Open Library of Humanities

Towards Language Sensitivity and Diversity in the Digital Humanities

Paul Joseph Spence, King's College London, UK, paul.spence@kcl.ac.uk
Renata Brandao, King's College London, UK, renata.brandao@kcl.ac.uk

Recent years have seen a growing focus on diversity in the digital humanities, and yet there has been rather less work on geolinguistic diversity, and the research which has been carried out often focuses on the structures of geographic representation in the field or has viewed 'language' as a technical or linguistic problem to solve. This article takes a different view, namely that we need to consider this diversity through multiple 'frames' of digitally-mediated language and culture, and that this is not just a question of epistemic justice or community manners, but that the digital humanities also need to address more actively challenges around global dynamics of digital multilingualism, transcultural exchange and geodiversity in its research agenda. This paper explores these questions through the prism of 'language indifference' in digital studies and, responding to Galina's call for better data on the state of geolinguistic diversity in DH (2014), it articulates possible frameworks for addressing this diversity in a strategic, programmatic and research-led manner. We conclude by exploring the role of a greater multilingual focus in what Liu calls 'the techne of diversity' in digital humanities (2020), and contend that the digital humanities has much to gain, and much to offer, in engaging more fully with the languages-related cultural challenges of our era.

Ces dernières années l'accent a été mis de plus en plus sur la diversité dans les sciences humaines numériques, et pourtant il y a plutôt eu moins de travaux sur la diversité geo linguistique, et les recherches qui ont été menées portent souvent sur les structures de la représentation géographique sur le terrain, ou estiment le 'langage' comme un problème technique ou linguistique à résoudre. Cet article adopte un point différent, à savoir que nous devons considérer cette diversité à travers plusieurs 'cadres' de culture et de language à médiation numérique, cela n'étant pas uniquement une question de justice ou de savoir-faire communautaire, mais que, dans son programme de recherches, les sciences humaines numériques doivent également relever plus activement les défis à la dynamique mondiale du multilinguisme numérique, aux échanges transculturels et à la geo diversité. Ce document explore ces questions à travers le prisme de 'l'indifférence linguistique' dans les études numériques et, en réponse à l'appel de Galina pour de meilleures données sur l'état de la diversité geo linguistique dans

DH (2014), il définit des systèmes possibles pour faire face à cette diversité de manière stratégique, programmatique et axée sur la recherche. Nous en concluons qu'en explorant le rôle d'une meilleure focalisation sur le multilinguisme dans les humanités numériques de ce que Liu appelle 'la tech de la diversité' (2020) et nous soutenons que les sciences humaines numériques ont beaucoup à gagner en s'engageant pleinement dans les défis culturels liés aux langues de notre époque.

1 Introduction

In her landmark article "Geographical and linguistic diversity in the Digital Humanities" published in 2014, Isabel Galina extended debates about diversity (then primarily focused on gender, class, ethnicity, race and sexual orientation) to examine how language and locale shape debates in the field of the digital humanities (DH). "Who is 'we'?," she asked, in assessing linguistic and regional initiatives which aim to challenge English language dominance in DH and to empower non-Anglophone communities in the field (Galina Russell 2014). The article is one of a series of articles over the last few years which have queried the geocultural makeup of DH, often putting the spotlight on its professional associations, and at least in geo-institutional terms it is clear that some progress has been made, with, for example, four new associations added to the Alliance of Digital Humanities Organizations (ADHO) in the six years since Galina's article (see appendix).

While there has been much attention to the geography of DH representation, there has been far less evidence of attention to geolinguistic diversity in this period. There have been numerous attempts to expand the linguistic focus of DH – through initiatives such as *DayofDH* in Spanish and Portuguese (Priani Saisó et al. 2014), *Digital Humanities Quarterly* journal special issues in Spanish, French and Portuguese, and Spanish/French editions of *The Programming Historian* – but while this has started to open up space for speakers of other languages beyond English, it does not constitute a major scholarly engagement with the implications of multilingualism and geocultural diversity for DH scholarship, nor indeed a substantive engagement with languages-focused research more widely.

Although there is a long history of engagement and overlap with some disciplines such as computational linguistics and philology, digital humanities engagement with languages-focused research fields has traditionally been, at best, uneven. This has started to change recently, as we have seen increasing attention to the digital mediation of Modern Languages and other language-based disciplines, a process which has generated overlapping debates about transcultural and translingual exchange. Nevertheless, we argue here that the digital humanities have not generally been as receptive as they might be to bidirectional intellectual interaction with fields such as Modern Languages, translation studies or minority / endangered languages archives, and we contend that the digital humanities would benefit from engaging more critically (and practically) with the dynamics of "language indifference" or "language insensitivity" which are a crucial barrier to digital multilingualism and geolinguistic diversity in the field.

Historically, debates about language diversity and awareness in DH have often focused on the politics of its scholarly communications. Important though that is, we maintain that a broader axis of interaction between DH and languages-focused fields is needed in order to enact change and we approach the topic through a number of epistemic frames and perspectives in order to examine key features of linguistic diversity in the field. We trace the contours of emerging (and overlapping) networks or communities of practice which might be loosely defined as "multilingual DH," "translingual/ transcultural digital scholarship" and "digital modern languages," before considering what consequences arise for both the strategic direction and scholarly agenda of the digital humanities, and proposing some theoretical and practical frameworks for modelling linguistic diversity in DH. We begin by briefly exploring two useful frames for approaching this topic: geolinguistic diversity and the concept of "language insensitivity" (sometimes called "language indifference").

2 Frames

2.1. Frame #1 – Linguistic and geocultural dynamics online

It is estimated that there are around 7,000 living languages in the world, but over 40% are endangered (Endangered Languages Project 2020) and in 2004 Graddol estimated that as many as 90% of spoken languages "may be doomed to extinction" (Graddol 2004, 1329).

Some have connected concern over the survival of human languages with ecological threats to human existence under the concept of *biocultural diversity* (Maffi 2005). In a European context, numerous documents underline the importance in safeguarding this "living heritage" of languages and cultures, both as a key element of regional identity and a driver for social and material wealth (Prys Jones 2013; The Network to Promote Linguistic Diversity (NPLD) 2015). In a similar vein, a series of studies in recent years has shown the individual and social benefits of multilingualism across various indicators, including cognitive performance in early years and defence against cognitive aging later in life (Bak and Mehmedbegovic 2017). While some warn against the dangers of "essentializing discourse" around discourses of linguistic endangerment and propose attention to "a polynomic" model of linguistic identity" (Jaffe 2007, 57), there can be little doubt that linguistic divides play a significant role in cultural production and related knowledge practices.

In the academic sphere, little has changed in overall international policy terms since a 2001 Vienna Manifesto highlighted a series of measures for academic institutions to address the "cost of monolingualism" (Cillia, Krumm, and Wodak 2001). In his exploration of "scientific monolingualism" and the underlying threat of *epistemicide*,

or "killing of knowledge systems," Forsdick emphasises the importance of multilingual knowledge flows, the value of translation and the wider scope for research which language-aware multicultural research teams enable (Forsdick 2018, 75–76). The OPERAS Multilingualism White Paper in 2018 affirms that "the choice of a language system often implies the choice of a frame of references, of a methodology, of a school" (Delfim Leão et al. 2018). In a similar vein, a 2016 article by Matt Pickles on the BBC website asked if the "dominance of English [could] harm global scholarship" (Pickles 2016). Quoting several academics, the article raised a number of different challenges, including different rhetorical cultures, limitations in the scope and quality of research caused by ignoring work in other languages, the epistemological bias caused by using only one language and the *de facto* regional "gatekeeping" effects this has on knowledge flows. This is by no means a new topic in the digital humanities – in "Digital Humanities and the Geopolitics of knowledge"; for example, Domenico Fiormonte suggests that some of the standards and technocultural codes popular in Anglophone DH pose a threat to biocultural diversity (Fiormonte 2017) – and yet there have been few concerted efforts by DH as a field to gain a wider and more detailed understanding of the underlying dynamics of diversity, in particular in relation to the *linguistic-cultural* element of that diversity, where "there is little known data available" (Galina Russell 2014: 308).

