Journal of Spatial Information Science
Number 28 (2024), pp. 103–127

doi:10.5311/JOSIS.2024.28.323

Research Article

Textual geolocation in Hebrew: mapping challenges via natural place description analysis

Tal Bauman*[,1], Tzuf Paz-Argaman*[,2], Itai Mondshine[2], Reut Tsarfaty[2], Itzhak Omer[3], and Sagi Dalyot[1]

[1]Mapping and Geo-Information Engineering, The Technion, Israel
[2]Department of Computer Science, Bar-Ilan University, Israel
[3]Department of Geography and Human Environment, Tel Aviv University, Israel

Received: September 15, 2023; returned: October 25, 2023; revised: December 3, 2023; accepted: January 8, 2024.

Abstract: Describing where a place is situated is an innate communication skill that relies on spatial cognition, spatial reasoning, and linguistic systems. Accordingly, textual geolocation, a task for retrieving the coordinates of a place from linguistic descriptions, requires computerized spatial inference and natural language understanding. Yet, machine-based textual geolocation is currently limited, mainly due to the lack of rich geo-textual datasets necessitated to train natural language models that, in-turn, cannot adequately interpret the language-based expressions. These limitations are intensified in morphologically rich and resource-poor languages, such as Hebrew. This study aims to analyze and understand the linguistic systems used for place descriptions in Hebrew, later to be used to train machine learning natural language models. A novel crowdsourced geo-textual dataset is developed, composed of 5,695 written place descriptions provided by 1,554 native Hebrew speakers. All place descriptions rely on memory only, which increases spatial vagueness and requires referring expression resolution. Qualitative linguistic analysis of place descriptions shows that geospatial reasoning is greatly used in Hebrew, while empirical analysis with textual geolocation engines indicates that literal descriptions pose challenges for existing methods, as they require real understanding of space and geospatial references and cannot simply be geolocated by matching gazetteer with textual geo-entity extractions. The findings offer improved understanding of the challenges entailed in natural language processing of Hebrew geolocation, contributing to formalizing computerized systems used in future machine learning models for complex geographic information retrieval tasks.

*Equal contribution.

Keywords: textual geolocation; geographic information retrieval; Hebrew; natural language processing; spatial cognition and reasoning

1 Introduction

Geolocation is the process of determining the geographic position (i.e., coordinates) of a physical object. Commonly, geolocation uses various analytical methods that rely on sensory data, such as Wi-Fi fingerprinting and satellite navigation systems trilateration. Textual geolocation, a task frequently used in Geographic Information Retrieval (GIR), aims at inferring the geographic position of a place or a physical entity based on linguistic descriptions and systems. Most Named-Entity Recognition (NER) and entity linking textual geolocation engines today rely on gazetteer matching, meaning that the place—or object—is explicitly mentioned in the text, and thus can be retrieved easily by querying the geodatabase. Still, in many cases, the place is not explicitly mentioned, and thus gazetteer matching will fail. Accordingly, computerized textual geolocation requires models to understand spatial language and spatial reasoning to be able to geo-reference textual linguistic terms to the environment [20]. Cognitive mapping of the environment [42], a method that may be employed in geolocation, is mostly obtained from various different sources, such as visual sensors [2, 49] or symbolic world representations, such as maps [1, 34]; it is less commonly obtained using text descriptions, mainly since it requires rich geo-textual datasets used in natural language (NL) models to adequately interpret the expressions.

Natural language processing (NLP) aims at enabling computers to understand, interpret, and manipulate texts and NL data to perform tasks. In NLP, the term semantic parsing refers to the process of translating a NL sentence into a formal representation that captures its complete meaning. Data-driven methods are a common approach in NLP for semantic parsing, that is based on machine learning techniques and requires data for training a model that can directly predict the desired output from the given input. In this context, the textual geolocation task poses numerous challenges.

Let us consider, for example, the place description: "Our meeting place is east of Rabin Square, very close to the square, and two buildings from the pharmacy; when facing Ben Gurion Boulevard, the place will be on your right". Some challenges are generic to NL understanding, such as anaphora resolution; for example, in the above sentence "It will be on your right," what does the word "it" refer to? Advanced NLP models already exist for handling non-spatial NL challenges, such as anaphora resolution [31], especially in English. NER and entity linking are two examples for resolving the location of named entities, such as 'Rabin Square'. However, geolocating linguistic terms that involve spatial expressions without the explicit mention of a place name still pose a significant challenge, including: (1) spatial terms, such as cardinal direction, e.g., "east of"; (2) spatial numerical reasoning, e.g., "two buildings away from the pharmacy"; (3) geo-referencing generic entities, e.g., "pharmacy"; (4) egocentric spatial relation, e.g., "It will be on your right"; and (5) ambiguity, e.g., how close is "very close to the square"? To overcome such challenges, we need to understand spatial relations between entities that are mentioned in the text, while georeferencing the mentioning of these entities to their corresponding entities in the environment. To do so, we need a corpus that presents a textual geolocation task; meaning, having textual expressions of places where about that are linked to their physical location.

However, existing corpora available for textual geolocation tasks suffer from low-resolution limitations, as they can only provide geolocation within a range of several dozen kilometers [27]. Furthermore, these corpora lack spatial orientation and predominantly focus on English descriptions [38, 48, 50, 51]. There is a void in corpora for low-resource languages, such as Hebrew, a Semitic morphologically-rich language (MRL) that is infamously difficult to parse [45]. As such, processing Hebrew texts presents unique challenges that are not evident when processing texts in English. For example, the single Hebrew word וכשבשדרת (ukhshebashderat) corresponds with six word-tokens in English: and when in the avenue of.

To investigate the challenges of textual geolocation in the Hebrew language, we first need a rich geo-textual corpus, which currently does not exist, to formalize missing NLP capacities. The key challenge in data collection for such a corpus is the designing of a scalable process for obtaining georeferenced place descriptions. The process must (1) reflect the natural way in which people describe place locations, to create an interesting case for spatial cognitive and reasoning research; and (2) be interpreted and geolocated by humans in a relatively simple manner, yet still pose a challenge for current NLP models. Among others, this means creating a task that cannot be resolved by NER or fuzzy matching of the goal location and nearby landmarks to a Gazetteer list. To do so, we first designed a crowdsourced assignment based on realistic scenarios of place descriptions that were provided by people relying on their memory of the environment. The sparse representation captured by human memory affects the place descriptions, which results in increased spatial vagueness and related expression resolution [18], leading to disambiguation between the described entity and other similar entities in the environment. Participants were asked to provide written texts in which they describe the whereabouts of familiar places, (a) without explicitly naming the places, while (b) allowing other participants to locate these places on a map based on the provided description only. Later, this corpus was analyzed to map the geolocation challenges for handling natural place descriptions in Hebrew. We also collected general demographics and supplementary information of the participants, to study how different groups of people describe the same places. The following research questions guided our methodology: (a) How do place type, city features, and spatial knowledge influence the usage of Hebrew place description linguistics (e.g., named entities, spatial language terms, prepositions, adjectives, and conjunctions)? (b) What is the most important spatial knowledge type machines should learn to enable accurate textual geolocation? (c) What are the technological gaps of existing textual GIR services, and what is required from Hebrew NLP models to enhance their performance? Our research results provide better understanding of the challenges entailed in Hebrew NLP, thus contributing to formalizing future machine learning models for complex geolocation tasks.

2 Background and related work

2.1 Human spatial language

Human spatial language has fascinated researchers since the early 1960s, having examined a variety of topics, such as how people remember and describe locations. When answering a "Where?" question, people must communicate the location by composing a spatial expression that refers to that place [33, 43]. This is known as a referring expression, i.e., a textual expression that uniquely identifies a particular entity [9]. Referencing occurs hierarchically in place descriptions, starting with the most prominent features surrounding the place and finishing with the less-known features that are mostly closer to the place [28]. Humans differentiate between the following five characteristics of physical features: nodes, paths, edges, districts, and landmarks, that compose the image of the city [30]. As such, referencing is related to bodily and sensory experiences that are acquired by traversing and exploring the environment. Moreover, the importance of certain features in a city is determined through a combination of visual, semantic, and structural characteristics [36], who found, for example, that the visual characteristics of a district are generally weaker than its semantic and structural ones. [44] show that structural properties allow streets to be experienced as facilitating travel in the city, since they are embedded in the ground; landmark buildings, on the other hand, are recalled by people based on their unique visual or semantic characteristics, such as how their facades are detailed and the type of businesses that reside within them. For instance, when observing New York, it becomes apparent that 5th Avenue lacks a distinctive visual identity in comparison to the Empire State Building, which boasts a unique and visually striking form. Furthermore, a compelling correlation emerges between the significance of a spatial feature (referred to as salient) within a specific environment and the intensity of the associated experience, ultimately influencing the formation of memories. The Eiffel Tower, for example, can be easily seen in the urban landscape and is able to create an experience unlike any other in Paris, thereby leaving a lasting impression on the viewer.

