



Christoph Purschke*

Crowdscapes. Participatory research and the collaborative (re)construction of linguistic landscapes with *Lingscape*

<https://doi.org/10.1515/lingvan-2019-0032>

Received May 18, 2019; accepted January 14, 2020

Abstract: Public signage is a central element of the socio-pragmatic organization of everyday practice. The survey of signs and written language in the public sphere has developed into a vital branch of sociolinguistics called “linguistic landscapes”. The paper introduces a participatory research project, Lingscape, that focuses on the documentation and analysis of linguistic landscapes worldwide. Making use of a dedicated mobile research app, the project aims at creating awareness for the semiotic complexity and social relevance of public signage. After the discussion of two basic functions of public signage and two fundamental pragmatic conditions of sign perception, the text briefly introduces the app and project workflow. This is followed by the results of an analysis of the Vienna linguistic landscape using a large dataset collected with the app. This includes a quantitative evaluation of the contributions by different user groups to the (re)construction of a Vienna “crowdscape” as well as a qualitative investigation of the presence and social status of Austrian German in said crowdscape. Finally, the text presents an in-depth discussion about challenges for participatory research in practical work that focuses on the objectives of participatory research and the practical necessities against the backdrop of academic practice.

Keywords: Crowdsourcing; Citizen Science; Linguistic Landscape; participatory research; multilingualism.

1 Signage in everyday practice

Linguistic landscapes (LL) are a core feature of everyday social practice: moving through public space, people pass by and interact with countless signs and written language on a daily basis. Public signage conveys rich social information about a given location or society, i.e., through the presence/absence and hierarchization of languages or through different translation modes for depicted information (see Gorter 2013 or van Mensel et al. 2016 for an introduction to LL research). Generally speaking, there are two basic characteristics of signs in the public sphere:

- They serve a practical purpose in providing *socio-pragmatic orientation* by conveying information, giving instructions, regulating practice or addressing a particular audience.
- They are specifically symbolically charged and thus provide information about the *socio-semiotic structuring* of a given space in terms of cultural, linguistic, political or social frameworks for practice.

In everyday life it is common for people to only consciously interact with signage to a limited extent, especially in places they are familiar with, because most signs they encounter, for example on the way to work, are either irrelevant for what they are doing at that moment or simply don’t attract their attention (any more). For other groups of people, such as tourists, newcomers, or those who struggle to access public information (e.g., due to lacking language skills or illiteracy), interaction with the linguistic landscape may pose a significant challenge, whether for purposes of practical orientation (“finding things”) or social integration (“abiding by the rules”).

Christoph Purschke: University of Luxembourg, Faculty of Humanities, Education and Social Sciences, Esch-sur-Alzette, Luxembourg, E-mail: christoph.purschke@uni.lu



Figure 1: An example of a very salient sign at Luxembourg airport.

There are two basic pragmatic conditions which can trigger a conscious interaction with public signage (see Purschke 2014 for a theoretical discussion):

- *practical relevance*, i.e., if the signage fulfills a specific purpose for which it is then *pertinent* (= practically relevant) within the context of a current action plan. Examples include looking for directions or being in need of a pharmacy, or someone who cannot otherwise access information easily.
- *contextual conspicuity*, i.e., when the particular nature of the signage’s semiotic, visual, or spatial characteristics renders it *salient* (= contextually conspicuous) within the realm of current expectations. Examples include eye-catching lettering or design, or a choice of placement that attracts attention (see Figure 1).

Establishing these two basic characteristics of signage in public and the two pragmatic conditions of sign perception allows one a promising starting point for researching LL in the context of a participatory research project. Public signage affects human action in the lifeworld, that is, the negotiation and structuring of everyday practice by *actors using symbols* (such as, signs and languages). Signage also affects the semiotic structure of this lifeworld, that is, the intelligibility of this world and the order of everyday practice through *signs as symbolic actors* (and substitutes for actors). By addressing both aspects of everyday practice, a participatory LL project can create awareness of the (socio-)semiotic structure and complexity of signage – or the public domain as a *semiotic space*. Furthermore, such a project can help foster awareness of the (socio-)pragmatic function and relevance of signage – or the public as an *actioning space*.

Against this backdrop, the project *Lingscape – Citizen science meets linguistic landscaping* intends to create a participatory and interactive research platform for the study of LL around the world.

2 Lingscape – A participatory LL project

The *Lingscape* project, an initiative by Christoph Purschke and Peter Gilles, is hosted at the University of Luxembourg.¹ It focuses on the collaborative documentation and analysis of signage in the public sphere using a dedicated mobile research application and crowdsourcing technology. Methodologically, the project builds on a citizen science framework: participants actively contribute to all aspects of project work (besides data

¹ See <https://lingscape.uni.lu>.