2.2. Frame #2 – Language insensitivity/language indifference

Some, such as Forsdick, have criticized "the unmarked monolingual assumptions of numerous academic disciplines and non-academic sectors," arguing that language insensitivity is a phenomenon which needs to be addressed throughout academic scholarship as a whole (Forsdick 2017). Similarly, a Transnationalizing Modern Languages project report in 2018 argued that we need to counter "language indifference," that we "need to make the work of language and of translation more visible" and that "we need to stress that languages are not neutral but deeply connected to the cultural, political and economic dimensions of social life". Finally, they argue that we "should pay attention to languages across our educational systems and in our everyday practices" (Burdett et al. 2018). From a digital humanities perspective then, language sensitivity here amounts to far more than the simple awareness that multilingualism exists as a box to tick; it foregrounds how languages (plural) are bound up with notions of culture and community in forming the sense of location and perspective in our work.

DH suffers from these "monolingual assumptions" as much as most other fields and while there have been some initiatives such as the Global Outlook DH "Translation

toolkit"– which provides recommendations for multilingual conference etiquette and represents a linguistically sensitive counter-model for globally inclusive scholarly communications – there is a lot to improve in DH's internal practice (Dacos 2013). These critical contributions are important elements of a wider debate about DH's political infrastructure and scholarly communications. However, we would argue that policy-based responses to this question only deal with part of the problem and that we need to think of language sensitivity – and language diversity – not just as a key factor in the field's communication practices, but also as a fundamental feature of the digital humanities research agenda. In other words, here we do not merely posit language indifference as a design fault in DH's scholarly communications, but rather, in our view, overcoming it is a crucial research and pedagogical challenge for DH.

In calling for a campaign against language indifference, Loredana Polezzi has argued that it is precisely the "pervasiveness" of language that leads to its frequent invisibility, but that language is also "always plural, always a place of difference" (Polezzi, forthcoming). In this context, we believe that it is vitally important for digital research to get away from (1) seeing "languages" in narrowly linguistic terms, separated from their cultural context and (2) viewing this as a purely "technical" problem, which can be resolved by digital tools alone.

3. Perspectives on "language" in DH

A central proposition of this article is that language sensitivity and diversity in digital studies start with awareness of the scope of the challenges, and this in turn depends on visibility – of language communities, their languages and their cultures. One central obstacle in addressing digital language diversity is that the disciplines and practices on which it depends are subject to a high degree of fragmentation. This section explores how and where digital language diversity is in focus for the digital humanities, principally through six perspectives: (1) global DH debates and advocacy; (2) geolinguistic communities in DH; (3) language technologies and associated practices/research; (4) the field of Modern Languages; (5) sociolinguistic research into multilingualism online; and (6) multilingual DH infrastructure initiatives. The first perspective is embedded in wider analysis of global divides, the second is defined by geocultural and linguistic positionality, the third, fourth and fifth respectively have languages, cultures and multilingualism as their object of study; and the sixth focuses on efforts to foster digital geolinguistic/cultural diversity in digital ecosystems. The very diverse set of perspectives at play here make the topic challenging in terms of scope and terminology, but we believe that it is important to consider all six in analysing the future implications for language diversity.

3.1. Perspective #1 – Global DH

While there has been modest coverage of geolinguistic diversity in digital humanities which has focused largely, as we have seen, on the field's own scholarly communications, there have not been many initiatives to chart or promote this diversity from international DH institutions. The activity of ADHO's Multilingual Multicultural committee has so far been largely limited to translating the Call for Papers for the annual ADHO "DH" conference, and the main impetus for linguistic transformation has come from Global Outlook DH (GODH), which in addition to the afore-mentioned Translation toolkit also fostered the "DH Whisperers" initiative to encourage informal translation at DH events (Gil 2014; Gil and Ortega 2016; Del Río Riande et al. 2020; Ortega 2014). There have also been similar projects elsewhere, such as the "RedHD in Translation" initiative, which aimed to provide "flash" translation of Spanish language DH scholarship (Ortega 2019, 183).

More recently these activities have led to other expressions of global and multilingual digital scholarship such as the Force 11 Open, Multilingual and Global Scholarly Communication (OMG) working group, which challenges the "global scholarly communication community [to] develop more openly and equitably "trans-lationships" (translational relationships) across cultures, languages, regions, boundaries, disciplines and worldviews" (Del Río Riande, Lujano, and O'Donnell 2020) and the Open Methods platform, which follows a languages-focused approach to its curation of digital humanities methods and tools in order to champion multilingual and multicultural identities in DH (Open Methods Languages 2017). The toolkits which emerge from these initiatives may need to aim for "universality" or for a more regional focus, depending on context.

3.2. Perspective #2 – Geolinguistic communities

Regional DH professional associations display a variety of different relationships to linguistic identity, but generally speaking the relationship has so far been implied rather than overt, at least when it comes to the formal member organisations of the Alliance of Digital Humanities Organizations (ADHO). Of the ten member organisations currently listed on the ADHO website, only CSDH/SCHN (Canada), EADH (Europe), Humanistica (francophone focus) and JADH (Japan) explicitly address linguistic diversity in their "About" pages, although others refer to local or "indigenous" knowledge. Only one of these is specifically defined by a linguistic focus: Humanistica has existed since 2014 (Humanistica: Présentation 2014) as a professional association which both unites and promotes francophone DH, irrespective of its geographic location, through a number of activities which include the francophone journal "Humanités numériques" ("Digital Humanities"). At European level, there is also a German-language association,

"Digital Humanities im deutschsprachigen raum". It goes without saying that these geolinguistic communities are to some extent shaped by colonial or other historical currents, a condition which shapes their engagement with the power dynamics of specific languages.

More recently, a series of events in Africa and Europe, some with a strong geolinguistic focus, have led to the constitution of the Network for Digital Humanities in Africa in 2020 (Network for Digital Humanities in Africa 2020a). A panel featuring multiple African researchers at the DH2019 conference in Utrecht (in the Netherlands) highlighted the challenges in situating African languages in the global digital landscape and a follow-up forum at the virtual DH2020 conference sought to consolidate this within a broader drive to promote African DH scholarship, by drawing on language communities and seeking to facilitate mechanisms for wider access, in Africa, to digital data for African and other languages (Network for Digital Humanities in Africa 2019 2019; Steyn et al. 2020; Network for Digital Humanities in Africa 2020b ; SADiLaR 2020).

Community-led geolinguistic initiatives like these (with strategic support from funders and other DH organisations where required) undoubtedly provide one very effective approach to addressing linguistic diversity in DH, but we now turn to approaches focusing on languages, cultures and multilingualism as the object of study.

3.3. Perspective #3 – Language technologies

Language technologies (viewed both through research and professional practice) are an obvious place to start in discussing digital multilingualism as an object of study, and we do not wish to underestimate the very important role these fields will play in promoting digital language diversity. Ambitious visions for digital multilingual academic and commercial infrastructure at European level such as CLARIN's Language Resource Switchboard (CLARIN 2020) and the European Language Grid project (2020) can offer effective responses in resource-rich language environments. In low resource language contexts, often marginalised by advanced computational approaches, language technology approaches need to take into account: the importance of speech technologies, adaptive NLP strategies, graphical/multimodal interfaces and community engagement in areas with poor literacy, low availability of written data and an absence of language standardisation or no script (Joshi et al. 2019).

3.4. Perspective #4 – DHML

One key reference point for this article has been the emerging body of research into interactions between the field of modern (foreign) languages (ML/MFL) and digital culture, and the concept of "ML-inflected Digital Humanities" (or "DHML") proposed

by Pitman and Taylor in a *Digital Humanities Quarterly* journal article in 2017 (Pitman and Taylor 2017). We will continue to explore the concept of DHML elsewhere as part of ongoing research, but in our view a fundamental element in this discussion is the "very fertile ground of difference, of cultural and linguistic difference and otherness" which Phipps and Gonzalez suggest may be a key area of potential for greater theorisation by fields such as the Modern Languages (Phipps and Gonzalez 2004, 43). DH–ML collaborations are most fruitful when they involve what Ortega calls "narratives of cultural encounter and cultural mixing" (Ortega 2018, 19) and when they expose transnational/cross-border dynamics, which clearly require a multilingual (and translingual) sensibility. This perspective invites "broader conceptions of culture and cultural representations" in our engagement with digital methods and ecosystems (Arriaga 2020). More broadly, this agenda foregrounds the cultural aspects of language education and research which DH responses (often focusing on narrowly technical or primarily linguistic aspects) tend to overlook. This is the challenge which a programme of activities co-convened by Spence and Wells under the "Digital Modern Languages" label attempts to address. This programme includes a *seminar series* (Digital Modern Languages 2019a), *blog series* (Digital Modern Languages 2019b), a *discussion list* with over 400 members (JISCMail " Digital Modern Languages List" 2019) and a *publication section* in Liverpool University's open access platform Modern Languages Open (Liverpool University Press 2019).