Studies show that human spatial language has typical patterns, i.e., 'building blocks' for verbal spatial language [23,35], including: (a) concise language with as few syllables as possible, and with a preference for descriptions that include names and spatial relations, rather than quantitative geometric indices; (b) a hierarchical description with a bottom-up structure; and (c) the use of salient features as reference points [16]. Spatial inference is necessary for achieving accurate textual geolocation and enabling the processing of place descriptions. Spatial inference is derived through spatial-quantitative techniques, which evaluate connections and associations between entities and in different hierarchies. As shown by [37], different methods are applicable for different types of spatial entities (point, line, and polygon) that differ in their computational complexity and in the spatial knowledge that is produced [10,11,52]. Moreover, since languages differ in how they can be used to describe space, such differences can be addressed as means for exploring relationships between language and thought [8, 13, 29]. Levinson [29], for example, shows that non-linguistic cognition mirrors the lexical systems that exist in the local language, and that language influences how people memorize, think, and reason about spatial relations and directions. E.g., languages like Guugu Yimithirr and Tamil do not use at all words that reflect an egocentric or relative frame of reference, like “The cat is to the left of the tree”, in which the description is relative to the individual’s perspective. Instead, they use only words that reflect allocentric or absolute frame of reference based on cardinal directions, like “The cat is in front of the tree”, that does not depend on any individual’s perspective. Thus, people who speak these languages know where they are with respect to the world. Therefore, their spatial orientation is much better than those who frequently rely on a relative frame of reference. As such, the English language cannot be the sole focus when investigating human spatial language, specifically for addressing textual geolocation challenges and building computerized models that perform spatial tasks [4].

2.2 Text-Based geolocation

The term textual geolocation relates to the task of retrieving the coordinates of a place that is described in the text. One of the most practical applications of textual geolocation is GIR, enabling users to browse and search for content through a text-based geospatial interface, using applications such as MetaCarta's geographic text search [3] and NewsStand [41]. Since a significant portion of internet content incorporates a flavor of geospatial referencing [17], datasets currently used for textual geolocation predominantly rely on open-source resources like Wikipedia articles [50], [51] and tweets [38, 48]. In their study, Krause & Cohen [26] focus on deriving geolocations for Wikipedia pages. They propose a four-step process leveraging textual and categorical data, demonstrating improved precision-recall trade-offs compared to text-only approaches. Furthermore, Krause and Cohen [27] propose an approach to assign real-world locations to documents for GIR, where their method surpasses other baseline methods by determining latitude-longitude coordinates in the range of several dozen kilometers for relevant Wikipedia articles. The authors' approach can also detect—with low spatial resolution—geolocation errors in Wikipedia articles, propose approximate coordinates for non-location articles and expand possibilities for geographic retrievals.

[37], for example, collected place descriptions via the "Tell-Us-Where" website to gain better understanding on how people describe places. By utilizing the global navigation satellite system (GNSS) embedded in users' mobile phones for precise location retrieval, participants were prompted to provide textual descriptions of their current whereabouts, emphasizing their visual observations instead of relying on their memory. Yet, the participants' demographic data were not collected, nor was information about the described location type, where both are important to investigate the linguistic building blocks of different place descriptions.

Another important task in the field of GIR is text-based navigation, which involves the ability to interpret textual instructions and navigate accordingly. Unlike existing geolocation datasets, text-based navigation tasks are crowdsourced and offer detailed navigation descriptions. These tasks heavily rely on environmental knowledge derived from various sources, including visual sensors [2] and symbolic representations of the world, such as maps [1, 34]; thus, they are less relevant to textual geolocation.

2.3 Natural language processing in Hebrew

Hebrew is a resource-poor and challenging language to process, especially due to its limited users (10M), rendering it less attractive for developers and therefore resulting in fewer resources. Further difficulties arise from its Semitic structure and characteristics, including it being an MRL [45]. As explained by [45], languages vary in how much information is encoded in their morphology, with Hebrew falling into the MRL category at the far end of the spectrum. Hebrew, therefore, encodes significant amounts of information, especially in relation to syntactic units and relations expressed at the word level, where each word token in Hebrew may consist of multiple lexical and morpheme units, i.e., functional units that fill a particular role. As such, to process Hebrew texts, the word tokens must first be segmented into their constituting morphemes. At the same time, as shown by [46,], even raw unvocalized Hebrew (commonly used) word tokens tend to be highly ambiguous, making even that difficult to segment. For example, the word token מבנה can be interpreted in at least two different ways, according to segmentation and context: (1) mivneh, i.e., building; and (2) mibnah, from her son.

Moreover, Hebrew has a rich inflection of verbs, nouns, and prepositions that alter the person (first, second, or third), number, and gender of the subject, as well as the object of the verb—and that can be incorporated in the verb as a suffix. As shown by [32], inflections in Hebrew may generate twenty to forty different morphological structures for each word, thereby increasing the number of possible interpretations of strings of letters in Hebrew. Although Hebrew is defined by having a subject-verb-object sentence structure, it is more flexible than English in relation to word order, and many variations are possible for a sentence [7, 14]. Hebrew is also unique among Semitic languages in that it is a revived language that only functioned as a religious language until the 19th century. Modern Hebrew is therefore susceptible to rapid changes, with a fast increase in new words that are needed for modern pragmatic needs and following intensive exposure to European languages [53]. Recently, researchers have been working on bringing Hebrew up-to-date with current advances in NLP, collecting datasets [5, 6, 12, 24], and developing large pre-trained language models [5, 6, 12, 24, 39]. More specifically, datasets for textual geolocation tasks do not currently exist in Hebrew. The sole geospatial-oriented dataset that exists in Hebrew is the map-based navigation task [22]; yet, it contains only 32 dialogues, which is inadequate for modern NLP data-driven models. Moreover, these recent downstream applications have rarely (if at all) been put to the test for examining their applicability.

2.4 Summary

In conclusion, three main research gaps are evident to solve textual geolocation in Hebrew. First, no dataset (corpus) exists, designed specifically to solve the above task, thus theoretical knowledge on place and language is missing (e.g., spatial reasoning, use of vocabulary). Second, the Hebrew language imposes specific challenges that require customized NLP modeling, which English-based models cannot solve. Finally, existing textual geolocation engines rely heavily on open-source datasets, which lack the possibility for accurate geolocation (e.g., < 1 km), since linked textual expressions and physical locations do not exist. To solve the Hebrew text-based geolocation challenge, we designed the HeGeL crowdsourced dataset that relies on people's spatial memories to capture the cognitive representations of the environment as it is reflected in the spatial language. This will provide an important step to produce knowledge and possible solutions to the abovementioned gaps.

3 Methodology

3.1 Corpus collection design

The HeGeL corpus data collection was designed by the authors together with NLP researchers and spatial cognition experts in a way that it can provide data to map the abovementioned challenges. Namely, we aimed to study how people naturally describe the location of a place from memory without the assistance of visual aids, such as maps. The dataset was created using GIRit, which was developed for this research, serving as an online assignment application comprised of three sections: (1) Participants' demographics and supplementary information; (2) Task 1: Writing a description for a given place; and (3) Task 2: Geolocating a place based on a given description.

The GIRit online assignment was distributed to participants by a surveying company[1]. As a place can be described differently by diverse groups of people, we collected the participants' age, gender, education, perceived navigation level, and city of residence (Figure 1A). For perceived navigation knowledge level, the participants were asked to provide a score on a scale of 0 (not at all) to 5 (excellent). To avoid prenotions of the correct way to describe a place, participants were first asked to complete Task 1 (Figure 1B), and only then to complete Task 2 (Figure 1C).

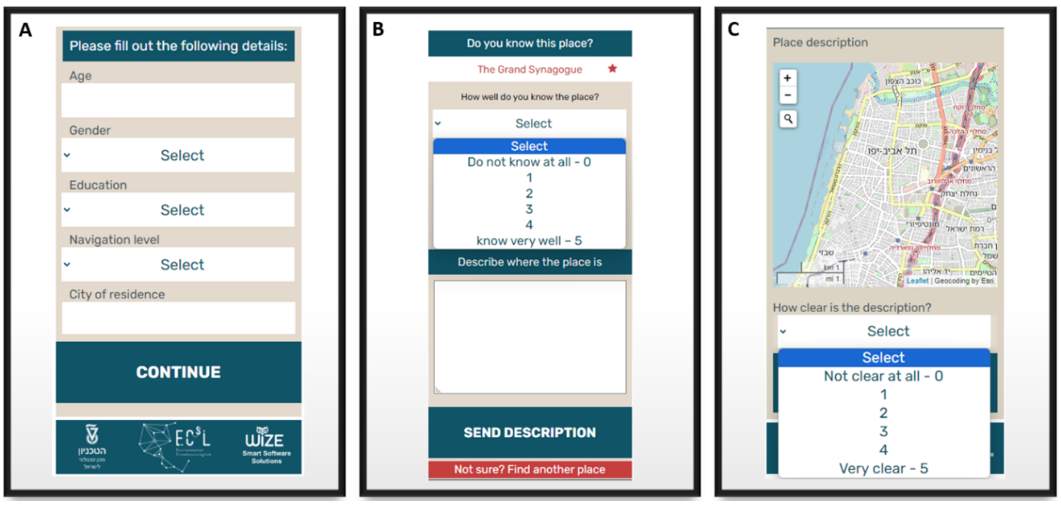

Figure 1: Screenshots from the GIRit online assignment application (text is translated to English): (A) participant's demographics and characteristics; (B) written place description (Task 1); and (C) geolocation (validation) (Task 2).

