In-Situ Communication and Labeling of Places

Falko Schmid[1] and Colin Kuntzsch[2]

Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition,
University of Bremen, P.O. Box 330 440, 28334 Bremen, Germany
[1]schmid@sfbtr8.uni-bremen.de
[2]kuntzsch@informatik.uni-bremen.de

Abstract. An increasing number of applications is based on the concept of *personally meaningful places* detected in individual trajectory data. This class of applications will only be accepted if the individual effort of labeling places is reduced to a minimum and the proposed place labels are meaningful. To allow for either the automatic generation of place concepts and corresponding labels, or the maintenance of a collaborative place database, we require better understanding of how people conceptualize their familiar environments in-situ, thus while being *in* the place. In this paper we present the results of an in-situ place labeling study. Our results suggest that the harmonization of diverse labels for personally meaningful places is possible and the operators for the automatic computation of place names are seizable.

1 Introduction

Place is recognized as a central concept in geographic information sciences. The literature on place is rich and the history of trying to tackle the essentials of place is long. But as soon as we want to operationalize place for applications, we recognize that there are no obvious operators we can use, even the most basic concepts are unclear and far from empirical evidence. For computational purposes, the literature on place is as unseizable as the concept itself seems to be in general. Place is usually understood as a conceptual partitioning of space. However, there is nothing like a prototypical place or size of a place, as the partitioning is applied on many granularities of space, from the earth and beyond to a corner in a room and below. According to Relph (1976), place consists of three components: physical setting, thus the locale of a place, activities performed at a place and the meanings of a place to the public and the individual. This stresses the commonsense that place is not just an address or a point drawn on a map - place is always a construct created by the interplay of actual environmental setting, individual and public, as well as the experiences and activities established at the spatial partitioning. This theory is taken a step further towards operationalizability by earlier work of Lynch (1960). In his work he identified the structural elements in urban environment that influence the creation of place concepts. In the areas of space syntax (see e.g. Hillier & Hanson (1984)) the syntactic construction of neighborhoods (which are assumed to cover larger regions than places (see Cresswell (2004); Agarwal (2004, 2005)) is proposed by Dalton (2007). An individual perspective of place offers Seamon (via Cresswell (2004)). Seamon proposes spatio-temporal construction and anchoring of places by "everyday movement in space", i.e. routines or habits that anchor meaning to physical places. Following Seamon, places are the result of movement and activities performed at locations. Although not explicitly investigating place, Hagerstrand (1970) created the foundations of time-geography as a toolkit to analyze spatio-temporal life paths, the containers of individually constructed places. With the increasing availability of GPS sensors, researchers got more and more interested in analyzing trajectories according to the spatio-temporal patterns as suggested by Seamon and Hagerstrand. Especially in the fields of Location Based Services there are many application scenarios grounded in a personal experience of space. Marmasse (1999) propose a place detection to prompt users with location based To-Do lists. Ashbrook & Starner (2003) propose a place detection algorithm to forecast future locations based on past place visits. Liao et al. (2007) developed a framework to infer routines from GPS data to support cognitively impaired persons in public transportation systems. Bicocchi et al. (2007) develops an automated travel diaries. In Schmid & Richter (2006), we developed a fine grained place detection algorithm for familiarity estimation for personalized wayfinding assistance (Schmid (2008)). These examples stress the need to detect "personally meaningful places" to create personalized Location Based Services. They usually propose that users have to label the detected place accordingly. Hightower (2003) proposes to automatically generate labels for places, as manual approaches would not scale. Both positions are critical for applications: manual labeling will under circumstances not scale (as not every possible place will be labeled) and the assigned labels will not be meaningful to everybody. When we have a look at studies about spatial communication, we can notice the strong dependency from intention, mutual level of familiarity, and context on the choice of labels. In Weilenmann & Leuchovius (2004) the authors report on a study of analyzing mobile phone calls with respect to spatial communication. E.g., to describe their current location, subjects chose different levels of granularities and referred entities. The particular choice is depends on needs to obfuscate the real location or to clarify a location by a personal place of mutual knowledge (e.g. "I am home", "the place we met last time"). Duckham & Kulik (2005) propose a formal model for place obfuscation for Location Based Services, such a service requires transparent meaningful place names if a user needs to control the spatial extends of the obfuscation model. Weilenmann and Leuchovius also noticed, that subjects switched the description for the same place or made use of comparably rough descriptors without problematic consequences.

Zhou et al. (2005a) report about a diary study of subjects keeping track of the places they visited. Zhou et al. identified that peoples' communication practice highly depends on the purpose of communication, the mutual familiarity or the assumed familiarity with an area, as well as privacy issues. Although there seems to be good understanding of what a place is and how it can be described unambiguously on different levels of granularity, other studies show a high individual heterogeneity. When ever we use the term *heterogeneity* in this paper, it has always to be understood as the selection of *different* labels for the *same* or a *similar* place [1]. In Lovelace et al. (1999) tried to identify which spatial entities people refer to in familiar and unfamiliar routes. An interesting observation of this study was the strong heterogeneity of the addressed landmarks across the subjects. 31 subjects mentioned 119 different landmarks along the unfamiliar route, but only 16 had been mentioned by 30% or more of the subjects. It is worth to mention that the study took place on a university campus, a highly structured environment with a usually common spatial vocabulary. This study is a good example for heterogenous naming of spatial entities. [2] People are usually able to find a common name for a place, but so far it is unclear which entities they address when they describe a place in-situ (all place studies known to the authors are ex-situ or do test multiple subjects under the same conditions) and up to which degree this process can be automated.