We do not suggest for a minute here that Modern Languages as a field is the sole repository of knowledge about intercultural dynamics and translation, but instead simply argue that its role will increasingly help DH to address translingual and transcultural challenges going forward and to meet other challenges beyond the scope of this article, such as the "international classroom" and multilingual pedagogies.

3.5. Perspective #5 – Sociolinguistic research into multilingualism online

Much of the research currently taking place into multilingualism online is being carried out by sociolinguists to study the relationship between linguistic performance in people's online and offline practices, identity management and sociocultural production. Again, DH would benefit from closer collaboration with this field, which has done important work in many areas DH could usefully draw on in articulating its own multilingual frameworks including: the language choices made by multilingual speakers in different contexts; multilingual affordances and barriers in digital ecosystems; the "multilingual practices" of monolingual subjects; the ways in which multilingualism is conceptualised and implemented by digital media companies; and the different ideological, technical and social practices these entail (Danet and Herring 2007; Lee 2016).

3.6. Perspective #6 – Multilingual DH

At least in an Anglophone context, discussion around multilingual DH has largely centred around the activities of the Global Outlook Digital Humanities group (Global Outlook::Digital Humanities 2020.), although in recent years initiatives such as the Multilingual DH network (Multilingual DH 2020), workshops for Right-to-Left (RTL) languages and cultures (NYU Abu Dhabi Winter Institute in Digital Humanities2020) and Non-Latin Script (NLS) workshops (Lee and Wagner 2019) have started to explore the practical and thematic challenges of geolinguistic diversity in digital research, in particular in areas such as digital literacies, script/text representation, OCR, data curation, NLP and visualisation. While these initiatives still currently tend to be primarily focused on resource-rich languages, they provide a welcome impetus to the drive for translingual research in the field.

We will explore these examples in greater detail later, but we wish to end this section by emphasising that engaging with language sensitivity and diversity in digital research inevitably involves multiple epistemic, social and technical perspectives, and that it should be addressed in different stages: firstly, in gaining better awareness of the scope of the challenges; then designing and articulating new models for multilingual and "language-sensitive" research; enabling "language-sensitive" research methods and infrastructures; consolidating geolinguistically diverse communities of practice; and finally, by articulating DH-specific roles in combatting "language indifference" in digital research.

4. Towards a languages-centric agenda for DH

4.1. Improving understanding of the dynamics of linguistic diversity in digital research

Contrary to what many assume in the Anglophone world, monolingualism is the exception in the world, and may one day be viewed as "peculiar" in historical terms (Wallraff 2000). This view is supported by new developments in human conflict, economic migration, diasporic communities, and information & communication practices which foster increasingly transnational, and translingual human relations. However, while language practices are increasingly subject to the effects of "superdiversity" (Androutsopoulos and Juffermans 2014), global language diversity is under attack and lower resource languages are under severe threat of extinction. It is important to note here that while the dominance of English is undoubtedly the primary linguistic challenge here, it is part of a wider dynamic. The Anglophone world is not alone in fomenting mistaken assumptions of nation-bound normative monolingualism, and we can also observe a much wider consolidation globally of what Abram de Swaan calls "supercentral" languages (such as Spanish/Castilian, French,

Arabic and Mandarin Chinese) – with English as the "hypercentral" language (Swaan 2002). This is a phenomenon which will increasingly become a major challenge for DH and scholarly communications as a whole. All of this has major implications for how DH engages with linguistic diversity and understands knowledge flows in the future, implications which are currently under-researched in DH and digital research more generally.

Despite growing sensitivity to the issue on digital platforms, linguistic diversity is not well represented in digital environments overall and it is likely that only a few hundred languages are actively in use on web pages to a significant degree. Nevertheless, there is evidence of a higher profile for linguistic diversity in some mass/digital "mediascapes" (Androutsopoulos 2007, 207), and this varies across different tools and media; for example, Prado suggests that informal media such as blogs or messaging tools present greater linguistic diversity than formal digital media (Maaya Network 2012). Similarly, Cuncliffe points to the more "embedded" nature of social media in everyday life as favouring low resource languages (Cunliffe 2019, 451). While there is some excellent qualitative ethnographic work in this area, there is a shortage of studies and tools which might help us get a sense of the scale of linguistic diversity in digital media, and projects or initiatives committed to this task have not always enjoyed the stability they deserve. This makes it difficult to come to hard conclusions about the degree of linguistic diversity in digital communication overall, and so, for now at least, we lack crucial detail about the linguistic flows in digitally-mediated knowledge production, but there can be little doubt that most instruments of global knowledge production strongly favour first English, and then a small number of high resource languages. Given this landscape, a key challenge for the digital humanities in the next few years will be to better understand and analyse these dynamics, and to then design strategies to disrupt digital monolingualism accordingly.

As part of our research into linguistic coverage in DH infrastructures on the AHRC-funded "Language Acts & Worldmaking" project (2020), we surveyed an array of DH projects and repositories, examining which languages feature and how information about languages is represented. Language-based research has been a central part of the digital humanities (DH) since its inception, but in spite of the significant proportion of projects with a strong languages focus in DH catalogues (84 out of 794 projects listed in the currently defunct DH Commons portal according to our research), we found that their relationship is generally under-articulated in comparison to other cognate disciplines, such as English or History. At present, and despite the valuable work of corpus-based research infrastructures such as CLARIN, it is generally not easy to discover the nature and extent of the language focus of DH research online, even on

otherwise excellent and implicitly language-rich resources such as the EADH project list (EADH – The European Association for Digital Humanities 2019). Our study found that on the EADH list, while there is often decent multilingual coverage (especially in language-oriented studies), the actual linguistic coverage in the list as a whole seems to favour a very small number of languages. Analysing the actual language of presentation on the list of 197 projects we were able to access, we found 72% used English, 19.8% German, and only three other languages hit over 5% (2019). One starting point, then, would be to achieve greater recognition for, and organisation of, multilingual resources and methods in DH as a whole, a task funders, academic institutions and professional associations in DH can all contribute to.

While attention to linguistic diversity in its own scholarly communication practices is important, it is equally important for DH to make a greater contribution to linguistic diversity as a research topic. There are many areas where DH could make a greater impact, and one of these is in analysing and designing critical infrastructure for low resource languages. Indeed, an article by Nick Thieberger in the *Digital Scholarship in the Humanities* journal in 2017 argues that making information about the world's small languages more freely available should be a digital humanities project – promoting greater visibility for language repositories, developing initiatives to connect endangered language resources and elevating the value of oral evidence (Thieberger 2017).

Digital humanists are already active in some momentous programmes to protect endangered, low resourced or heritage languages (Álvarez Sánchez 2018). Numerous projects in Mexico attempt to conserve the richness of pre-Hispanic language families such as Nahuatl, Zapotec and Maya (Gutiérrez Vasques 2018 and Vocabulario En Lengua Zapoteca 2015). Initiatives such as the CLARIN Knowledge-Centre for linguistic diversity and language documentation offer expertise on data, methods and tools which facilitate a wealth of digital research opportunities (CLARIN K-Centre CKLD. 2017). Nevertheless, many of the world's most endangered languages are orally based or do not have a meaningful written trajectory, a situation which poses a particular challenge for the text-biased toolset of DH. And even for larger languages with a strong textual tradition, the DH tools which do exist often have limited language support and are based fundamentally on European language paradigms.

