



Tilburg Papers in Culture Studies

Paper 104

Digital ethnography

by

Piia Varis ©

p.k.varis@tilburguniversity.edu

© August 2014

DIGITAL ETHNOGRAPHY

Piia Varis

ABSTRACT

This chapter outlines digital ethnography as an approach to studying digital communication. Ethnography, as a holistic approach to societies and cultures, can make a substantial contribution to the study of present-day digital communication environments and our digital culture(s), and there is indeed already a burgeoning literature on ethnographically oriented research on digital cultures and communication.

The chapter builds on the fundamental idea that ethnography is not a method, but an approach, and will start with a brief review of the ‘pre-digital’ ethnographic tradition and its basic tenets. Next, critical issues for ethnographic research on digital environments will be discussed; socio-technological developments such as the recent rise of social media and increasing ‘de-computerisation’ (i.e. the increased use of smart phones and other mobile devices for accessing the internet) have raised new issues to be tackled by ethnographers. This entails addressing important notions such as ‘participant observation’ or engagement, and ‘lurking’ (i.e. ‘invisible observation’), as well as related ethical implications. After a discussion of recent and current research on digital communication, methodological choices in doing ethnographic research – which is essentially methodologically flexible – will be addressed.

Finally, the chapter will position digital ethnography as an approach in the broader context of recent discussions within internet studies, more specifically on ‘big data’ methods which have become increasingly popular partly because online data can be easily collected in large amounts. The chapter will conclude with critical reflections on the ‘big data’ discussion and argue for the value of ethnography in future research on digital environments.

INTRODUCTION

Ethnographic research on online practices and communications, and on offline practices shaped by digitalisation, has become increasingly popular in the recent years with the growing influence and presence of the internet in people's everyday lives. This research takes a myriad of forms, appearing within different disciplines and under several different labels such as 'digital ethnography' (Murthy 2008), 'virtual ethnography' (Hine 2000), 'cyberethnography' (Robinson & Schulz 2009), 'discourse-centred online ethnography' (Androutopoulos 2008), 'internet ethnography' (boyd 2008; Sade-Beck 2004), 'ethnography on the internet' (Beaulieu 2004), 'ethnography of virtual spaces' (Burrell 2009), 'ethnographic research on the internet' (Garcia et al. 2009), 'internet-related ethnography' (Postill and Pink 2012) and 'netnography' (Kozinets 2009).

The common denominator for these studies is that they all include some kind of online data, and they all employ (a particular version or understanding of) ethnography in the research process. This is basically where the commonalities end; so diverse is the field – if such a field can even clearly be identified – of ethnographic research on digital culture and practices. This is not least because of the various types of data and environments covered in research on digital communication – social network sites, blogs, forums, gaming environments, websites, dating sites, wikis etc. – but also due to seemingly different understandings of what exactly 'ethnography' is, ranging from limiting it to specific techniques or data collection methods (mainly observation and interviews) to seeing it as an approach rather than a set of techniques. This chapter builds on this latter understanding of ethnography; that is, ethnography is not reduced to the employment of certain techniques, but seen as an approach to studying (digital) culture with specific epistemological claims (see e.g. Blommaert & Dong 2009).

Digitalisation and online communications provide researchers with unprecedented opportunities for accessing and examining people's communicative repertoires – the complexities of the 'global', the 'local' and the 'translocal', and the ways in which people make (globally) circulating semiotic materials part of their own communicative repertoires, can all be traced online (see e.g. Georgakopoulou 2013; Leppänen et al. 2013; Rymes 2012; Varis & Wang 2011). Digital ethnography is one approach for capturing the shape and nature of such communicative practices; this chapter discusses its main principles, current

applications and future directions, building on a brief review of ‘traditional’ or ‘pre-digital’ ethnography.

HISTORICAL PERSPECTIVES

Research on technologically mediated communication has come afar from what Androutsopoulos (2008: 1; 2006) identified as the ‘first wave’, where “(...) the focus was on features and strategies that are (assumed to be) specific to new media; the effects of communications technologies on language were given priority over other contextual factors.” In this early research, as Androutsopoulos (ibid.) points out, “The data were often randomly collected and detached from their discursive and social contexts, and generalisations were organised around media-related distinctions such as language of emails, newsgroups, etc.”. The focus was thus on reified end products, texts and pieces of language, rather than production and uptake of discourse as socially meaningful, context-specific activity.