2 Motivation

Personalized wayfinding assistance (PWA) as introduced by Schmid (2008, 2009) requires place labels to describe routes within a personal frame of reference. PWA is based on a spatial familiarity estimation by analyzing movements of users. By analyzing the trajectories with the place detection approach in Schmid & Richter (2006) a spatial user profile, consisting of meaningful places and paths is compiled. By means of the profile it computes routes along personally meaningful places and paths to generate cognitively ergonomic wayfinding assistance. In order to generate meaningful assistance, PWA requires labels (i.e. names) and spatial concepts (e.g. spatial extends, borders, membership functions, etc.) for the places a user visits and knows. Like in e.g. Marmasse (1999) the user will have to enter place names up to a certain extend. But to minimize the effort for the individual, we can either try to automatically generate place labels, as postulated by Hightower (2003) or we can set up a collaboratively maintained database for place names. In the first case we could just compute a label as it is required, in the latter case previously labeled places can be offered to users visiting the same places. Users can then adopt, reuse, or alter a label without much individual interaction. However, both solutions require a deeper understanding of place naming under ambiguous conditions and with respect to different granularities. When we generate user profiles with GPS sensors and mobile devices,

[1] As places are hard to define, the identity relation, thus the identification of the the same place is most certainly equally hard. We assume a place to be the same place if the same label or concept for a place is expressed at a geographically similar position and has a similar spatial scope.

[2] However, this study is not an in-situ study, all names for landmarks have been gathered with photographs or ex-situ, when the subjects have accomplished the given tasks. Ex-situ always incorporates a certain degree of reflection, a process where places are mentally pre-selected according to "mental" salience. We can assume that photographs bias the selection of references, as well as the ex-situ labeling will result in different results than asking "Where are you at the moment?".

we always have to face uncertainty. The usual positioning accuracy of mobile devices (carried e.g in pockets) are far from unambiguity. I.e., the computed places are not necessarily situated at the sensed location and a label would be attached to a wrong place or queried for the wrong position. Furthermore, places are extended spatial regions, but so far it is unclear how large these regions are in the context of "personally meaningful places" as required by a number of applications. It is further unclear what labels people choose for places in their familiar environment without being in an explicit communication situation. It is unknown how homogeneous or heterogeneous the set of assigned labels across multiple persons is and if they can be harmonized by a possibly higher-level concept of a place.

2.1 Understanding of Place

As indicated in Section 1, places exist on any scale and granularity and the selection of a particular concept is highly context dependent. In this paper we constrain ourselves on places as communicable (urban) units within Location Based Services requiring the widely used notion of "personally meaningful places". Intuitively, we have an idea what such a place can be. However, computers as the backbones of Location Based Services need operators and functions to harmonize heterogeneous labels for such places, in order to generate labels that are meaningful for humans, and all this on a granularity which is plausible in the given application context (like PWA, diaries, meeting assistants, etc). Only if we know more about how people conceptualize familiar places while they are in-situ (many scenarios require this condition, either for labeling of places or the location dependent communication of place knowledge) we can develop applications that simulate this conceptualization and foster the semantic access to spatial information.

2.2 Related Work

The aimed granularity is finer than existing approaches of modeling and generating places from available data. Up to the knowledge of the authors, there exist no approach for the computation of place names for a dense urban environment. Grothe & Schaab (2008) propose a machine learning approach for the identification of spatial footprints, thus shape approximations, for regions based on available Geo-tagged resources. The approach is interesting to identify concepts of large scale geographic regions. It requires a relatively large amount of tagged media and consequently will work well for popular places or regions. However, for all places without tags, the approach will fail. Unfortunately, most places on the granularity we address will not be tagged (e.g. living areas), whilst other will be tagged by plenty of users (e.g. touristic places). Schockaert et al. (2005) propose an automatic method to derive fuzzy spatial footprints by consulting gazetteers with complex phrases triggering constraint analysis including bordering regions. Twaroch et al. (2008) investigate on the mining of cognitively plausible place names from social networks to create alternatives in gazetteers. However, these approaches do not investigate on places that are on the granularity of personally meaningful places. They are either on the granularity of large geographic regions (e.g. alps), districts within cities, or few very specific (commonly known) places. None of the approaches do consider the in-situ labeling of places and how it relates to the surrounding environment. Additionally, they to not try to foster a bottom-up labeling and concept construction approach. This means it does not guide labeling constructively and is not flanked with additional analysis of the underlying geographic structures. It is known that spatial structures influence the conceptualization of space and it is possible to operationalize spatial properties (e.g. Lynch, 1960; Hillier & Hanson, 1984). Dalton (2007) utilizes these syntactic properties to automatically compute regions and borders for neighborhoods. This is an interesting approach to predetermine spatial regions, also on different level of granularities. However, this is a purely syntactic approach with no link to situated human concepts. We still do not know how people really refer to it and which elements they use to anchor a concept in the environment. A synthesis of both worlds, the syntactic analysis of spatial properties and the semantic partitioning and labeling seems to be one promising direction to take in the future.

3 The Study

From the perspective of PWA, we are interested in how people refer to places in-situ in their familiar environment and if similar labels can be automatically computed. However, we could not find evidence on how homogeneous or heterogeneous the selection and the assignment of names for places are across multiple persons, the possible spatial extents, and granularities of places. To move a step towards answering these questions, we set-up an explorative study. The goal of our study was the examination of following hypotheses:

3.1 Hypotheses

1. The location of a person clearly determines the selection of a place name.
2. The labeling of places in familiar environments is homogeneous across people and it is possible to develop computational models of place on one level of granularity (in our case a spatial region of the size of typical positioning uncertainty in dense urban environments).
3. Place labels allow for the computation of semantic higher-level concepts (coarser granularity of a concept of place), which can be utilized in context-aware service configuration and communication.