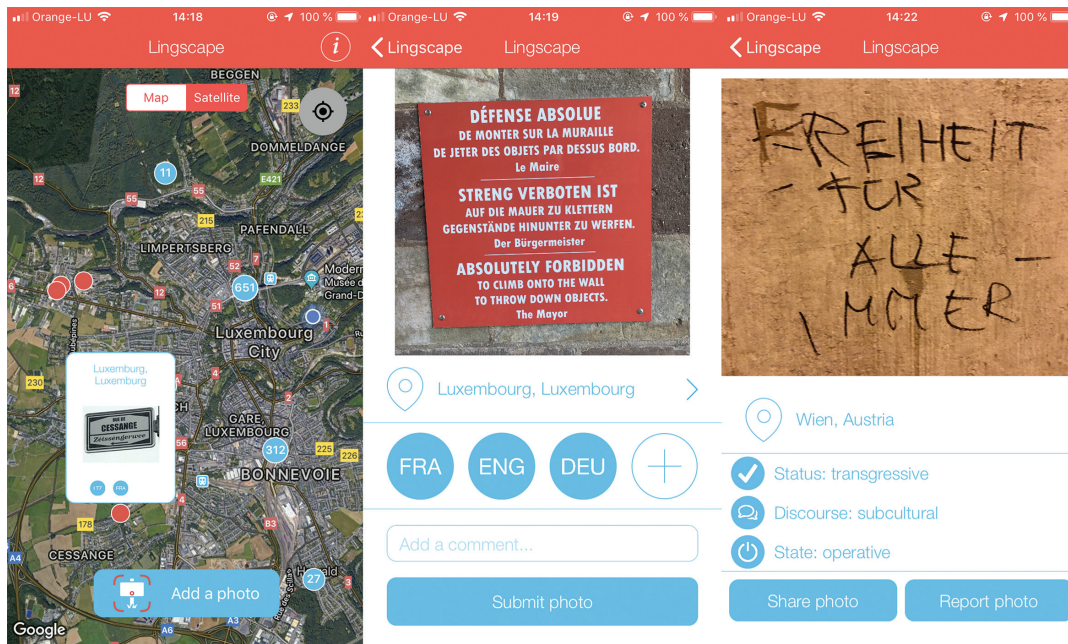


Figure 2: The three main functions of the mobile research app *Lingscape*: map viewer (left), upload function (center), advanced mode (right).

collection), including data processing and analysis as well as dissemination of results and technical development. Such an approach is in line with the overall goals of the citizen science movement that strives for an *opening* (in terms of public participation in research activities), *democratization* (in relation to a shared authority between citizens and scientists) and *social embedding* (in the form of a societal engagement of researchers) of academic research (see Irwin 1995; Strasser et al. 2019).

All project work is based on the mobile research app *Lingscape* which was first released in fall 2016 and can be downloaded worldwide and free of charge on Android and iOS devices. The app offers three main functions (see Figure 2): a *map viewer* to explore all uploads and metadata; an *upload function* for contributions, i.e., a guided process by which participants can choose, adjust, and annotate photos; and an *advanced mode* for sub-projects, offering freely customizable taxonomies and annotation categories. App usage and data access are open and anonymous: no personal login is required and the app does not collect any personal information. All uploads are instantly published on the map and moderated ex-post to avoid misuse and inappropriate material (for a detailed description of app functions and design, see Purschke 2017a).

To facilitate app usage and thus contributions by the general public, the app uses a basic data scheme that requires very little information about a photo (i.e., location, visible languages, time stamp). Advanced annotation possibilities (apart from comments) are currently only used in dedicated sub-projects in need of an in-depth analysis of sign characteristics, such as script types, language hierarchization strategies and sign material. Project leaders can access and administer project data via a web frontend. Participants can easily join a project by entering a project password in the app settings screen. Project data are visible to the public by default, but visibility of sub-projects can be restricted to participants for data collection purposes. Additionally, all public photos are accessible in an interactive online map that offers dynamic analysis widgets (using the map service CARTO).²

To date (August 2019), the *Lingscape* project has collected more than 18,000 publicly available photos contributed by more than 1000 unique participants – including data from more than 100 sub-projects carried out by project partners worldwide, among them LL researchers, teachers and private initiatives. There are in addition more than 5000 photos in private sub-projects currently hidden from the public. All collected

² See <https://lingscape.carto.com>.

data are stored on a dedicated server at the University of Luxembourg and are processed in three different ways: *manually* (ex-post moderation and administration); *computationally* (transformation and transfer to the CARTO map); and *geostatistically* (filtering and analysis of data based on metadata).

3 Vienna – A linguistic crowdscape

To illustrate the scientific potential of crowdsourced data collected through the *Lingscape* project, the following section presents results from an analysis of the Vienna linguistic landscape (for a comparison of the multilingual make-up of Vienna and Luxembourg, see Purschke forthcoming). Starting from an analysis of the distribution, frequency and presence of languages in the Vienna LL, the discussion then turns to the contributions by different types of users before moving on to a qualitative analysis of instances of Austrian German in the dataset. Given that the data stem from the contributions made by a multitude of participants and therefore largely reflect the participants' personal (salience- or pertinence-based) choices (see Comber et al. 2016), the dataset only represents a small subsample of all signs available in the Vienna LL. To account for that fact, I will refer to this kind of collaborative (re)construction of a given LL as a *linguistic crowdscape*.

