THE ISTROX PROJECT: FROM DONATION TO COMMUNITY ENGAGEMENT

ADINA BRĂDEANU[1]

Abstract. The article provides an overview of the main steps in the development and the follow-up of the ISTROX pilot project developed within the Faculty of Linguistics, Philology, and Phonetics at the University of Oxford. Planned for 2018–2020, and extended to July 2021 in the context of the Covid pandemic, ISTROX aimed to prepare the ground for a larger envisaged project exploring the history of the Istro-Romanian language (vlaški and žejanski) and community. At the core of ISTROX is a previously unpublished body of sound recordings of Istro-Romanian made in Istria during the 1960s by the Oxford linguist Tony Hurren and donated to the University in the 2000s by his wife, Vera Hurren. Part of the approach of ISTROX entailed initiating the online sourcing of linguistic information from the very small community of remaining speakers of 'Istro-Romanian' worldwide. As the development of the project overlapped with the Covid pandemic, the ISTROX team had to redesign the relationship between the offline and the online elements of the project – a process which is being reoriented towards the offline in post-pandemic context.

Keywords: Crowdsourcing, community-sourcing, participant engagement, pandemic research, Istro-Romanian.

1. INTRODUCTION

In this article I trace and describe the main phases in the development of the project *ISTROX: The Istro-Romanian Language and the Oxford University Hurren Donation* – an interdisciplinary pilot project which combined linguistic research and online community-sourcing to prepare the ground for a larger project aimed at exploring the history of the Istro-Romanian language and community. I follow the timeline of the project, identify its main inflection points, and contextualize some of our decisions and challenges, to take stock of what we achieved during the pilot phase and the two short-term follow-up projects developed after its completion.

My responsibilities as part of ISTROX had to do with contributing to the overall vision of the project, developing the funding bids and ethical approval documentation, planning the structure of the two digital components of ISTROX (the project website and the community sourcing site), liaising with the digital technologists involved in building them, and collaborating with my linguist colleagues (Martin Maiden and Oana Uță

[1] Taylor Instituion Library, The Bodleian Libraries, University of Oxford, adina.bradeanu@bodleian.ox.ac.uk

Bărbulescu) in the production of the content required by these structures; taking part in a research fieldtrip to Istria; conducting the online research needed to prepare specific segments of the project; developing strategies of online and offline outreach and community engagement.

I first came across the Hurren Donation in my capacity as a Subject Consultant for the Romanian collection at the Taylor Institution Library (part of the Bodleian Libraries and Oxford University's centre for the study of Modern European languages and literatures, other than English). In 2010 and 2017, Vera Hurren donated to Oxford University a large amount of unpublished material connected with her late husband's research on Istro-Romanian language. Tony (Antony) Hurren (1933-2006) was an English linguist who, during the 1960s, set out on an almost decade-long series of annual trips to the Istrian villages to study the little-known Istro-Romanian language. In 1972 he completed, at the University of Oxford, a doctoral thesis entitled *A linguistic description of Istro-Rumanian*. The Hurren Donation comprised: approximately thirty hours of sound recordings (on reel-to-reel and cassette tape) made by Hurren in several Istro-Romanian villages in the summers of 1966 and 1967; eight field notebooks written during his research visits, containing transcriptions and/or translations of some of the audio material, as well as linguistic questionnaires, information about the speakers/participants etc.; a bound, but unpublished, grammar of Istro-Romanian (*Istro-Romanian: a functionalist phonology and grammar*, 1999); fifty photographs taken during Hurren's research visits; miscellaneous notes and offprints of articles by Hurren on the Istro-Romanian language.

2. PRELIMINARY CONSIDERATIONS: 'CROWD' VS. 'COMMUNITY'-SOURCING

The ISTROX project started several years after the first donation and has grown organically around it ever since. After an initial assessment by Martin Maiden, it became clear that the material donated to Taylor Library constituted an exceptional repository of information about the language. Although Tony Hurren drew on it for his doctoral thesis, for his outline grammar of the language, and for some published research articles on the Istro-Romanian aspectual system, the recordings had never been published, and their existence was virtually unknown to the wider world.

The digitization of analogue holdings is a precondition for opening-up new access routes to collections and archives. This was our first priority for making this material available to specialists and the Istro-Romanian linguistic community. We had the audio recordings digitised in 2012, with funding provided by the Taylorian Library. We produced two sets of digital recordings: one for the library, the other for the Research Centre for Romance Linguistics. Further assessment of the donated material continued over the following years, in particular after Vera Hurren donated a second batch of materials. It emerged that Hurren gathered recordings from about 40 informants, covering nearly all the villages in which Istro-Romanian was spoken, and thereby capturing material for a description of the major linguistic subdivisions. He worked with a representative sample of speakers, of all ages, which also made the material into a unique documentation of the history of the IR community. There were not only responses to linguistic questionnaires, but also folktales, accounts of local traditions, and autobiographical remarks and stories. To

identify correctly which informant was speaking, and which dialect was represented, this material required essential preliminary cataloguing work, to match narrations on the original reel-to-reel and cassette recordings with information contained on their covers and with the material partially transcribed by Hurren in the notebooks. The notebooks were a useful guide to the contents of the audio recordings, as they presented in written form the linguistic questionnaires used, and even transcribed some of the recordings. However, these notebooks did not include references to all the recorded material that we had; there were many gaps and discrepancies between the recordings and the transcribed material.

It became clear that our immediate priority was to provide an account of the content of the recordings, in the form of fairly literal translations tagged to the relevant portions of the recordings. As the tapes were unlikely to be easily intelligible to anybody other than a native speaker of the language, we decided that a potential project around the Hurren recordings would have to tap into the linguistic and cultural knowledge of the remaining speakers of the language. My linguist colleagues were interested in engaging the knowledge of the last remaining speakers of the language – not only those based in Croatia, but also those living in the diaspora, from Italy, to the US and Australia. Thus, we decided to consider what is generically described as (online) "crowd-sourcing" as a possible method to approach the Hurren material. We needed a pilot project that would allow us to start work on the donated recordings, test our intentions to source linguistic information from the members of this community in the long term, and, during this process, learn more about the online Istro-Romanian community worldwide.

My previous experience with heritage material in museum and visual archive environments, and my current work in academic librarianship, have made me aware of the increased popularity of collaborative work and lay "crowd" involvement in Humanities research and so-called GLAM ("galleries, libraries, archives, and museums") institutions, where researchers and curators migrate archival material from analogue to digital, place it in online environments, and invite members of the public to engage with it and produce new knowledge. This kind of research is often labelled "crowdsourcing", a term originally coined with reference to the business environment and defined as "the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call" (Howe 2006 in Schenk and Guittard 2011: 94). In recent years, academia and the cultural heritage sector have adopted and developed the corporate model of crowdsourcing by adapting it to their own needs. While the business model aims mainly to obtain a reduction of labour costs, academic and cultural heritage crowdsourcing aim to produce new knowledge and engage specific sections of the public with existing archives or collections around the world (Dunn and Hedges 2014). As pointed out by Ridge (2014: 4), while the term 'crowdsourcing' has proven a "useful shorthand" in the recent years, it also suffered the fate of other buzzwords, having its boundaries constantly being pushed by many; this was possibly in response to a certain vagueness of the term and the exploitative potential carried by the idea of uncompensated labour.[2] Hence, a growing number of projects emerging in academic or

[2] Owens (2012) writes that, when considered in relation to contexts other than the business one, there are "two problems with crowdsourcing: crowd and sourcing" – where both the idea of the undifferentiated "crowd" and the practice of using it to "source" information are seen as problematic. Dunn and Hedges (2014:232) also remark that the early use of the term in a business context becomes senzitive when borrowed into academic or cultural heritage contexts, as it 'conjures notions of exploitation, manipulation and amateurisation: terms at odds with the aims of academia".

heritage environments whose authors resorted to alternative terms to signal that the said "crowd" was neither large nor anonymous. One of these emerging terms has been "community"(-sourcing), which we also decided to employ for our project. Apart from the small number of remaining speakers worldwide, which hardly qualifies them as a "crowd" in the sense proper to crowdsourcing environments, the reason why we opted for "community-sourcing" has more to do with how we conceptualize the group of people with whom we would like to work and the terms of this envisaged long-term collaboration. We approach the Istro-Romanian "community" in the sense suggested by Anderson (1983), as consisting of those who speak or understand vlaški and žejanski, or who regard themselves as being part of the group called "Istro-Romanians", irrespective of their geographic location. While we have decided to set up an online segment of the project in order to mobilize the remotest speakers and collect as many translations as possible, our main aim goes beyond that of strictly creating "content" towards the idea of building a community around the project[3], with the intention to engage it in different forms of "meaningful participation" (Ridge 2014), whether online or offline, during the subsequent stages of the project.