3.2 Design of the Study

We designed the study to gather place descriptions of subjects being present within a partially familiar environment in order to get insights about where a place is referred to as a place and in the (forced) case of ambiguous situations which place of the possible choices is selected.

Selection of Places We intended to introduce "undefinedness" of the places to be labeled. I.e., we wanted the participants not to label a specific entity like a building, but an area that is plausible in an positioning context of a Location Based Service application. [3] This allowed us to create ambiguity, with the intention to force people to select a reference out of multiple choices. We covered a range of place classes:

- **Places in structured environments**: places usually have clear names and functions. We covered a large area of the Bremen University campus (region A in Figure 1).
- **Places in less structured environments**: places and buildings are not clearly assigned with names or host a well-known functionality. We decided to cover a part of the "Technologie Park" (region B in Figure 1). In this area there are several companies, cafes, restaurants, spin-offs, external research institutes and a museum area.
- **Places in natural/unstructured environments**: such as places in parks or forests. A forest and recreation area is found in region C in Figure 1. Here we selected places containing either natural features (water bodies), infrastructures (bridges) or recreational objects (playground, horse stable, restaurant).

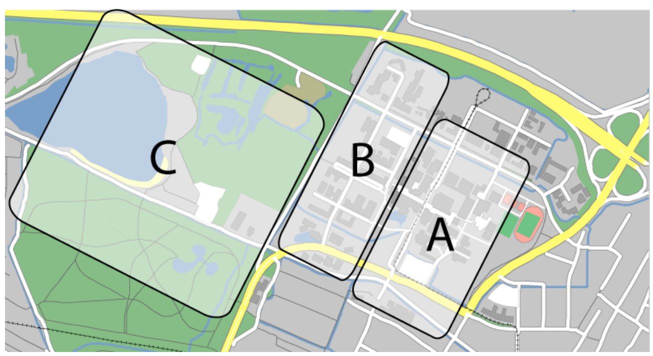

Fig. 1. The different environments within the covered region: A is the main university campus (structured environment), B is a part of the Technologie Park area (semi-structured environment) and C is the recreational area (unstructured environment).

[3] The motivation was to simulate positioning uncertainty within urban environments: when a position is estimated by a e.g. a GPS sensor, it always introduces a certain amount of uncertainty (e.g. 50 meters). When a place is now labeled with a positioning uncertainty of 50 meters, every entity within 50 meters can be the potentially really addressed entity (if we can reduce the label to a single entity at all). If we now add the uncertainty to the query as well (a user queries for a place name and has low positioning quality), the offered place name can be far from the actual location the user is currently at.

3.3 Place Communication

The communication of the pre-selected places is problematic: most modalities either bias subjects by means of the chosen medium (like an annotated map) or distract from the main task by applying navigation and positioning tasks. We decided not to use an annotated map, as we expected influence on the choice of labels. Guiding subjects and placing them at a particular location will bias them to label the pre-selected position.

For these reasons we designed maps without any labels and with no route to follow. For each subsequent pair of places we designed a map with two regions, one for the current place and one for the next place to navigate to. We did not incorporate a route, as the route would determine a particular approaching of a place. The density of the places was relatively high, such that there were only few alternatives at all.

Place Visualization We used regions of different shapes and sizes, each of them included multiple plausible references. The diameters of the regions ranged from 50-100 meters (which is plausible to the positioning uncertainty of low-cost GPS sensors). The subjects navigated freely inside the region and selected and labeled those places they thought the region represents. In the following we will call these regions *place regions*.

Fig. 2. Map cutouts from the place navigation map. The left map with the crosses was the first version of the place communication, the middle and right map the improved version using the place regions as place communicators. The right map contains the subsequent place of the map in the middle.

Place Visualization Modification In the first three trials we annotated the place with a cross, but always stated, that this cross was only marking a fuzzy location. The subjects should place themselves somewhere in the area around the cross at a location they thought they could label properly. It turned out, that this representation strongly determined the choice of labels: subjects tried to position themselves as accurate as possible at the location they thought the cross marked and tried to label this position as exact as possible. This lead to completely unlikely labels (”the second rightmost lantern in front of XY”’). As this problem could be clearly observed for all place-marks and across the first three subjects, we decided to alter the map and not to use the three data sets in our analysis. The usage of regions turned out to work as intended: the subjects now navigated freely in the region and selected places autonomously.

Place Region Modification We had to modify one place region due to an unexpected construction place: region 7.2 (see Figure 3) had to be altered, as the old region 7.1 was not accessible anymore.

Fig. 3. All 13 places which had to be labeled.

3.4 Subjects and Procedure

After discarding the data from the first three trials, we ran the study with 10 subjects, all were either students of higher semesters or scientific staff. They have been familiar with the university campus for three to six years. Both students and staff had a different background of studies and employment (computer science, law, biology, chemical engineering). The subjects walked the course of approximately 3.5 kilometers length in about two hours. The participants walked the course of approximately 3.5 km length in about 2 hours. There are variations as each subject took a slightly different route due to the experimental setup: the subjects carried a folder containing a stack of maps, each illustrating two subsequent regions (see figure 2). After successful navigation to the illustrated place region they had to turn the map to see the next map with the next pair of regions. I.e., they always only knew were they have been to and what the next region would be. They could never optimize their path according to the future regions and select places according to that. When our subjects entered a region, they were asked to select the label that would describe the place best for themselves. They had to place themselves at the position and mark it with a GPS waypoint; the conductor did the same. Each participant was tracked doubly (participant's GPS, conductor). The participants were equipped with a mobile phone containing a GPS tracking software. The conductor had a dedicated GPS device and walked just alongside the subject (estimated distance < 1 meter).