Where initiatives with a "languages" dimension do exist, moreover, they predominantly focus on how "languages" will be transformed by "digital," often implying that languages are subject to inevitable and unidirectional "digital disruption." We would argue that there is a pressing need to explore the relationship between "digital" and "languages" from the opposite perspective, namely, to gain a better understanding of how digital research projects are shaped (and disrupted)

by their specific linguistic and cultural contexts. So, for example, achieving a fuller awareness of the role of languages in DH involves developing a deeper and more nuanced understanding of how linguistic/cultural diversity challenges the digital research ecosystems we create and use.

We are far from a mature understanding of language variation in digital research environments, but a brief survey will demonstrate the range of factors that influence this variation, which include linguistic, cultural, technical and academic aspects. From a practical perspective for example, Thomas Mullaney contrasts the "vibrant" digitally mediated environments available to even non-expert users in Western European/ American contexts, who "can download off-the-shelf analytical platforms and data corpora, and venture into new and cutting-edge research questions," with the "context of Asian Studies, [where] we find an environment in which many of the most basic elements of DH research remain underdeveloped or non-existent" (Mullaney 2016).

As noted earlier, there is a broad range of research on multilingualism online in the fields of applied and socio-linguistics, and this provides some useful general social and linguistic context to discussions of digital multilingualism. It does not, however, typically address the kinds of methods, critical infrastructure and content underpinning digital humanities research, and DH would benefit from similar studies examining such question as: "What influences cultural preferences for specific tools and ecosystems in DH?," or "How do differences in formal and informal media channel usage in different locales affect what language DH researchers use in particular situations?"

Turning our gaze to the way that language-focused academic fields have interacted with DH, the form of digital scholarship in a given area is also firmly shaped by the history, culture and epistemological assumptions of the field itself and its predominant subjects of research. To give a few examples (in an Anglophone context), we see: a certain emphasis on visual arts and social contestation in digitally mediated Latin American studies; a tradition of cultural heritage and historical databases in Chinese studies; an interartistic and intermedial focus in Italian studies; and a focus on political, social and cultural aspects of digital media in Asian Studies. That is not to say that these are the only manifestations of each digital + language/culture pairing, but differences exist, and to some extent, the history and current institutional location of a given field (e.g. as "modern languages" or "areas studies" / within the humanities or social sciences) have a marked effect on its identity and composition.

Providing a full account of the challenges in fairly representing even the world's most widely spoken languages is well beyond the scope of this article, but we need contributions on the scale and depth of Danet and Herring's edited volume on "The Multilingual Internet," which aimed to chart the state of language, culture and

communication online in 2007 (Danet and Herring 2007), or the Net.Lang publication "Towards the Multilingual Cyberspace," which surveyed multilingual technology, digital spaces, inclusivity and internet governance in 2012 (Maaya Network 2012). A thorough review of the current state of "Multilingual DH" would do much to advance the agenda of linguistic sensitivity and diversity in digital research. Such contributions might include case studies by language or language family, landscape studies of the underlying digital infrastructure which shapes our communication practices, evaluation of the "multilingual readiness" of DH tools, best practices for multilingual DH platform design, analysis of DH multilingual data dynamics or competitions to design linguistically inclusive solutions to language challenges.

Having considered some of the underlying dynamics of languages and language disciplines in digital ecosystems, we now suggest practical ways in which we might foster linguistic diversity within DH's research agenda and consider how these might reshape DH's research practices.

4.2. Articulating a framework for "languages-sensitive" DH research

How might we articulate a wider conceptual framework to effectively address linguistic sensitivity and diversity in the digital humanities? Formal theoretical or practical work in this area is relatively scarce, but the fields of language documentation and minority language studies offer wider frameworks which can serve as useful starting points. The META-NET ("Multilingual Europe Technology Alliance") network of excellence has carried out extensive surveys of languages technologies in a European context, in order to foster "the technological foundations of a multilingual European information society" (Multilingual Europe Technology Alliance 2012). Its 32 volumes in a Language White Paper series examine the digital-readiness of European languages, whether widely spoken or not (META-NET White Paper Series by META Multilingual Europe Technology Alliance 2012), and a cross-language comparison is made of the degrees of support for Machine Translation, Speech Processing, Text Analysis and Speech & Text Resources across different European languages. While the landscape has inevitably shifted to some extent since the white papers were published in 2012, the series summary confirms that a small number of languages, and in particular English, enjoy a high degree of overall digital support, while most languages included in the survey are in a much weaker position, a finding which is unlikely to have changed much.

Acquiring even indicative data about language usage in digital contexts is notoriously challenging (Pimienta 2017) but a languages-aware approach to digital research would clearly benefit from a better understanding of the different dynamics which affect language usage, and this ideally needs a common set of headings with

which to examine them. Pimienta's analysis of languages and cultures on the Internet, based on earlier work by Daniel Prado and others, contemplates six key indicators (internet users, content, internet usage, traffic, interfaces/translation availability and information society indexes) in order to guide four over-arching macro-indicators created to capture the status of languages online (Pimienta 2017). Another initiative, the "Digital Language Diversity Project," aims to "advance the sustainability of Europe's regional and minority languages in the digital world" by providing a set of analysis and training tools, complemented by recommendations, a "Digital Language Survival Kit" and a "roadmap to digital language diversity" aimed at policy makers and other stakeholders (The Digital Language Diversity Project 2019). Recommendations are grouped under three headings: *Digital Capacity* (Digital Literacy; Character Encoding, Input and Output Methods; Availability of Language Resources); *Digital Presence and Use* (Use for E-Communication; Use on Social Media; Availability of Internet Media; Wikipedia) and; *Digital Performance* (Availability of Internet Services; Localised Social Network; Localised Software: Operating Systems and Basic Software; Machine Translation Services; Dedicated Internet Top-Level Domain). Another interesting case study comes in a report into the digital health of the Basque language (whose health is generally speaking very positive considering the relatively small number of speakers of that language in a European context) that makes twelve recommendations, which include attention to content, technical development plans, localisation of digital media, advocacy and policies towards open knowledge (Consejero Asesor del Euskera, Viceconsejería de Política Lingüística). These examples give us several pointers to thinking about frameworks for monitoring digital linguistic diversity, but a fully DH-focused treatment of the subject would have to draw on wider research carried out by linguistics/language-documentation resources such as the CLARIN Virtual Language Observatory, Linguistic Data Consortium, the Linguistic Data Consortium and the ELRA Catalogue of Language Resources, among others.

What kind of a framework do we need in order to assess linguistic diversity in the digital humanities? These frameworks generally have a much wider scope than the kind of approach envisioned in this article, but researchers studying the digital status of less resourced languages, including the Kurdish language and Gaelic in Scotland, have proposed an approach based on the Basic Language Resource Kit (ELDA – Evaluations and Language Resources Distribution Agency 2011) to frame the "DH Readiness" of less-resourced and minority language communities. This framework includes six basic components: maturity of DH research (in a given target community), status of DH education, digital media status, digital visibility and computability of the language, DH tools and the existence of digitised resources (Hassani, Turajlić, and Taljanović 2019).

This approach still emphasises the "digital readiness" of languages. How might we turn the paradigm on its head and assess the "language-readiness" of DH? How might a wider framework for language diversity in DH look, and what kinds of areas would it need to cover? In our view such a framework would need to be a bottom-up initiative drawing on geographically, linguistically and thematically representative voices, but here we tentatively propose some strategic areas where the DH community could foster linguistic diversity in its research practices:

- **Analysis and monitoring of geolinguistic diversity in DH.** As noted already, in order to properly study global knowledge flows with any serious intent, the digital humanities need to improve their understanding of how this diversity operates across key indicators such as content, discovery mechanisms, tools and community. A set of benchmarking terms would be helpful here, as would periodic studies to review the state of the art. This requires analysis of the social and technical incentives and blockages which operate on the dynamics of linguistic diversity – digital scholarship is subject to different forces in different locales. A deeper analysis of work already carried out in this area would help to facilitate more multilingual awareness in research design and would help DH make a valuable contribution in defining requirements for counter-hegemonic models in digital research ecologies.

