



SPACE REF: A Corpus of Street-Level Geographic Descriptions

JANA GÖTZE & JOHAN BOYE

SUMMARY

Humans **reason about space primarily in a qualitative way** and refer to objects and relations between objects when talking about space. In contrast, **geographic systems represent space quantitatively**. Understanding the relationship between these two representations is necessary for automatic wayfinding systems to give more natural and more easily understandable route directions. This article describes SPACE REF, a corpus of street-level geographic descriptions. Pedestrians are walking a route in a real urban environment, describing their actions. Their **position** is automatically logged, their **speech** is manually transcribed, and their **references to objects** are manually annotated with respect to a crowdsourced **geographic database**. We describe how the data was collected and annotated, and how it has been used in the context of creating resources for an automatic **pedestrian wayfinding** system.

Wizard-of-Oz Data Collection

Task:

Describe your way to the system so that it can retrace your path without knowing where you are

Material:

- Map
- Phone app / headset

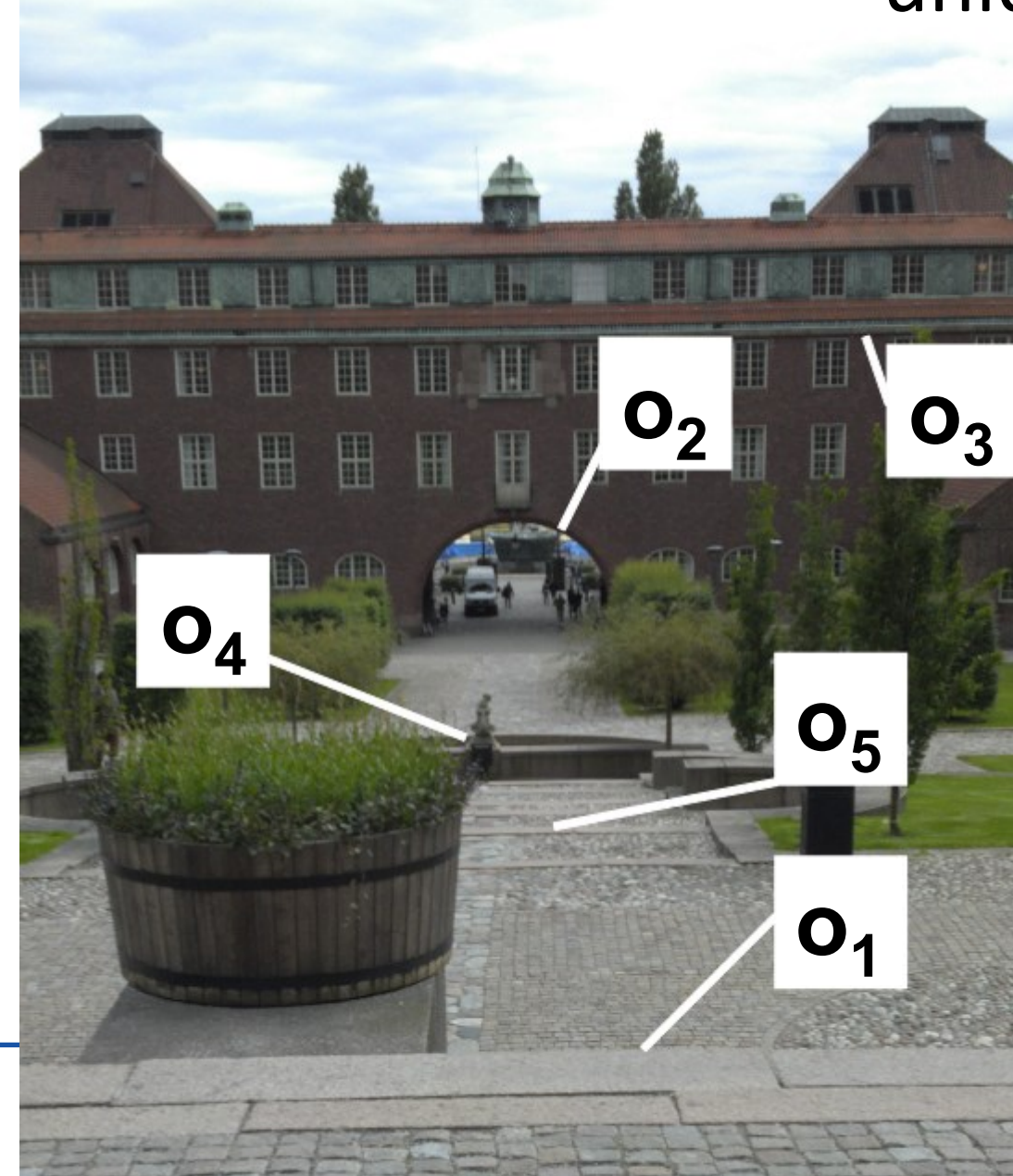
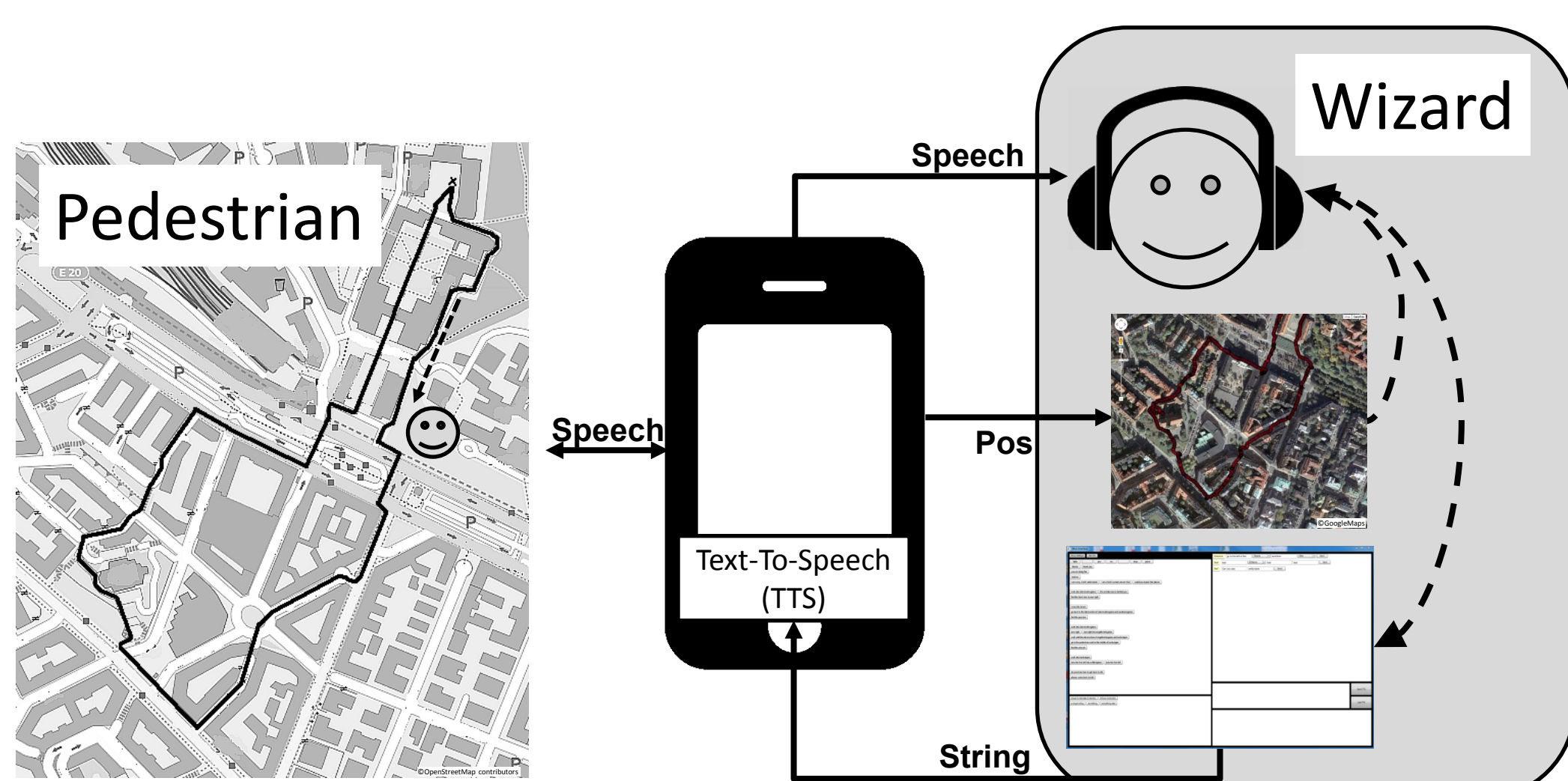
Recordings / Data:

- Transcribed speech
- GPS coordinates and path

Annotation:

- Mentioned objects (OSM snapshot)

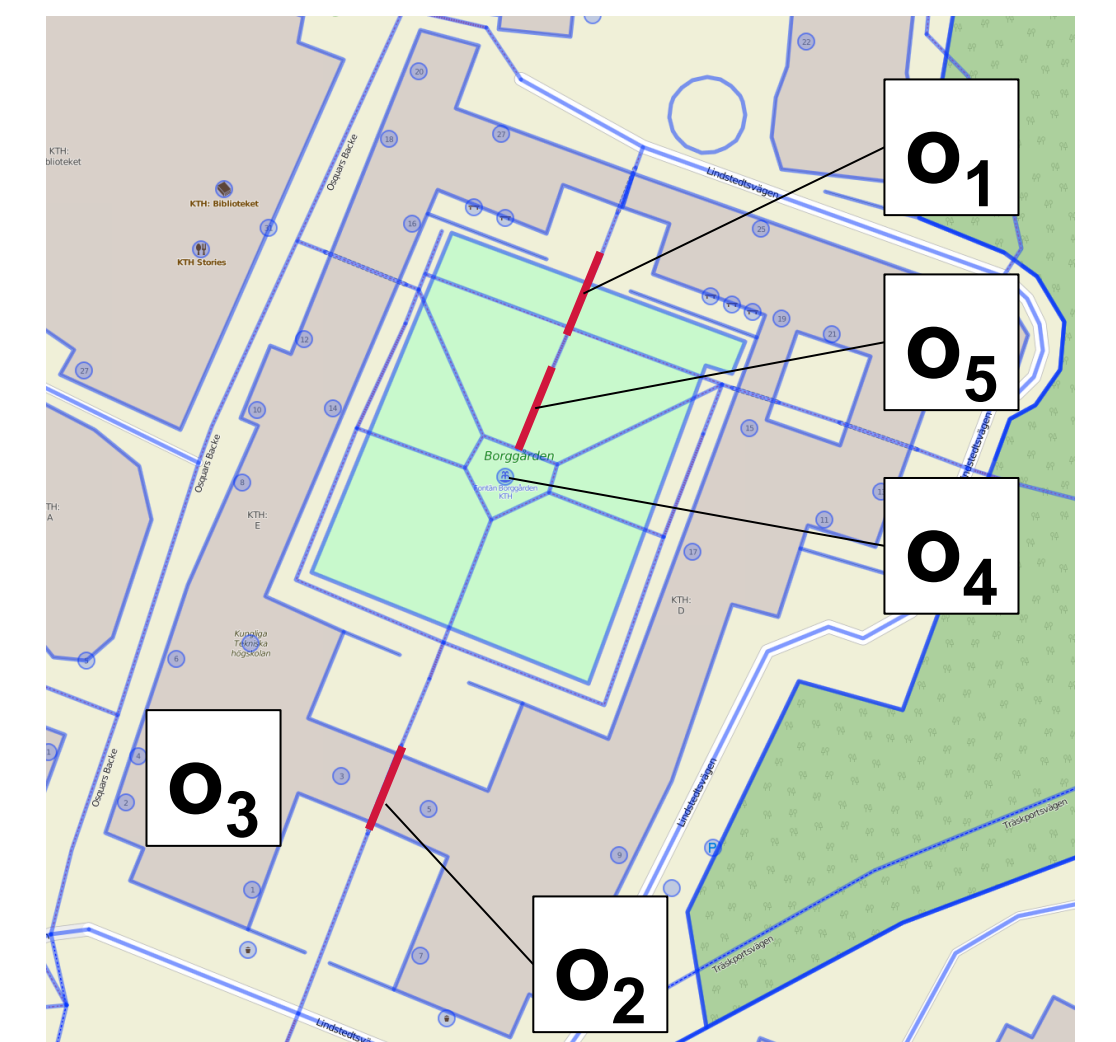
Number of participants	11
Male / female participants	9 / 2
Average age	27.4
Familiarity with the area (1- to 6+)	4.4
Voice application usage (1- to 6+)	1.8
Walking time	5h 44m
Utterances	1,676
Referring expressions	1,323
Unique referents per participant	54.7
REs without referent on the map	58 (4%)



