of spatial relations that are categorized by their type, *e.g.*, topologic, orientation and distance (Fig. 3).

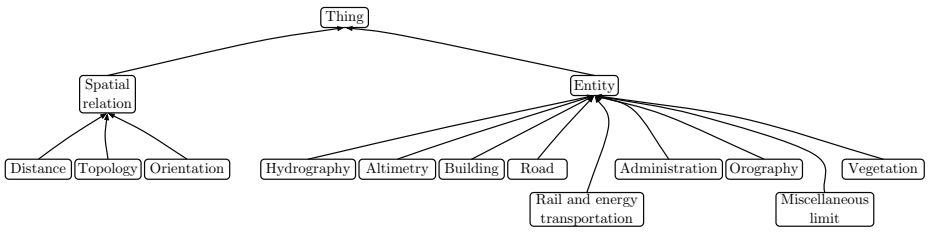

Fig. 3 Top level concepts of the application ontology

5 Schematization of a verbal description

We introduce a schematization approach the objective of which is to facilitate the understanding of the concepts and spatial structures that emerge from the verbal description of a scenery. A verbal description is made of a corpus of sentences. A sentence contains several concepts associated using relations that underline the structure of the scene. We characterize a sentence by concepts, *i.e.*, forms and landmarks modeled as entities, and relationships that relate them. We consider that an environmental scene is ordered by two orthogonal dimensions: the rhythm produced by the timeline of the sentences generated by the observer [16],

and proximity spaces that locate the entities relatively to his/her location. The rhythm of the verbal description is modeled by a *linguistic view*, and proximity spaces by a *proximity spaces view*. The approach is completed by a *semantic view* that takes into account the ontological characteristics of the represented scene, and directional properties modeled using a *directional cones view*. The principles and properties of these views are developed in the following subsections.

5.1 Linguistic view

The linguistic view provides a semantic schematization oriented to the modeling of the properties of the entities that appear in a verbal description. Concepts and spatial relations are represented by their corresponding lexemes, and associated to semantics derived from the application ontology. The linguistic view combines a conceptual diagram with a framework structured by proximity spaces and the rhythm of the description.

The linguistic view gives a representation of the verbal description of the environment, composed of entities related by relationships, and associated to lexemes of the verbal description (Fig. 4). This view is structured by two dimensions, the temporal and spatial ones. The temporal dimension is given by the ordering of the entities and relationships, and the ordering of sentences represented by the vertical bars that illustrate the end of a sentence and the beginning of another one. The spatial dimension is delineated by the limits of the proximity spaces. The linguistic view models the entities identified in the verbal description according to the temporal ordering of the sentences in which they appear, and materializes the spatial relations between them. This timeline reflects a cognitive order of importance, *i.e.* important or salient entities are firstly perceived and quoted [22,28].

The semantics exhibited by the verbal description of a scenery can be closely schematized by a musical score composed of several spatial entities the layout of which constitutes rhythm. According to Bar Yosef, the properties of the musical time organizations and the space concepts are analogous when time is conceived as an axis that transforms time organization and space concept into an analogy between two spaces, the first one-dimensional and the second two- or three-dimensional [2]. In the eyes of what is achieved in the musical domain, the interest of such a metaphor is to better understand the phenomenon of rhythm of a verbal description, and its harmony that reveals the respective role of the entities of the scenery. This representation can be analyzed along a melodic dimension that follows the timeline. This should favor the study of the entities that compose a verbal description and the spatial arrangements that emerge.

The representation of a verbal description by a linguistic view is achieved by a semi-automated data processing that retains the sentence structure. This process is achieved with the Tinki parser, a semi-automatic parser used in natural language processing [26]. The text resulting from the verbal description of an environmental scene is filtered in order to keep the relevant information only. Let us consider the example of the sentence *"I'm on a footpath that runs along a castle that was certainly constructed during the Middle Ages"*. Since the noun phrase qualifying the *"castle"* is not directly pertinent for our modeling purposes, it is not retained. A linguistic analysis is applied, taking into account co-references. A co-reference corresponds to a relation between a pronoun and its antecedent. Let us consider the following sentence *"On the left, we can see a castle and a valley on the right of it"* which becomes *"On the left, there is a castle. A valley is on the right of the castle"* after processing. This representation outlines the entities quoted several times in the description (or entities with initial co-references such as the one labelled *"castle"* in the previous example), this

reinforcing their role in the scene description. This confirms previous studies that outline the fact that linguistic salience is generally linked to physical or visual salience [22,4].