As of April 2019, there were 2689 photos from Vienna in the dataset (see Table 1). The vast majority of signs contain just one or two languages.³ The data show a high number of monolingual signs for Vienna (70.2%), and only a few contributions with more than two languages per sign. Taken together, mono- to quadrilingual signs account for more than 99% of the collected data, including the 202 signs with missing language labels (these signs have been checked manually and show no deviation from the pattern). Plotting the distribution of mono- and multilingual signs on a map of the city (see Figure 3) reveals that, at least in our dataset, signs with three or four languages seem to be limited to individual streets with a large amount of shops and restaurants.

In relation to the presence and frequency of languages in the Vienna crowdscape, the analysis reveals a strong dominance of German and a prevalence of English signs in the sample (see Table 2): while German is present in 84.4% and English in 26.6% of all signs, the next most frequent languages of French and Italian appear only sparsely in the dataset, mostly in the context of restaurant menus. Given the sociolinguistic make-up of Vienna, this result is not surprising, as one would expect similar results in many other cities in the German-speaking area. This result also illustrates a general problem with crowdsourced data, however: the composition of the dataset entirely relies on the participants' personal choices. Eastern European languages, for example, such as Croatian, Czech or Slovenian, are completely missing in the dataset despite their relatively strong presence in the Vienna LL; in a recent sub-project called “Südslawische Sprachlandschaften in Wien” [South Slavic LL in Vienna] led by Katharina Tyran, a group of students focused exclusively on these varieties, thus changing the multilingual make-up of the Vienna crowdscape entirely.

Table 1: Language statistics for the Vienna crowdscape.

Languages per sign	Vienna
1	1888 (70.2%)
2	511 (19.0%)
3	64 (2.4%)
4	9 (0.3%)
5+	15 (0.6%)
Signs missing labels	202 (7.5%)
Total	2689
Total 1–4 languages	2472 (91.9%)
Average languages per sign	1.27

³ The term *sign* represents one photo in the Lingscape database. In practice, some photos contain several different signs, e.g., a collection of transgressive stickers on the back of a street sign.

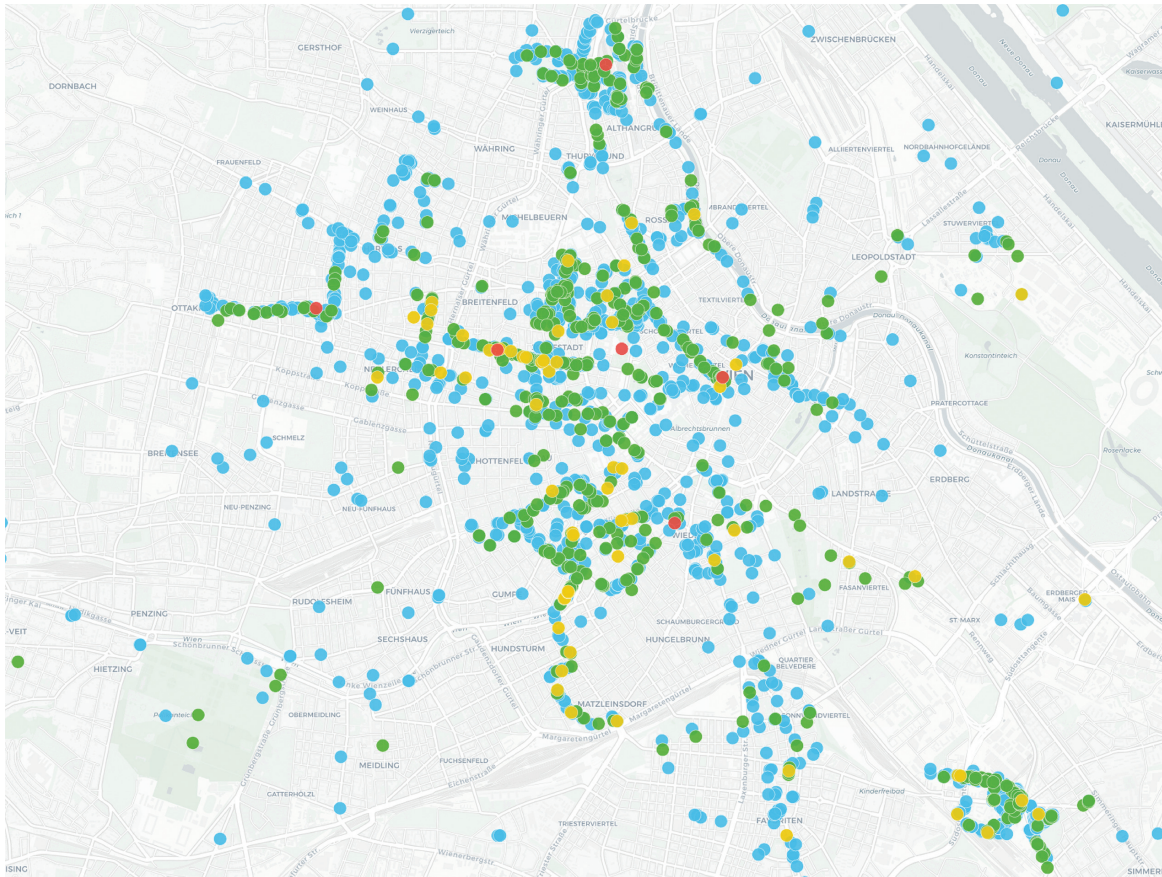


Figure 3: Mono- and multilingual signs in the Vienna crowdscape: monolingual (blue), bilingual (green), trilingual (yellow), quadrilingual (red).