Labeling and GPS Tracking At each place we asked the subjects to answer a set of questions verbally. The conductor wrote the answers to files. Answers have never been corrected, no hints or feedback has been given at any point. The questions at every place asked for

- place names (multiple mentions were allowed)
- neighboring places in vicinity (multiple mentions were allowed)
- judgment of familiarity

After the subject completed the course and labeled all places, they had to answer a questionnaire containing questions about

- demographic information
- the subject of study/profession (determination of familiar area on the campus)
- their usual travel behavior within the regions A, B, C (see Figure 1)
- assumed labeling behavior depending on input modality

3.5 Limitations and Scope of the Study

We are aware that one limiting property of the study is the number of subjects, which can be considered as relatively low. However, when we think of labeling in a realistic setting, there will be only few places with a high number of labels and a large number of places with only very few labels. Insofar, the number of participants reflects something in between.

Another point is the predefinition of our place regions. However a more "natural" design of partitioning of space requires a large number of participants even for very few places. Our study setup allows us to simulate positioning within a certain degree of accuracy at specific locations. As discussed earlier, all of the regions contained entities of different kinds, which also can be expected to result in according heterogeneous labels for the respective regions.

This study is a place labeling study with the aim to identify the potential to harmonize heterogeneous place labels for locations which are potentially the same, similar or neighbored places. From our everyday life we know that people use different names for the same places or the same names for different places. This property causes sometimes some confusion, but is very helpful to talk about places in terms of regions or groups of entities. This property enables a new paradigm of place communication: moving away from the coordinates/distance based concept towards a pre-computation of regions and names. These regions can be addressed by the same label and is still uniquely understandable across groups of persons. However, not much is known about this property so far and we try to investigate on the first basic results covering this question.

4 Analysis

We received a total of 127 primary in-situ place labels (one is missing due to the construction place, two due to undetected navigation errors) of 10 subjects at 13 places. We call the first mention of a label for a place "primary", as we assume that this is the label subjects would instinctively use to describe a place (remember that we allowed multiple labels). All in all we received 175 labels of which 60 were unique labels. Of those 60 different labels, 28 have been chosen only a single time.

4.1 Homogeneity

A computational model of place, thus the automatic generation of place concepts would ideally identify the commonsense concept of a required place. The same holds for a bottom-up repository of place labels: only if we overcome coordinate based labeling and move toward region based labeling, we can find common identifiers for places. In this section we analyze the harmonization potential of diverse labels within same regions. We compared the number of unique labels in each region and computed the most common labels for each region as well. The results are summarized in Table 1. An overview over the of choice of alternative labels for place regions is given in Figure 4.

I	II	III	IV
1	7	Unibad (40%)	Unibad (60%)
2	5	Cartesium (70%)	Cartesium (70%)
3	3	Haltestelle NW1 (50%)	Haltestelle NW1 (75%)
4	10	MZH (50%)	MZH (50%)
5	5	Boulevard (50%)	Mensa (60%)
6	7	3* (each 30%)	3* (each 30%)
7.1	2	2* (each 50%)	2* (each 50%)
7.2	6	MPI (40%)	2* (each 40%)
8	9	Wiener/Fahrenheitstr. (40%)	Wiener/Fahrenheitstr. (40%)
9	4	Universum (100%)	Universum (100%)
10	2	Haus am Walde (100%)	Haus am Walde (100%)
11	4	2* (each 40%)	3* (each 40%)
12	5	Uni-See (80%)	Uni-See (80%)
13	3	2* (40%)	Haus am Walde (50%)

Table 1. The table shows the agreement on labels for each region. Column I is the region number, column II shows the diversity of unique labels from primary and alternative choices. Column III and IV show the most common label for each region and its relative frequency among primary labels (III) and among all labels (IV). Cells with asterisks show the number of different labels of equal mentions (a complete list of labels would be too comprehensive). Regions 3 (8 data sets), 7.1 (4 data sets) and 7.2 (5 data sets) differ from the maximum number (10) of data sets per region.

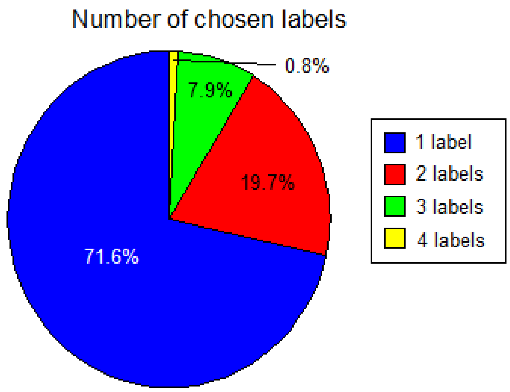

Fig. 4. The distribution of the numbers of chosen labels during single place labeling tasks.

From Table 1 we see the high number of unique place labels for all regions. The average number of unique labels in region C (3.5) strongly differs from those in regions A (6) and B (5.6), indicating a quantifiable measure for the density of well-known places in a region. However, every region has either a single most dominant label or a small set of labels common among subjects. Half of the subjects (55%) independently selected the same label, even though they could position themselves inside the rather coarse place regions, which, in all cases, contained a multitude of plausible entities. Considering this amount increases even further to about 60% when we also consider alternative labels. 28 of all 60 unique labels were only mentioned once.