Once the linguistic process is achieved, the identified entities are associated to one-to-many proximity spaces. This implies a specification of the distance and directional relations that qualify the location of entities. Let us consider the example *"I'm on a footpath. In front of me, there is a little valley with the castle on the left of it and at the horizon, I can distinguish a mountain range"*. As the *space of the body* corresponds to the area where entities can be directly apprehended, the *"footpath"* is directly located in the *space of the body*. On the contrary, as language can generate several interpretations, the *"valley"* and the *"castle"* can be located into either the *experienced* or *distant proximity spaces*. Lastly, specific terms such as *"at the horizon, I can distinguish"* illustrate specific cases where the entities are located in the *space at the horizon*. Since the application ontology also integrates distance relations and the associated linguistic terms used by the panel of participants, the allocation of a particular proximity space to an entity is semi-automatically computed.

The principles of the modeling approach being introduced, we hereafter present its formal representation. Let $\mathscr{P}$ be the set of sentences composing a verbal description, $\mathscr{D}$ the set of verbal descriptions, $\mathscr{U}$ the set of elementary units composing a sentence, $\mathscr{E}$ the set of entities of an environmental scene, and $\mathscr{R}$ the set of spatial relations including the null element $\emptyset$. A verbal description D is formalized as an ordered set of sentences $p_i \in \mathscr{P}$, *i.e.*, $D = [p_1, p_2, \ldots, p_n]$ where D $\in \mathscr{D}$ and n ≥ 1. A sentence p_i is an ordered set of elementary units $u_i \in \mathscr{U}$, *i.e.*, $\forall i \in [1, \ldots, m]$ with $m \geq$ n, $p_i = [u_1, u_2, \ldots, u_m]$. An elementary unit u_i is a triplet such as $u_i = [e_j, r_k, e_l]$ with $e_j, e_k \in \mathscr{E}$, and $r_k \in \mathscr{R}$.

As the observer implicitly acts as a spatial reference, he/she refers to his/her location at least once during the description of the surroundings, *i.e.*, $\forall D \in \mathscr{D}, \exists e_i \in \mathscr{E}, e_i = observer$. Three classes of spatial relations are used by the panelist to locate entities. Let $\mathscr{T}$ = {directional, distance, topological} be the set of possible types for the relations, then the function $f_{relation}$ that associates a type to a relation is given by

$$\begin{aligned} f_{relation} : \mathscr{R} \setminus \{\emptyset\} &\rightarrow \mathscr{T} \\ r_i \quad &\mapsto type \end{aligned} \tag{1}$$

Spatial entities are linked by relationships, the function h that models the relation r_k associating two entities e_i and e_j is given by

$$\begin{aligned} h : \quad \mathscr{E}^2 \quad &\rightarrow \mathscr{R} \\ (e_i, e_j) &\mapsto r_k, \text{such as } [e_i, r_k, e_j] \in \mathscr{U}. \end{aligned} \tag{2}$$

Let $\mathscr{S}$ = {body, acc, dist, hor} be the set of proximity spaces, $\mathfrak{P}(\mathscr{S})$ the power set of $\mathscr{S}$ *i.e.* the set of subsets of $\mathscr{S}$, and $e_i \in \mathscr{E}$. The function $f_{space}(e_i)$ that locates the entities is given by

$$\begin{aligned} f_{space} : \mathscr{E} &\rightarrow \mathfrak{P}(\mathscr{S}) \setminus \{\} \\ e_i &\mapsto \{s_i\} \text{ such as } s_i \in \mathscr{S} \end{aligned} \tag{3}$$

With respect to the example given by Fig. 4, we have $f_{space}(footpath) = \{obs\}$ and $f_{space}(castle) = \{acc, dist\}$. It is immediate to note that at most four proximity spaces can be associated to a given entity, *i.e.*, $\forall i, 1 \leq Card(f_{space}(e_i)) \leq 4$ where $Card()$ is the cardinality operator. Also, when an entity is located over several proximity spaces, this entity fulfils a contiguity constraint defined as follows

$$2 \leq Card(f_{space}(e_i)) \leq 4 \Rightarrow \bigcap f_{space}(e_i) \neq \emptyset \tag{4}$$

Let us denote $f_{occurrence}(e_i, j)$ the function that returns the number of occurrences of an entity e_i in the sentence j

$$\begin{aligned} f_{occurrence} : \mathscr{E} * \mathbb{N}^* &\rightarrow \mathbb{N}, \text{ with } \mathbb{N}^* \text{ the set of natural numbers excluding 0} \\ (e_i, j) &\mapsto k \end{aligned} \tag{5}$$

It is also immediate to note that each entity quoted in the verbal description appears at least one time in a sentence, *i.e.* $\forall e_i \in \mathscr{E}, \exists j \in \mathbb{N}$ such as $f_{occurrence}(e_i, j) \geq 1$. Finally, $f_{sentence}$ is defined as the function that associates a set of ordering sentence numbers $\{j\}$ to an entity e_i. A non-empty set corresponds to the presence of the entity e_i in the sentence(s) $\{j\}$

$$\begin{aligned} f_{sentence} : \mathscr{E} &\rightarrow \mathfrak{P}(\mathbb{N}) \\ e_i &\mapsto \{j\} \text{ such as } f_{occurrence}(e_i, j) \geq 1. \end{aligned} \tag{6}$$