Table 2: Language frequency in the Vienna crowdscape.

Languages per sign	German	English	French	Italian
1	1546 (81.9%)	226 (12.0%)	26 (1.4%)	20 (1.1%)
2	484 (94.7%)	372 (72.8%)	30 (5.9%)	34 (6.7%)
3	60 (93.8%)	52 (81.3%)	22 (34.4%)	14 (21.9%)
4	7 (77.8%)	7 (77.8%)	5 (55.6%)	5 (55.6%)
Total (of labeled signs)	2097 (84.8%)	657 (26.6%)	95 (3.4%)	73 (3.0%)

Against the backdrop of such limitations, and the overall linguistic make-up of the Vienna crowdscape, it is interesting to analyze the contributions of different groups of users to the overall image. Although participation in *Lingscape* is anonymous, contributions by individual users can be grouped by a device-specific technical identifier that is transmitted with every upload (due to legal requirements). Prior research (see Purschke forthcoming) has already identified different *user types* (in terms of active months and number of uploads) as well as personal *spatial orientation strategies*: while most participants (called *casual users*) contribute to data collection for only a very limited amount of time (1–2 months) and add only a few signs to the map (5–6 photos), there are also smaller groups of *regular* (5 months and 90 uploads on average) and *power users* (10 months and 500 uploads on average). Likewise, users show different spatial orientation patterns in the LL in focusing on specific areas or streets, by strolling in a specific area or covering the city exhaustively.

As can be seen in Table 3 is below now, all three groups contribute similarly to the dataset in terms of total uploads, but while there are more than 100 participants responsible for the 1082 uploads of the casual

Table 3: Contributions to the Vienna crowdscape by user type.

Languages per sign	Casual users	Regular users	Power users
1	707	659	522
2	238	193	80
3	44	14	6
4	6	3	0
Signs missing labels	73	129	0
Average languages per sign	1.37	1.10	1.15
Users total	101	12	1
Uploads total	1082	999	608

user group, the power user group is represented by one sole participant, contributing 608 photos alone to the crowdscape.

In terms of the number of languages per sign, the greatest diversity is expressed in the casual user group, and the least in the power user group. The reason for this becomes apparent when visualizing contributions to the Vienna crowdscape by user group (see Figure 4): while photos from casual users are scattered across the entire city, regular users seem to concentrate on specific areas and individual streets in their contributions (e.g., as part of a specialized sub-project), whereas the single power user focuses on three specific residential neighborhoods (with limited linguistic diversity).

While these results are primarily quantitative, the dataset can also be used to take a closer look at the societal role of individual languages or language varieties from a qualitative point of view. In the case of the Vienna crowdscape, for example, the presence and strategic use of Austrian German in public signage is an interesting case study. In our dataset, Austrian German variants are present on 165 photos representing 7.7% of all German signs in the sample (see Table 4; for the sake of this analysis, family and street names have been left out). These variants appear almost exclusively in monolingual (i.e., German) contexts. The individual instances of Austrian German can be assigned to different *author domains*, representing types of social actors, and *discourse types*, reflecting different kinds of practical purposes of signs (see Reh 2004).

Most of the signs containing Austrian German variants in the dataset originate from authors representing either *economic* or *institutional* actors (e.g., supermarkets or local administration), while only a minority of signs are sourced from *private* authors. While *commercial* signs are exclusively of an economic kind, namely advertisements, business signage or in the context of gastronomy (see Figure 5, left), Austrian German is used in very different discursive contexts in institutional signage, i.e., giving instructions to regulate certain aspects of practice (*regulatory discourse*), organizing (interaction with) public infrastructure (*infrastructural discourse*), or giving information to the public (*informatory discourse*, see Figure 5, middle). Private signs, on the other hand, serve a variety of purposes, i.e., personal expression (*expressive discourse*), political protest (*political discourse*, see Figure 5, right), or addressing certain subcultures (*subcultural discourse*).

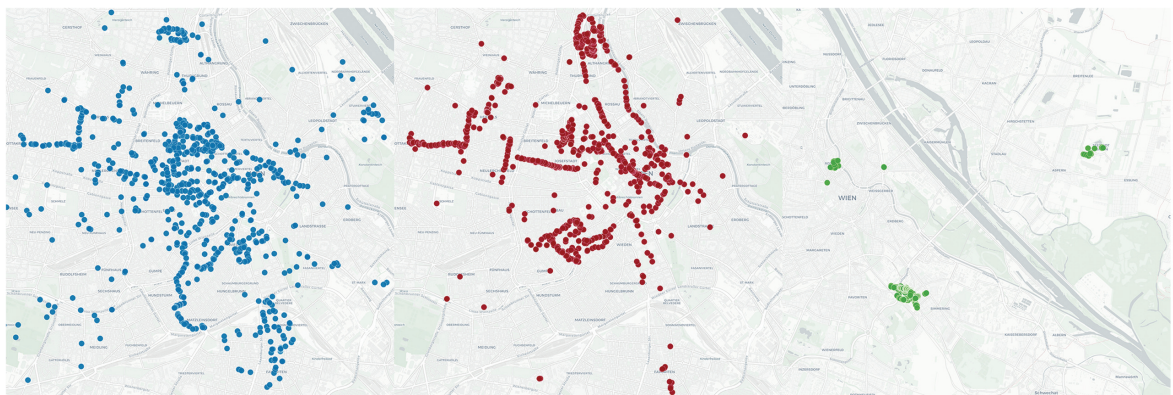
**Figure 4:** Contributions to the Vienna crowdscape by user type: casual users (left), regular users (center), power users (right).