4.2 Proximity

To gain insight to the influence of labeling location and the choice of labels, we measured and compared distances between the GPS location (presumably already including typical positioning errors of GPS) and the physical boundary (e.g. wall of a building, shore at a lake) of the primary referenced entities as well as nearby places known to the respective subject (see Section 3.4).

Subjects in most cases placed themselves towards the referred entities. However in most cases subjects did not minimize the distance between the labeling location and referenced entity. For each subject and each place region we compared the distances of the recorded labeling position for the primary label to the referred entity with the distances to other places referred to in the secondary (alternative) labels for the place region.

13 of the 127 primary labels referred to entities outside the place regions. As they are not analyzable with respect to the place regions, we excluded them from this analysis. We did the same for additional 14 labels, as they referred to rather unseizable concepts with no clearly computable boundaries ("South from the Kuhgraben heading towards Wiener Strasse"). Among the remaining 100

Fig. 5. The distribution of measured coordinates of labeling locations inside the 13 place regions. A positioning behavior towards certain (usually also labeled) entities is observable.

primary labels only ten labels referred to entities that were more distant than alternative labels.

4.3 Visibility

Proximity usually implies the visibility of entities. There were only few occasions where referred entities were not visible at all from the respective points of view. In region 6 three subjects referred to the central university cafeteria (which itself proved very important among subjects), while they at the same did not name a large visible building (Studentenwohnheim) which would be assumed as highly salient by a usual classification.

4.4 Saliency

For each region we classified spatial entities according to saliency (base area and height, as available from our geo-data). Other spatial features where selected according to their clear out-sticking from the "background", e.g. the lake or the playground in the forest.

Table 2 shows the (classified) most salient entities for each place region, according to computable entities and their properties (size and height) contained in available geo-data at the time of the analysis. The symbol > expresses "more salient than" relation between two given entities A and B (based on values of height h and the base area b) Entities located outside a place region are marked with a single asterisk. Entities on a granularity below building level were excluded from this view, as well as the large region-type entities. All references not explicitly listed are summarized as "others".

The results listed in Table 2 show that the correlation between saliency and the actually labeled entity is in some regions low. Even when we expand the

region ID	order of saliency & rate of reference
1	Sportturm (30%) > Unibad (40%) > others (30%)
2	MZH (10%) > GW2 (10%), VWG (10%) > Cartesium (60%), SFG (0%) > others (10%)
3	MZH* (0%) > NW1 (0%) > others (100%)
4	MZH (50%) > Glashalle (10%) > others (40%)
5	Studentenwohnheim (0%) > NW1 (0%) > Library (0%) > Glashalle (0%) > Mensa (40%) > others (60%)
6	ZHG (20%) > others (80%)
7.1	MZH* (50%) > Mensa (0%) > others (50%)
7.2	MPI (40%) > others (60%)
8**	others (100%)
9	Universum (100%) > Chocoladium (0%) > GW1 (0%) > others (0%)
10	Haus am Walde (100%) > others (0%)
11	Pavillon* (30%) > Stadtwaldsee (20%) > others (50%)
12	Uni-See (80%) > Spielplatz (10%) > others (10%)
13	Haus am Walde (40%) > Reiterhof (40%) > others (20%)

Table 2. An overview over salient spatial entities, ordered by the saliency. The ratio of the primary labels at the respective place regions are listed in brackets.

observation to the next salient entities, the number of references to not salient entities still can be considered as high. There are even examples where the most salient, clearly visible entities have been entirely ignored by all subjects. Figure 6 shows the ratio of the selection of salient/non-salient entities for place labels. The selection of salient entities correlates with the self-reported familiarity: The number of references to the most salient entity in each region increased as individual familiarity with the predefined regions decreased. Personal experience or "social saliency" seems to be dominant factors in-situ labeling.

4.5 Label Granularity

The dominant labeled entities were buildings (54,3% of all in-situ-labels). The reason for the selection of buildings as labels is possibly due to the respectively structured environment of a campus, but clearly can be observed in the other two regions as well. This observation raises the question, if the situated conceptualization of place in urban environments is equivalent to buildings. Of course this is not the general case and we observed different classes of labels:

- **Sub-Building-Level**: a label refers to a functional/logic unit inside a building (a company inside a bureau building, a cafeteria inside a building). 7 of the primary 127 (5,5%) labels can be assigned to this granularity.
- **Building-Level**: a label refers to an entire building without addressing finer structures. This was the dominant granularity for self-reported well-known regions. 69 labels directly referred to buildings (54,3%).
- **Transportation**: 9 (7,1%) labels referred to nodes of public transportation (bus and tram stations).

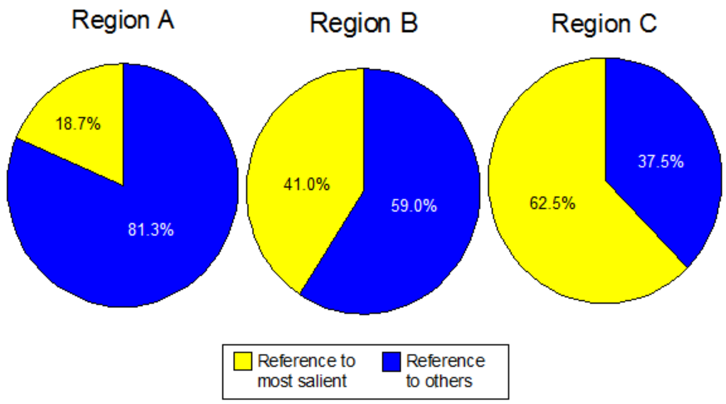

Fig. 6. The diagram shows the numbers of references to the most salient spatial entities for regions A, B and C. The interrelation between the amount of saliency-based references and the structuredness of the surrounding environment is observable.