The verbal description of Fig. 4 has three sentences with $f_{sentence}(footpath) = \{1\}$ and $f_{sentence}(castle) = \{1, 2\}$. The constraint $\forall i, 1 \leq Card(f_{sentence}(e_i)) \leq n$, where n is the number of sentences of the verbal description, illustrates the fact that a given entity exists at least in one sentence, and at most in all sentences of the verbal description. The example illustrated in Fig. 4 shows evidence of a progressive description from nearby to distant spaces. A landmark, the castle, is present in two different sentences, thus playing a prominent role on the relative location of the other landmarks of the description. This example confirms that the boundaries between the distant and the experienced spaces are difficult to evaluate in most cases.

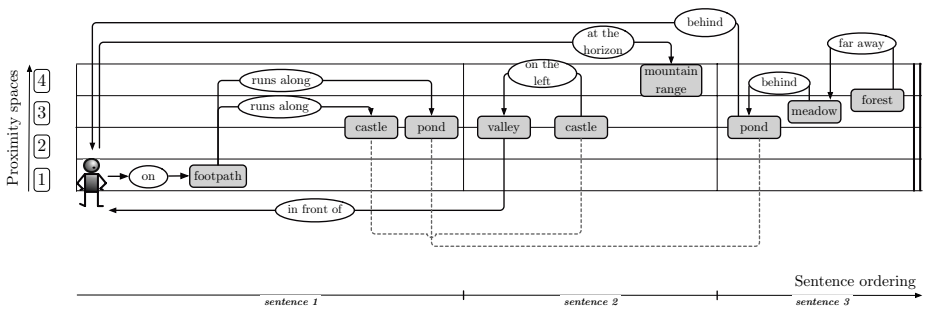

Fig. 4 Linguistic view of a verbal description

5.2 Semantic view

The objective of the semantic view is to provide a summarized representation of the entities and spatial relations identified in a verbal description. A semantic view is defined by a finite and connected graph G where nodes model named entities, and edges the named

relations (Fig. 5). The semantic graph makes the difference between the node referencing the observer, and the ones that reference spatial entities quoted in the verbal description and represented by the linguistic view. More formally, the semantic graph G of the scene description is given by the pair of elements $G = (V, E)$. The elements of V are the nodes of the graph G, the elements of E are the labelled edges. Each triplet u_i corresponds to a subgraph G_i where e_j and e_k are the vertices and r_k the edge.

This view outlines the entities that play a central role in the description (*e.g.*, observer, pond and footpath) and the outliers (*e.g.*, forest). It also reveals the diversity of the terms used for the entities and relations derived from the ontology, and thus the variety of the elements identified in the landscape.

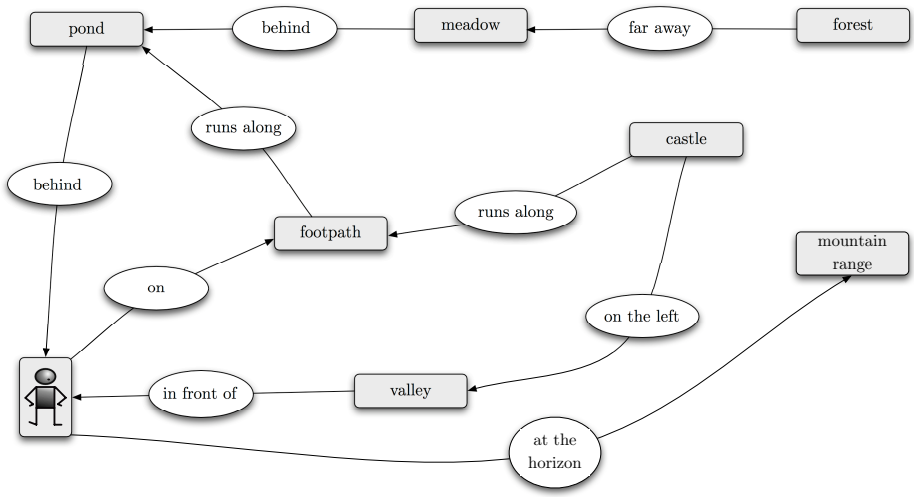

Fig. 5 Semantic view of a verbal description

5.3 Proximity spaces view

The integration of proximity spaces within the semantic view gives another interpretation of a scene description (cf. Fig. 6). Let $\mathscr{E}_{body}$ be the set of entities located in the space of the body, $\mathscr{E}_{exp}$ the set of entities located in the experienced space, $\mathscr{E}_{dist}$ the set of entities located in the distant space, and $\mathscr{E}_{hor}$ the set of entities located in the space at the horizon. The proximity spaces are defined as

$$\begin{aligned}
\mathscr{E}_{body} &= \{e_i \text{ such as } f_{space}(e_i) = \{s_j \text{ such as } \exists s_k \text{ such as } s_k = body\}\} \\
\mathscr{E}_{exp} &= \{e_i \text{ such as } f_{space}(e_i) = \{s_j \text{ such as } \exists s_k \text{ such as } s_k = exp\}\} \\
\mathscr{E}_{dist} &= \{e_i \text{ such as } f_{space}(e_i) = \{s_j \text{ such as } \exists s_k \text{ such as } s_k = dist\}\} \\
\mathscr{E}_{hor} &= \{e_i \text{ such as } f_{space}(e_i) = \{s_j \text{ such as } \exists s_k \text{ such as } s_k = hor\}\}
\end{aligned} \tag{7}$$