To put our heritage material into circulation, we needed to upload, on a customized online platform, a selection of audio samples whose linguistic meaning or structure was problematic, and invite the Istro-Romanian speakers worldwide to help us interpret them by listening to the audio clips, commenting on them, and responding to our questions. While most crowdsourcing projects emerging from academic and cultural heritage environments aim to attract as wide a "crowd" as possible and, from that, carve out a smaller "community" of invested participants, we were to start from an existing, albeit drastically limited, number of speakers of Istro-Romanian worldwide and hope to identify an understandably very limited number of speakers invested enough in their linguistic identity so as to devote some time helping us interpret the problematic segments of the Hurren recordings.

3. SHAPING ISTROX UNDER GDPR

The project was to have two phases, each of which corresponded roughly to one year of half-time work on the project. The first phase involved mainly indexing and cross-referencing the audio and written material; selecting the audio snippets to be used for further linguistic analysis; tracing Hurren's original subjects, or their surviving relatives in Istria, to obtain their consent to handle the archival material; building the project website. The second phase entailed building the online infrastructure and launching the community-sourcing process, and a second, more extensive field trip to Istria, focused on linguistic research.

Practices developed as part of projects are driven by mechanisms such as funding and by accountability structures such as research ethics which are built into the design of these projects. One of the challenges that we faced while shaping our funding proposal was ethical approval. The implementation, on 25 May 2018, of the new General Data Protection Regulation (GDPR) impacted the way in which universities and funding bodies addressed

[3] In their seminal Scoping Study about crowd-sourcing in the humanities, conducted for the United Kingdom's Arts and Humanities Research Council (AHRC), Dunn and Hedges (2012: 18) pointed out that while academic need is the engine for crowdsourcing practices in both sciences and humanities, in the case of the latter what is described as crowd-sourcing is "sometimes seen as a means of community building, and thus increasing the wider impact of academic research activities".

data protection as part of their policies. Our award of funding[4] was conditional on an ethics approval which had just changed due to the new GDPR and came with specific requirements for the projects that involved archive material. Before applying for funding for ISTROX, we had obtained Mrs. Hurren's permission to use the unpublished material connected with her husband's research on Istro-Romanian, and also to place it in the public domain whenever we considered that necessary. However, as this heritage material involved prior research with human participants, obtaining ethical approval from Oxford University's Central University Research Ethics Committee (CUREC) was conditional on having also obtained the permission of Hurren's interviewees or of their surviving relatives.

The Hurren material had been produced in the mid to late 1960s, decades before GDPR. Making it available online entailed circulating the voices, photographs, and possibly the names of a number of potentially living individuals whose rights and freedoms were guaranteed by the new GDPR. Therefore, in the first instance, we had to obtain the consent of all those involved in Hurren's original research (or of their relatives). Since Hurren's notes mentioned the names and villages of all the participants, we planned to trace them or their families and obtain the appropriate consent. We were also aware that that might not be possible in some cases, largely due to the substantial emigration from that area. In those cases, we had to weigh the individuals' right to privacy against the broader aim of our project, which was to archive this material and make it available for historical research purposes. Part of that process was a question regarding the expectations of the original research participants at the time when they were recorded and how they might react to finding their personal data made public as part of our project.

There were two main aspects which entitled us to be optimistic about such a scenario. First, the fact that we intended to use the Hurren material in accordance with the original purpose – linguistic research – for which Hurren collected it. Second, the knowledge that those who participate in linguistic research often do so with pride in their command of the language and ask specifically to be identified for their contribution. That was even more the case when dealing with a language which is on the verge of extinction: existing websites with Istro-Romanian content testify that the remaining speakers of the language are strongly invested in their linguistic identity. In addition to this, we were confident that the material was not sensitive, as it concentrated on the structure of the language rather than on the content of speech. A large portion of the material consisted of linguistic questionnaires or participants recounting traditional tales, with very few segments that could be considered personal or autobiographical in nature. Thus, we started off with the assumption that the former participants in Hurren's research or their relatives would be happy to see them(selves) credited for their contribution.

Ethical approval was granted by CUREC depending on the types of interactions between the researchers and the participants: different types of requirements applied and specific documentation was needed in view of obtaining approval, respectively, for real-life interactions with the research participants (e.g. research *in situ*), and for online research environments such as the platform on which we planned to upload a selection from Hurren's recordings. The latter case posed questions such as what exactly constituted valid

[4] ISTROX was funded by Oxford University's John Fell Fund. See https://innovation.ox.ac.uk/award-details/john-fell-fund/

consent in a given online environment and how could the researchers behind a project or another make sure that their participants have fully understood the issues involved. To secure ethical approval sooner, we decided to split the project in two and seek approval only for the first phase of our research – which involved our first field trip to Croatia – and then prepare a second round of documentation for the phase which involved online community-sourcing. By doing this, we obtained permission to travel to Istria and get ethical clearance for our archival material. This delayed our initial plan by several months, as we had to prepare and conduct our first field trip before we were able to start building any of the two online structures entailed by the project: the project website and the community-sourcing site. While obtaining ethical approval is an essential part of all research that involves human subjects and personal data not available in the public domain, complying with the new GDPR regulations took substantial time and effort: producing all the bilingual material (English and Croatian) required by the application for the first ethical approval – that is, the templates for the recruitment and advertising materials, written information sheets and oral information scripts; written consent forms or oral information scripts – took us more than a month, and the entire process of approval for the first phase of the project took around two months.[5] We obtained approval for the first year of the project at the end of August 2018 and embarked on the project on 1 September 2018.

The first phase of the project involved a range of linguistic tasks, such as producing a description and index of the contents of the sound recordings; segmenting the recordings into coherent "snippets" to be further used for linguistic analysis; producing a detailed description and index of the contents of Hurren's field notebooks; correlating the written materials with the sound recordings. I was involved in (co-)writing and editing the content for the project's website, I worked with a developer to build it, and I was part of our first field trip to Croatia, to trace Hurren's subjects and obtain their consent for our future use of the Hurren material.

The second phase involved setting up the online infrastructure and starting the process of community-sourcing, and assembling a community of committed participants around the project. By the time we reached that phase, we had obtained ethical clearance for most of the Hurren material. However, to start the second phase, the ethical approval board required us to produce a detailed description of the types of online interactions created by the online platform that we intended to use for community-sourcing and of all the written material that mediated those interactions, in order to assess them from a data protection perspective. The platform that we planned to use was the crowdsourcing site Zooniverse.[6] Once we started preparing the documentation required to obtain ethical

[5] This process made us reflect on how, in specific cases, existing ethics legislation within universities can hamper research. Ethics is a crucial aspect of the work with human subjects, yet sometimes the assumptions behind current regulations – such as that regarding the presumption of a desire for anonymity by research participants – may need to be reconsidered for certain categories of subjects, such as the remining speakers of a severely endangered language who are usually proud of their ethno-linguistic identity and would often appreciate being mentioned in connection with specific projects. See Cassell and Jacobs (1987) for a series of case studies of informed consent from anthropological fieldwork.

[6] Zooniverse (https://www.zooniverse.org/) is well-known as one of the pioneers in the field of so-called 'citizen science', with projects such as the early Galaxy Zoo, where some 200 000 volunteers made 100 million classifications of images of galaxies between 2007 and 2009.

clearance for this phase of the project, we realized that we would only be able to detail the future online interactions between us and our potential participants after we finished setting up our site on Zooniverse. That was because, in order to obtain what we needed, we had to introduce some new elements to Zooniverse – a platform which had been established in 2009 for a different approach to crowdsourcing as "citizen science". The thorniest aspects of this process had to do with the fact that – unlike conventional crowdsourcing, which channels the contribution of large bodies of anonymous volunteers – we wanted to build a system able to gather linguistic knowledge from a rather small number of individuals for whom we needed to collect a minimal amount of socio-demographic data. Once this had been done, the second process of ethical approval took one month. By that time the Covid pandemic was causing delays in the project. We launched the site at the beginning of August 2020. By that time, due to the disruptions caused by the pandemic, our project had been extended to 31 July 2021.