- **Syntactic Level**: a label refers to syntactic constructs created from a spatial relation between entities (crossing of two streets, an open space between two or more known places, etc). Those kind of labels can only be understood when the referenced entities are already known. This level was often chosen when subjects estimated low familiarity with a location. 8 (6,3%) labels referred to entities of this category.
- **Region-Level**: a label refers to a region of a certain, vaguely defined spatial extent without a clearly defined borderline. Subjects only resorted to this strategy of labeling at place region 11, where there is no other known place to refer to in sight. Instead of making a reference to known nearby places (e.g. "Haus am Walde"), some subjects decided to refer to the entire forest. 13 (10,2%) labels were references to regions.
- **Natural Features**: a label refers to natural features; in our study just water bodies have been mentioned. 13 labels fell in this category (10,2%).
- **Others**: 8 (6,3%) labels referred to entities like mail boxes, sculptures, or stairs.

A noticeable observation is the high percentage of building-type labels in region C (similar to the urban regions A and B). The presence of a single well-known spatial entity in an otherwise unstructured region seems to attract a high number of references and implies a high homogeneity among place labels (see Section 4.1).

4.6 Region Blindness

Our assumption was that people at some point introduce higher-level concepts to describe places. In Weilenmann & Leuchovius (2004) the authors report that

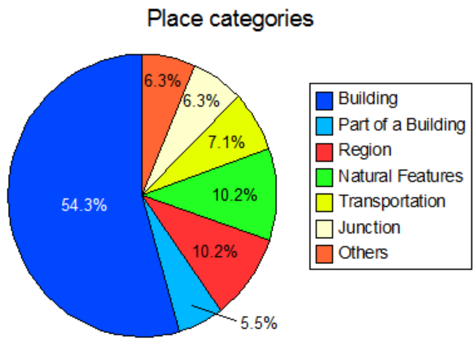

Fig. 7. The diagram shows the distribution of identified classes of labels. References to buildings are dominant, all other categories are used similarly often.

the selection of spatial granularity was adjusted very flexibly. Especially for rather unfamiliar regions, we expected the choice of labels like "Technologie Park", "Natural Sciences Part of Campus", or even "University". But our study strongly suggests that pure in-situ labeling of locations, thus the choice and assignment of a spatial label within the environment, is not the right modality to gather region based or hierarchical labels. Only at few points of the experiment subjects chose labels of coarser granularity. Rather the opposite strategy could be observed. When subjects were in a self-reported unfamiliar region, we observed two dominant labeling strategies. The choice of strategy among the subjects was consistent: every subject used either the strategy of "place extension" or "syntactic place determination", no subject mixed those two strategies over the place regions.

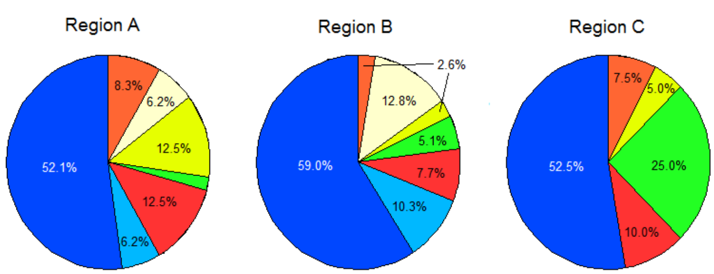

Fig. 8. The diagrams show the distribution of label classes over regions A,B and C. The same legend as in Figure 7 is effective. It is noticeable that the amount of references to buildings does not vastly differ between the three regions.

– **Place Extension**: Subjects selected known neighboring places to describe the current place region. The choice of the place was not necessarily driven by spatial closeness, sometimes producing references to buildings/objects far outside the current place region (in some cases referencing buildings more than 100m away from the current location); visibility seems to affect the choice of the label.
– **Syntactic Place Determination**: The other group of subjects selected syntactic elements from the environment to describe the place (e.g. street names at junctions, bus stop names).

There are two exceptions: the entity "Boulevard" in region A and the "Stadtwald" in region C (both in Figure 1. Both exceptions can be explained quite well: the "Boulevard" is a strongly connecting network link on the campus. It is a campus spanning bridge-like construction that connects main facilities of general university life (Mensa/Cafeteria, library, a shopping and dining facility). It is different from other entities on the campus and seems to be recognized as a self-contained place. The "Stadtwald" is a forest with the usual undefinedness and lack of unique structuring elements; at many points there is simply no better way to determine one's location than referring to the label "Stadtwald".

5 Discussion

Our first hypothesis assumes that the labeling location has direct influence on the choice of the label. It is not surprising that the results in section 4.2 clearly show that there is a strong correlation between the position and the choice of labels, as the referred entity in the majority of cases was the closest entity of its class. Visibility also has a great impact on the choice of labels. Only in few cases subjects decided to label occluded entities. 79.8% of all labels referred to the respectively most proximate entity (see Section 4.2), 97.6% of labels referred to an entity at least partially visible (see Section 4.3). In 77.2% of all labels both proximity and visibility were given at the same time. The influence of saliency the labels resulted in less clear results, in numerous cases subjects did not even mention the most salient feature in a region (see Section 4.4). This can be explained by the individual experienec with the environment. However this finding challenges salience based landmark models (see e.g. Elias (2003); Nothegger et al. (2004); Winter (2003)): cognitive inspired place models are not necesserily based on visual salience or structural importance, but on individual meaningfulness of places. This property has clear functional aspects: hub-like places, i.e., places where many people meet temporally (e.g. "Cafeteria") seem to be mentally more salient than visually salient entities. I.e., the most salient reference for a certain region might be the entity with the strongest impact on many people's daily routines. These results also have impact on the map offered to a user during labeling: depending on the granularity of concepts and the extents of places we want to gather, the application has to consider the possible visibility situation and the established experience of a user. An application will have to include or highlight plausible entities in the environment to streamline and harmonize the labeling of place regions.