3.2 Task 1: Written place description

The dataset included 167 places, which the participants could choose to describe, all located in Israel's three largest cities: Tel Aviv, Haifa, and Jerusalem (Figure 2). As these cities differ in shape, morphology, physical features, and topography, they have potential effects on the legibility and imageability of urban components, and therefore also on the words and terms used for place description. These differences can be expressed in the use of different physical features and prepositions, e.g., frequent use of the object landmark, use of the prepositions "above" or "below" in hilly terrain that characterize Haifa and Jerusalem, and the use of cardinal directions, such as "North off", that characterizes Tel Aviv; these enabled the gathering of richer descriptions needed for achieving a comprehensive corpus. By including three different cities in the assignment, we were able to generate the three splits required for NLP tasks: training, developing, and testing. To ensure diverse human-generated textual descriptions, places were chosen based on their type, location in the city, distinct geometry, size (area), and context. Place types included 13 squares, 56 compounds,

[1]The study was approved by the Ethics Committee at the authors' affiliated academic institution.

five street markets (traditional bazaars on street and pedestrian paths), 50 buildings, 21 parks, 15 neighborhoods, five boulevards, and two bridges (Figure 3). All are urban objects with different functions and experiences, such that they should produce different cognitive representations and hence different place descriptions related to scales, referencing frames and levels of spatial knowledge.

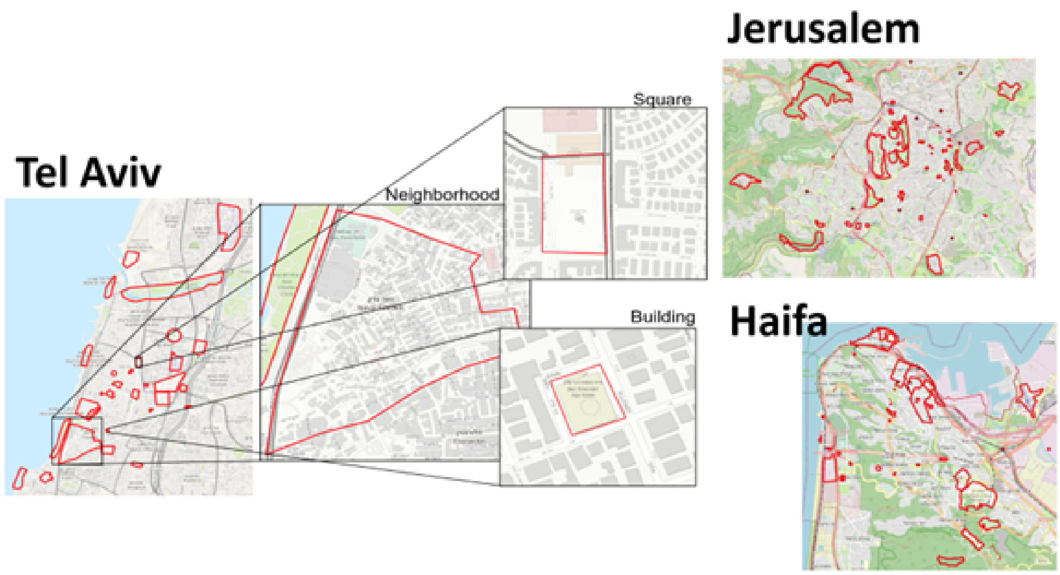

Figure 2: The 167 places (red polygons): Tel Aviv (left), Jerusalem (top right), and Haifa (bottom right).

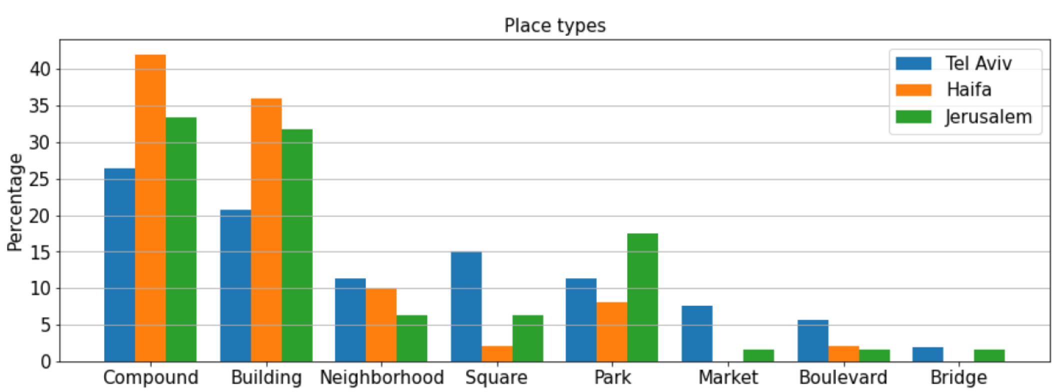

Figure 3: Place type percentage per city: Tel Aviv, Haifa, and Jerusalem.

As a means of simplicity and clarity, we framed Task 1 as a treasure hunt, since games have been found to encourage participant engagement [25]; we also provided the participants with an example prior to their embarking on the tasks. The place name was presented in writing, not on a map. Prior to inserting the written descriptions for each place, the participants were asked how familiar they are with the place on a scale of 1 (not at all) to 5 (very much so). If participants marked 1 or 2, they automatically received an alternative place to describe. We used this familiarity score to analyze whether there were significant features associated with the familiarity score. The participants inserted a description in free text that should serve as a treasure hunt game for a follower to find (translated example from the survey: "The destination is south of the Nation Buildings, north of the Israel Museum and east of Leyada High School"). This process was repeated, whereby each participant was asked to describe 10-15 different places. To ensure a balanced dataset with an equal number of descriptions per place, we developed a process that allocated the place for the user in ascending order, according to the existing place-description count.

To increase description clarity and the likelihood of correct geolocation (in Task 2), the description length was set to a minimum of six words [37]. Moreover, participants were asked not to explicitly mention the place name—as well as of nearby landmarks—in the description to address the NLP challenge at hand. This process was done to prevent descriptions that can serve as the solved problem in geolocation that uses named entities. Once participants explicitly stated such entity name (identified according to a rule-based function using fuzzy string matching[2] we developed), a message appeared on the screen asking them to alter their textual description (the original description was saved for future research).

3.3 Task 2: Geolocation – place description verification

This task aimed at verifying that place locations can be retrieved by treasure hunters according to the descriptions provided by other participants; meaning that the treasure hunters read the descriptions and are able to realize where the treasure is hidden/geolocated. Participants mark the communicated location using an interactive online map based on OpenStreetMap (OSM), which visualizes street networks, points of interest, and spatio-textual labels. The map allows participants to mentally process the place description using geolocating mentioned entities and performing spatial reasoning. Each treasure hunter repeats this process for 10-15 different descriptions, where each place description is given to at least two independent participants to geolocate. After marking the place location on the map, participants had to rate the clarity of the description on a scale of 0 (not clear at all) to 5 (very clear). These scores enabled us to analyze correlations between people's actual retrieval performance and their perceptions of their understanding. As a preliminary task, participants (treasure hunters) were asked to mark a well-known place (its name was given) on the map using the maximum zoom level; those who failed this task did not proceed to complete Task 2. This preliminary task was done to ensure participants have basic map skills and that they do not mark wrong places on purpose.

[2]Fuzzy string matching is a technique of finding strings that approximately match a pattern (rather than exactly). The closeness of a match (i.e., the edit distance) is measured in terms of the number of primitive operations necessary for converting the string into an exact match.

3.4 Geolocation retrieval error

To measure the textual geolocating performance of Task 2, we calculated a retrieval error, which we referred to as the shortest Euclidean distance between the OSM's coordinates of the place shape (point, line, and polygon)—true physical position, and the location (point coordinates) that the participant marked on the OSM map. To come up with a threshold that defines a valid retrieval (acceptable geolocation), we used all the location markings made by users who stated that the description clarity was 4 or 5. Depicted in Figure 4 for all 23,852 validations, we can see that there is a strong correlation, Spearman's rank correlation of -0.34 with p-value<0.001, between the retrieval error distance and description clarity, where the calculated median value of 4 was 291 meters. Accordingly, we defined a 300-meter threshold as a valid retrieval error. Verified geolocation—a place that is retrieved by a human based on its textual description—is considered if at least one of its two verifications is less than the 300-meter threshold.