4. FIELD TRIP TO ISTRIA

In April-May 2019 – that is, before launching on Zooniverse – I travelled with Maiden to Croatia, to meet the surviving informants from the original Hurren study and obtain their consent to use the Hurren material. By that time, Oana Uță had cross-referenced Hurren's notebooks and audio recordings, therefore we had a comprehensive list of 36 participants whom Hurren had recorded in the 1960s, together with the names, village of residence, and approximate age of those informants in 1967. Approximately half of them would have been over 100 years old at the time of our field trip, therefore we inferred that we would need to find their surviving relatives and obtain their consent to use material provided by those participants.

During our ten days in Istria, we met around thirty people, of whom nine took part in Tony Hurren's original research and the rest were descendants of the remaining participants; to those we added several other people of Istro-Romanian descent suggested by our initial contacts in the area.[7] We sought the permission of those that we met to make Hurren's photographs and recordings publicly accessible online (with the option of full identification or pseudonymization). We also sought their assent to further assist us in the linguistic interpretation of the recorded material and possibly to answer other linguistic and biographic questions at a later date. We informed other members of the Istro-Romanian community with whom we established contact about our project, including about its online element, and obtained their permission to contact those interested once the project was ready to start. We suggested that they inform their diasporic relatives about our project and put us in contact with them wherever possible, to enable us to build up a pool of potential online participants for the second phase of the project. The number of new contacts that we could make during this fieldtrip was relatively small (under twenty), because we aimed mainly to obtain ethical approval for the Hurren material.

[7] We thank Dr Robert Doričić and Mrs Viviana Brkarić for their generous assistance in putting us in touch with members of the Istro-Romanian community. We also thank the members of the community in Žejane and Šušnjevica for receiving us so kindly and for giving so freely of their time to help us.

This initial phase of the project involved many elderly subjects for whom knowledge of rights to privacy, according to current laws, and of informed consent forms, was unfamiliar. While we aimed to obtain written consent in all cases (the option recommended by Oxford's CUREC), the written consent forms that we had prepared according to CUREC regulations proved to be too complex, therefore we adopted alternative forms of oral consent.[8] All but one of those that we met consented to us using the Hurren material; one former participant in Hurren's research allowed us to use the recordings but preferred not to have the audio and photo material including himself and his deceased mother placed online.[9]

Although it did not include any linguistic interviewing, this field trip was essential because it allowed us to establish a connection with the community – in particular with the surviving subjects of Hurren's initial interviews. On our return, it took us another few weeks to clear most of the remaining recordings and images by phone and email.[10] We made the material available online after having obtained the consent of those who appear in the archive photographs, with only a few exceptions, where we were assured by the former participants in Hurren's research that those who appear in the photographs have either died, leaving no known relatives, or moved outside Croatia, there being no known means of contacting them or their descendants.[11]

5. COMMUNITY-SOURCING IN PRACTICE: ISTROX ON ZOONIVERSE

ISTROX entailed the building of two online structures, both of them bilingual (English and Croatian). The project website (https://istrox.ling-phil.ox.ac.uk) provides information about the language and the community, the donation, the project, and the community-sourcing element of the project, and includes samples of ethically cleared audio and photo material. The community-sourcing site (www.istrox.uk) is based on the online platform Zooniverse and linked to the project website.[12] While ISTROX was in planning

[8] The oral consent process involves the researchers producing a short description of the scope and context around the research project, plus details of the type of consent required, then asking for the oral consent of the participants and filling in a so-called "Researcher record of oral consent". Alternatively, the participant is invited to read the full description of the research project and types of consent involved, and to sign the form. In this case, the Written Consent Form that we had with us, in the Croatian and English languages, was two pages long; we realised that reading that document was somewhat alienating to many of the people involved and therefore we switched to a form of oral consent.

[9] The subject kindly reversed this decision when we met him again in 2022.

[10] We thank Dr Ana Horvat for her help with the phone calls to Croatia as well as for producing a Croatian version for all the English material that we uploaded on our online sites.

[11] When we uploaded the material on Zooniverse, we included a Copyright Notice and Take-Down Policy in which we invited those who felt that our online sites contained something that should not be there to get in touch with us, and promised to remove the respective material (see https://www.zooniverse.org/projects/sandorkruk/istrox/about/research).

[12] In the initial phase of the project, we used the project website to recruit participants for our Zooniverse site: on the upper right-hand side of the landing page of the website, we placed a "get involved" button which offered the speakers of the language the option to assent to be contacted by us if they were interested to take part in our research. Initially, the button led to a page that offered information about our intentions to involve those members of the IR community who were willing to work with us, irrespective of their geographical location. Once we launched the Zooniverse site, we linked it directly to our project page.

phase (2017), Zooniverse was already one of the largest citizen science web portals in the world, with a community of over two million volunteers and more than 250 large crowdsourcing projects from various fields completed. However, it was not the size of the body of volunteers amassed by Zooniverse that attracted us to the platform, given that we were planning to use it to work with a specific, and very small, community defined by the ability to speak or understand a given language. What attracted us to Zooniverse was the chance to use an already existing online infrastructure which made the process more time and cost effective than any other options available: as we were working on a pilot project, we had neither the time, nor the means to build an online structure from scratch. When ISTROX was in planning phase, Zooniverse could not yet integrate sound: all the projects launched on the platform were either image or text-based. However, we were aware that the Zooniverse developers were working on the sound capabilities of the platform in order to open it to audio-based projects.[13]

The project template created and made available by Zooniverse is called Project Builder.[14] It provides the structure of the pages developed by each project created on the platform, with self-explanatory sections on which the project teams can input all the necessary information about the project: name, avatar, description of the project and its methodology, information about the team behind it, a FAQ page, and a description of the so-called "workflows" (the sequence of tasks volunteers are invited to perform) entailed by the project. The template landing page of the Zooniverse projects also includes an announcement banner placed strategically at the top of the project's landing page, and a "call to action" for potential volunteers which is displayed on the main page of the project. For ISTROX, this reads as "Can you help us translate the Hurren audio-recordings of the Istro-Romanian language?".

On the site we uploaded a selection of 563 short audio clips with durations ranging from several seconds to several minutes, selected from the Hurren audio material and prepared by Uță and Maiden. The clips are displayed as two "workflows", one for each language variety; each workflow is further organized in micro-clusters, where the audio-clips are ordered according to their linguistic difficulty, starting from those considered less challenging for a native speaker and advancing towards those regarded as the most difficult. Both workflows involve a sequence of simple tasks: listen to an audio clip, offer a translation for it and rate its difficulty. Participants are invited to start with the variety that they speak better and, once they have finished work on that, to continue with the other variety. In addition to this, there is also the option to add contextual knowledge by commenting either on the audio clip or on the photo attached to a specific bundle of audio clips. At the end of each sequence of listen / translate / rate, participants have the option to click on a button called TALK, which opens a note in a special section of our message board where they can comment on any aspect that they find relevant about the material that they were exposed to. Once this happens, we find the note archived in a special section of the message board and have the option to get back to the participant to continue the conversation about it.

[13] We would like to thank Professor Chris Lintott for meeting us during the preliminary phase of our project, and to Dr. Sandor Kruk for offering his time and technical expertise to help us build the ISTROX site on Zooniverse.

[14] See Zooniverse's Project Builder here: https://www.zooniverse.org/lab. Prior to starting work on the platform, I attended the Digital Humanities at Oxford Summer School (July 2018), to learn about crowdsourcing practices for research purposes and get used to the Project Builder.

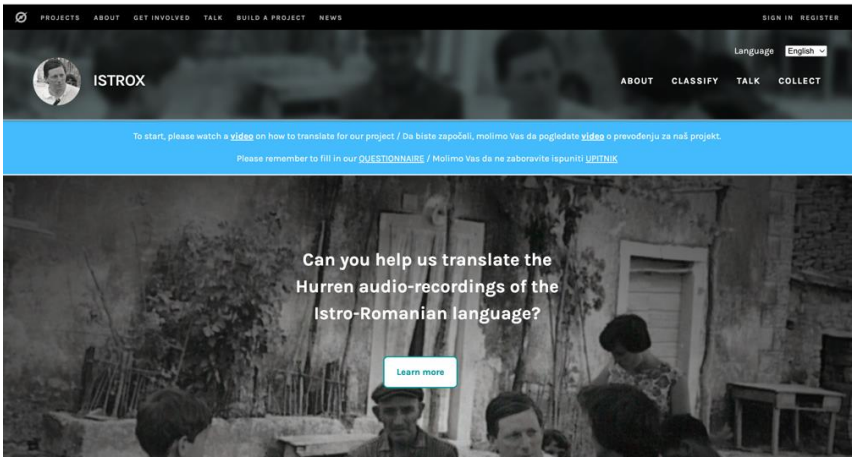

Figure 1. ISTROX on Zooniverse, main page (English interface).