Broadly, then, the difference between the earlier and later research is a difference between the study of ‘things’ and the study of contextualised ‘actions’ – or the study of texts, and cultural practices (cf. Shifman’s 2011 study on YouTube memes for such a conclusion; also Hine 2000). This largely corresponds with the two phases in social research on technologically mediated communication identified by Hine (2013 [2005]: 7): the first one was characterised by experimental research, and the latter one by “(...) growing application of naturalistic approaches to online phenomena and the subsequent claiming of the Internet as a cultural context”, with ethnographic research increasingly applied. Indeed, Hine (ibid.: 8) suggests that “(...) our knowledge of the Internet as a cultural context is intrinsically tied up with the application of ethnography.”

Digital ethnography as an approach of course builds on ‘pre-digital’ ethnography. Ethnography, with its roots in anthropology, takes as its object of interest the very lived reality of people, of which it aims to produce detailed and situated accounts – in the words of Geertz (1973), ‘thick descriptions’. As such, ethnography is the approach of choice against generalisation and narrow assumptions regarding the universality of digital experience in general (Coleman 2010), or, in terms of language use, against the kinds of sweeping statements on for instance ‘the language of emails’ produced in the first wave of research on technologically mediated communication. Seen in the perspective of long-term developments

on the study of technologically mediated communication, research under the umbrella ‘digital ethnography’ has deepened our understanding of locally specific digital practices. Using the internet, and using language and other semiotic means in doing so, are locally situated experiences and entail locally specific practices, platforms and semiotisations, and ethnography has precisely the means of capturing this, taking the task of understanding informants’ life-worlds and their situated practices and lived local realities. To this end, ethnographic fieldwork is essentially a *learning process* where research is guided by experience gathered in the field; it is a mode of discovery and learning (Blommaert & Dong 2009: Velghe 2011) – as Dell Hymes (1996: 13) put it,

It [ethnography] is continuous with ordinary life. Much of what we seek to find out in ethnography is knowledge that others already have. Our ability to learn ethnographically is an extension of what every human being must do, that is, learn the meanings, norms, patterns of a way of life.

While ethnography assumes such a holistic position, the approach, digital and otherwise, is often reduced to specific methodologies and procedures (fieldwork with participant observation and interviews in many cases as *the methods*). As Blommaert and Dong (2009) point out, however, ethnography is not only a complex of fieldwork techniques. It has its origins in anthropology, and “These anthropological roots provide a specific direction to ethnography, one that situates language deeply and inextricably in social life and offers a particular and distinct ontology and epistemology to ethnography” (ibid.). From an ethnographic perspective, studying language means studying society and larger-scale socio-cultural processes, and making a distinction between the linguistic and the non-linguistic is seen as a fundamentally artificial one.

Digitalisation has offered scholars of language and communication, also ethnographers, with opportunities to easily collect, store and sort (e.g. by ‘tagging’ contents in electronic databases) ‘logs’ of interaction, i.e. “characters, words, utterances, messages, exchanges, threads, archives, etc.” (Herring 2004; see also Androutsopoulos 2008). While online environments provide opportunities for easy collection of huge amounts of data, from an ethnographic perspective this becomes problematic if the material is taken out of its context – a ‘log’ of communication only serves as ethnographic data if it is understood in its context. This is where more recent, ethnographically informed research on digital communications

dramatically differs from the early studies on technologically mediated communication and their de-contextual analyses. Ethnographically speaking, context is an interactional achievement, and contexts should be investigated rather than assumed (e.g. Blommaert 2007). This is perhaps particularly important in today's complex world of globalisation, translocal communication environments and complex online-offline dynamics, where pre-digital presuppositions regarding contexts and communicators are often invalid.