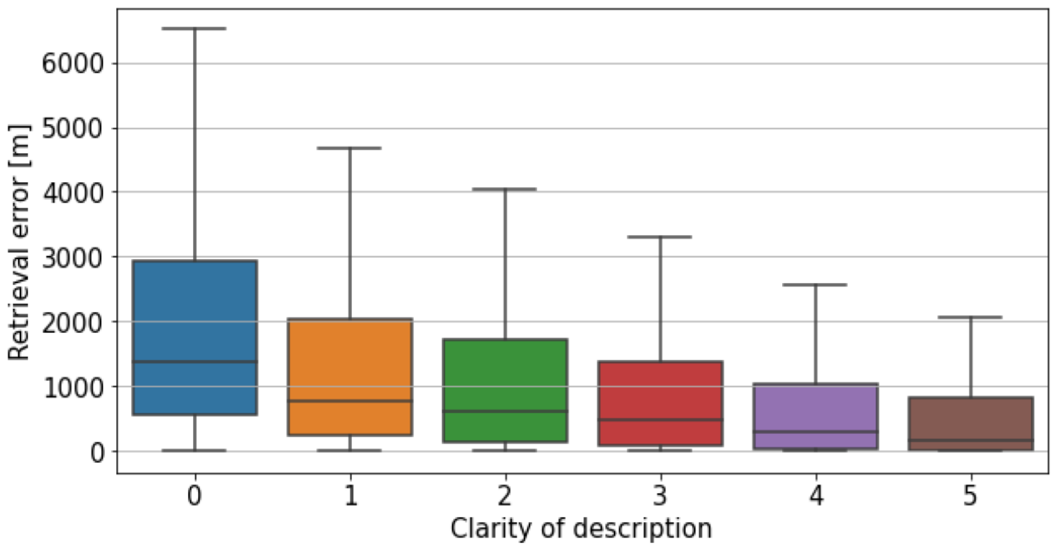

Figure 4: Boxplot of retrieval error value (Y-axis) by clarity of description (X-axis).

4 HeGeL statistics, analysis, and experiments

4.1 Descriptive data statistics

HeGel contains 5,695 verified descriptions, including 2,141 in Tel Aviv, 1,440 in Haifa, and 2,114 in Jerusalem. 1,554 native Hebrew speakers participated in the online assignment (Task 1), providing 10,946 place descriptions, which were then validated by 2,073 participants (Task 2). Table 1 depicts the demographic characteristics of participants. Most participants were aged 20-50, with an almost equal number of males and females for the three cities; most participants had graduated from high school or held an academic diploma, and most stated that they possess good navigation skills (values 3-5 on a 0-5 scale).

Characteristic	**Tel Aviv%** (n=637) 41.0%	**Haifa%** (n=397) 25.5%	**Jerusalem%** (n=520) 33.5%	**Total%** (n=1,554) 41.0%
Gender %				
Male	54.8	51.7	54.4	53.7
Female	45.2	48.3	45.6	46.3
Age groups %				
18-22	7.1	6.8	14.8	9.7
23-29	13.0	16.6	24.8	17.9
30-39	21.8	22.9	25.8	23.4
40-49	24.0	19.1	16.5	20.2
50-72	34.1	34.3	18.1	28.8
Education %				
None	5.5	6.3	6.7	6.1
High school	28.4	33.5	34.2	31.6
Bachelor's degree	33.0	29.7	28.1	30.5
Master's degree	20.4	15.9	15.2	17.5
PhD	0.5	1.0	1.5	1.0
Practical engineer	5.5	7.1	4.8	5.7
Diploma	6.8	6.5	9.4	7.6
Navigation level %				
0 – Not good	1.3	1.3	1.9	1.5
1	2.8	2.3	2.9	2.7
2	6.0	6.5	7.7	6.7
3	22.8	25.7	26.3	24.7
4	34.2	34.5	30.4	33.0
5 – Very good	33.0	29.7	30.8	31.4

Table 1: Demographic characteristics of participants.

One of the research questions of this study is concerned with how people differ in describing places. Accordingly, we analyzed several relationships between the place descriptions' characteristics (e.g., word count and clarity) and the participants' demographics. One example is presented in Figure 5 (top), demonstrating a clear correlation between the description clarity and the retrieval error (x-axis). The comparison between Haifa and Tel Aviv reveals an interesting aspect derived from their respective topography: Haifa's hilly and intricate terrain poses a greater challenge for geolocation compared to the mostly flat and well-structured morphology of Tel Aviv. Figure 5 (bottom) shows a less distinct trend that correlates the number of used words and the description clarity. Also, a subtle distinction in word count exists between males and females, where females were able to effectively describe place locations with fewer words.

4.2 Linguistic analysis

Figure 6 depicts the word cloud of the HeGeL corpus. By using three cities, not only are city-specific named entities used by participants, such as תל אביב (Tel Aviv), but also much richer spatial language terms, such as קרוב (Karov, meaning near), prepositions, such as ליד

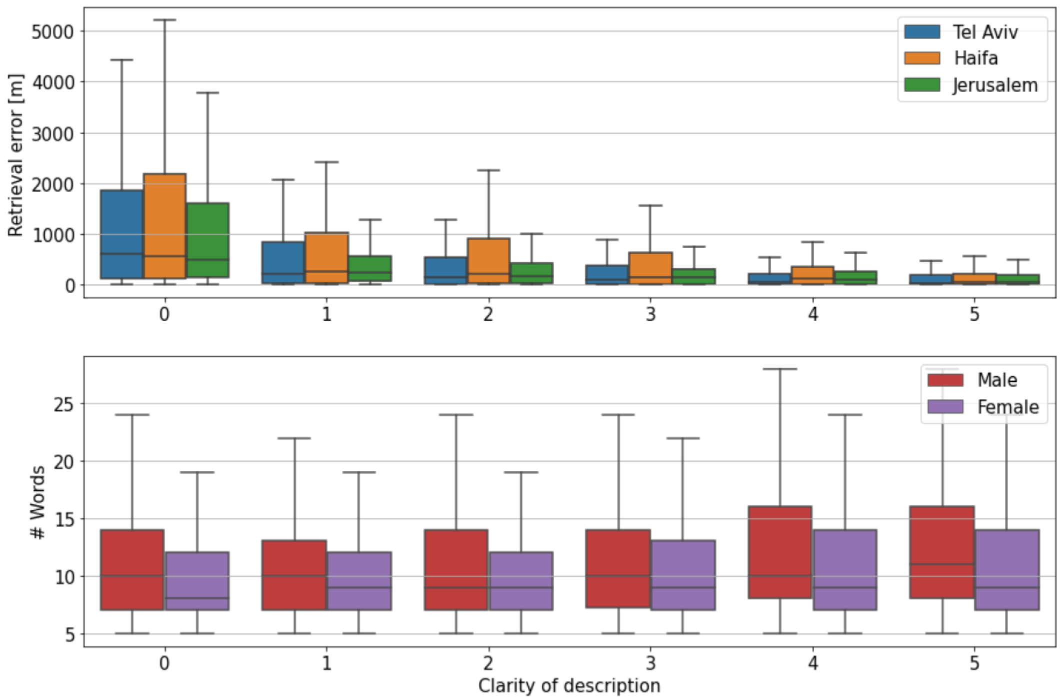

Figure 5: Top: Boxplot representation of retrieval error by region; Bottom: Boxplot representation of word count by gender. x-axis in both images depict the description clarity as chosen by the treasure hunter (follower).

(le-yad, meaning next to), city entities, such as רחוב (rehov, meaning street), adjectives that describe visual aspects of entities, such as גדול (gadol, meaning big), and conjunctions, such as או (oh, meaning or).

Figure 7 presents a Venn Diagram of the logical relationship between the three sets of city-based used vocabularies (formed from unique lemmas[3]). Surprisingly, only 15.07% of the entire vocabulary is used in all three cities, containing spatial language terms, such as ”between”, and city entities, such as ”street”. Almost half the lemmas in the three vocabularies (i.e., corresponding to the three cities) contain city-specific lemmas: 48.6% of the Tel Aviv vocabulary, 40.65% of the Haifa vocabulary, and 49.3% of the Jerusalem vocabulary. Words as “overlook”, “Wadi”, “ridge” and “tunnel” appear both in Haifa and Jerusalem, both having distinct topographies, where Tel- Aviv and Haifa, both situated on the seaside, have words such as “Sea”, “Beach” and “Port”. As HeGeL enables a city-split setup (i.e., conducting training on one city and then testing it on a different previously unseen city), the city-specific named entities present an out-of-vocabulary (OOV) challenge for NLP models (i.e., encountering new words that had not previously been seen during the training process.)