- **Promotion of languages-sensitive and multilingual practices in DH research**. While there have been some interesting experiments in multilingualism, DH could do more to address multilingual practices, for example by actively promoting positive models for multilingualism in its scholarly communications or supporting multilingual journals and dissemination practices. It is not uncommon to hear DH researchers cite the preference of some non-native English speakers to communicate in English to reach a wider audience. This is of course fine but does not represent the experience of a high proportion of researchers, who are excluded by such assumptions, and ignores differences in linguistic behaviour according to specific platforms or socio-technical incentives and barriers. The "online world […] is very nearly a monoculture, an echo chamber where the planet's few dominant cultures talk among themselves" argued Perlin in 2014, in an article title "The Internet, where languages go to die?" (Perlin 2014). How can DH challenge that view? More explicit and ambitious promotion of a "languages" agenda will help ensure that DH does not become an unwitting servant to monolingualism and hyper/super central language hegemonies.

- **Bidirectional collaboration with language-focused fields.** The digital humanities offer new paradigms for language-based research design, but DH is currently under-theorised in relation to language-focused research in areas such as modern (foreign) languages, sociolinguistic research on multilingualism, translation studies or language pedagogies. Greater partnership with Modern Languages and other languages-focused researchers/practitioners would help to redress the predominantly unidirectional relationship between DH and languages-based research agendas described earlier. There are many ways to address this, including funder policy (increased focus on languages/linguistic diversity), DH professional association strategy and the research community itself (bottom-up initiatives).

- **Guidance, case studies and training**. We have argued that visibility is a key challenge for languages in DH and digital studies as a whole. The problem is that many researchers do not currently have time or incentives to make the extra effort to be more linguistically inclusive. Guidelines, case studies and (mostly light touch) training for DH researchers would go some way to helping them to get over this "academic cultural" hurdle.

4.3. Multilingualism as DH research

Up to this point, we have considered some of the general conditions required to make the digital humanities more "language sensitive" and linguistically diverse. In this final section, we are going to zone in on two of DH's historic areas of strength in research: critical approaches to infrastructure design and digital methods.

So far, we have suggested that key challenges for DH are (1) to develop greater critical sensitivity to its own multilingual practices and (2) to design frameworks to promote geolinguistic diversity in both its scholarly communications and its research. Here we focus on multilingual DH initiatives which are generating disruptive models to overcome digital monolingualism in DH practice and we point to potential implications for future DH research directions.

4.3.1. Multilingualism and infrastructure in DH

Even with the best intentions, DH research infrastructure strongly favours anglophone research and content at present. The pervasive nature of English language in digital culture often means that "for most Anglophone scholars in primarily-Anglophone countries, the discrepancy between English and every other language is all but invisible," argues Dombrowski (Forthcoming b).

At present, digital humanities methods and infrastructure offer poor support for the languages of the world as a whole, and in the case of Non-Latin Script languages, they sometimes do not function at all. A workshop at the DH2019 conference in Utrecht brought together multilingual DH practitioners focused on research involving non-Latin scripts (NLS) (Towards Multilingualism in Digital Humanities 2019), an area which has seen some notable advances in the DH in recent years (KITAB Project 2019; Ho and De Weerdt 2014) but the implications of which are poorly understood in Anglophone DH as a whole. The workshop addressed challenges in a whole range of areas, including multilingual data curation, Optical Character Recognition (OCR), character/sign recognition, digital research ecosystem design, markup, metadata, data/text mining, Named Entity Recognition (NER) and machine translation. This initiative represents part of an ongoing dialogue between DH researchers and library/cultural heritage sector partners aimed at raising awareness of NLS challenges in the design and maintenance of digital scholarship infrastructure, and together with digital studies in Right-To-Left (RTL) languages and cultures, it represents a sustained attempt to broaden language coverage in advanced DH research.

In part inspired by this 2019 NLS workshop, the recently established Multilingual DH network is another attempt to address the "lack of robust tools for working with non-Latin scripts" and the general bias against languages other than English in digital humanities infrastructure. Representing a community-driven effort to draw together "good practices for working with multi-lingual and multi-script data," the network provides multilingual and language-specific resources, a GitHub project for Multilingual NLP and a forthcoming "living" NLS DH handbook (Multilingual DH 2020).

Projects such as these serve both to bolster the visibility of languages other than English in anglophone digital humanities and to offer practical scaffolding for future multilingual development, but what wider lessons do they bring for DH infrastructure more broadly? Firstly, they highlight the need for more support in challenging linguistic assumptions and better understanding the implications of these for DH methods and infrastructure design. Secondly, they illuminate the potential for far greater collaboration across languages. Mirroring the fragmentation in Language Technology – even at European level where there is substantial political support for this area – more can be done: to sustain DH infrastructure in and across different languages; make it easier to adapt existing DH tools for languages other than English; and to elevate the visibility of non-English tools in global DH settings. Thirdly, they highlight the need for greater linguistic labelling in DH infrastructure. Influenced by the "Bender

Rule" – which proposes that we "state the name of the language that is studied, even if it's English" and which its originator Emily Bender argues is a crucial condition for expanding linguistic coverage (in her case, in the field of Natural Language Processing) (Bender 2019) – Dombrowski has signalled the importance of surfacing our language usage in DH (Dombrowski Forthcoming a), and this is particularly important for infrastructure. What languages does a given infrastructure operate in? What languages will a particular tool be useful for? At present, this information is largely absent from the DH research ecosystem, making language diversity less tractable to DH research design. Finally, they demonstrate the indispensable role that speaker communities can play. In his wider review of the "value proposition" in Digital Language Diversity, Benjamin highlights the importance of carrying out research to understand attitudes and behaviours towards multilingualism in order to foster this diversity, and to ensure greater agency for stakeholders from speaker communities in linguistic infrastructure design. DH would benefit from similar studies exploring the incentives and barriers to diversification of its research infrastructure, which has a marked influence on the shape of its linguistic flows (Benjamin 2016).

4.3.2. Situating DH in multilingual research and debates

What does this all mean for DH research agendas of the future? Alan Liu explores multilingual challenges briefly in broader work focusing on diversity which suggests that DH's "unique, as opposed to follow-on, contribution" to cultural criticism may be "the *techne* of diversity" (Liu 2018). We need "new paradigms" and "platforms" for diversity, he argues, and in an expanded version of the essay he goes on to propose that we replace the "big tent metaphor" for DH with what might be usefully defined as a "diversity stack," a multi-faceted approach to diversity, which without relegating the social and cultural dimensions, offers highly technical answers to multilingual, multimedia, corpora-based, chronotypical and identity-related diversity challenges in DH (Liu 2020). If we take up this call for a new "fused techno-ideological apparatus ... that can do urgently needed work" (Ibid, 135), how might we, then, envisage a more substantial response from DH to debates about linguistic and geocultural deficits and divides? What might be uniquely "DH-like" about such a response?

Liu's treatment of multilingual DH foregrounds the current difficulties in effectively carrying out cross-lingual research, drawing on the work of Lee and Dilley in English/ Latin topic modelling, and Mimno et al.'s "polylingual topic models" and neural-network translation-generated "interlingua" – "machine-generated, emergent, and transitional language forms that are a kind of pure comparatism" – as examples of

the kind of computationally-driven methods which will allow us to productively query mixed-language collections (136–137) and in so doing "mine the mathesis of difference and similarity" (145). Cross-lingual DH research is still relatively uncommon, but there are promising developments in: cross-language corpus-building exercises for literary texts (Distant Reading 2020); cross-lingual approaches to "distant, deep, and close reading" in translated cultural production based on "an exploratory model for collecting, processing, and visualising data from across languages" (King's Digital Lab 2020); and cross-lingual methods to research events at scale through textual and visual evidence (CLEOPATRA 2020).

As Liu says, "the digital humanities need to solve the language problem" (Liu 2018) but we would argue that while advanced computational models can certainly make an important contribution to this challenge, we should not forget that an effective DH response will need to include a combination of social, cultural and technical dynamics, which should be as much driven by language disciplines and professions as computational perspectives. In particular, we would suggest that the digital humanities would benefit from playing a more significant role in the languages-related cultural questions of our era in future. DH as a whole has, for example, been notably mute in its response to the series of cultural battles which operate behind "set-piece" polar oppositions in debates about topics such as human versus machine translation, or human languages versus computing languages in the linguistic-cultural sphere. These debates are too often driven by the simplistic discourse of inevitable technological disruption, which attempts to paper over complexities such as uneven digital language support, orality and the cultural perspective, and DH is potentially well placed to offer alternative multilingual reorientations for digital architectures, methods and content.