Table 4: Author domains and discourse types for instances of Austrian German in the Vienna crowdscape.

Author domain	Discourse type	Description	Frequency
Economic	Commercial	Advertisements, gastronomy and business signage	69
Institutional	Infrastructural	Communication regarding public infrastructure, e.g., street signs or instructions on public trash cans	34
	Informatory	Announcements about public events, opening hours or elections	16
	Regulatory	Messages for regulating public behavior, e.g., prohibition signs	13
Private	Expressive	Messages about social, cultural or private issues	19
	Political	Communications regarding political issues and protest	6
	Subcultural	Messages targeting certain subcultures, e.g., skateboarding, hip-hop or soccer	8

**Figure 5:** Examples of Austrian German variants in the Vienna crowdscape: commercial (left), informatory (center), and political (right) discourse.

The patterns of use of regional variants in the Vienna LL illustrate the clear role of Austrian German as a sociocultural identifier, providing relevant insight into the debate concerning the linguistic status and (socio)symbolic value of Austrian German as a national variety of German, as opposed to Germany and Switzerland. The results indicate that, while Austrian German is present in the Vienna LL, its use is mainly restricted to economic and institutional purposes. The case study of the Vienna crowdscape ultimately demonstrates how qualitative analysis of a given LL can contribute to a better understanding of how the public sphere, as a complex and dynamic socio-semiotic space, is structured, maintained and perceived by the actors that move through it.

4 Potential and shortcomings of participatory research

With the rise of crowdsourcing as a means of data collection and processing (see Brabham 2013), the growing number of app-based projects within linguistics (see Leemann et al. 2016) and the general tendency to include participant perspectives in academic research projects (see Riesch and Potter 2014), it has become possible to collect large data samples from a broad demographic spectrum of participants with relatively little effort. A number of concerns have arisen, however, regarding the methodological and empirical aspects of such initiatives, particularly with respect to control over data quality and the reliability of user-generated content (see Wang et al. 2016). Neither of these concerns are yet to be suitably addressed in linguistics.

Within the context of the *Lingscape* project, however, we have tried to relevantly address concerns related to the methodological aspects of crowdsourcing (see Purschke 2017a), the theoretical implications of a citizen

science approach (see Porschke 2017b) and the use of an app as a digital teaching and learning resource (see Porschke 2018). These discussions have revealed both the innovative potential and the shortcomings of a participatory approach to linguistic landscapes research. The following section will discuss the relative strengths and weaknesses of this approach in addition to detailing the practical aspects of participatory research.

4.1 Participatory research vs. academic practice

A citizen science approach opens up new possibilities for linguistic research by integrating non-specialists into academic practice. In the context of a participatory LL project, the value of this approach is evident in the diversity of personal perspectives on public signage, allowing for the collaborative (re)construction of linguistic crowdsapes, as opposed to a (limited) expert-view driven by a specific research interest. The analysis of user participation and the deeper insight into the multilingual make-up of Vienna demonstrate the potential of such user-generated content. At the same time, participatory research requires a variety of activities that make it necessary to rethink many established routines of research as an academic practice and habitus. For example, a lot of time has to be invested in the preparation of materials (e.g., leaflets, stickers, websites), outreach to the participant community (e.g., via social media, see Entringer et al. in this volume) and dissemination of results to the greater public, in addition to the infrastructural aspects, such as app development and technical maintenance of servers. For small-scale projects like *Lingscape* operating without support by research assistants, these factors pose a substantial challenge. Moreover, conducting a participatory research project contributes to a reinterpretation of the role of academic researchers in light of participant engagement and research that is embedded in and motivated by societal practice (see Chevalier and Buckles 2013). While participatory research activities empower citizens to take an active role in knowledge production, the role of trained researchers in these projects largely revolves around activities such as providing guidance and training, troubleshooting problems, hosting public events or organizing exchanges between citizen scientists, the scientific community and societal stakeholders.