Our second hypothesis assumes that place names across multiple persons for places of the size of "personally meaningful places" can be harmonized. We find strong support of this assumption from homogeneity analysis in section 4.1 - subjects are able to identify a number of different labels which sum up to a common name for a place (for the given region). Just as in the study of Lovelace et al. (1999), we observed that the number of different primary labels for a place region is in some cases high; However, although not the individually first choice, there are always alternative labels which are meaningful to most subjects and foster the harmonization. We can also observe and support the assumption of applications dealing with personally meaningful places: in-situ labels often reflect individually meaningful concepts. All subjects agreed that they would change the labeling behavior if the input modality would be constrained. I.e., in an application built on user-driven place labels, we can expect a inertial labeling behavior and as long as users can use previously defined labels for a place. The application itself has either to propose plausible candidates (which we will discuss in the next paragraphs), or has to induce the choice of high-level labels that are meaningful to a potentially larger group of persons. As we have seen from the visualization problems in the map material (see 3.3- the visualization of places has strong influence on the nature of labels. Depending on the kind of labels an application requires, it has to initiate and communicate the labeling process.

Our third hypothesis assumed that people will make use of different spatial granularities to describe a place. However, we observed a preference for entities with a clearly identifiable function or name. References to buildings or places inside buildings made up about 60% of all labels. Only about 10% of the primary labels referred to region concepts often in situations where the region as such is the only really adressable entity, e.g. the forrest instead of a bunch of trees. Interestingly, most of the region type labels could have been derived from existing geo-data as well (e.g. "Stadtwald", "Wiener Strasse"). Subjects never made attemps to disambiguate the location within the region by referring to close known elements (e.g "in the forest close to Haus Am Walde"). I.e., in-situ labels in familiar environments do not offer significant possibilities to deduct spatial semantic hierarchies. The concept place in the sense of "I am currently here" seems only to reflect the immediate environment or the closest known reference, usually on a granularity of clearly distinguishable entities. In an urban environment it is the granularity of buildings (at least for naming places) and sometimes functional units within buildings, in a natural environment like a forest, it is the forest. In contrast to the expectation of our third hypothesis, people tend either to expand entities on the granularity of buildings, or to fall back to elements on a finer level of granularity (street names, junctions). Only reflected or ex-situ communication, as verbal communication with other persons or visual communication in form of maps seem to introduce the expression of region concepts Weilenmann & Leuchovius (2004); Montello et al. (2003). When we want to retrieve human-centered hierarchical spatial information, we have to facilitate the expression of them - either by dialogs or by a careful pre-selection of possible regions. The combination of semantic and syntactic Dalton (2007) approaches is promising: there are clear correlations between the regions annotated in maps and the structural logic of space syntax. These observations imply two consequences: We cannot expect semantic higher-level concepts from situated labeling. I.e., we have to identify other sources of available information or have to constructively facilitate the fostering of region concepts. However, spatial entities on the granularity of buildings are obviously suitable and behavioral valid concepts for places in urban familiar environment. However, the strong environment dependency also implies that only if we really can assign a name or a function to a building, it is a meaningful reference. In residential areas where only few buildings with public meanings are present, street sections, natural features and the few public buildings are most likely suitable references (see Section 4.5).

6 Towards a Computational Model of Place

Applications utilizing personally meaningful places benefit from maintaining a collaborative repository of places, or from the automatic computation of places and places labels. Our study suggests that both cases are implementable. Users can find homogeneous names for place regions of rather large size (in our case 50-100 meters) and the labeling of places is in many cases functionally seizable. There are many established methods to compute the isovist-visibility of entities (see e.g. Batty (2001)) to select visible entities for a specific location (estimation), and computing the most proximate entity is relatively straightforward. The dominant usage of building labels support straightforward place interpretations as postulated in various approaches and the place name study of Zhou et al. (2005b). Zhou et al. found that people often rely on businesses to describe places, a source which is accessible via business directories. For residential areas we can generate labels with respect to rather structural elements (street names or nodes of public transportation). However, our subjects preferred hub-like places to classically visually salient places. The harmonization of bottom-up place labels can underly the same rationals: we can compute plausible areas based on visibility and proximity analysis and attach the labels to those regions instead of coordinates. The representation of the environment has great influence on the choice of labels and should be applied to foster labeling as required: the proposition of concept borders and commonsense regions but as well as of semantic regions. We could not observe any indication of introducing region concepts without map-like representations.

Acknowledgments

We gratefully acknowledge funding granted by the Deutsche Forschungsgemeinschaft (DFG) and the Deutscher Akademischer Austauschdienst (DAAD).

Bibliography

Agarwal, P. (2004). Contested nature of place: Knowledge mapping for resolving ontological distinctions between geographical concepts. In M. J. Egenhofer, C. Freksa, & H. J. Miller (eds.), *GIScience*, vol. 3234 of *Lecture Notes in Computer Science*, (pp. 1–21). Springer.