In "Other worlds, other DHs: Notes towards a DH accent," Roopika Risam has identified one crucial challenge as being the recognition of "both local specificity and global coherence in DH" (Risam 2017, 378) a theme she pursues further in the book "New Digital Worlds" (Risam 2018). In that book, Risam uses postcolonial perspectives to examine code studies, interface design and content management and calls on us to resist the logic/dynamics of digital universalism, a critique Priani Saisó applies to defending the importance of regional epistemologies within (in this case Latin American) DH practice (Priani Saisó 2019). One important manifestation of this community-informed practice is in locally-situated Human–Computer Interaction (HCI) (Risam 2018, 12) and Escobar Varela has argued that we need to take an emic approach to constructing user experience (UX) in DH research through user interfaces (UI) which embed the rhetorical, gestural and visual conventions of given cultural

communities, particularly in the case of under-represented cultural subjects. Escobar Varela's question "Can we design emic interfaces for intercultural exchange?" offers a compelling challenge in this regard (Escobar Varela 2020).

DH has an important part to play both in analysing (and influencing) the geolinguistic affordances of new technologies and in shaping the design of computational tools for multilingual (and translingual) content. The need to articulate a specific DH response to these questions was behind our organisation, with Naomi Wells, of the "Disrupting Digital Monolingualism" workshop (June 2020), an event we will report on separately, and we believe that digital multilingualism will be one of the key challenges for the digital humanities in the coming years.

5. Conclusions

In her study on geographical and linguistic diversity in the digital humanities, Galina Russell noted the lack of hard data to support arguments regarding the dominance of anglophone instantiations of DH in its scholarly communication, arguing that this "hampers the possibility of effective benchmarking in order to propose effective solutions" (Galina Russell 2014, 308). At the same time, "languages and cultures" (in the plural) in DH tend to be under-articulated and fragmented across disciplinary perspectives. In this article we have proposed steps towards creating a broad framework for examining geolinguistic diversity in DH, argued for connecting this to discussion around "language sensitivity" in the field — in an approach combining a number of languages-focused fields and approaches —and we have proposed several ways in which DH can engage with contemporary academic, and *public*, debates about the relationship between languages, cultures, communities and technology. This is a broad debate, and so impossible to capture every perspective — we have not, for example, explored multilingual pedagogies in DH, the role of inter-cultural competence or the international classroom — but we have presented the case for a strategy and outlook which we believe will foster valued and meaningful multilingual interactions in the digital humanities.

In this article, we have grounded ongoing debates in DH about its multilingual identity in wider research into digital language diversity. We have identified multiple disciplinary perspectives which we believe come to bear, and we have proposed some general frameworks for thinking about multilingual policy in DH, fostering language sensitivity, building linguistic diversity into its infrastructure and engaging with wider contemporary debates about the relationship between languages and digital culture.

A key challenge for DH, therefore, is both to develop an over-arching critical awareness of its multilingual practices in relation to data, methods, tools and infrastructure, and to craft counter-models which instantiate linguistic and geocultural positionality and inclusivity. We somehow need to replace monolingual assumptions and English as the "default setting" (Dombrowski Forthcoming a), with new models for digital scholarship which are "language-sensitive," by design and from the start.

Appendix

Data on DH associations mentioned in article

The new associations added since 2014 are: Digital Humanities Association of Southern Africa (DHASA); Humanistica, L'association francophone des humanités numériques/digitales (Humanistica); Red de Humanidades Digitales (RedHD); Taiwanese Association for Digital Humanities (TADH). Analysis based on data from https://adho.org/. In the same period Italian, Czech, German language, Nordic and Russian DH associations have joined a European cluster with Associate or Partner status connected to the European Association for Digital Humanities (EADH).

Competing Interests

The authors have no competing interests to declare.

Contributors

Authorial

Authorship is alphabetical after the drafting author and principal technical lead. Author contributions, described using the CASRAI CredIT typology, are as follows:

Paul Spence – King's College London – pjs https://orcid.org/0000-0001-9236-2727

Renata Brandão – King's College London – rfb https://orcid.org/0000-0003-1074-3740

Authors are listed in descending order by significance of contribution. The corresponding author is pjs

Conceptualisation: pjs

Investigation: pjs and rfb

Data Curation: rfb

Writing – Original Draft Preparation: pjs

Writing – Review & Editing: pjs, rfb

Editorial

Section/Copy/Layout editor

Shahina Parvin, The Journal Incubator, University of Lethbridge, Canada

References

Álvarez Sánchez, Adriana. 2018. "Lenguas, patrimonio cultural y educación en proyectos digitales." In *Humanidades Digitales: lengua, texto, patrimonio y datos*, edited by Isabel Galina Russell, Miriam Peña Pimentel, Ernesto Priani Saisó, José Francisco Barrón Tovar, David Domínguez Herbón, and Adriana Álvarez Sánchez, 183–208. Ciudad de México: Bonilla Artigas Editores.

Androutsopoulos, Jannis. 2007. "Bilingualism in the Mass Media and on the Internet." In *Bilingualism: A Social Approach*, edited by Monica Heller, 207–230. Palgrave Advances in Linguistics. London: Palgrave Macmillan UK. DOI: https://doi.org/10.1057/9780230596047_10

Androutsopoulos, Jannis, and **Kasper Juffermans.** 2014. "Digital Language Practices in Superdiversity: Introduction." *Discourse, Context & Media*, Digital language practices in superdiversity, 4–5 (June): 1–6. DOI: https://doi.org/10.1016/j.dcm.2014.08.002

Arriaga, Eduard. 2020. "Expanded Digital Infrastructure in Teaching and Research Afro-Latin American and Hispanic Digital Cultures." Presented at the MLA Convention, Seattle, January 9.

Bak, Thomas, and **Dina Mehmedbegovic.** 2017. "Healthy Linguistic Diet: The Value of Linguistic Diversity and Language Learning across the Lifespan." http://www.meits.org/files/policy_papers/uploads/Bak_Mehmedbegovic_JLSP_2017.pdf.

Bender, Emily. 2019. "The #BenderRule: On Naming the Languages We Study and Why It Matters." *The Gradient* (blog). 15 September. Accessed June 14, 2021. https://thegradient.pub/the-benderrule-on-naming-the-languages-we-study-and-why-it-matters/.

Benjamin, Martin. 2016. "Digital Language Diversity: Seeking the Value Proposition." In 2nd Workshop on Collaboration and Computing for Under-Resourced Languages, Portoroz, 23 May. Slovenia.

Burdett, Charles, Jennifer Burns, Derek Duncan, and **Loredana Polezzi.** 2018. "Transnationalizing Modern Languages: Reframing Language Education." University of Bristol: Policy Bristol. Accessed June 15, 2021. http://www.bristol.ac.uk/media-library/sites/policybristol/PolicyBristol-report-35-Sept18-transnationalizing-modern-languages.pdf.

Cillia, Rudolf de, Hans-Jürgen Krumm, and **Ruth Wodak,** eds. 2001. "Vienna Manifesto 'The Cost of Monolingualism.'" Austrian Academy of Sciences. Accessed June 15, 2021. https://www.oeaw.ac.at/fileadmin/NEWS/2001/manifesto_e.pdf.

CLARIN. 2020. "Language Resource Switchboard." Accessed October 15. https://www.clarin.eu/content/language-resource-switchboard.

CLARIN K-Centre CKLD. 2017. "CKLD: CLARIN Knowledge-Centre for Linguistic Diversity and Language Documentation." Accessed October 15, 2020. https://ckld.uni-koeln.de/.

CLEOPATRA. 2020. "Cross-Lingual Event-Centric Analytics Research Academy. About." Accessed October 16. http://cleopatra-project.eu/index.php/about/.

Consejero Asesor del Euskera, Viceconsejería de Política Lingüística. 2016. "Euskarazko IKTak: gomendioak herri-aginteentzat/Las TIC en euskera: recomendaciones para los poderes públicos." Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia / Servicio Central de Publicaciones del Gobierno Vasco. Vitoria-Gasteiz. 1 September.