The proximity spaces view illustrates a summarized representation of the entities associated to the proximity spaces. Fig. 6 illustrates the prominent role of the distant and

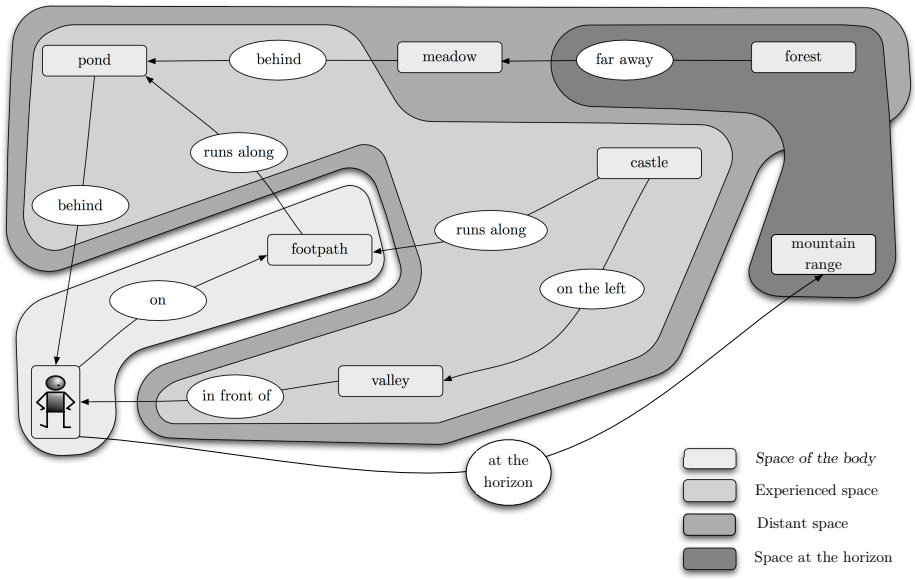

Fig. 6 Proximity spaces representation

experienced spaces in the example considered. It is also worth noting the relations that connect entities between different proximity spaces (*e.g.*, footpath), or within a given proximity space (*e.g.*, pond). Overall, this view provides a representation of the structure of the landscape that results from the verbal description, of the respective importance of the proximity spaces identified, and the relative depth of field of the image of the landscape perceived by the observer.

5.4 Directional cones view

Humans also tend to structure space using bodily directions that relate the perceived entities to their location in space. The way these directional relations organize space implicitly generates a partition of space, often represented using conceptual directional cones [33]. A cone-based partition emerges from the directional relations identified. The proximity spaces view is enriched by the directional relations and an orientation-based structure of space. We retain a cone-based partition with four possible directions : front, back, right, and left (cf. Fig. 8).

A directional cone is detected when at least one directional relation exists between the observer and an entity of the environmental scene. The number of directional cones can vary from two (front - back or right - left) to four (front, back, right and left), depending on the directional relations used between the observer and an entity or a group of entities to describe the surroundings. More formally, let $<_t$ be a temporal ordering operator of the terms of the description, *i.e.* entities and relations of elementary units, $\mathscr{E}_{front}$ the set of entities located in the directional cone in front of the observer, $\mathscr{E}_{back}$ the set of entities located in the back cone, $\mathscr{E}_{left}$ the set of entities located in the left cone, $\mathscr{E}_{right}$ the set of entities located in the right

cone, and x a directional relation such as x $\in \{front, back, left, right\} \subset \mathscr{R}$. We define $\mathscr{E}_x$ such as

$$\mathscr{E}_x = \{e_m \text{ such as } e_n <_t e_i \leq_t e_m \text{ with } (f_{relation}(h(e_n, e_i)) = directional \wedge \\ e_n = observer) \wedge ((h(e_n, e_i)) = x\} \quad (8)$$

Since the application ontology makes the distinction between the directional relations and the others, the process of detection of a new directional cone is automatically done when a directional relation is specified between the observer and an entity or a group of entities. The user should finally manually adjust the number of directional cones and their composition. The example of Fig. 7 illustrates the roles played by directional relations in the verbal descriptions. It appears here that directional relations tend to reference distant entities in order to make their location more precise. Conversely, they are barely used for nearby entities.