Agarwal, P. (2005). Operationalising 'sense of place' as a cognitive operator for semantics in place-based ontologies. In A. G. Cohn & D. M. Mark (eds.), *COSIT*, vol. 3693 of *Lecture Notes in Computer Science*, (pp. 96–114). Springer.

Ashbrook, D. & Starner, T. (2003). Using gps to learn significant locations and predict movement across multiple users. *Personal Ubiquitous Comput.*, 7(5):275–286.

Batty, M. (2001). Exploring isovist fields: space and shape in architectural and urban morphology. *Environment and Planning B: Planning and Design*, 28(1):123–150.

Bicocchi, N., Castelli, G., Mamei, M., Rosi, A., & Zambonelli, F. (2007). Supporting location-aware services for mobile users with the whereabouts diary. In *MOBILWARE '08: Proceedings of the 1st international conference on MOBILe Wireless MiddleWARE, Operating Systems, and Applications*, (pp. 1–6). ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Cresswell, T. (2004). *Place: a short introduction*. Wiley-Blackwell.

Dalton, N. S. (2007). Is neighbourhood measurable? In *Proceedings of the 6th International Space Syntax Symposium, Istambul*.

Duckham, M. & Kulik, L. (2005). A formal model of obfuscation and negotiation for location privacy. In *Pervasive*, (pp. 152–170).

Elias, B. (2003). Extracting landmarks with data mining methods. In *COSIT 2003, LNCS 2825*, (pp. 398–412). Springer.

Grothe, C. & Schaab, J. (2008). An evaluation of kernel density estimation and support vector machines for automated generation of footprints for imprecise regions from geotags. In W. Stephan, K. Werner, & Krueger (eds.), *International Workshop on Computational Models of Place PLACE08*, (pp. 15–28). Melbourne, Australia: University of Melbourne.

Hagerstrand, T. (1970). What about people in regional sciences? *Papers in Regional Sciences*, 24:7–24.

Hightower, J. (2003). From position to place. In *Proceedings of The 2003 Workshop on Location-Aware Computing*, (pp. 10–12). Part of the 2003 Ubiquitous Computing Conference.

Hillier, B. & Hanson, J. (1984). *The social logic of space*. Cambridge: Cambridge University Press.

Liao, L., Patterson, D. J., Fox, D., & Kautz, H. (2007). Learning and inferring transportation routines. *Artif. Intell.*, 171(5-6):311–331.

Lovelace, K. L., Hegarty, M., & Montello, D. R. (1999). Elements of good route directions in familiar and unfamiliar environments. In *COSIT '99: Proceedings of the International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*, (pp. 65–82). London, UK: Springer-Verlag.

Lynch, K. (1960). *The image of the city*. Cambridge: MIT Press.

Marmasse, N. (1999). commotion: a context-aware communication system. In *CHI '99: CHI '99 extended abstracts on Human factors in computing systems*, (pp. 320–321). New York, NY, USA: ACM.

Montello, D., Goodchild, M., Gottsegen, J., & Fohl, P. (2003). Where's downtown? behavioral methods for determining referents of vague spatial queries. *Spatial Cognition and Computation*, 3(2):185–204.

Nothegger, C., Winter, S., & Raubal, M. (2004). Computation of the salience of features. *Spatial Cognition and Computation*, 4:113136.

Relph, E. (1976). *Place and Placelessness*. Routledge Kegan & Paul.

Schmid, F. (2008). Knowledge based wayfinding maps for small display cartography. *Journal of Location Based Services*, 2(1):57–83.

Schmid, F. (2009). Enhancing the accessibility of maps with personal frames of reference. In J. A. Jacko (ed.), *Proceedings of the 13th International Conference on Human-Computer Interaction, Part III: Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction*, vol. 5612 of *Lecture Notes in Computer Science*. Springer; Berlin.

Schmid, F. & Richter, K.-F. (2006). Extracting places from location data streams. In *UbiGIS 2006 - Second International Workshop on Ubiquitous Geographical Information Services*.

Schockaert, S., De Cock, M., & Kerre, E. E. (2005). Automatic acquisition of fuzzy footprints. In *On the Move to Meaningful Internet Systems 2005: OTM 2005 Workshops, OTM Confederated International Workshops and Posters (SeBGIS 2005)*, vol. 3762 of *LNCS*, (pp. 1077–1086).

Twaroch, F. A., Smart, P. D., & Jones, C. B. (2008). Mining the web to detect place names. In *GIR '08: Proceeding of the 2nd international workshop on Geographic information retrieval*, (pp. 43–44). New York, NY, USA: ACM.

Weilenmann, A. H. & Leuchovius, P. (2004). "i'm waiting where we met last time": exploring everyday positioning practices to inform design. In *NordiCHI '04: Proceedings of the third Nordic conference on Human-computer interaction*, (pp. 33–42). New York, NY, USA: ACM.

Winter, S. (2003). Route adaptive selection of salient features. In W. Kuhn, M. Worboys, & S. Timpf (eds.), *COSIT 2003, LNCS 2825*, (pp. 349–361). Berlin Heidelberg: Springer-Verlang.

Zhou, C., Ludford, P., Frankowski, D., & Terveen, L. (2005a). An experiment in discovering personally meaningful places from location data. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, (pp. 2029–2032). New York, NY, USA: ACM.

Zhou, C., Ludford, P., Frankowski, D., & Terveen, L. (2005b). Talking about place: An experiment in how people describe places. In *In Proc. Pervasive, Short Paper*.