[3]The canonical or dictionary form of a word.

Figure 6: Word cloud of the 100 most used words in the HeGeL corpus: in Hebrew (top) and translated to English (bottom).

Table 2 presents the results of a linguistic phenomenon analysis of the HeGeL dataset for 25 random descriptions, addressing the spatial knowledge and reasoning that are needed for NLP modeling. We analyzed the frequency of the five types of elements in a city [30] and of the three types of spatial knowledge [15, 21, 40], and additional spatial properties. Each example presented in the table is annotated with the presence and count of each phenomenon, whereby n is the number of instructions out of the 25 that contain at least one

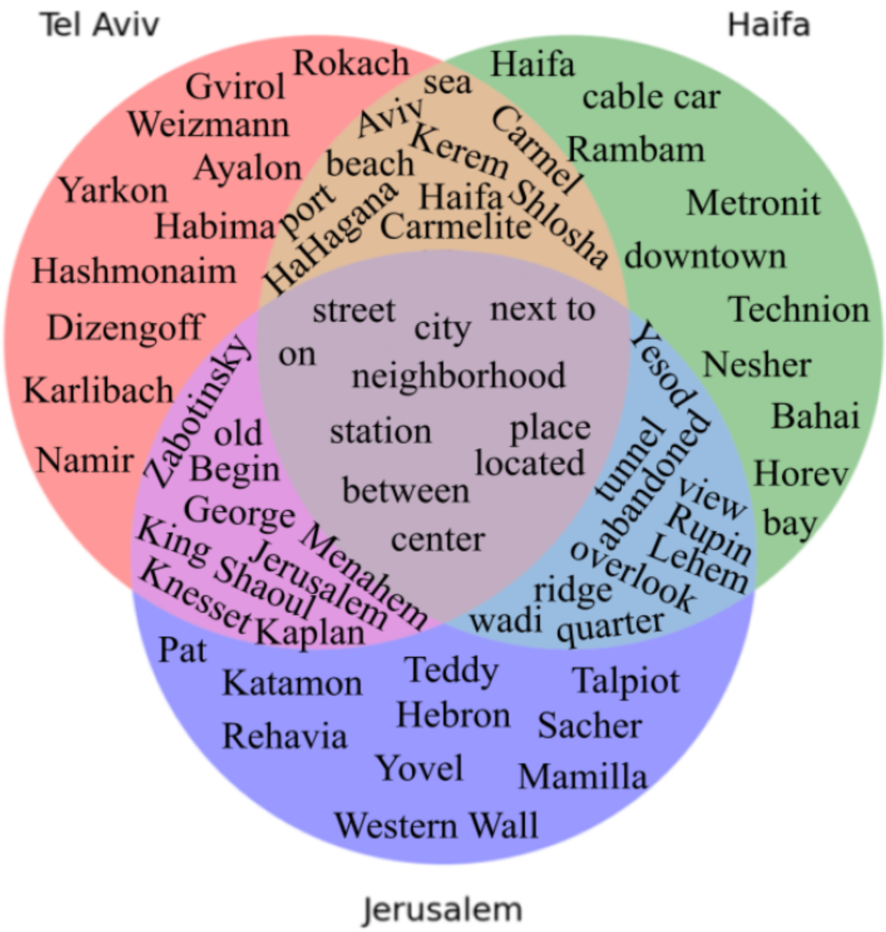

Figure 7: Venn diagram of the top-10 used words.

example of the phenomenon, and m is the mean number of times each phenomenon appears in each of the 25 instructions. Results show that all examples in the HeGeL dataset include references to unique entities, and almost half include cardinal directions, which are non-trivial to use in place descriptions, as the descriptions were based solely on memory, not on map views. The frequent use of cardinal directions, as well as the use of survey knowledge, suggests that any developed NLP model should not only represent a local view of the goal, or possible routes, but also take into consideration the full region, while mimicking people's map-like view of the environment. Learning a representation that captures the entire map involves abundant semantic and spatial information, thereby presenting a challenge in NLP.

Phenomenon		Descriptions n	m	Examples from HeGeL	Examples translated into English
Type of city element [30]	Edge	9	0.6	כשמגיעים ליפו מתקדמים לכיוון הים ...	When reaching Jaffa, one should go toward the sea...
	Node	10	0.44	... כמספר דקות הליכה מכיכר השעון ...	...a few minutes' walk from the Clock Square...
	Landmark	20	1.08	... ליד שוק לוינסקי	...near Levinski Market
	District	9	0.4	דרום העיר ליד ...	South of the city next to...
	Path	17	0.76	על רחוב קרליבך ...	On Carlebach Street...
Spatial knowledge [40]	Landmarks	8	-	צמוד לים בתל אביב יפו	Next to the sea in Tel Aviv-Jaffa
	Route	5	-	עוברים את עזריאלי על מנחם בגין ופונים ימינה ...	Passing Azrieli on Menachem Begin and then turn right...
	Survey	12	-	דרום העיר ליד שוק לוינסקי ...	South of the city near Levinski Market
Reference to unique entity		25	2.32	... באמצע רחוב דיזנגוף	...in the middle of Dizengoff Street
Cardinal direction		11	0.76	דרומית לשרונה ...	South of Sarona...
Coreference		4	0.16	... תמשיך קצת מערבה וזה ...	...continue west for a bit and it's...

Table 2: Linguistic analysis of 25 randomly sampled place descriptions from HeGeL.

A quantitative analysis[4] of the descriptions included in the HeGeL dataset, depicted in Table 3, presents an impressive sized vocabulary, with 6,663 unique lemmas and 9,207 unique word tokens. There is some mention of named entities in the dataset, yet (as per our constraint during Task 1 of nearby landmarks) these are scarce and mostly refer to salient landmarks. The prevalence of prepositions in the descriptions align with expectations for spatial language, particularly in languages that utilize this strategy for expressing spatial relations, such as English and Hebrew. However, the limited use of verbs can be attributed to the predominantly non-route-based nature of the place descriptions.

Feature	Average per description	Median per description	SD per description	Unique in the corpus
Number of lemmas	12.97	11	7.30	6,663
Number of word tokens	11.74	9	7.46	9,207
Number of named entities	0.62	0	0.84	3,490
Number of prepositions	2.53	2	1.81	14,256
Number of verbs	0.55	0	0.94	3,152

Table 3: Quantitative analysis of HeGeL.

Using the Welch's analysis of variance (ANOVA) test, we can determine whether the mean number of words differs between groups, in this case—the different place types de-

[4]We used the following automatic tools: NEMO (https://github.com/OnlpLab/NEMO) and YAP (https://github.com/OnlpLab/yap).

scribed in Task 1, based on description features (number of words, named entities, verbs, lemmas, and prepositions) and verification values (clarity of description and retrieval error). Table 4 summarizes the p-values corresponding to the Welch's ANOVA tests performed, as well as the False Discovery Rate (FDR) corrected p-values. Four features were found to have a significantly ($p<0.001$) different distribution between place type descriptions: number of named entities, number of prepositions, retrieval error, and clarity of description score. The findings demonstrate variations in the utilization of named entities and prepositions when individuals describe different types of places. Furthermore, the clarity of these descriptions and their retrieval error (derived from Task 2 results) also differ across various types of places, meaning that different place types will require different use of retrieval methods for accurate geolocation.

Feature	**p-value**	**np2**	**FDR corrected p-value**
Task 1: Place description (linguistic features)			
Number of Words	0.827	0.001	0.827
Number of named entities	< 0.001	0.013	$\mathbf{< 0.001}$
Number of prepositions	0.011	0.003	**0.020**
Number of lemmas	0.764	0.001	0.827
Number of verbs	0.205	0.002	0.287
Task 2: Human verification			
Retrieval error	0.001	0.004	**0.002**
Clearness score	<0.001	0.009	$\mathbf{< 0.001}$

Table 4: Correlations of place types with linguistic and verification features.

Figure 8 illustrates two key distributions derived from the above analysis. The clarity of description distribution (Figure 8, top) reveals distinct patterns, with notable concentrations of clearer descriptions observed for the categories 'Square' and 'Bridge' compared to other types. The retrieval error distribution (Figure 8, bottom), a nuanced perspective emerges as 'Square', 'Boulevard', and 'Building' exhibit a higher likelihood of achieving superior retrieval accuracy (distribution peak on the far left side). These findings provide valuable insights into the challenges associated with location retrieval of place descriptions and the impact of place types on the accuracy of retrieval mechanisms. Future research will further explore the effects of prepositions and named entities on retrieval error across different place types.