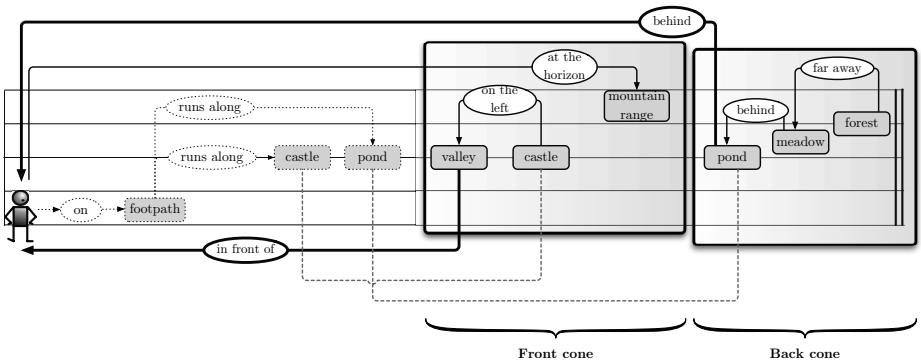

Fig. 7 Directional cones search

This directional-based view clearly makes a difference between the spaces in front and behind the observer. It provides another characterization of the verbal description. This allows a preliminary spatialization of the identified entities thanks to an observer-centered structure of the scenery. The scene that results from the perception of an environmental scene is then structured by the proximity spaces, and the directional cones that provide an observer-centered reference for the location of entities.

6 Conceptual map of an environmental scene

The successive views provide a semantic foundation for the derivation of a conceptual map that should reconcile them within a representation that integrates the proximity spaces, directional cones, entities, relationships and sentences of the verbal description. We define such a "conceptual map" as an abstraction that characterizes the mental mapping of a verbal description. It should provide a bridge between the linguistic characteristics of a verbal description, its mental abstraction, and the semantics of the scene observed.

Fig. 9 gives an example of such a conceptual map where space is structured into two areas, *i.e.* the front and back of the observer. The conceptual map that results from this

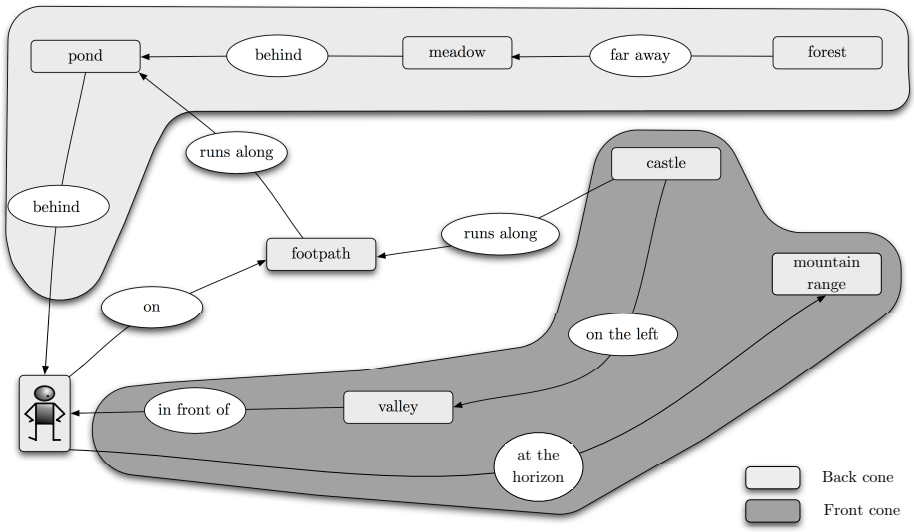

Fig. 8 Directional cones representation

example shows evidence of a close relationship between the different sentences described and the entities identified. This reflects and generates a sort of continuity of the description, this being reflected by the intertwining of the salient features that emerge.

These conceptual maps allow comparison of descriptions either made by different observers, or resulting from different landscapes. Let us consider another example of verbal description, the one of an observer walking along a seacoast of Brittany, France and illustrated by Fig. 10: *"Behind me, there is a huge lighthouse and four houses nearby. I'm on a footpath, and on my right there is a grassy hill. On my left is the Ocean"*. The conceptual map that emerges from this verbal description shows a clear separation between the different parts of the scene. The observer mainly uses bodily-centered directions to qualify the location of the entities such as the *"footpath"*, the *"lighthouse"* and the *"ocean"*. There are a few connections between the entities, this reflecting a landscape with a clear separation between the different entities identified by the observer, the sentences and the entities, and the cone-based regions of the scene. It is also worth noting the small number of entities identified, this illustrating the relative flatness of the landscape. Last, the *"lighthouse"* clearly denotes a salient role as several entities are spatially related to it.