CRITICAL ISSUES AND TOPICS

Context and contextualisation are indeed a critical issue in digital ethnography, not least due to the fact that in today's complex world, we are increasingly encountering polycentric environments in which little, if anything, can be taken for granted (see e.g. Blommaert & Rampton 2011). Also, compared to 'traditional' ethnography, there are new types of issues related to contextualisation that ethnographers of digital culture and communication need to address.

boyd (2008)'s ethnographic study established certain technical properties – persistence (semiotic material online is automatically recorded and archived), searchability (semiotic material can be accessed through search), replicability (digital content, made of bits, can be duplicated) and scalability (the potential spread and visibility of semiotic material is great) – that shape interactions in online networks. It is particularly the last two characteristics of digital communication, replicability and scalability, that bear upon the ethnographically important notion of indexicality – links between signs and the macro-level of socio-cultural contexts and meanings (see e.g. Silverstein 2003) – and contextualisation in online communication. Thanks to the technical properties of replicability and scalability, linguistic and more broadly semiotic material is quickly and easily mobilised, recontextualised and resemiotised (see e.g. Georgakopoulou 2013; Leppänen et al. 2013; Rymes 2012), making often for complex and unpredictable uses, reuses, trajectories and uptake.

Contextualisation of digital communication is also shaped by other kinds of digital affordances, many of which platform-specific ones. Marwick and boyd (2010) discuss the notion of 'context collapse' to refer to the idea that, in networked online environments such as social network sites, people's networks potentially include people from different spheres

of life (family, friends, co-workers, people one has only met online, people one has not been in offline contact with for years etc.). In such conditions, the uptake for which communications have been designed may not be clear or transparent at all. Such contextual complexities potentially shape people's communicative practices and need to be ethnographically established.

We may also mistakenly reduce the context of digital communications to seemingly self-evident abstractions such as 'Facebook', while what 'Facebook' means for people is by no means a consistent or static thing, but is a media ideological construct shaped by, among other things, the way in which users view this medium in relation to other media (see 'Current contributions and research' below for a discussion of these notions). Hence, the online environments studied cannot be taken as self-explanatory contexts, but need to be investigated for locally specific meanings and appropriations. Further, Facebook for instance as 'context' only makes sense if we see its features as essentially linked to the commercialisation of the Web and the way in which the shape of commercial platforms such as Facebook or YouTube, amongst others, influences semiotic activity. While the shape of any platform does not *determine* the way in which people will use it for their communicative purposes, the design of the site will influence interactions. "Defaults" in digital environments, as van Dijck (2013: 32) points out, "are not just technical but also ideological maneuverings (...) Algorithms, protocols, and defaults profoundly shape the cultural experiences of people active on social media platforms." These coded structures, she (ibid.: 20) maintains, "are profoundly altering the nature of our connections, creations, and interaction. Buttons that impose 'sharing' and 'following' as social values have effects in cultural practices." This is yet a further contextual layer that digital ethnographies of communication need to investigate.

Online-offline dynamics is another, contextually important issue for ethnographers of digital communication. It is becoming increasingly difficult, if not impossible, to make clear-cut distinctions between what is 'online' and what is 'offline', especially with the recent 'de-computerisation', i.e. the emergence and increasing popularity of mobile technologies (smartphones, tablets) with internet access. Understandings of space and place – and indeed, understandings of what constitutes ethnographic 'field' – have also been complicated by the fact that mobility is increasingly not confined to physical movement: rather, what we often see is akin to what Raymond Williams (1974) described as 'mobile privatisation' – a development where subjects are increasingly mobile, yet 'private', i.e. (socially) mobile in a

cocoon, in a way. The example of the car, for instance, has been used to describe this seeming paradox: while the car can take its driver anywhere, the driver is self-sealed, isolated in a private space. The same goes for internet use, and also the smartphone and other mobile devices – these are vehicles of mobility, yet their users do not necessarily have to physically go anywhere in order to be mobile, to be ‘elsewhere’, or to experience a change of context without physically changing context.

Indeed, a lot of the activity that we now see as taking place simply ‘online’ is, thanks to mobile technologies, linked to and influenced by all kinds of offline environments, situations and practices: people do not only sit at home at their desk PC – they produce what we see as our ‘online’ data in trains, shops, bathrooms, airports, classrooms, restaurants, cars, meeting rooms, concerts and conferences; there are tweets sent from toilet seats and selfies posted from shopping centres. The ‘finished’ communicative products that researchers collect online can thus be shaped not only by the immediately observable *online* context, but also by the *offline* context in which the digital activity has taken place. This may introduce a further normative layer on communication, depending on what kind of digital communication is expected and ‘acceptable’ in a specific physical, offline context. Such normative understandings are visible in normative public and lay discourses regarding digital activities – debates on whether it is, for instance, ‘acceptable’ to post selfies at funerals, or to be browsing and updating on social media while having an ‘offline’, face-to-face conversation. Similarly, broader socio-cultural issues such as internet censorship can heavily influence communications at least in certain parts of the world (e.g. Varis & Wang 2011). The contexts for online activity, thus, are in fact layered and polycentric, and it may be necessary to attend to further layers of context than what is visible on the screen.