4.3 Spatial knowledge analysis

We conducted a qualitative analysis by randomly choosing 20 place descriptions from each of the six self-reported navigation levels. For each place description, we manually classified the spatial knowledge type—survey, route, or landmark—used by those studying the acquisition of spatial knowledge, as well as place descriptions and spatial navigation [40]. This classification was conducted while considering the asymmetric inclusion relations between the three, as follows:

1. Survey knowledge: a description of a spatial object/place that includes spatial relations in terms of cardinal directions. It should be noted that we refer in this case to the notion of 'absolute' frame of reference, as suggested by Levinson (1996), according to which a) the object is described with respect to cardinal or external axis (e.g., "the

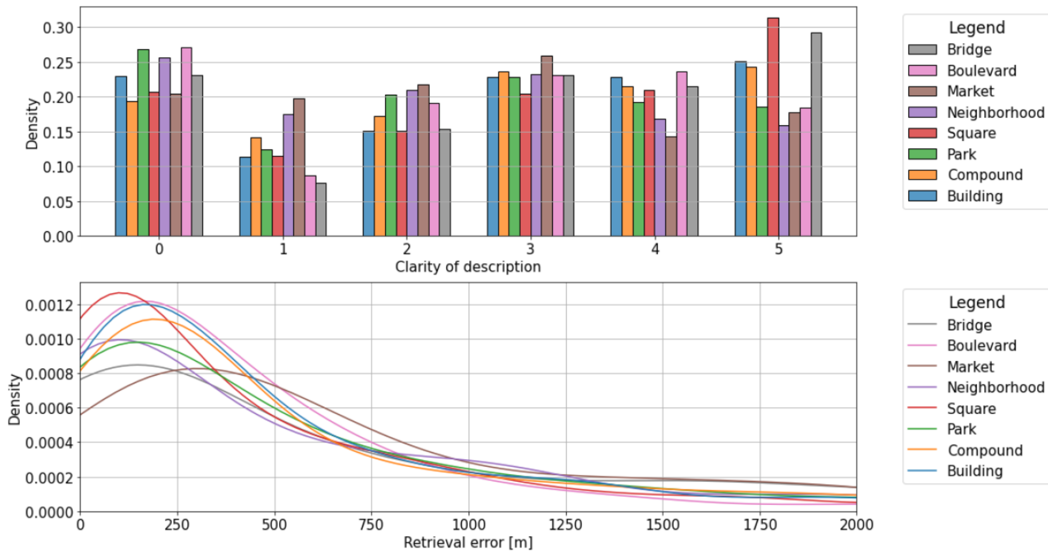

Figure 8: Correlation of place type and clarity of description (top) and retrieval error (bottom).

building is north to the park"), as opposed to b) 'relative' frame of reference, in which the description is relative to the viewpoint of the one describing (e.g. "the building is to the left of the park"), and to c) 'intrinsic' frame of reference, where the description depends on the properties of the object (e.g., "the building at the front of the park"). The last two reference frames reflect other knowledge types. Thus, cardinal directions are exclusive to the survey knowledge type. Meaning that the description cannot be based solely on route knowledge and/or on landmark knowledge. For example, "The south part of the city near Levinski Market," or "The neighborhood is located between downtown Haifa and the Carmel, on the eastern side."

2. Route knowledge: the description includes information on how to get to a given spatial object/place with explicit reference to spatial relations to other object/s, but with no reference to cardinal relations. Meaning that the description cannot be based solely on landmark knowledge. For example, "Passing Azrieli on Menachem Begin and then turning right," or "At the third station of the light rail that leaves the Central Station."
3. Landmark knowledge: the description refers to a spatial object/place by specifying its name and/or features, with no reference to spatial relations (of any kind) with other objects. Meaning that the description can only be based on this type of knowledge. For example, "Next to the sea in Tel Aviv-Jaffa," or "Residential buildings with lots of offices."

Figure 9 depicts the distribution of the analyzed 120 descriptions in terms of navigation level (top) and retrieval error (bottom) according to spatial knowledge type. Figure 9 (top) shows that participants who have route knowledge are more uniformly distributed in terms of their perceived navigation level, while those with other types of spatial knowledge are normally distributed. Figure 9 (bottom) shows that participants who have survey knowledge are more uniformly distributed in terms of the retrieval error, while those with other types of spatial knowledge follow a retrieval error distribution with relatively much fewer errors. This could be related to differences in spatial knowledge types, whereby with landmark and route knowledge the use of more specific (less abstract) spatial features at the local scale tend to be more accurate than survey knowledge in terms of place descriptions. Hence, survey knowledge, the highest level of knowledge, might be more challenging in the context of textual geolocation. The disparity in retrieval error of the different spatial knowledge likely stems from their distinct approaches to describe places. While survey knowledge primarily relies on cardinal directions and places less emphasis on observable objects in the area—adding ambiguity to place retrieval, the other knowledge types (route and landmarks) employ landmarks as crucial reference points. Accordingly, when prioritizing efforts in measuring machine capabilities for GIR processes, it is beneficial to first solve the technological gaps associated with route knowledge and landmark knowledge, making them an excellent foundation for future machine learning endeavors.

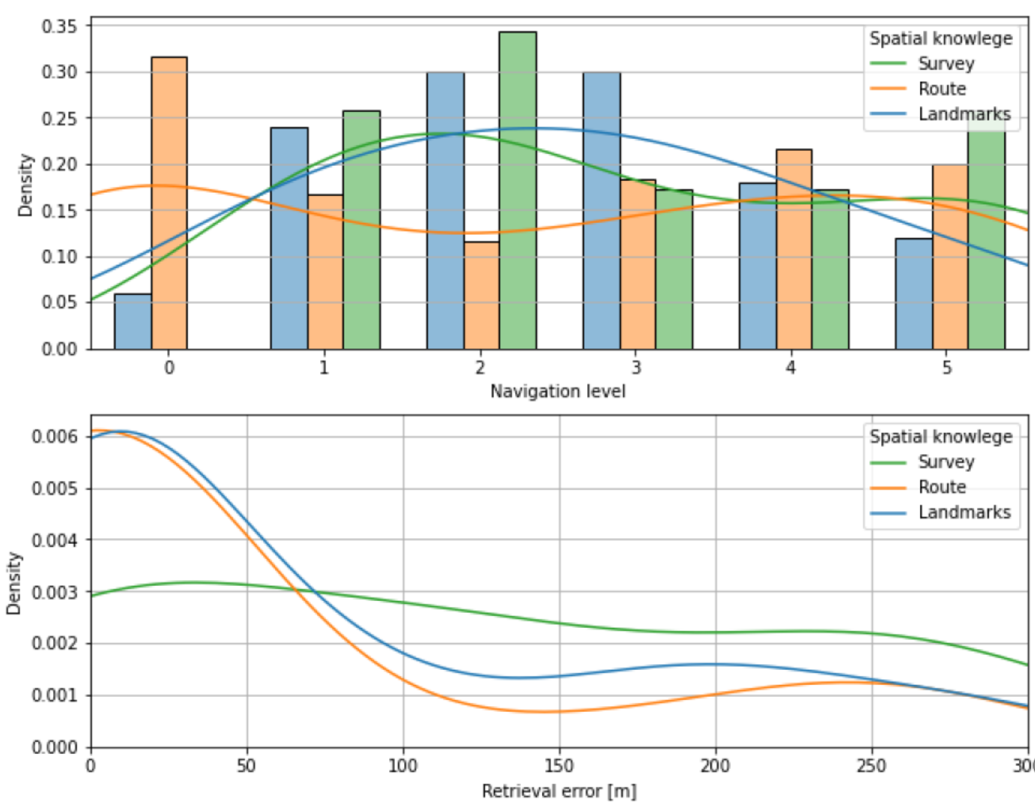

Figure 9: Navigation level (top) and retrieval error (bottom) by spatial knowledge type.

4.4 Geolocation task

Our research focuses on the geolocation of places using colloquial language and memory-based spatial descriptions. The objective is to generate an output consisting of a pair of (x, y) coordinates that accurately pinpoints the described place's location. We evaluated different baseline models for the textual geolocation task on the HeGeL dataset. We used three evaluation metrics based on the distance between the coordinates generated by the model and the known place location: mean, median, and task completion (TC) accuracy, i.e., the percentage of place descriptions that are correctly located within the 300-meter threshold. We tested the data with an NER approach: could recognizing entities in a text and retrieving their corresponding coordinates be sufficient to solve textual geolocation tasks? To this end, we used Google Maps API[5] to produce two baseline models:

1. Google Maps API Query. Here, we queried the application programming interface (API) with the full raw text descriptions from HeGeL as input, and with no prepossessing.
2. Oracle NER. Here, we prepossessed the descriptions by dividing the input text into n-grams (1-5 grams), and then queried all single n-grams on the Google Maps API Query. We used the Google Maps API Query model (baseline model I) output as the candidate prediction, whereby the final prediction is the closest place to the gold state.