7 Conclusion

The research presented in this paper introduces a language-based and cognitive approach that models the verbal description of a landscape scene. A verbal description is modeled by a qualitative, structural and proximity-based representation that reflects its semantics. The structural properties of a verbal description are linked to several semantic views that attempt to reflect the user's perception, and rely on a semantic-based analysis. Semantics, proximity-spaces and directional cone-based views together provide a step towards a *"conceptual map"*, conceived as an abstraction that characterizes an environmental scene from its verbal description. The model is based on the spatial structure, the rhythmic organization

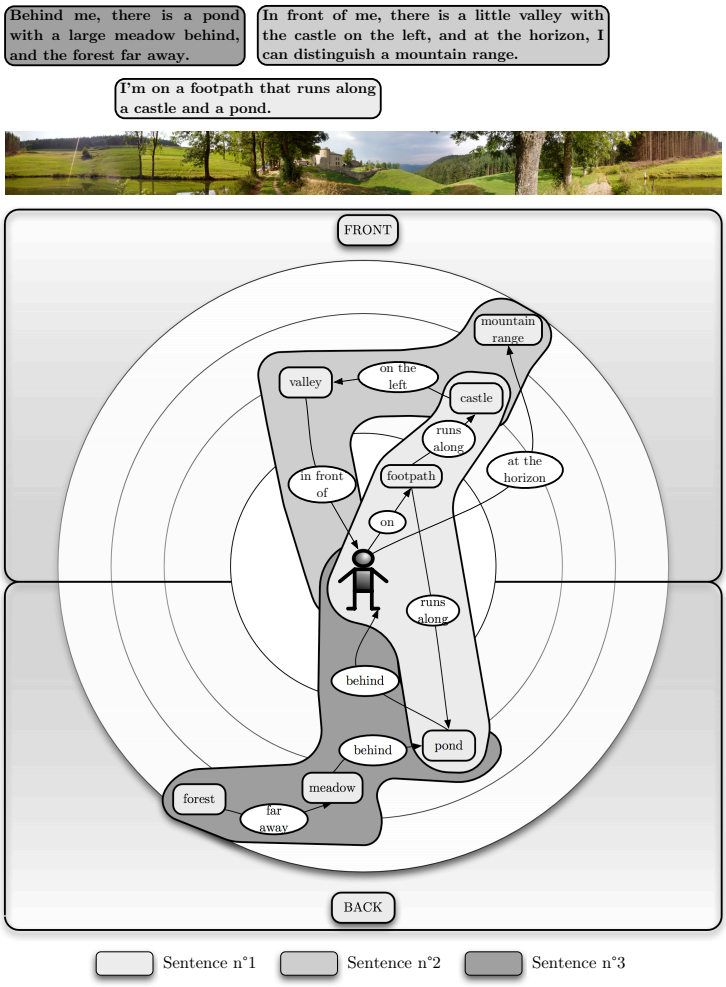

Fig. 9 Conceptual map of a semi-natural landscape

of space, the diversity, ordering and salience of the entities identified by the observer, and reflect the main characteristics and entities of the environment. Such a model qualifies and characterizes natural landscapes, and provides a framework for the analysis of the properties of the verbal descriptions made by different observers, and cross-comparisons of different landscape descriptions. Although experienced in natural landscapes, this modeling approach could potentially be applied to urban environments.

Further work concerns an evaluation of the properties of the salient entities identified in a description of a given landscape and the design of a prototype that should provide a preliminary resource for the development of a mapping between a scenery description of an observer and a conventional GIS representation. It should provide a support for the georeferencing of the observer from the analysis of the salient and structural components of a scenery description.