Becoming a volunteer for any Zooniverse project involves registering on the platform by choosing an ID and a password. Since we need basic demographic and linguistic data about our contributors – unlike most Zooniverse projects, which rely on anonymous contributors – we incorporated into the registration process an invitation to fill in a short questionnaire linked to the announcement banner, which opens separately in Google Forms. The questionnaire is optional, yet in the opening of the form we make the case for our need to learn some basic biographic details about our participants and their relationship with the language.[15]

Using the template available on Zooniverse, we built a tutorial which explains the process that one needs to undertake before starting on the site. The tutorial starts automatically when a participant first lands on the site, then it remains available on the upper right side of the screen for consultation whenever necessary. It explains the process that one needs to undertake to become a volunteer for the project and works as a concise reminder of the most significant features of the site: 1. volunteers can decide to work in the language (English or Croatian) that they feel more comfortable in; 2. the material is divided into clusters of clips where the translation of one cluster should take no longer than app. 15 minutes; 3. after each group the participants can decide to continue or stop and come back another time; 4. if they decide to leave and return later, they need to be logged in, to make sure that the system remembers where they left and presents them with the following audio clips.

[15] The questionnaire consists of 26 questions, which refer, among other things, to the place of birth and current residence of the participant, the number of years spent living away from the birthplace (whether within or outside Croatia), their age, gender, education, occupation, the language variety and other languages spoken by the participant and by their parents / grandparents / children, the frequency and type of encounters in which they use this language variety. See the English-language questionnaire at https://tinyurl.com/bde39k8j. Both the English and the Croatian questionnaires are linked to the blue band at the top of our Zooniverse page. See Figure 1.

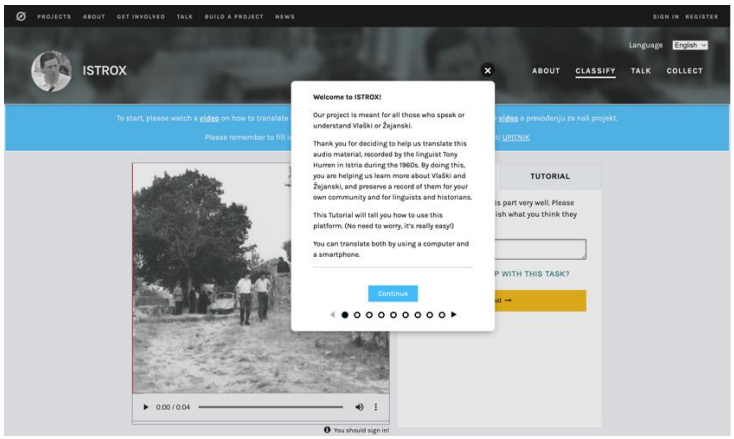

Figure 2. Page from the Zooniverse tutorial (English interface).

Apart from aiming to design tasks that are as clear as possible, we wanted to contextualize the clips selected for translation. Community-sourcing practices are known to open up access to previously inaccessible heritage material, yet they often do this in fragmented ways, by presenting in succession only snippets of text, image, or sound, in order to make these more manageable for the volunteers. Before uploading the material on the site, we had to decide how much context we needed to provide our potential participants with in order to increase their engagement and support the quality of their contributions, while keeping the whole process as simple and straightforward as possible. Thus, alongside each clip that we uploaded for translation, we also included a longer contextualizing version that provides information about the speech context and helps the participants understand better the meaning of the shorter clip they are invited to translate: when they start translating on the site, volunteers are provided firstly with the short clip, but are also given the option to listen to the longer one. While the clip offered for translation opens on the site at a simple click on a "play" button, the second one is available under the audio clip through an "i" (information), and opens into a separate window, where the participant needs to click on 'full audio'. See Figure 3.

Once a short clip is translated, participants are asked to rank the difficulty of the translation (five options, "very easy" to "impossible"), and then they can choose to continue by translating another clip or to add a comment to their translation. The comments may refer to the audio clips or to the photographs that we included on the site from among those donated to us alongside the audio recordings. While we built our Zooniverse site with a focus on the audio material, we decided to include the photographs, on the one hand, with an engagement purpose, so as to build more context around Hurren's research and the time that he spent in Istria; on the other hand, to give our participants the option to include any commentaries that they might have about one image or another. These images are only generically connected to the (bundles of) audio clips with which they are paired online: that is, a recording in *žejanski* will be paired with a photo taken in Žejane, with no other correspondence between the person recorded and the one featured in a photo or another; we made this aspect clear in the informative material that accompanies the project on the site.

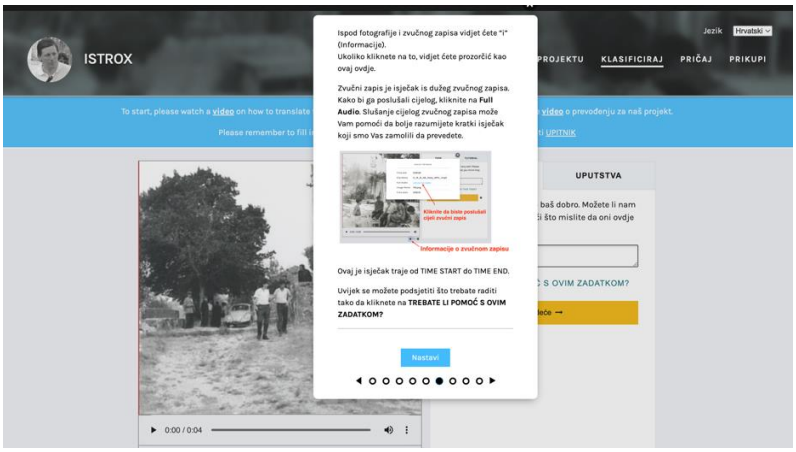

Figure 3. Page from the Zooniverse tutorial: contextualizing the audio clips (Croatian interface).

6. RESEARCH IN A TIME OF PANDEMIC

After January 2020, the circumstances created by the Covid pandemic called for a reorientation from physical fieldwork towards online research. The role played by information gleaned online gradually became more central to the project. Luckily, the Zooniverse element of the project ended up fitting the new context where many offline research activities have been halted due to the need for physical distancing and isolation, and most research had to move online. One of our initial objectives had been to "activate" the members of the Istro-Romanian community based in the diaspora, in addition to those based in Croatia. At the beginning of the pandemic we found ourselves managing an online infrastructure for a community sourcing project that targeted a very small, globally scattered community about which we knew very little and whose members were still to find out about us. As the pandemic advanced, Zooniverse and other crowdsourcing platforms started reporting an increased engagement from their participants (Cohen 2020, Viellieux 2020). However, this referred mostly to "conventional" crowd-sourcing projects where technologically proficient people who had already spent time contributing online in the pre-pandemic time found themselves stuck at home with more time available and therefore started spending more time on various online platforms. Our site on Zooniverse – which, by design, cannot be found by browsing the projects on the platform[16] – did not catalyse a similar enthusiasm.

[16] At Zooniverse, a live project can exist in three different states: 1. Public Project 2. Review Project 3. Zooniverse Project. The latter two categories describe projects that want to be findable on the platform as they rely on the "crowd" of volunteers that Zooniverse has amassed so far. The first category – which applies to ISTROX – is for projects that come with their own communities of volunteers and only want those members to get involved in classifying / translating / commenting on the material uploaded. These projects are not included in the newsletters sent regularly by Zooniverse to its (millions of) volunteers; instead, the administrators of the project spread the news about the

Considered in general, the practice of crowdsourcing entails a significant degree of serendipity. When building the infrastructure for such a project – even one that counts on the anonymous "crowd" and not on a very small community, as in our case – it is difficult to predict the size and dynamic of the said crowd as nobody can foresee what kind of community will be formed and what will be achieved in time (Dunn and Hedges 2012: 14). As participation in such projects is elective, attaining a critical mass of contributions cannot be guaranteed (Schenk and Guittard 2011: 104). Researchers have tried to understand the motivations behind participant involvement in such projects: it appears that, unlike business crowdsourcing, which often involves small-scale material recompense for input, the academic practice of sourcing knowledge from the public is rooted in more intrinsic motivations which are almost always concerned directly with the project or activity's subject area (Dunn and Hedges 2012: 9-10). Writing about the psychology of participant motivation, Owens (2014: 276-7) remarks that "people identify and support causes and projects that provide them with a sense of purpose. They establish and sustain their identity and sense of self through their actions [...] Projects that can tap into these identities and purposes while providing meaning to people's lives are projects that – far from exploiting people – can provide a way for them to connect with each other and make meaningful contributions to the public good."