4.2 Co-creation vs. crowdsourcing

Within the field of citizen science, there are different methodological models regarding the level of citizen engagement and project workflows (see Schrögel and Kolleck 2019). Bonney et al. (2009: 17–18) distinguish between three levels of participation in this context: *contributory projects* that make use of citizen-generated content but foresee no further involvement, *collaborative projects* designed by scientists in which citizens may also participate in project design, data processing or dissemination of findings, and *co-created projects* which are designed jointly by scientists and citizens and involve the public in all steps of project work. With *Lingscape* we have aimed for a co-created project that combines aspects of crowdsourcing, outreach activities and joint project work. App development follows feedback from the participant community, concerning for example the addition of new languages and analytical descriptors or the implementation of new features. Sub-projects are free to pursue their specific research interests using the *Lingscape* platform and to shape the workflow according to their needs. A prominent difficulty when working with the general public, however, is that of “motivation asymmetry” between the participants and the researchers, whereby citizen engagement, interest and motivation can be difficult to maintain over the long-term (see Füchslin et al. 2019 for a discussion of personal attitudes towards engagement in citizen science). Moreover, the quality and quantity of participant contributions largely depend on different types of motives, such as collective, norm-oriented and intrinsic motives as well as the reputation of the project (see Nov et al. 2014 for a discussion of these factors), which researchers thus need to address in order to maintain participation.

4.3 Collaboration vs. communication

As a consequence, a large proportion of project work consists of outreach activities (e.g., media appearances) and participant recruitment (e.g., via social media or public science fairs). Occasions for concrete collaboration (e.g., collaborative LL exploration walks or thematic workshops; see Entringer et al. in this volume), on

the other hand, are often difficult to realize and sometimes poorly attended. One specific challenge in relation to project communication concerns the many available social media platforms and their technical specifications and community practices. Facebook is still the most widely used social network in Luxembourg, so in the first two years of the *Lingscape* project, a Facebook app page was the main method of community-building. Following some changes to the Facebook news algorithm in 2018 (prioritizing private messages), however, it became more difficult to maintain visibility for app pages and other advertising content. Other (image-centered) platforms, such as Instagram, provide an ideal ecosystem for a project like *Lingscape*, but require constant campaigning to successfully build and maintain a participant community, including a high rate of posts, content embedding (via a multitude of hashtags), and prettification of messages. As a consequence, project communication in social media is currently concentrated on Twitter (accepting a certain “academic bias” in public outreach). Additionally, we have started to experiment with different blogging platforms to collect fieldwork reports from sub-projects (via the project website) and offer in-depth analyses of individual photos (on tumblr).

4.4 Free exploration vs. methodological approach

One of the most important notions of citizen science involves a “shared authority” (Frisch 1990) between citizens and scientists, i.e., the possibility to develop and implement research topics collaboratively between researchers and project participants. This approach has become more and more popular in the last 30 years and has successfully resulted in novel insights in fields as different as astronomy, economics and medical research (see Garbarino and Mason 2016). As project participants are not trained experts in the field of research, however, even if they develop a certain amount of thematic expertise over the course of the project, an open research platform such as *Lingscape* bears the risk of a heterogeneous dataset and “faulty” contributions; quality and reliability of crowdsourced data are two of the main concerns in relation to participatory research (see Lewandowski and Specht 2015). As an illustration, one participant in the southwest of the United States persistently tagged bilingual signs as supposedly showing French lettering next to English. While these annotations were clearly incorrect (all the signs contained Spanish instead of French), the contributions are nevertheless a representation of the participant’s perception of the LL; “correcting” them would therefore constitute tampering with the data. In order to minimize such instances while still allowing participants to freely explore the LL, the in-app tutorial only provides basic instructions about the task to fulfill, while further information about the social semiotics of signs can be found on the project website. In addition, we are experimenting with automated language extraction from photos using AI technology to compare participants’ perceptions of the LL with factual sign content.

4.5 User orientation vs. data optimization

The example of the participant from the United States is part of a deeper consideration between, on the one hand, an open, motivating, user-friendly application and, on the other hand, a clean, representative, fully annotated dataset. Our primary interest in development is to cultivate an enjoyable application that is easily accessible to lower the inhibition threshold for contributions (see Sun 2016). At the same time, in-depth analysis of public signage requires additional information about the material, social and linguistic characteristics of the signs. As a consequence (and compromise), we introduced the advanced mode allowing for comprehensive and customizable annotations of photos. The basic version of the app, which addresses the broader public, requires very little information during the upload process.

4.6 Gamification vs. intrinsic motivation

Another method of fostering participation would be the implementation of gamification elements to the app, e.g., rankings, high scores, levels or badges. Research on gamification in crowdsourcing tasks (see Morschheuser et al. 2016) highlights the positive effects of such incentives on participant motivation and

performance, albeit with contrasting evidence regarding the negative impact of gamification in educational settings (see Toda et al. 2018). In the context of the *Lingscape* project, the main reason for not implementing such reward mechanisms relates to a central motivation of the study: to create awareness of the semiotic complexity and social richness of public signage. While adding competitive elements or personal rewards to the app might increase the number of submissions and active months per participant, these measures would inevitably lead to a shift of motivation from intrinsic and *task-oriented* to extrinsic and *reward-oriented*.

4.7 Personalization vs. data protection

Another motivation against adding gamification elements is the difficult balance between personalization of the user experience and data protection requirements. Gamification requires a form of personalized access to the app, whether in the form of a personal account, nickname or unique identifier. As a result of the combination of strict privacy policies in the EU in general and ethical requirements for research at the University of Luxembourg in particular, we opted for an anonymous user model in *Lingscape*. Participants can contribute to the project without disclosing any personal information – except for the aforementioned device-specific (albeit anonymous) identifier that must be included in transmissions for legal reasons (in order to report misuse liable to prosecution to local authorities).⁴ As a consequence, the possibilities for personalization within the app are very limited, despite the negative effect this might have on user engagement.