Cunliffe, Daniel. 2019. "Minority Languages and Social Media." In *The Palgrave Handbook of Minority Languages and Communities*, edited by Gabrielle Hogan-Brun and Bernadette O'Rourke, 451–480. London, UK: Palgrave Macmillan. DOI: https://doi.org/10.1057/978-1-137-54066-9_18

Dacos, Marin. 2013. "La Stratégie Du Sauna Finlandais." Blogo-numericus. Accessed June 16, 2021. http://bn.hypotheses.org/11138.

Danet, Brenda, and **Susan C. Herring**, eds. 2007. *The Multilingual Internet: Language, Culture, and Communication Online*. Oxford; New York: Oxford University Press.

Del Río Riande, Gimena, Alex Gil, Daniel O'Donnell, and **Élika Ortega.** 2020. "The Translation Toolkit." *The Translation Toolkit.* Accessed October 15. https://go-dh.github.io/translation-toolkit/about/.

Del Río Riande, Gimena, Ivonne Lujano, and **Daniel O'Donnell.** 2020. "Open, Multilingual and Global Scholarly Communication (OMG)." FORCE11. May 12. Accessed October 15, 2021. https://www.force11.org/group/open-multilingual-and-global-scholarly-communication-omg.

Delfim Leão, Marina Angelaki, Andrea Bertino, Suzanne Dumouchel, and **Frédéric Vidal.** 2018. "OPERAS Multilingualism White Paper." *OPERAS White Papers*, July 30. DOI: https://doi.org/10.5281/zenodo.1324026

Digital Modern Languages. 2019a. "About." Accessed October 16, 2020. https://digitalmodernlanguages.wordpress.com/about/.

---. 2019 b. "Digital Modern Languages Blogs." 16 March. Accessed October 15, 2020. https://digitalmodernlanguages.wordpress.com/blog-posts/.

Distant Reading. 2020. "ELTeC – Distant Reading for European Literary History." Accessed October 15. https://www.distant-reading.net/eltec/.

Dombrowski, Quinn. Forthcoming a. "Language Is Not a Default Setting: Countering Digital Humanities' English Problem." *Debates in the Digital Humanities 2021.* Minneapolis, Minnesota: University of Minnesota Press.

---. Forthcoming b. "Non-English DH Means Resilience". *Debates in DH: Pedagogy*. Minneapolis, Minnesota: University of Minnesota Press.

EADH – The European Association for Digital Humanities. 2019. "Project" Accessed October 15, 2020. https://eadh.org/projects.

ELDA – Evaluations and Language Resources Distribution Agency. 2011. "BLARK: Basic LAnguage Resource Kit." Accessed October 15, 2020. http://www.blark.org/.

Endangered Languages Project. 2020. Accessed October 15. http://www.endangeredlanguages.com/about/.

Escobar Varela, Miguel. 2020. "Emic Interfaces: UX Design for Cultural Specificity." Presented at the 2020 Global DH Symposium, April 24. Accessed October 15. https://www.youtube.com/watch?v=tCOVPtKwQG8.

European Language Grid. 2020. "European Language Grid." Accessed October 15. https://www.european-language-grid.eu/.

Fiormonte, Domenico. 2017. "Digital Humanities and the Geopolitics of Knowledge." *Digital Studies/Le Champ Numérique* 7 (1). DOI: https://doi.org/10.16995/dscn.274

Forsdick, Charles. 2017. "Translating Cultures, Translating Research." *Multilingualism: Empowering Individuals, Transforming Societies (MEITS)*. DOI: https://doi.org/10.17863/CAM.16840

---. 2018. "Science and Languages." In *Languages after Brexit: How the UK Speaks to the World*, edited by Michael Kelly, 1st ed., 75–83. Cham: Palgrave Macmillan. DOI: https://doi.org/10.1007/978-3-319-65169-9_7

Galina Russell, Isabel. 2014. "Geographical and Linguistic Diversity in the Digital Humanities." *Literary and Linguistic Computing* 29 (3): 307–316. DOI: https://doi.org/10.1093/llc/fqu005

Gil, Alex. 2014. "The (Digital) Library of Babel." Accessed October 15, 2020. http://elotroalex.com/digital-library-babel/.

Gil, Alex, and **Élika Ortega.** 2016. "Global Outlooks in Digital Humanities: Multilingual Practices and Minimal Computing." In *Doing Digital Humanities: Practice, Training, Research*, edited by Constance Crompton, Richard J. Lane, and Ray Siemens, 22–34. New York, NY: Routledge.

Global Outlook:Digital Humanities. 2020. Accessed October 15. http://www.globaloutlookdh.org/.

Graddol, David. 2004. "The Future of Language." *Science* 303 (5662): 1329–1331. DOI: https://doi.org/10.1126/science.1096546

Gutiérrez Vasques, and **María Ximena.** 2018. "Corpus paralelo español-náhuatl y su uso en las tecnologías del lenguaje humano." In *Humanidades Digitales: lengua, texto, patrimonio y datos*, edited by Isabel Galina Russell, Miriam Peña Pimentel, Ernesto Priani Saisó, José Francisco Barrón Tovar, David Domínguez Herbón, and Adriana Álvarez Sánchez, 157–181. Ciudad de México: Bonilla Artigas Editores.

Hassani, Hossein, Emir Turajlić, and **Kemal Taljanović.** 2019. "Digital Humanities Readiness Assessment Framework: DHuRAF." *ArXiv:1902.06532 [Cs]*. Accessed October 15, 2020. http://arxiv.org/abs/1902.06532.

Ho, Hou Leong Brent, and **Hilde de Weerdt.** 2014. "MARKUS. Text Analysis and Reading Platform." Communication and Empire: Chinese Empires in Comparative Perspective. Accessed June 17, 2020. https://dh.chinese-empires.eu/markus/beta/.

Humanistica. June 16, 2014. "Humanistica: Présentation." Accessed October 15, 2020. http://www.humanisti.ca/presentation/.

Jaffe, Alexandra. 2007. "Discourses of Endangerment: Contexts and Consequences of Essentializing Discourses." In *Discourses of Endangerment: Ideology and Interest in the Defence of Languages*,

edited by Alexandre Duchene and Monica Heller, Illustrated Edition, 57–75. London; New York: Continuum International Publishing Group Ltd.

JISCMail "Digital Modern Languages List." 2019. Accessed October 15, 2020. https://www.jiscmail.ac.uk/cgi-.bin/webadmin?A0=DIGITALMODERNLANGUAGES.

Joshi, Pratik, Christain Barnes, Sebastin Santy, Simran Khanuja, Sanket Shah, Anirudh Srinivasan, Satwik Bhattamishra, Sunayana Sitaram, Monojit Choudhury, and **Kalika Bali.** 2019. "Unsung Challenges of Building and Deploying Language Technologies for Low Resource Language Communities." *ArXiv:1912.03457 [Cs].* Accessed October 15, 2020. http://arxiv.org/abs/1912.03457.

King's Digital Lab. 2020. "DRaL | Distant Reading across Languages." Accessed October 15. https://www.kdl.kcl.ac.uk/our-work/distant-reading/.

KITAB Project. 2019. "Knowledge Information Technology and the Arabic Books." Accessed October 15, 2020. http://kitab-project.org/.

Language Acts & Worldmaking. 2020. "Disrupting Digital Monolingualism Workshop." June 16-17, 2020. Accessed October, 15. https://languageacts.org/digital-mediations/event/disrupting-digital-monolingualism/aims-and-themes/.

Lee, Carmen. 2016. *Multilingualism Online.* 1st edition. London; New York: Routledge.

Lee, Martin, and **Cosima Wagner,** eds. 2019. "Towards Multilingualism In Digital Humanities: Achievements, Failures And Good Practices In DH Projects With Non-Latin Scripts (Workshop)." Multilingual DH. Utrecht July 8. Accessed October 15, 2020. https://multilingualdh.org/en/dh2019/.

Liu, Alan. 2018. "Digital Humanities Diversity as Technical Problem." *Alan Liu blog.* DOI: https://doi.org/doi:10.21972/G21T07

---. 2020. "Toward a Diversity Stack: Digital Humanities and Diversity as Technical Problem." *PMLA* 135 (1): 130–51. DOI: https://doi.org/10.1632/pmla.2020.135.1.130

Liverpool University Press. 2019. "Modern Languages Open: Launch of the Digital Modern Languages Section." *Blog.* November 13, 2019. Accessed October 15, 2020. https://liverpooluniversitypress.blog/2019/11/13/modern-languages-open-launch-of-the-digital-modern-languages-section/.