It should be noted that these are not competitive models, but rather models that provide us with the insight that geolocating based on NER alone is insufficient for handling the spatial reasoning presented in HeGeL. Table 5 depicts the results of this analysis, as we show that our task is not solvable even with adequate resolution of the Google Maps API. Human performance provides the upper bound performance, while the Google Maps API Query provides the lower bound. Moreover, the TC accuracy shows that the human agreement rate is 66.67%, meaning that overall, most places are geolocated within 300 meters. We hypothesize that the low agreement rate for the textual geolocation task is due to the reliance of human participants on their memory when providing place descriptions.

Model	Mean Distance [m]	Median Distance [m]	TC Accuracy
Google Maps API Query	2,804	850	27.66%
Oracle NER*	2,366	497	37.79%
HUMAN	579	314	66.92%

Table 5: Baselines for the textual geolocation task with the HeGeL corpus.
* Oracle NER is not a competitive model, but a higher bound on a NER approach.

5 Conclusions and discussion

In this paper, we set out to investigate several aspects of human place descriptions in Hebrew, aiming to address our research questions and shed light on the intricacies of this linguistic phenomenon. We first designed and implemented a pioneering survey, in which we assembled a corpus—HeGel—that contains data that does not exist elsewhere. HeGeL

[5]http://code.google.com/apis/maps/

provides an infrastructure for developing—and later testing—state-of-the-art Hebrew NLP models, primarily designed for textual geolocation. To the best of our knowledge, this is the only crowdsourced textual geolocation corpus available in any language, as the Tell-Us-Where dataset [37] is no longer available. Unlike other open-source corpora, such as Wikipedia and Twitter, the HeGeL place descriptions are spatial-oriented and can be geolocated with high spatial-resolution, as required for realistic city-level geolocation tasks. Since HeGeL provides place descriptions that rely on memory, it reflects people's cognitive representations of the environment, thus allowing analyzing their level of spatial knowledge. The descriptions capture data that depicts the natural way people describe places, while providing a dataset that creates an open NLP challenge.

Our linguistic analysis showed that almost half of the HeGeL vocabulary is city-specific language and named entities, presenting an OOV challenge for developing NLP models. Empirical analysis showed that any developed textual geolocation method will require a real understanding and modelling of space and geospatial references. The analysis indicates that Hebrew place descriptions frequently contain cardinal directions and rely on survey knowledge. Given that human textual geolocation struggles with place descriptions that depend on survey knowledge, it is crucial for future NLP models to incorporate a comprehensive understanding of the surrounding environment. Additionally, the HeGeL setup that includes descriptions collected from three cities that show physical, morphological, and topographical differences will allow future spatial cognitive research to investigate the effect of these differences on place descriptions.

Our analysis revealed that different types of places exhibit variations in the choice of words and the description structure. For example, descriptions of natural places tend to use vocabulary related to topography and morphology. Additionally, irrespective of the place and city type, we identified certain predominant word types, indicating commonalities in Hebrew language usage across diverse places. Preliminary analysis showed that there are dependencies between user characteristics and place description, requiring further research to identify these dependencies, allowing the development of tuned NLP models, and hence customized GIR services.

The effect of spatial knowledge on place descriptions is evident in our study. We observe that individuals with diverse levels of spatial knowledge describe a place in distinct ways, which directly impacts the follower's ability to comprehend and accurately locate the described place. Additionally, when considering the improvement of machine capabilities in textual geolocation, it is advantageous to first focus on route knowledge and landmark knowledge, since both yielded better human retrieval results, indicating their effectiveness in guiding location-based tasks. Moreover, our research shows that despite current advancements, the existing textual geolocation search engines still face challenges in providing precise and reliable results. To allow better results, this research identifies several algorithmic gaps, among them, understanding contextual cues (e.g., spatial knowledge usage) and mimicking (modeling) people's map-like view of the environment.

In future surveys, we plan to implement the Santa Barbara Sense of Direction Scale [19], replacing the rather subjective navigation level question with a more comprehensive measure that will allow us to gain better understandings on the relation of human place descriptions and spatial knowledge. We also plan to use our data-collection methodology to collect additional parallel corpora in other languages that represent linguistic diversity relevant for NLP models, including Semitic and non-Semitic languages (e.g., Arabic and English, respectively). Parallel corpora collected with the same method will allow the investigation of differences between languages and cultures when conveying place descriptions with respect to spatial cognition aspects. For example, frame of reference is used in comparable navigation tasks. Such comparisons will also support the development of data-driven multilingual NLP models for the textual geolocation task that has interdisciplinary potential contribution in this respect.

This study provides valuable insights for textual geolocation, informing various domains, including linguistics, spatial cognition, and GIR, guiding future research and practical applications in these fields. The HeGeL dataset is a critical addition to the challenging NLP benchmark that serves for the ongoing spatial cognition research. We believe that some Hebrew-related challenges—and solutions—are also relevant to other languages, thus contributing with theoretical and practical knowledge in formalizing computerized systems for complex geographic information retrieval tasks.

Data and code availability statement

The data and code that support the findings of this study are available with the identifiers at the private link.

References

[1] ANDERSON, A. H., BADER, M., BARD, E. G., BOYLE, E., DOHERTY, G., GARROD, S., ISARD, S., KOWTKO, J., MCALLISTER, J., MILLER, J., SOTILLO, C., THOMPSON, H. S., AND WEINERT, R. The hcrc map task corpus. *Language and speech 34*, 4 (1991), 351–366. doi:10.1177/002383099103400404.

[2] ANDERSON, P., WU, Q., TENEY, D., BRUCE, J., JOHNSON, M., SUNDERHAUF, N., REID, I., GOULD, S., AND VAN DEN HENGEL, A. Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), IEEE, pp. 3674—3683. doi:10.1109/CVPR.2018.00387.

[3] ANDOGAH, G. *Geographically constrained information retrieval*. University Library Groningen, 2010.

[4] BALDRIDGE, J., BEDRAX-WEISS, T., LUONG, D., NARAYANAN, S., PANG, B., PEREIRA, F., SORICUT, R., TSENG, M., AND ZHANG, Y. Points, paths, and playscapes: Large-scale spatial language understanding tasks set in the real world. In *Proceedings of the First International Workshop on Spatial Language Understanding* (2018), Association for Computational Linguistics, pp. 46–52. doi:10.18653/v1/W18-1406.

[5] BAREKET, D., AND TSARFATY, R. Neural modeling for named entities and morphology (NEMO2). *Transactions of the Association for Computational Linguistics 9* (2021), 909–928. doi:10.1162/tacl_a_00404.

[6] BAUMAN, T., OMER, I., AND DALYOT, S. *Constructing Place Representations from Human-Generated Descriptions in Hebrew*, vol. 13238 of *Lecture notes in computer science*. Springer International Publishing, 2022, pp. 51–60. doi:10.1007/978-3-031-06245-2_5.

[7] BERMAN, R. A., AND BOLOZKY, S. Modern Hebrew structure. *Modern Hebrew structure* (1978).

[8] CHOI, S. Influence of language-specific input on spatial cognition: Categories of containment. *First language 26*, 2 (2006), 207–232. doi:10.1177/0142723706060748.

[9] DALE, R. *Generating referring expressions: Constructing descriptions in a domain of objects and processes*. 1992.

[10] DU, S., FENG, C.-C., AND GUO, L. Integrative representation and inference of qualitative locations about points, lines, and polygons. *International Journal of Geographical Information Science 29*, 6 (2015), 980–1006. doi:10.1080/13658816.2015.1004333.

[11] DYLLA, F., LEE, J. H., MOSSAKOWSKI, T., SCHNEIDER, T., DELDEN, A. V., VEN, J. V. D., AND WOLTER, D. A survey of qualitative spatial and temporal calculi. *ACM Computing Surveys 50*, 1 (2017), 1–39. doi:10.1145/3038927.

[12] EYAL, M., NOGA, H., AHARONI, R., SZPEKTOR, I., AND TSARFATY, R. Multilingual sequence-to-sequence models for Hebrew NLP. *arXiv* (2022). doi:10.48550/arxiv.2212.09682.

[13] GLEITMAN, L., AND PAPAFRAGOU, A. 32. relations between language and thought.

[14] GLINERT, L. *The Grammar of Modern Hebrew*. Cambridge University Press, 2004.

[15] GOLLEDGE, R. G. Place recognition and wayfinding: Making sense of space. *Geoforum 23*, 2 (1992), 199–214.

[16] GONG, Y., WU, L., LIN, Y., AND LIU, Y. Probability issues in locality descriptions based on Voronoi neighbor relationship. *Journal of Visual Languages & Computing 23*, 4 (2012), 213–222. doi:10.1016/j.jvlc.2012.04.002.

[17] HAHMANN, S., AND BURGHARDT, D. How much information is geospatially referenced? networks and cognition. *International Journal of Geographical Information Science 27*, 6 (2013), 1171–1189. doi:10.1080/13658816.2012.743664.