4.8 Open access vs. commercial interest

One important pillar of the research rationale behind the *Lingscape* project (see Purschke 2017b) is the commitment to open and transparent research practices. This directly affects all technical procedures, data usage and project communication, but also entails a disclosure of the analytical process that creates knowledge. By openly addressing (explicit and implicit) hierarchies in relation to the creation and dissemination of knowledge, we hope to contribute to a critical re-evaluation and further development of the methodological and theoretical foundations of socially responsible research. Another aspect concerns open access to all project-related resources. This is not difficult to implement for data, information materials and project publications (except with regards to copyright restrictions by publishers), but is more challenging with regards to the code used in the apps and web frontend, as a result of working with a commercial software studio that owns the rights to the code and products. The transfer of the project into an open-source repository is currently the subject of negotiations but may be hindered by financial considerations from the developers (i.e., loss of earnings).

4.9 Self-financing vs. third-party funds

The funding basis for participatory research such as this naturally also plays an important role, not only for app development and the successful implementation of projects, but also for achieving defined goals and standards, e.g., in relation to an open-access policy. Funding schemes for citizen science and public outreach initiatives are available on the national and European level but can be difficult to acquire in this line of research, especially when external funding is only needed for app development and maintenance. Fortunately, costs for the *Lingscape* project have so far been entirely covered by internal funds from the institute for Luxembourgish language and literature at the University of Luxembourg. This also makes it possible to further develop the app regardless of project duration or expected outcomes. However, bigger sums for expansions and running costs, i.e., for maintaining and updating the app and server, are a constant challenge for self-financed projects.

⁴ Members of the *Lingscape* project can under no circumstances identify app users using the device-specific universal identifier. In case of misuse liable to prosecution, however, the identifier must be reported to local authorities so that the mobile service provider can match the identifier with a customer for law enforcement purposes.

5 Outlook

The aim of this article was to discuss practical aspects of project work in the context of the participatory LL project *Lingscape* and to demonstrate the potential of user-generated data for academic research. The manifold challenges of participatory research have been discussed and the many necessary compromises revealed which were required for a feasible and successful implementation of crowdsourcing and citizen science, the methodological pillars of the *Lingscape* project. In terms of the potential for quantitative and qualitative analyses of crowdsourced data, the results attest to the manifold insights into participatory research, participant behavior and the socio-pragmatic meaning of signs. On the other hand, the results also demonstrate some of the problems in relation to crowdsourced data, such as the dependence of the collected data on the decisions and interests of the participants. Irrespective of such difficulties, the empirical work in this project offers exciting opportunities to explore linguistic landscapes collaboratively worldwide with an interested public and at the same time to foster awareness for cultural complexity and linguistic diversity in public signage.

One interesting path to follow in this regard is the use of the app in educational settings, whether as part of classroom activities or in university projects (see Gorter 2018). In November 2018, a pilot project was carried out in collaboration with German teachers at the *Deutsche Höhere Privatschule* (DHPS) in Windhoek, Namibia. Two groups of students (11th grade) explored the surroundings of their school in order to document and analyze linguistic diversity against the backdrop of the complex sociocultural situation in Namibia (including colonial history and present-day language policy). Starting from a discussion about the role of Namibian German in society and based on the teaching concept “Language in the city” (see Purschke 2018), data collection revealed a predominance of English in the Windhoek LL, which then led to a problematization of its societal role as a “neutral” official language with no ties to a specific population group (see Frydman 2011). The students identified problems in relation to the accessibility of information (for people from more rural areas), a lack of identification (mostly in the older generation) and a general tension between official language policy and the historically grown cultural diversity as contributing factors. As a result of the success of this pilot project, one future line of work in the *Lingscape* project will focus on the development and implementation of teaching materials that will enable teachers and students to explore LL and critically reflect on the roles of public signage as a rich sociocultural resource and useful analytical lens. In doing so, the project may contribute to a better understanding of, and conscious way of interacting with, linguistic and cultural complexity in everyday life.

6 Data statement

Data and code for this paper is open access and available via our project GitHub repository at: <https://github.com/questoph/lingscape>. The quantitative analysis was performed using a *Jupyter Notebook* with the *Python* packages *Pandas*, *NumPy* and *Scikit-learn* for statistical analysis and *Matplotlib*, *Seaborn* and *CartoFrames* for plotting/map creation.

References

- Bonney, Rick, Heidi Ballard, Rebecca Jordan, Ellen McCallie, Tina Phillips, Jennifer Shirk & Candie C. Wilderman. 2009. *Public participation in scientific research: Defining the field and assessing its potential for informal science education*. A CAISE Inquiry Group Report. Washington, DC: Center for Advancement of Informal Science Education (CAISE).
- Brabham, Daren C. 2013. *Crowdsourcing*. Cambridge: The MIT Press.
- Chevalier, Jacques & Daniel Buckles. 2013. *Participatory action research: Theory and methods for engaged inquiry*. London: Routledge.
- Comber, Alexis, Peter Mooney, Ross S. Purves, Duccio Rocchini & Ariane Walz. 2016. Crowdsourcing: It matters who the crowd are. The Impacts of between Group Variations in Recording Land Cover. *PLoS One* 11(7). e0158329.
- Frisch, Michael. 1990. *A shared authority: Essays on the craft and meaning of oral and public history*. Albany: SUNY Press.