We have been in touch with members of the community who were genuinely interested in the preservation of their linguistic and cultural identity and who expressed a strong enthusiasm for our project. However, recruiting participants developed very slowly and online participation through Zooniverse remained low. We started by snowballing in two ways: on the one hand, from the small pool of initial informants encountered during fieldwork in Istria in 2019; on the other hand, from a smaller number of second and third generation members of the community from Australia and North America who had contacted us through our website, expressing an interest to get involved.

The pandemic brought different types of setbacks to the project. When the *Prizma* programme on Croatian national television (HRT), which focuses on Croatia's ethnic and linguistic minorities, contacted us, we discussed the future production of two TV programmes which were meant to inform the community in Istria about our project and introduce them to our method of collecting data online. HRT have produced one programme filmed exclusively online and broadcast on 3 October 2020 and then rebroadcast by regional media.[17] A second programme, planned to coincide with our second fieldwork in Istria, had to be cancelled when the fieldwork itself, planned for spring 2020, also had to be cancelled.

---

project by sharing the project's URL with those that they want to involve. In this sense – of not sharing Zooniverse's own crowd of volunteers – ISTROX is indeed not an official Zooniverse project, the reason being our intention to work with a specific linguistic community. Hence, ISTROX is not visible when browsing Zooniverse: participants need to have the link to the project either from us or from other participants who consider them suitable as potential volunteers. One consequence of this status is a possibly confusing message mid-way through our main page, on a thin, barely visible, horizontal band which reads "This project has been built using the Zooniverse Project Builder but is not yet an official Zooniverse project. Queries and issues relating to this project directed at the Zooniverse Team may not receive any response." The band and message are part of Zooniverse's template and cannot be deleted as it is meant to clarify the status of each project hosted by the platform.

[17] See https://www.youtube.com/watch?v=tIeGNwsaERs

We had been aware from the planning stage that, given the paucity of remaining speakers of Istro-Romanian, building even a tiny online community of contributors would take time and effort – in particular because any potential participant in our project had to combine two types of abilities: speak or understand *žejanski* or *vlaški*, and be sufficiently "computer-savvy" to understand the instructions on our site and advance comfortably through the "workflows". Our engagement plans were likely to be further hampered by the kind of demographic that we envisaged for our project: drawing on existing knowledge about the remaining speakers on the IR communities, we reckoned that the majority of our contributors would be, with few exceptions, from the 50+ group. We have placed this awareness within the larger context of what has been described as the "grey divide" as regards technological competence (Morris et al., 2007). Drawing on existing research on the said "grey" participants in crowdsourcing projects[18], we assumed a low to moderate web literacy and prepared for a limited lack of concurring skills (language + ICT) that would enable these to contribute to our project. As Ridge (2014: 4) points out, "the more specialized the skill, knowledge or equipment required, the more strongly a 'crowd-sifting' effect operates as individuals unable to acquire the necessary attributes fall out from the pool of potential participants". In our case, in addition to the limited number of remaining speakers worldwide, age and technological competence emerged as two significant barriers to online participation. Undertaking the tasks required by our Zooniverse site involves a combination of the following: interest in the project, speaking or understanding the language, ownership of a computer or smartphone, technological competence to understand the instructions and navigate through the site, and availability of time. While we were aware that only a small subset of the remaining speakers would fit all of the above, we approached online community-sourcing as part of a larger pool of (online and offline) methods to engage the community, rather than as a unique method of gathering linguistic information. What we could not predict was that all offline, face to face research would be brought to a halt by the pandemic, and therefore we would have to make do thorough online communication only.

Building a successful community-sourcing project requires a constant effort to ensure both an initial and an ongoing participation from the community. A project that depends on contributor effort is usually a combination of casual participants and so-called "super-contributors" (Causer and Wallace 2012); the latter, for one reason or another, put significantly more time than others into their translation or classification as part of a given project. As a small percentage of contributors often do a large percentage of the "crowd" or "community" work, the success of such projects usually depends not on the total number of contributors but on "finding pockets of enthusiasm and expertise for specific areas" (Dunn

[18] Currently there is a gap in our understanding regarding the relation developed by more senior people with different types of online "crowd" work. Most research on this type of online work refers to business environments, therefore its applicability to research contexts is yet to be tested. Existing research suggests that those who engage in this kind of online work tend to be in young or middle adulthood. Older adults are under-represented as regards online activities, most likely because they need to overcome accessibility issues and tend to need specific online support systems such as discussion groups. Studies examining the age-based digital divide have identified a number of key factors that hinder older users' engagement with the internet and digital technologies, from physical and mental challenges, visual or auditory problems, mistrust and security, as well as concerns regarding privacy and the sharing of personal details online. (Dibeltulo *et al*., 2020: 811).

and Hedges 2014: 244). While we have identified a small number of members of the community strongly committed to their linguistic identity, neither of these have become "super-translators" yet, though some of them have already contributed on Zooniverse. As online translation is entirely elective, we tried to adapt to the mode of communication preferred by each participant, with no pressure put on any of them to return to our site for more translation work. The questionnaires attached to Zooniverse, which our existing contributors have filled in prior to starting translation work, have helped us collect basic socio-demographic data about our participants so far. Like other researchers who adapted their work to the pandemic context, we conducted preliminary qualitative interviews on the Zoom platform with some of these contributors from Australia and the USA, as well as with others who contacted us without expressing and interest to contribute on Zooniverse. Through these interviews – or, sometimes, just recorded conversations – we not only built rapport with our interlocutors[19] but also gathered background knowledge on the history of those communities and the present life of their members with a view to future stages of our research. One-to-one meetings were also held with several speakers who expressed an interest in focusing more specifically on language practice.[20] While international air travel was suspended at the time, travel within Croatia was possible during specific time intervals, and therefore a set of linguistic interviews were conducted "by proxy", through the Croatian linguist Ana Horvat with live supervision through Zoom by Maiden and Uță.

The ISTROX pilot project formally ended in July 2021. In the final segment of this article I discuss two short-term follow-up projects developed under the "ISTROX" umbrella after the completion of the pilot project.

7. FOLLOW-UP PROJECTS

Scholarship on crowdsourcing suggests that most projects that entail community-involvement change, often dramatically, after their launch, in response to their encounter with different types of "crowds" (Ridge 2014:7). Thus, the subsequent stages of such projects often involve amends meant to increase the clarity of communication and the quality of the user experience offered by the online platforms hosting the projects. As Ridge remarks, this acquired awareness of various types of revisions needed by these projects also reverberates on the level of the funding needs of the projects: while funding bodies conventionally regard the launch of an online site as an end in itself, the launch of a crowdsourcing site is, in fact, only the beginning of a process of networking, communication and moderation which continues to require resources and involvement of the team in the long term.[21] User communities, in particular very small ones such as the

[19] See Archibald et al. (2019) for the usefulness of Zoom for qualitative research due to ease of access, time-effectiveness, and convenience for busy work schedules. Although we had to accommodate different time zones between the UK, USA and Australia, this has been manageable given the small number of interviewees (under ten).

[20] For instance, there have been extensive conversations with an elderly contributor in Australia which were organized online by Uță and Maiden.

[21] *Transcribe Bentham* is a useful example of the way in which, due to the nature of funding, the project was unable to benefit from many of its recruitment strategies, as it had to scale back just as

Istro-Romanian community, require constant communication on an individual and group level through updates, reminders, and messages of encouragement, as well as constant snowballing in search of new members of the community interested to try out the site or to become involved in the project on a one-to-one basis.

This is where we found ourselves in July 2021. Though we had been in touch with around fifty potential participants[22] in the months prior to our launch online, only eight had registered on Zooniverse and contributed translations by 1 August 2021. We had been aware from the very beginning that those with whom we had been in touch preferred to engage with us either face to face (when that was to become possible, post-pandemic) or by Skype or Zoom, yet we could not afford to cease our engagement work abruptly while in between projects, particularly when considering the tiny size of the global community. We therefore sought short-term funding meant to help us avoid loss of momentum between projects and continue to develop ties with the community.