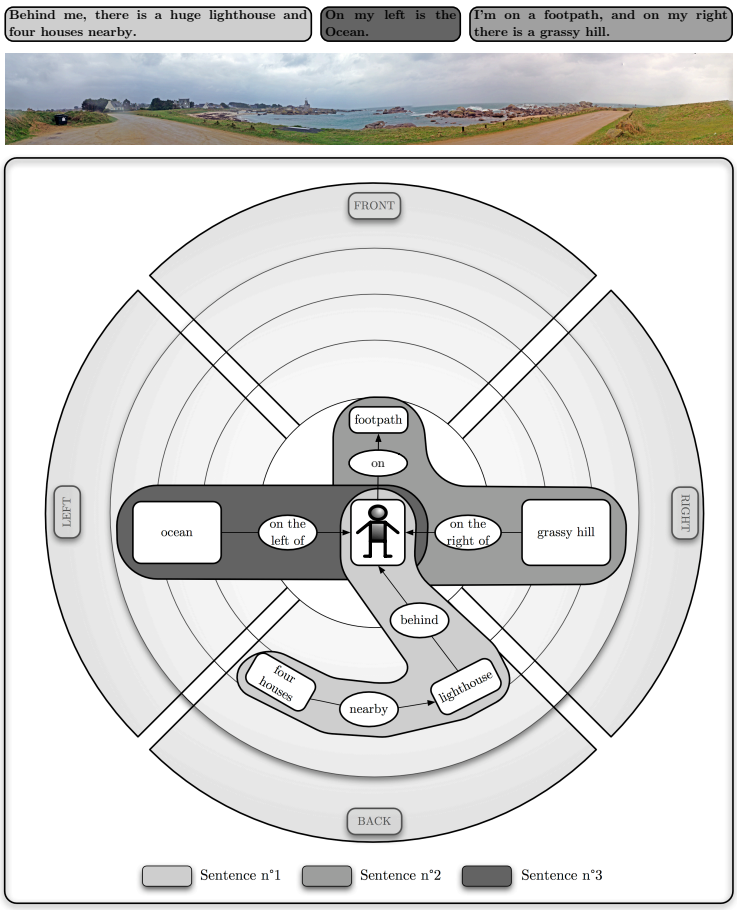

Fig. 10 Conceptual map of a maritime environment

Acknowledgements The authors gratefully thank the anonymous reviewers for their helpful comments, and Alcino Ferreira for his careful language reviewing and comments.

References

1. Appleyard, D., Gerson, M., Lintell, M.: Livable Streets. University of California Press, Berkeley (1981)
2. Bar-Yosef, A.: Musical time organization and space concept: A model of cross-cultural analogy. Ethnomusicology **45**(3), 423–442 (2001)
3. Beard, R., Volpe, M.: Handbook of Word-Formation, *Studies in Natural Language and Linguistic Theory*, vol. 64, chap. Lexeme-Morpheme Base Morphology, p. 464. Springer Netherlands (2005)
4. Carlson, L.A.: On the "whats" and "hows" of "where": The role of salience in spatial descriptions. In: C. Freksa, N.S. Newcombe, P. Gärdenfors, S. Wölfl (eds.) Spatial Cognition VI. Learning, Reasoning, and Talking about Space, *LNCS*, vol. 5248, pp. 4–6. Springer Heidelberg (2008)
5. Chevriaux, Y., Saux, E., Claramunt, C.: A landform-based approach for the representation of terrain silhouettes. In: C. Shahabi, O. Boucelma (eds.) ACMGIS '05: Proceedings of the 13th annual ACM

international workshop on Geographic Information Systems, pp. 260–266. ACM Press, New York, NY, USA (2005)
6. Couclelis, H.: Worlds of information: the geographic metaphor in the visualization of complex information. Cartography and Geographic Information Systems **25**(4), 209–220 (1998)
7. Deleuze, G., Guattari, F.: Mille Plateaux. Critique. Editions de Minuit (1980)
8. Egenhofer, M., Mark, D.: Naive Geography. In: A.U. Frankand, W. Kuhnand (eds.) COSIT'95: Conference on Spatial Information Theory, *LNCS*, vol. 988, pp. 1–15. Springer Heidelberg (1995)
9. Frank, A.U.: Spatiotemporal Databases: The Chorochronos Approach, chap. Ontology for Spatio-Temporal Databases, pp. 9–78. Springer-Verlag (2007)
10. Freksa, C.: Using Orientation Information for Qualitative Spatial Reasoning. In: A.U. Frank, I. Campari, U. Formentini (eds.) Proceedings of the International Conference GIS - From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning on Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, *LNCS*, vol. 639, pp. 162–178. Springer Heidelberg (1992)
11. Granö, J.: Reine Geographie. No. 2 in 202. Acta Geographica (1929). Tr. as Pure Geography, reprinted in English from Johns Hopkins Press in 1997
12. Gruber, T.: A translation approach to portable ontology specifications. Knowledge Acquisition **5**(2), 199–220 (1993)
13. Guarino, N.: Formal ontology in information systems. In: N. Guarino (ed.) Formal Ontology in Information Systems, vol. 46, pp. 3–15. IOS Press, Trento, Italy (1998)
14. Herskovits, A.: Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English. Cambridge University Press (1986)
15. Hirtle, S., Jonides, J.: Evidence of hierarchies in cognitive maps. Memory & Cognition **13**(3), 208–217 (1985)
16. Hohne, H., Coker, C., Levinson, S., Rabiner, L.: On temporal alignment of sentences of natural and synthetic speech. Acoustics, Speech and Signal Processing, IEEE Transactions on Acoustics, Speech, and Signal Processing **38**(12), 807–813 (1983)
17. Institut Géographique National: BD TOPO : version 2. Descriptif de contenu (2008). URL `http://professionnels.ign.fr/DISPLAY/000/506/560/5065604/BDORTHO_specification.pdf`
18. Johnston, R.: Geographical Information Systems: Principles, Techniques, Management and Applications, vol. 1, chap. Geography and GIS, p. 404. Wiley and Sons, Inc (1999)
19. Kaplan, S.: Perception and Landscape: Conceptions and Misconceptions. In: G.H. lsner, R.C. Smardon (eds.) roceedings of our national landscape: a conference on applied techniques for analysis and management of the visual resource, pp. 241–248. Berkeley, CA. (1979)
20. Kuipers, B.: Modeling spatial knowledge. Cognitive Science **2**(2), 129–153 (1978)
21. Kuipers, B.: The "map in the head" metaphor. Environment and Behavior **14**(2), 202–220 (1982)
22. Landragin, F.: La saillance comme point de départ pour l'interprétation et la génération. Journée d'étude de l'Association pour le Traitement Automatique des Langues (ATALA) (2003)
23. Le Yaouanc, J.M., Saux, E., Claramunt, C.: A semantic and language based model of landscape scenes. In: M.P. Il-Yeol Song (ed.) ER 2008 Workshops, *LNCS*, vol. 5232, pp. 334–343. Springer-Verlag (2008)
24. Levelt, W.: Linearization in describing spatial networks. Processes, Beliefs, and Questions. Essays on Formal Semantics of Natural Language and Natural Language Processing pp. 199–220 (1982)
25. Levinson, S.: Frames of reference and Molyneux's question: cross-linguistic evidence. Language and Space pp. 109–169 (1996)
26. Ligozat, G., Nowak, J., Schmitt, D.: From language to pictorial representations. In: Wydawnictwo Poznańskie (ed.) Proc. of the Language and Technology Conference. Poznan, Poland (2007)
27. Lynch, K.: The Image of the City. M.I.T. Press, Cambridge (1960)
28. Mainwaring, S.D., Tversky, B., Ohgishi, M., Schiano, D.J.: Descriptions of simple spatial scenes in English and Japanese. Spatial Cognition and Computation **3**(1), 3–42 (2003)
29. Meitner, M.: Scenic beauty of river views in the Grand Canyon: relating perceptual judgments to locations. Landscape and Urban Planning **68**(1), 3–13 (2004)
30. Montello, D.: The geometry of environmental knowledge. In: A.U. Frank, I. Campari, U. Formentini (eds.) Proceedings of the International Conference GIS - From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning on Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, vol. 639, pp. 136–152. Springer-Verlag, London, UK (1992)
31. Montello, D.: Scale and multiple psychologies of space. In: A.U. Frank, I. Campari (eds.) Spatial Information Theory: A Theoretical Basis for GIS, *LNCS*, vol. 716, pp. 312–321. Springer Heidelberg (1993)
32. Muller, P.: Eléments d'une théorie du mouvement pour la formalisation du raisonnement spatio-temporel de sens commun. (1998). Unpublished PhD report, Université Paul Sabatier de Toulouse
33. Peuquet, D.J., Ci-Xiang, Z.: An algorithm to determine the directional relationship between arbitrarily-shaped polygons in the plane. Pattern Recognition **20**(1), 65–74 (1987)