- Frydman, Jenna. 2011. A critical analysis of Namibia's English-only language policy. In Eyamba G. Bokamba, Ryan K. Shosted & Bezza Tesfaw Ayalew (eds.), *Selected Proceedings of the 40th Annual Conference on African Linguistics: African languages and linguistics today*, 178–189. Somerville, MA: Cascadia Proceedings Project.
- Füchslin, Tobias, Mike Schäfer & Julia Metag. 2019. Who wants to be a citizen scientist? Identifying the potential of citizen science and target segments in Switzerland. *Public Understanding of Science* 28(6). 652–668.
- Garbarino, Jeanne & Christopher Mason. 2016. The power of engaging citizen scientists for scientific progress. *Journal of Microbiology & Biology Education* 17(1). 7–12.
- Gorter, Durk. 2013. Linguistic landscapes in a multilingual world. *ARAL – Annual Review of Applied Linguistics* 33. 190–212.
- Gorter, Durk. 2018. Linguistic landscapes and trends in the study of schoolsapes. *Linguistics & Education* 44. 80–85.
- Irwin, Alan. 1995. *Citizen science: A study of people, expertise and sustainable development*. London: Routledge.
- Leemann, Adrian, Marie-José Kolly, Ross Purves, David Britain & Elvira Glaser. 2016. Crowdsourcing language change with smartphone applications. *PLoS One* 11(1). e0143060.
- Lewandowski, Eva & Hannah Specht. 2015. Influence of volunteer and project characteristics on data quality of biological surveys. *Conservation Biology* 29(3). 713–723.
- Morschheuser, Benedikt, Jonna Hamari & Juho Koivisto. 2016. Gamification in crowdsourcing: A review. *Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS), Hawaii, USA, January 5–8*, 4375–4384. Washington, DC: IEEE Computer Society.
- Nov, Oded, Ofer Arazy & David Anderson. 2014. Scientists@Home: What drives the quantity and quality of online citizen science participation? *PLoS One* 9(4). e90375.
- Purschke, Christoph. 2014. “I remember it as if it was interesting.” Zur Theorie von Salienz und Pertinenz. *Linguistik Online* 66(4). 31–50.
- Purschke, Christoph. 2017a. Crowdsourcing the linguistic landscape of a multilingual country. Introducing Lingscape in Luxembourg. *Linguistik Online* 85(6). 181–202.
- Purschke, Christoph. 2017b. (T)Apping the linguistic landscape. Methodological challenges and the scientific potential of a citizen-science approach to the study of social semiotics. *Linguistic Landscape* 3(3). 246–266.
- Purschke, Christoph. 2018. Sprachliche Vielfalt entdecken mit der Lingscape-App. *Der Deutschunterricht*, 4/2018. 70–75.
- Purschke, Christoph. forthcoming. Exploring the linguistic landscape of cities through crowdsourced data. In Stanley Brunn & Roland Kehrein (eds.), *Handbook of the changing world language map*. Heidelberg: Springer.
- Reh, Mechthild. 2004. Multilingual writing: A reader-oriented typology – with examples from Lira Municipality (Uganda). *International Journal of the Sociology of Language* 170. 1–41.
- Riesch, Hauke & Clive Potter. 2014. Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. *Public Understanding of Science* 23(1). 107–120.
- Schrögel, Philipp & Alma Kolleck. 2019. The many faces of participation in science: Literature review and proposal for a three-dimensional framework. *Science & Technology Studies* 32(2). 77–99.
- Strasser, Bruno, Jérôme Baurdy, Dana Mahr, Gabriela Sanchez & Elise Tancoigne. 2019. “Citizen Science?” Rethinking science and public participation. *Science & Technology Studies* 32(2). 52–76.
- Sun, Jun. 2016. Tool choice in innovation diffusion: A human activity readiness theory. *Computers in Human Behavior* 59. 283–294.
- Toda, Armando, Pedro Henrique Dias Valle & Seiji Isotani. 2018. The dark side of gamification: An overview of negative effects of gamification in education. In Alexandra I. Cristea, Ig Ibert Bittencourt & Fernanda Lima (eds.), *Higher education for all: From challenges to novel technology-enhanced solutions*, 143–156. Heidelberg: Springer.
- van Mensel, Luk, Mieke Vandenbroucke & Robert Blackwood. 2016. Linguistic landscapes. In Ofelia García, Nelson Flores & Max Spotti (eds.), *Oxford handbook of language and society*, 423–449. Oxford: Oxford University Press.
- Wang, Yufeng, Xueyu Jia, Qun Jin & Jianhua Ma. 2016. Mobile crowdsourcing: Framework, challenges, and solutions. *Concurrency and Computation: Practice and Experience* 29(3). 1–17.