A small Public Engagement with Research (PER) grant funded internally through the university allowed me to work one day per week, for six months, solely on the 'engagement' element of the project: through this segment of the project, titled *Diasporic Istro-Romanians online*, we aimed, on the one hand, to keep in touch with our existing contacts and develop new ones; on the other hand, we wanted to learn more about the online forms of aggregation of the worldwide IR community, both at present and in a historical perspective, to make ISTROX more visible among the members of this community, and to think creatively about other ways to approach them. Apart from maintaining contact with the community, I used this time to learn about the preferred online spaces of communication of the community and about the small diasporic networks and spaces of physical encounter for local communities in the US and Australia (many of which had, sadly, closed during the pandemic). As part of our attempt to improve our communication strategies, we produced two bi-lingual leaflets with basic information, relevant links and contacts for the project; we circulated these, as PDFs, to the press, the relevant diasporic associations, and among known members of the community, encouraging the recipients to forward them to friends and relatives, to spread the word about the project. Through online consultations with our existing contributors on Zooniverse, we learned that prospective participants found the tutorials available on our site too difficult to navigate through. In response to that, we recorded two video tutorials, in English and Croatian languages, which we linked to our Zooniverse site, alongside the questionnaires. See Figure 1.

As media attention is known to lead to spikes in uptake of project participation (Causer and Wallace 2012; Dunn and Hedges 2012), we also got in touch with Croatian media and had around ten articles about ISTROX published in the Croatian press, ranging from HINA, the national press agency, to local and regional publications and radio stations (*Total Croatia News*; *Novilist*; *Croatia Week*; *Glas Istre*; *Radio Rijeka*). These articles have been shared on social media by most publications, and then re-shared on individual pages

it had started to gain momentum following the impact of international press attention. See Causer and Wallace (2012).

[22] These included both former participants in Hurren's research, their relatives, and speakers who contacted us after learning about the project. We had been in various stages of communication with each of these whether though our project page on Facebook, through Zoom/Skype, or by email.

or on the Facebook community page "Istro-Romanians". In response to the press and to the individual email messages that we sent out by snowballing from newly acquired contacts, we increased our number of contacts that we had within the community worldwide by around 40 people; the number of Zooniverse contributors increased to 17.

We also wanted to test the feasibility of "teaming up" different generations of members of the IR community: that is, younger, web-literate non-speakers of Istro-Romanian with older, and possibly less web proficient, speakers of the language. Until then, we had focused on diasporic Istro-Romanians who still spoke the language, which had often proven challenging due to their advanced age and, sometimes, limited web literacy. In time, we had become aware that the web users of Istro-Romanian descent who were present online were not necessarily those who still spoke or understood the language. Consequently, drawing inspiration from scholarship on older adult involvement with cultural heritage online (Leist 2013, Dibeltulo et al. 2020)[23], we tried to encourage a cross-generational engagement by involving members of the community who no longer spoke Istro-Romanian but had elderly relatives who did: that is, work with the former so as to assist the latter as they went online to listen to, and comment on our recordings. So far, the "teaming up" idea has only worked sporadically in a few specific households in Australia and the US, where non-speakers who expressed interest in our project have been willing to act as 'bridges' between our project and their senior relatives who are speakers of the language. We cannot assess yet how successful this collaboration across generations can be in the long term, but we plan to continue to highlight this as an option and allow it to develop whenever it fits organically with specific family or neighbourhood contexts.

At the time of writing (spring 2022), we are developing another short-term project meant to help us keep in touch with the community: *Istro-Romanians: Linguistic Heritage in Online Conversations* is a five-month project funded by Oxford's Research Centre for the Humanities (TORCH) aimed at helping us develop new contacts, primarily in Croatia, by working with a Croatian partner – the Lapidarium Museum (Novigrad – Cittanova), a regional museum around 70km from the Istro-Romanian villages.[24] With this, we are moving towards a new stage of the project as we start collaborating with an external

[23] One way to approach the online engagement around ISTROX is by drawing on the experience of cultural heritage institutions from the GLAM sector who invest resources in digitizing their collections with the purpose of increasing public access to them through digital platforms. These digitized collections encourage an active online user participation yet also pose accessibility challenges, such as the increasing divide between the younger, and digitally literate, segments of the population, and those aging potential users who lack ICT skills and therefore are less likely to engage with this kind of material. When discussing the capacity of older adults to cross the so-called "grey divide" and engage with online cultural heritage, Dibeltulo et al. (2020) suggest that it is possible to increase their digital engagement with online historical material and also to further younger generations' involvement in shared cultural heritage in an online context if these two groups are given the opportunity to work together and particularly if that collaboration involves some face-to-face interaction. This supports Anja K. Leist's (2013) prior research findings on the key role of so-called "moderators" from the younger generation to help novice users from an older generation with their daily engagement in digital environments.

[24] See a presentation of the project on the TORCH website: https://www.torch.ox.ac.uk/istro-romanians-linguistic-heritage-in-online-conversations. We thank Dr. Jerica Ziherl, the director of the Lapidarium Museum, for her interest in collaborating with our project.

institution able to help us expand our reach into the community in Croatia, beyond the traditional villages. Like the aforementioned 'teaming-up' idea, this partnership is an attempt at a creative response to the observed technological reluctance of the remaining speakers of the language towards our Zooniverse site. Since January 2022, we have been sharing with Lapidarium, through our social media channels, a series of bilingual posts which include selected heritage audio and photo material accompanied by a range of questions addressed to the members of the IR community. These social media posts are our way of approaching the worldwide Istro-Romanian community on an online platform (Facebook) where its members are already present (as we have learned from our previous research). Through the potential responses from those reached by our posts, we hope not only to acquire linguistic information, but also to gain a sense of proficiency in the language within and beyond Croatia. While engagement on the project's Facebook page is still low, re-shares by members of the community are bringing an increasing visibility to the Hurren material and, ultimately, to the project: while the first clip that we posted on 18 March 2022 has had just over 100 views as of 30 April 2022, the second one, posted on 2 April 2022, has reached 300 views in the first two weeks after posting.[25] This series of social media posts is also a way of introducing Lapidarium into our envisaged "conversation" with the community, to prepare the ground for an exhibition that we are co-curating with the museum, opening in June 2022.

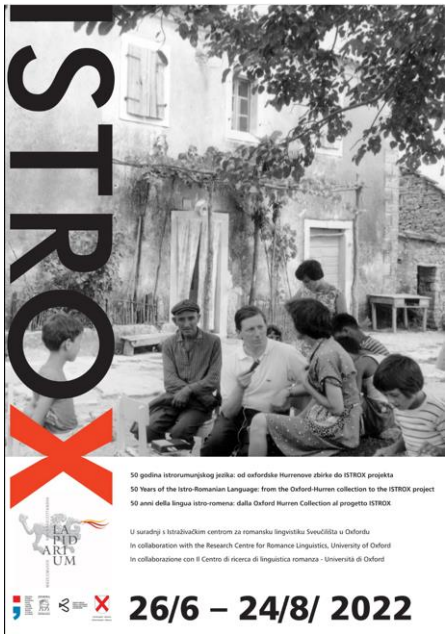

Figure 4. Poster of the exhibition at the Lapidarium Museum, Novigrad.

[25] First post (*Kostrčan, 1960. Linguist Tony Hurren interviews Valeria Jurman (Posidel)*: https://www.youtube.com/watch?v=OeApBzssM-w. Second post (*Žejane, 1967. Vesna Sanković ['Kljomina'], 8 years old, recorded by the linguist Tony Hurren):* https://www.youtube.com/watch?v=OeApBzssM-w

Entitled *50 Years of the Istro-Romanian Language: from the Oxford-Hurren collection to the ISTROX project*, the exhibition aims to familiarize the public with the history and current state of the Istro-Romanian community in the Istrian Peninsula and to provide an overview of the research developed at Oxford University into the community and their language. It draws on the photographs and audio recordings made by Tony Hurren in Žejane, Šušnjevica, and surrounding area, during the 1960s, and on material collected recently by our project, building on the Hurren archive and expanding on Hurren's research. For us, the exhibition is an important outreach element of the project – all the more so as it is an offline outcome which we planned for reluctantly while still during the pandemic, unsure whether the circumstances would allow this form of direct encounter with the community in Croatia.