OpenStreetMap (OSM)

- crowd-sourced geographic data (Odb Licence)
- good coverage of streets and buildings in urban areas
- data includes also semantic information about objects, e.g. their name or their type

Number of nodes	29,451
Number of ways	4,031
Number of unique tags	5,142
Number of unique tag keys	210
Average number of objects in candidate set	33



- Object representation:
 - nodes: latitude/longitude
 - ways: set of nodes
 - unique IDs

O₂ the arch

```
<way id="163"><nd ref="30"/><nd ref="35"/>
<tag k="highway" v="footway"/>
<tag k="layer" v="-1"/>
<tag k="tunnel" v="yes"/>
</way>
```

O₄ a fountain

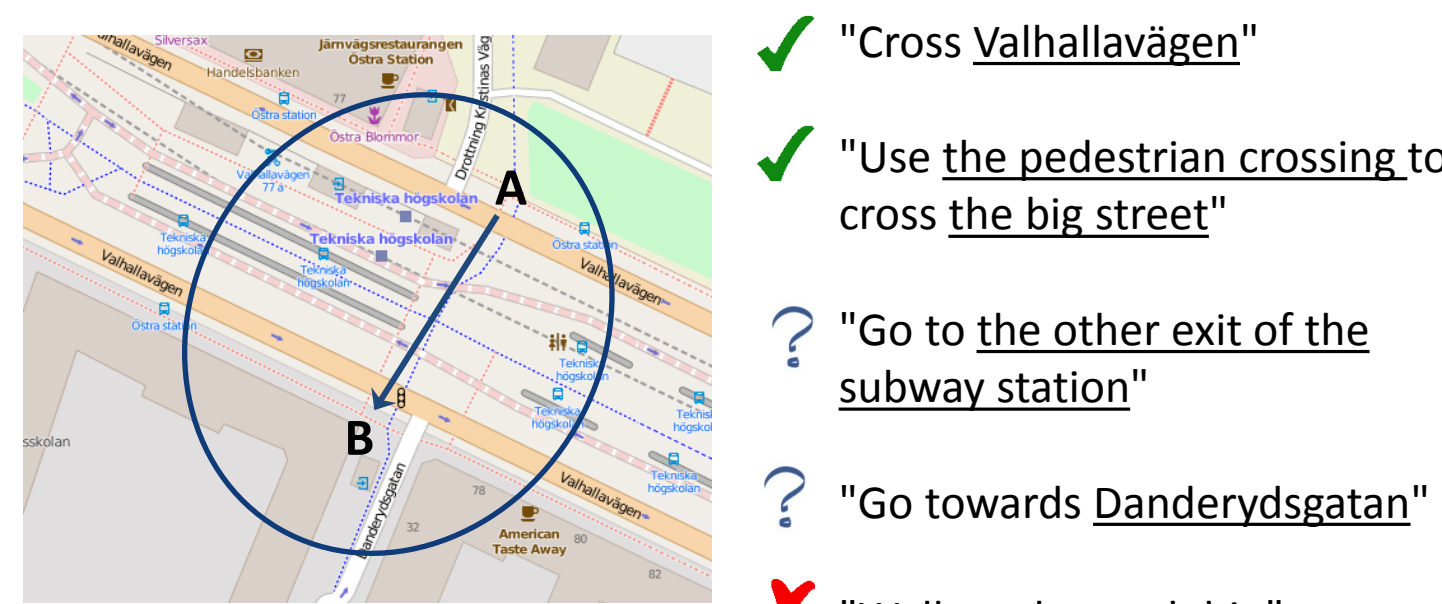
```
<node id="160" lat="59.3474432" lon="18.0737897">
<tag k="amenity" v="fountain"/>
<tag k="name" v="Fontän Borggården KTH"/>
</node>
```

Corpus Usage: Pedestrian wayfinding

Generation:

Learning about landmark usage

- derive salience models from pedestrians' preferences in using landmarks
- rank map objects



"I continue down the steps towards the arch at the bottom"
time : '2:1:10:14:41:8571'
latitude : "59.34787" longitude : "18.07406"
RE₁ : "down the steps" id="1" referent id="20680216"
RE₂ : "towards the arch at the bottom" id="2" referent id="163195369"