[18] HAYWARD, W. G., AND TARR, M. J. Spatial language and spatial representation. *Cognition 55*, 1 (1995), 39–84. doi:10.1016/0010-0277(94)00643-y.

[19] HEGARTY, M. Development of a self-report measure of environmental spatial ability. *Intelligence 30*, 5 (2002), 425–447. doi:10.1016/S0160-2896(02)00116-2.

[20] HU, R., FRIED, D., ROHRBACH, A., KLEIN, D., DARRELL, T., AND SAENKO, K. Are you looking? grounding to multiple modalities in vision-and-language navigation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (2019), Association for Computational Linguistics, pp. 6551–6557. doi:10.18653/v1/P19-1655.

[21] ISHIKAWA, T., AND MONTELLO, D. R. Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology 52*, 2 (2006), 93–129. doi:10.1016/j.cogpsych.2005.08.003.

[22] JACOB, A., LERNER, A., AND SILBER-VAROD, V. The map task corpus of the Open University of Israel (MaTaCOp). *Unpublished* (2015). doi:10.13140/rg.2.2.19362.56004.

[23] KARIMZADEH, M., PEZANOWSKI, S., MACEACHREN, A. M., AND WALLGRÜN, J. O. GeoTxt: A scalable geoparsing system for unstructured text geolocation. *Transactions in GIS 23*, 1 (2019), 118–136. doi:10.1111/tgis.12510.

[24] KEREN, O., AND LEVY, O. Parashoot: A Hebrew question answering dataset. In *Proceedings of the 3rd Workshop on Machine Reading for Question Answering* (2021), Association for Computational Linguistics, pp. 106–112. doi:10.18653/v1/2021.mrqa-1.11.

[25] KNIESTEDT, I., LEFTER, I., LUKOSCH, S., AND BRAZIER, F. M. Re-framing engagement for applied games: A conceptual framework. *Entertainment computing 41* (2022), 100475. doi:10.1016/j.entcom.2021.100475.

[26] KRAUSE, A., AND COHEN, S. Deriving geolocations in Wikipedia. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (2020), ACM, pp. 3293–3296. doi:10.1145/3340531.3417459.

[27] KRAUSE, A., AND COHEN, S. Geographic information retrieval using Wikipedia articles. In *Proceedings of the ACM Web Conference 2023* (2023), Y. Ding, J. Tang, J. Sequeda, L. Aroyo, C. Castillo, and G.-J. Houben, Eds., ACM, pp. 3331–3341. doi:10.1145/3543507.3583469.

[28] KUIPERS, B. *A hierarchy of qualitative representations for space*, vol. 1404 of *Lecture notes in computer science*. Springer Berlin Heidelberg, 1998, pp. 337–350. doi:10.1007/3-540-69342-4_16.

[29] LEVINSON, S. C. *Space in language and cognition: explorations in cognitive diversity*. Cambridge University Press, 2003. doi:10.1017/CBO9780511613609.

[30] LYNCH, K. The image of the environment. *The image of the city 11* (1960), 1–13.

[31] MITKOV, R. *Anaphora Resolution*. Oxford University Press, 2020. doi:10.1093/oxfordhb/9780199573691.013.36.

[32] ORNAN, U., AND GUTTER, I. Machine translation by semantic features. *Proceedings of the International Conference on Machine Translation and Multilingual Applications in the new Millennium: MT 2000* (2000).

[33] PARABONI, I., VAN DEEMTER, K., AND MASTHOFF, J. Generating referring expressions: making referents easy to identify. *Computational Linguistics 33*, 2 (2007), 229–254. doi:10.1162/coli.2007.33.2.229.

[34] PAZ-ARGAMAN, T., AND TSARFATY, R. RUN through the streets: A new dataset and baseline models for realistic urban navigation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (2019), Association for Computational Linguistics, pp. 6448–6454. doi:10.18653/v1/D19-1681.

[35] PLUMERT, J. M., SPALDING, T. L., AND NICHOLS-WHITEHEAD, P. Preferences for ascending and descending hierarchical organization in spatial communication. *Memory & Cognition 29*, 2 (2001), 274–284. doi:10.3758/bf03194921.

[36] RAUBAL, M., AND WINTER, S. *Enriching Wayfinding Instructions with Local Landmarks*, vol. 2478 of *Lecture notes in computer science*. Springer Berlin Heidelberg, 2002, pp. 243–259. doi:10.1007/3-540-45799-2_17.

[37] RICHTER, D., WINTER, S., RICHTER, K.-F., AND STIRLING, L. How people describe their place: Identifying predominant types of place descriptions. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information - GEOCROWD '12* (2012), M. Goodchild, D. Pfoser, and D. Sui, Eds., ACM Press, p. 30. doi:10.1145/2442952.2442959.

[38] ROLLER, S., SPERIOSU, M., RALLAPALLI, S., WING, B., AND BALDRIDGE, J. Supervised text-based geolocation using language models on an adaptive grid. *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning* (2012), 1500.

[39] SEKER, A., BANDEL, E., BAREKET, D., BRUSILOVSKY, I., GREENFELD, R., AND TSARFATY, R. AlephBERT: Language model pre-training and evaluation from sub-word to sentence level. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (2022), pp. 46–56.

[40] SIEGEL, A. W., AND WHITE, S. H. The development of spatial representations of large-scale environments. *Advances in child development and behavior 10* (1975), 9–55. doi:10.1016/S0065-2407(08)60007-5.

[41] TEITLER, B. E., LIEBERMAN, M. D., PANOZZO, D., SANKARANARAYANAN, J., SAMET, H., AND SPERLING, J. NewsStand: A new view on news. In *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems - GIS '08* (2008), ACM Press, p. 1. doi:10.1145/1463434.1463458.

[42] TOLMAN, E. C. Cognitive maps in rats and men. *Psychological Review 55*, 4 (1948), 189–208. doi:10.1037/h0061626.

[43] TOMKO, M., AND WINTER, S. Recursive construction of granular route directions. *Journal of Spatial Science 51*, 1 (2006), 101–115. doi:10.1080/14498596.2006.9635067.

[44] TOMKO, M., AND WINTER, S. Pragmatic construction of destination descriptions for urban environments. *Spatial Cognition & Computation 9*, 1 (2009), 1–29. doi:10.1080/13875860802427775.

[45] TSARFATY, R., BAREKET, D., KLEIN, S., AND SEKER, A. From SPMRL to NMRL: What did we learn (and unlearn) in a decade of parsing morphologically-rich languages (MRLs)? In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (2020), Association for Computational Linguistics, pp. 7396–7408. doi:10.18653/v1/2020.acl-main.660.

[46] TSARFATY, R., SADDE, S., KLEIN, S., AND SEKER, A. What's wrong with Hebrew NLP? and how to make it right. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations* (2019), Association for Computational Linguistics, pp. 259–264. doi:10.18653/v1/D19-3044.

[47] TSARFATY, R., SEDDAH, D., GOLDBERG, Y., KÜBLER, S., VERSLEY, Y., CANDITO, M., FOSTER, J., REHBEIN, I., AND TOUNSI, L. Statistical parsing of morphologically rich languages (spmrl) what, how and whither. *Proceedings of the NAACL HLT 2010 First Workshop on Statistical Parsing of Morphologically-Rich Languages* (2010), 1.

[48] WALLGRÜN, J. O., KARIMZADEH, M., MACEACHREN, A. M., AND PEZANOWSKI, S. GeoCorpora: building a corpus to test and train microblog geoparsers. *International Journal of Geographical Information Science 32*, 1 (2018), 1–29. doi:10.1080/13658816.2017.1368523.

[49] WANG, X., XIONG, W., WANG, H., AND WANG, W. Y. *Look Before You Leap: Bridging Model-Free and Model-Based Reinforcement Learning for Planned-Ahead Vision-and-Language Navigation*, vol. 11220 of *Lecture notes in computer science*. Springer International Publishing, 2018, pp. 38–55. doi:10.1007/978-3-030-01270-0_3.

[50] WING, B., AND BALDRIDGE, J. Simple supervised document geolocation with geodesic grids. *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies* (2011), 955.

[51] WING, B., AND BALDRIDGE, J. Hierarchical discriminative classification for text-based geolocation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2014), Association for Computational Linguistics, pp. 336–348. doi:10.3115/v1/D14-1039.

[52] YUAN, X., CROOKS, A., AND ZÜFLE, A. A thematic similarity network approach for analysis of places using volunteered geographic information. *ISPRS International Journal of Geo-Information 9*, 6 (2020), 385. doi:10.3390/ijgi9060385.

[53] ZITOUNI, I. *Natural language processing of semitic languages*. Theory and applications of natural language processing. Springer Berlin Heidelberg, 2014. doi:10.1007/978-3-642-45358-8.