Maaya Network. 2012. *NET.LANG: Towards the Multilingual Cyberspace.* Caen, France: C & F Editions. http://net-lang.net//externDisplayer/displayExtern/_path_/netlang_EN_pdfedition.pdf.

Maffi, Luisa. 2005. "Linguistic, Cultural, and Biological Diversity." *Annual Review of Anthropology* 34 (1): 599–617. DOI: https://doi.org/10.1146/annurev.anthro.34.081804.120437

META Multilingual Europe Technology Alliance. 2012. "META-NET White Paper Series." 2012. Accessed October 15, 2020. http://www.meta-net.eu/whitepapers/overview.

Mullaney, Thomas S. 2016. "Digital Humanities Q and A with Thomas S. Mullaney." *China Policy Institute: Analysis.* Accessed June 16, 2021. https://web.archive.org/web/20171213084933/https://cpianalysis.org/2016/06/06/digital-humanities-qa-with-tom-mullaney/.

Multilingual DH. 2020. Accessed 15 October. http://multilingualdh.org/.

Multilingual Europe Technology Alliance. 2012. "Mission." Accessed October 15, 2020. http://www.meta-net.eu/mission.

Network for Digital Humanities in Africa. 2019. "Outcomes". *Network for Digital Humanities in Africa* (blog). July 5. Accessed October 15, 2020. https://dhafrica.blog/lorentzworkshop/outcomes/.

Network for Digital Humanities in Africa. 2020a. "Africa Forum at DH2020 Virtual Event." June 10. Accessed October 16. https://dhafrica.blog/africa-forum-at-dh2020/.

Network for Digital Humanities in Africa. 2020b. "Resources & Past Events." July 8. Accessed October 15. https://dhafrica.blog/resources/.

NYU Abu Dhabi Winter Institute in Digital Humanities. 2020. "Schedule." Accessed October 15. https://wp.nyu.edu/widh/schedule/.

Open Methods Languages. 4 August 2017. «OpenMethods.» Accessed October 15, 2020. https://openmethods.dariah.eu/languages/.

Open World Research Initiative (OWRI). 2016. "Arts and Humanities Research Council." Accessed October 15, 2020. https://ahrc.ukri.org/research/fundedthemesandprogrammes/themes/owri/.

Ortega, Élika. 2014. "Whispering/Translating during DH2014: Five Things We Learned." 21 July. Accessed October 15, 2020. https://elikaortega.net/blog/2014/dhwhisperer/.

---. 2018. "MLA19 582 Roundtable. Digital Hispanisms." *Discussion – TC Digital Humanities – MLA Commons* (blog). December 27. Accessed October 15, 2020. https://mla.hcommons.org/groups/digital-humanities/forum/topic/mla19-582-roundtable-digital-hispanisms-2/.

---. 2019. "Zonas de Contacto: A Digital Humanities Ecology of Knowledges." In *Debates in the Digital Humanities 2019*, edited by Matthew K. Gold and Lauren F. Klein, 179–87. Minneapolis, Minnesota: University of Minnesota Press. DOI: https://doi.org/10.5749/j.ctvg251hk.18

Perlin, Ross. 2014. "The Internet, Where Languages Go to Die?" *Aljazeera America* (blog). 18 March. Accessed October 15, 2020. http://america.aljazeera.com/opinions/2014/3/language-linguisticstechnologyinternetdigitaldivideicann.html.

Phipps, Alison, and **Mike Gonzalez.** 2004. *Modern Languages: Learning and Teaching in an Intercultural Field*. London: Sage Publications Ltd. DOI: https://doi.org/10.4135/9781446221419

Pickles, Matt. 2016. "Could the Dominance of English Harm Global Scholarship?" *BBC News*, January 20, sec. Business. Accessed October 15, 2020. https://www.bbc.co.uk/news/business-35282235.

Pimienta, Daniel. 2017. "A New Alternative Approach for Linguistic Indicators in Cyberspace." Accessed October 15, 2020. http://funredes.org/lc2017/.

Pitman, Thea, and **Claire Taylor.** 2017. "Where's the ML in DH? And Where's the DH in ML? The Relationship between Modern Languages and Digital Humanities, and an Argument for a Critical DHML." *Digital Humanities Quarterly,* 11 (1). Accessed June 19, 2021. http://www.digitalhumanities.org/dhq/vol/11/1/000287/000287.html.

Polezzi, Loredana. Forthcoming. "Language Indifference." *Translating Cultures*.

Priani Saisó, Ernesto. 2019. "Codificación y buenas prácticas. Crítica a la delimitación de las humanidades digitales en América Latina." *Relaciones Estudios de Historia y Sociedad* 40 (158): 129–144. DOI: https://doi.org/10.24901/rehs.v40i158.343

Priani Saisó, Ernesto, Paul Spence, Isabel Galina Russell, Elena González-Blanco García, Daniel Alves, José Francisco Barrón Tovar, Marco Antonio Godínez Bustos, and **Maria Clara Paixão de Sousa.** 2014. "Las Humanidades Digitales En Español y Portugués. Un Estudio de Caso: DíaHD/DiaHD." *Anuario Americanista Europeo* 0 (12). Accessed June 19, 2020. https://dialnet.unirioja.es/servlet/articulo?codigo=5071434.

Prys Jones, Meirion. 2013. "Endangered Languages and Linguistic Diversity in the European Union." IP/B/CULT/IC/2013-030. European Parliament. Accessed June 19, 2021. https://www.europarl.europa.eu/RegData/etudes/note/join/2013/495851/IPOL-CULT_NT(2013)495851(SUM01)_EN.pdf.

Risam, Roopika. 2017. "Other Worlds, Other DHs: Notes towards a DH Accent." *Digital Scholarship in the Humanities* 32 (2): 377–84. DOI: https://doi.org/10.1093/llc/fqv063

---. 2018. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Evanston, Illinois: Northwestern University Press.

SADiLaR. 2020. "About SADiLaR." Accessed October 16. https://www.sadilar.org/index.php/en/about.

Spence, Paul. 2015. "Modern Languages Research and the Digital Humanities." *Paul Spence* (blog). July 13. Accessed June 19, 2020. http://www.paulspence.org/mlr-and-dh/.

Steyn, Juan, Tunde Ope-Davies, Felix Ameka, and **Sara Petrollino.** 2020. "Building New Global DH Communities: Africa and Beyond." Accessed October 15. https://dh2020.adho.org/wp-content/uploads/2020/07/509_BuildingnewglobalDHcommunitiesAfricaandbeyond.html.

Swaan, Abram De. 2002. *Words of the World: The Global Language System*. 1st edition. Cambridge, UK; Malden, MA: Polity Press.

The Digital Language Diversity Project. 2019. "Project." Accessed October 15, 2020. http://www.dldp.eu/en/content/project.

The Network to Promote Linguistic Diversity (NPLD). 2015. "The European Roadmap for Linguistic Diversity." https://www.npld.eu/wp-content/uploads/2018/11/Roadmap-in-English.pdf. Accessed October 15, 2020.

Thieberger, Nick. 2017. "What Remains to Be Done—Exposing Invisible Collections in the Other 7,000 Languages and Why It Is a DH Enterprise." *Digital Scholarship in the Humanities* 32 (2): 423–34. DOI: https://doi.org/10.1093/llc/fqw006

Towards Multilingualism in Digital Humanities. 2019 Workshop (Zenodo). Zenodo. Accessed October 15, 2020. https://zenodo.org/communities/multilingual-dh/?page=1&size=20.

Vocabulario En Lengua Zapoteca. 2015. "Diccionario Zapoteco-Español." Accessed October 15, 2020. http://www.iifilologicas.unam.mx/cordova/acerca.php.

Wallraff, Barbara. 2000. "Interview With David Graddol." *The Atlantic*. November 1. Accessed June 19, 2021. https://www.theatlantic.com/magazine/archive/2000/11/interview-with-david-graddol/378426/.