While Hurren's audio recordings have naturally been at the centre of ISTROX as a project in Linguistics, this forthcoming exhibition has temporarily changed the hierarchy of the different types of material that have been part of the Hurren donation. Museum exhibitions unfold in multi-dimensional environments where images and material culture are often primary elements of the display. While sound is often part of different types of exhibitions – as will be the case with ours – the objects and images are usually the ones called to convey the multi-modal environments meant to facilitate the visitors' engagement with different types of content. Thus, we started work on the exhibition by revisiting the two main bodies of images that we have so far and exploring potential connections between them. The basis of the exhibition is provided by the fifty analogue photographs taken by the Hurrens in Istria in the 1960s. To those we added a second body of digital images produced by the photographer Bojan Mrđenović in the spring of 2019, when he accompanied us during our trip to Istria and documented our meetings with Hurren's former interviewees or their relatives.

In addition to these two bodies of images, the exhibition will include other items that were part of the Hurren donation, most importantly Hurren's notebooks, but also a range of newly discovered items that Vera Hurren has kindly lent to us specifically for the exhibition. Among these are the couple's passports (with Yugoslavian visas from the 1960s) and the letters of recommendation that Tony Hurren carried with him, to ensure that he was accepted in the country and the region, both by the Yugoslav authorities and by the villagers themselves; these are signed by Oxford specialists such as J. Mulder (Hurren's D.Phil. supervisor at Oxford), but also by August Kovačec, at the time a young assistant lecturer in Romance linguistics at Zagreb University.

Figure 5. Opatija, May 2019. Dolores Posidel listens to a recording of her mother Valeria Jurman (Posidel).

Photographs capture the texture of the lives and social worlds of people and communities. Though interested to learn about the past of the community as documented by these photos, our approach to them during our field trip was primarily functional: before anything else, we wanted to identify the people and the locations that featured in the photos, to clear and index them for future use in the project. Part of the ethical clearance process involved inviting our interlocutors to listen to selected audio clips of themselves or their relatives, as recorded by Hurren, to make sure that the attribution of identity in Hurren's notes and on the covers of the reel-to-reel tapes was correct. Many of these sound-mediated encounters turned into emotional moments, as the former participants or surviving relatives listened to old recordings of themselves or of their long-gone relatives.

While in Istria, we also wanted to identify the people and the locations featured in the fifty photographs donated to the university alongside the recordings. While we were aware from the beginning that some of the people from the audio recordings may also appear in the photos, we could not identify any of them, or the locations of the images, before our field trip. We had the photos with us available for viewing both in digital format, on a laptop, and as a large A4 print; we used one format or another depending on the personal preferences of our interviewees and on the context of viewing: while the large prints were handier to flick through during a short encounter, on a Sunday morning, while leaving the church, the laptop screen allowed a better scrutiny of the unidentified faces, particularly in the case of the few group photographs where most faces remained blurred even after we had them enhanced.

Figure 6. Šušnjevica church, May 2019. People look at the Hurren photos.

Produced with a non-professional camera, these photos provide contexts for the social worlds encountered and the human interactions enabled by Hurren's research. They document field interview moments, encounters with villagers in Žejane, Šušnjevica, Kostrčan and elsewhere, as well as village scenes which, for a reason or another, were regarded at the time as relevant. A few of Hurren's interviewees feature in the photos more frequently than others – an indication that the couple spent more time and developed closer

relationships with them, which extended beyond the domain of the linguistic interview into the leisure time spent by the couple in Žejane and Šušnjevica.

Apart from adding visual texture to some of the voices recorded by Hurren, these photos also helped us grasp better the role played by Vera Hurren in her husband's research. Adrijana Baćac was only 16 when Hurren interviewed her. When we met her some fifty years later and we mentioned his name, she exclaimed with a smile: "Tony and Vera, Hampton Poyle!". She explained that she remembered "Hampton Poyle" (the name of village where the Hurrens lived) because, over the years, Vera had kept in touch with her by sending letters, postcards, and small presents, particularly around Christmas. Gracijela Bortul, who was also interviewed by Hurren when she was a teenager, had a similar story: at some point, she had sent the Hurrens a photograph taken after her communion ceremony, when she wore a veil which Vera Hurren had given her as a present; we found that photograph among those donated to us. See Figure 7 and 8.

It is telling that, at a time when women's assistance with the work of their male partners, whether academic or not, has been increasingly a topic of discussion in academia and history (Desvina 2021), these two women remembered not only the linguist, but also how discreetly indispensable his wife was in his research: starting from 1963, she accompanied him to Istria each summer, for years, driving him everywhere while in the field and making logistical arrangements, and keeping in touch with his former research participants throughout the years. Most importantly, she typed and often retyped his manuscripts, including his doctoral thesis, before donating his archive to the university in the recent years, thus making it available for future research.

Figure 7. Trieste, Year unknown. Raffaela Bortul, Gracijela Bortul, Dalla Belulović after Gracijela's communion.

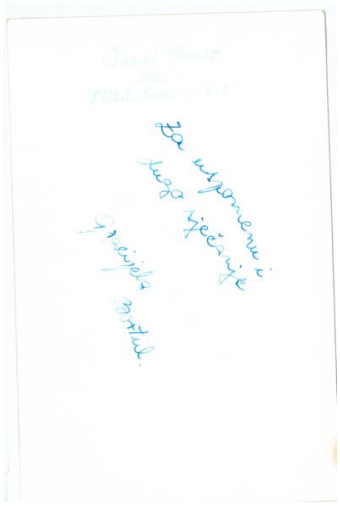

Figure 8. Handwritten note on the back of the photo: "Za uspomenu i dugo sjećanje, Gracijela Bortul" ('in memory and long remembrance, Gracijela Bortul')

Tony Hurren's research and profile as an Oxford scholar, Vera Hurren's role in her husband's work, the long-term relations that they developed with Tony's interviewees, followed by our own work building on the Hurren donation (from offline to online research and from the Croatian to the worldwide Istro-Romanian community) will provide the main narrative clusters of the exhibition. As part of our ongoing attempt to reach out to this community, the exhibition will open with a special preview for the members of the community from the Istrian villages, for whom we will arrange transport to Novigrad. While we have in mind in particular the more senior members of the community, perhaps less proficient online, for whom this offline opportunity to encounter the Hurren material is likely to feel more natural than through our Zooniverse site, we are ready to welcome members of the community off all ages, whether speakers or non-speakers of the language. We hope that this exhibition will mark the beginning of a return to offline encounters and interviews following the recent pandemic years – in particular with those members of the community who are not online.[26]

Figures 9, 10, 11. Novigrad, 26 June 2022. Exhibition opening at the Lapidarium Museum.

[26] I add this note after the opening of the exhibition on 26 June 2022. Around 60 members of the community attended the opening, 48 of them using the transport that we provided from the villages, others independently. The exhibition will remain open until 24 August 2022.

8. CONCLUSION

ISTROX has brought to light a virtually unknown body of historic sound recordings held in Oxford. One of our goals for this first phase of the project has been to stimulate public and scholarly engagement with the Hurren material which had been unavailable for decades. On the completion of the project we made all the original material available online through Oxford's ORA (The Oxford Research Archive)[27]. We also plan to make all the other outputs resulting from future stages of the project available on an open access basis – an aspect that is consistent with our attempt to produce this information with the voluntary contribution of the members of the community.[28]

As a pilot project, it aimed to help us, on the one hand, to start work on the Hurren material by indexing and transcribing it, and on the other hand, to learn about the Istro-Romanian community worldwide, establish a connection with as many as possible of its remining members, and test an online structure for long-term community-sourcing. The launch of the Zooniverse element of the project has been a first step towards assembling around the project a small group of members of the community, both from Istria and the diaspora, and both speakers and non-speakers of the language, who are willing to be involved in future phases of our research.

Given the very small number of remaining speakers worldwide, engagement work has been crucial throughout this lengthy initial phase and will remain so. The relationship between researcher teams and communities, particularly ones that are so globally scattered, move at a slow pace. While most literature on crowdsourcing touches on the amount of work and energy needed to keep a community sourcing project going, and to build and maintain a community's sense of connection with the project, we have learned from direct experience that there is a difference between large scale crowd-sourcing projects and (very) small scale community-centred projects such as ours; the latter require a different type of intellectual and emotional engagement, more oriented towards individuals and person by person communication. Not least, we also learnt about the amount of effort and flexibility needed to secure short-term funding during the transition phases of a project between different stages of long-term funding.