34. Shafer, E., Richards, T.: A comparison of viewer reactions to outdoor scenes and photographs of those scenes. In: D. Canter, T. Lee (eds.) Psychology and the Built Environment, pp. 71–79. John Wiley (1974)
35. Shuttleworth, S.: The use of photographs as an environment presentation medium in landscape studies. Journal of Environmental Management **11**(1), 61–76 (1980)
36. Smith, B., Mark, D.: Geographical categories: An ontological investigation. International Journal of Geographical Information Science **15**(7), 591–612 (2001)
37. Smith, B., Mark, D.M.: Do Mountains Exist? Towards an Ontology of Landforms. Environment and Planning B (Planning and Design) **30**(3), 411–427 (2003)
38. Smith, B., Varzi, A.: Fiat and bona fide boundaries: Towards an ontology of spatially extended objects. In: S.C. Hirtle, A.U. Frank (eds.) COSIT'97: Conference on Spatial Information Theory, *LNCS*, vol. 1329, pp. 103–119. Springer Heidelberg (1997)
39. Sorrows, M.E., Hirtle, S.C.: The nature of landmarks for real and electronic spaces. In: C. Freksa, D.M. Mark (eds.) COSIT'99: Conference on Spatial Information Theory, LNCS, pp. 37–50. Springer Heidelberg (1999)
40. Stewart, T., Middleton, P., Downton, M., Ely, D.: Judgments of photographs vs. field observations in studies of perception and judgment of the visual environment. Journal of Environmental Psychology **4**, 283–302 (1984)
41. Taylor, H.A., Tversky, B.: Perspective in spatial descriptions. Journal of Memory and Language **35**(3), 371–391 (1996)
42. Tolman, E.: Cognitive maps in rats and man. Psychological Review **55**(4), 189–208 (1948)
43. Tversky, B.: Cognitive maps, cognitive collages and spatial mental models. In: A.U. Frank, I. Campari (eds.) COSIT'93: Conference on Spatial Information Theory, *LNCS*, vol. 716, pp. 14–24. Springer Heidelberg (1993)
44. Tversky, B., Lee, P.U.: How space structures language. Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge **1404**, 157–175 (1998)
45. Williams, D.: On the elements of being. Review of Metaphysics **7**, 3–18 (1953). Reprinted in Metaphysics: The Big Questions, 1998