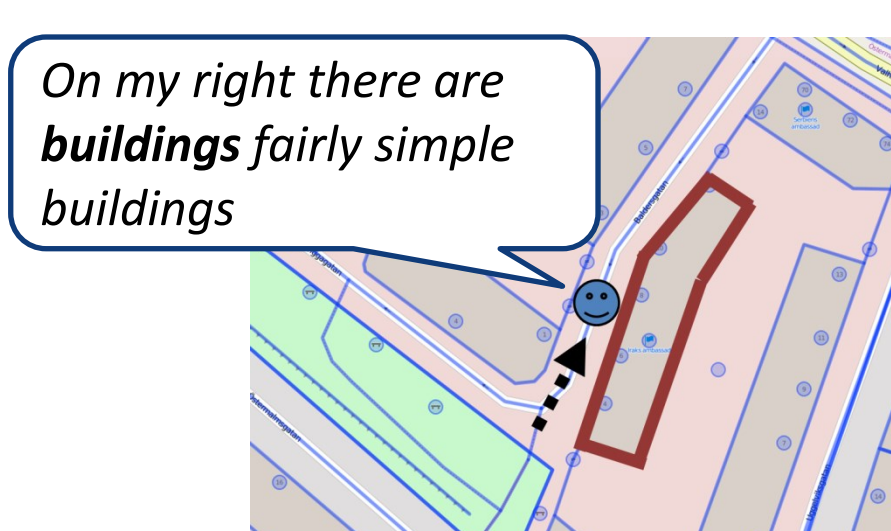
Other example utterances:

- "Okay you go uphill on the street Baldersgatan"
- "And then you take the first left"
- "You don't turn into this street so keep straight going ahead"
- "You should walk down the street until you reach a roundabout"
- "You can now see the church"
- "You should cross the road towards the university"
- "Now I come to a junction"
- "And once again I come to a pedestrian crossing"

Understanding:

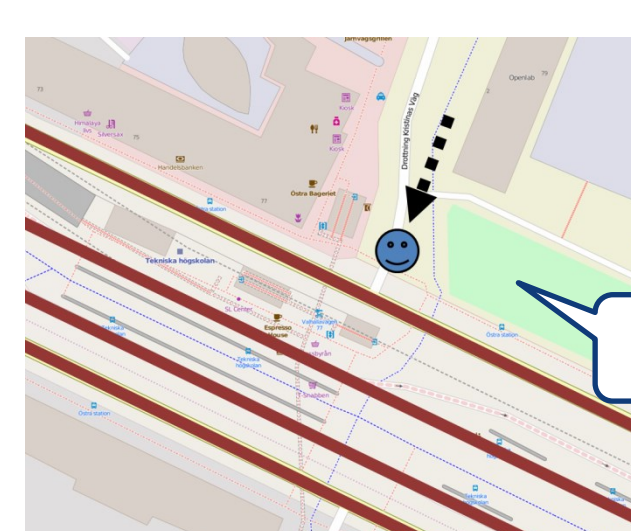
Resolving object references to map objects

- learn to connect words to object features



Case I: 1 object

```
<way id="321">
<nd ref="1"/> <nd ref="2"/> <nd ref="3"/>
<nd ref="4"/> <nd ref="5"/> <nd ref="6"/>
<nd ref="7"/> <tag k="building" v="yes"/>
</way>
```



Case II: >1 object

```
<way id="178">
<nd ref="10"/> <nd ref="11"/> <nd ref="12"/>
<tag k="foot" v="no"/>
<tag k="highway" v="secondary"/>
<tag k="name" v="Valhallavägen"/></way>

<way id="332">
<nd ref="20"/> <nd ref="21"/>
<tag k="highway" v="footway"/> </way>
```



Case III: no object

Cross the street

paper:



corpus:



<http://www.lrec-conf.org/proceedings/lrec2016/summaries/530.html>
<http://www.csc.kth.se/~jagoetze/data/spaceref>

KTH ROYAL INSTITUTE OF TECHNOLOGY

Jana Götz jagoetze@kth.se KTH Royal Institute of Technology
Johan Boye jboy@kth.se School of Computer Science and Communication
Department of Speech, Music and Hearing
100 44 Stockholm, Sweden

Acknowledgement:

This work was supported by the EU 7th Framework Programme (grant #270019) SPACEBOOK, and the Swedish Research Council (VR 2013-4854) "Personalized spatially aware dialogue systems"