We have integrated community-sourcing into a larger ecology of offline and online strategies to connect with the community, which included Zoom and Skype interviews, individual and group emails, media visibility, museum engagement. Flexibility and a

[27] We have deposited the original Hurren materials in ORA (the Oxford Research Archive), where they can be consulted and downloaded by researchers and members of the IR community. See https://ora.ox.ac.uk/objects/uuid:9c7e2da8-ae4a-434c-8dbe-589afbaa2cb6. There have been 158 downloads of the ISTROX dataset as of 28 June 2022.

[28] Research on crowdsourcing practices stresses that knowledge created through crowd-sourcing should be available on an open access basis. This has to do with the fact that the participants have volunteered to contribute, and limiting access to their contributions may be felt as devaluing those contributions. Also, the open access quality of the final products works as a reward structure, as most volunteers are proud of their achievements and like to show the final result to friends and family. See Dunn and Hedges (2012: 19). In the questionnaire that mediates the access to our Zooniverse site we mention our intention to acknowledge, whether by avatar or full name, all the contributors in the published materials resulting from our project and require their permission to be contacted in connection with this.

certain degree of experimentation have been essential while running the project. We have learned not to fixate too much on what we considered at the beginning to be the goal for an action or another that we developed: for instance, even when an action that we took did not result in attracting new participants on Zooniverse, it still increased the number of members of the community in our database who declared themselves willing to be further approached for face to face meetings, whether online or offline. Covid-19 prevented us from doing essential offline work, yet it led us to rethink and develop further the forms of online interaction that we had considered initially.

"Should academic or cultural crowdsourcing reflect the interest of the academy, or of the crowd itself?", ask Dunn and Hedges (2014: 234) in response to the oft-mentioned claim that crowdsourcing "empowers the crowd and democratises information". The said "empowerment" and "democratisation" do not come by default, they remind us; we should always look closer at the way a project is designed and executed. As we have just completed a stage that we considered preliminary to forthcoming research, it may be too early to comment on the above. However, we believe that there is shared purpose in our work and the preservation interests of the Istro-Romanian linguistic community. The Hurren audio material is not only significant for our research, but also meaningful for the speakers and non-speakers of the language who, through our online site, can encounter a version of the language that is no longer in use. We hope that, in the future, more members of the community will decide to contribute to our project or to mediate between our heritage material and their less "computer-savvy" relatives. Apart from helping us, this may also work for them as a way of bonding with their parents and grandparents, and also of learning about their linguistic heritage.

At the time of writing, participation on Zooniverse remains numerically small yet commensurate with the size of the community: 17 volunteers, of ages between 20s and 70s, and 1153 classifications completed. Of those 17 participants, 10 work on the Croatian interface and are based in Croatia, and 7 work on the English interface, with six based outside Croatia and one in Croatia; two of the ones working in English and based outside Croatia are translating with technical support from younger relatives. Despite their overt enthusiasm for the project, many remain reluctant to engage with our material on Zooniverse. While we are genuinely interested in increasing our number of online participants, what interests us even more is to increase the quality and intensity of their involvement. One aspect is to retain at least some of our existing Zooniverse participants by making sure that they return to the site from time to time to translate further and thus are exposed to a larger variety of recorded material. Another aspect is to extend their involvement through other segments of the project, both online and offline. Our expectations about the total number of potential contributors that we may be able to reach in the future are relatively modest: we believe that we cannot reach more than 50-60 participants, simply because the pool of people who possess the language and ICT skills required is itself rather small.

While the ISTROX pilot and the follow-up projects have come to an end, our Zooniverse site remains open for contributions. Although we cannot know how data collection will evolve in the future, we regard the Zooniverse element as a "long tail" type of engagement. Indeed, it is the long term that we are mostly concerned about. Crowd- or community-sourcing are not only able to make knowledge circulate better; apart from the quantifiable outcomes, projects of this kind also catalyse a range of intangible outcomes such as awareness or knowledge about a topic or another, which can be particularly

significant for those who participate (Dunn and Hedges 2012: 38). In our case, we hope that contributing to our project by listening and translating some of Hurren's recordings may bring to some participants, in particular those based outside Croatia, a renewed sense of belonging to a severely reduced community. Projects that involve crowdsourcing in the humanities are said to not only bring out new information that can lead to the creation of new knowledge, but also to help forge new communities whose engagement and "activation" – as pointed out by Dunn and Hedges (2012: 19-20) – may go beyond the scope or duration of the project itself. This is certainly what we hope to achieve in the long term, subject to further funding: create a virtual space in which the Istrian and diasporic communities, the young and the old, those who still speak or understand the language and those who do not, will be involved together in the preservation of their language with renewed interest in and enthusiasm for their shared linguistic and cultural identity. We thank all those who have contributed so far or will decide to work with us in the future.

REFERENCES

Anderson, B., 1983, *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, London, Verso.

Archibald, M., R. Ambagtsheer, M. Casey, M. Lawless, 2019, "Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants", *International Journal of Qualitative Methods*, 18, 1–8.

Cassell, J. and Jacobs, S., 1987, *Handbook on ethical issues in anthropology*, Arlington, VA, American Anthropological Association, Washington D.C.

Causer, T., V. Wallace, 2012, "Building A Volunteer Community: Results and Findings from Transcribe Bentham", *Digital Humanities Quarterly*, 6, 2, http://www. digitalhumanities.org/dhq/vol/6/2/000125/000125.html#p1 [accessed 10 June 2022]

Cohen, K., 2020, "Citizen Science in a Pandemic: A Fleeting Moment or New Normal?", [Blog] *CRASSH*, *https://www.crassh.cam.ac.uk/blog/citizen-science-in-a-pandemic-a-fleeting-moment-or-new-normal* [accessed 15 May 2022].

Dibeltulo, S., S. Culhane, D. Treveri Gennari, 2020, "Bridging the digital divide: Older adults' engagement with online cinema heritage", *Digital Scholarship in the Humanities*, 35, 4, 797–811.

Dresvina, J. (ed.), 2021, *Thanks for typing: Remembering Forgotten Women in History*, London, Bloomsbury Academic.

Dunn, S., M. Hedges, 2012, *Crowd-Sourcing Scoping Study: Engaging the Crowd with Humanities Research*, Arts and Humanities Research Council, https://stuartdunn.files.wordpress.com/2013/04/crowdsourcing-connected-communities.pdf [accessed 10 June 2022]

Dunn, S., M. Hedges, 2014, "How the Crowd Can Surprise Us: Humanities Crowdsourcing and the Creation of Knowledge", in: M. Ridge (ed), *Crowdsourcing our Cultural Heritage*, Farnham, Ashgate, 231–245.

Howe, J., 2006, "The rise of crowdsourcing", Wired, 14.

Leist, A., 2013, "Social Media Use of Older Adults: A Mini-Review", *Gerontology*, 59, 4, 378–384.

Morris, A., J. Goodman, and H. Brading, 2006, "Internet use and non-use: views of older users", *Universal Access in the Information Society*, 6, 1, 43–57.

Owens, T., 2012, "The Crowd and The Library", [Blog] *User Centered Digital Memory*, http://www.trevorowens.org/2012/05/the-crowd-and-the-library/ [accessed 20 May 2022].

Owens, T., 2014, **"**Making Crowdsourcing Compatible with the Missions and Values of Cultural Heritage Organizations", in: M. Ridge (ed), *Crowdsourcing our Cultural Heritage*, Farnham, Ashgate, 269–280.

Ridge, M., 2014, "Crowdsourcing our cultural heritage: Introduction", in: M. Ridge (ed.), *Crowdsourcing our Cultural Heritage*, Farnham, Ashgate, 1–13.

Ridge, M., (ed.), 2014, *Crowdsourcing our cultural heritage*, Farnham, Ashgate.

Schenk, E., C. Guittard, 2011, "Towards a characterization of crowdsourcing practices", *Journal of Innovation Economics & Management*, 7, 1, 93–107.

Viellieux, C., 2020, "Amidst the Pandemic, Citizens Create a Boost in Classifications", [Blog] *Snapshot Wisconsin- A Zooniverse project blog*, https://blog.snapshotwisconsin.org/2020/07/25/amidst-the-pandemic-citizens-create-a-boost-in-classifications [accessed 15 May 2022].