The other two technical properties shaping online interactions established by boyd (2008) – persistence and searchability – are also crucial both from the perspective of those communicating online, and ethnographers investigating such communications. To begin with, the fact that online materials can be easily traced and located has implications for how digital ethnographers present their analyses and refer to their informants. Anonymising data is of course one familiar step in ethnographic research in order to protect people’s privacy, and this also goes for usernames and avatars; while it might be argued that these are not people’s ‘real names’, such online means of self-representation should rather be seen as very real. The notion that these are not ‘real names’ seems to be based on an ideological understanding of

the internet as ‘less real’ than the offline world: usernames and avatars are very real to the people who use them to present themselves; they come with – often very well established and well-known – online identities; reputations are built on them, and recognising them, if people use the same identifiers for themselves across different contexts and platforms, potentially gives access to intricate worlds of online activity. When necessary, these ‘not real’ names should also then be protected.

The searchable nature of digital environments – or the issue of ‘googlability’ – poses problems perhaps in particular to ethnographers whose aim is to give justice to people’s own voices and present them ‘in their own words’. Even if data is anonymised and people’s names are changed, discourse is still searchable: while some platforms and sites are more easily trawled than others, and search engines do not reach every corner of the internet, there is still the possibility that simply by entering the online material quoted in a search engine, the data is easily connectable to accounts, usernames and activities. This is of course particularly alarming in cases where extremely sensitive material is being addressed, and people may be put at serious risk. It is the responsibility of ethnographers to see that they do not, for instance, jeopardise political activists in contexts where revealing their actions – or making it easier to establish their offline identities – might put them in danger, or that they are not inadvertently ‘outing’ people with stigmatised sexualities. Difficult compromises may have to be made, such as sacrificing ethnographic detail and accuracy in the reporting, but this does in no way prevent ethnographers from researching such ‘sensitive’ issues and environments.

A closely related issue is that of the differing understandings of what is ‘public’ and what ‘private’ and – still, regardless of extensive, global public debate on privacy and surveillance – lack of awareness of what is public and what private online, and how persistent online communications are. These broader concerns and individual understandings of what is ‘public’ and what ‘private’ will have an influence on the kind of interactions that are visible for researchers online – what people choose to make public about themselves – but also on ethical considerations. That is, while semiotic material may be publically available, this does not necessarily automatically mean that it can be used for research purposes, or that people behind the semiotic production accept that what they have entered online will become data unbeknownst to them. The classic example here is the blogger who sees their online writing as a *private* diary, not to be read by anybody else (not to mention to be used for research

purposes). This is just one illustration, and a rather extreme one at that, of the differing understandings of ‘public’ and ‘private’ online. While there is no consensus on ethical considerations or strict guidelines for ethnographers in this respect (apart from the ethics guidelines by the Association of Internet Researchers; see ethics.aoir.org), we can at least be cautioned against seeing the Web as one big, public database readily and voluntarily produced by ordinary internet users. Case-by-case considerations will have to be made, depending on, for instance, type of platform, the sensitivity of the issue investigated and possible harm caused to those whose communications are being studied.

With ethnography, the kinds of broader macro-contexts discussed here become part of the investigation; while it is the micro-level that often gains most of the attention from linguistic and discursive interrogations, digital ethnography maintains that the micro-level only makes sense when seen within the macro-level. Such contexts should be interrogated, and in the future studies addressing more complex data sets (both online and offline) will probably appear: that would be justifiable considering the way in which online and offline are in many cases inseparable.

CURRENT CONTRIBUTIONS AND RESEARCH

The fact that ‘online’ and ‘offline’ are difficult, even pointless, to disentangle is visible in the most recent linguistic ethnographic research. In these studies, social media and other online environments are not seen as separate contexts, detached from other spheres of life, and digital communication practices are seen in the wider sociolinguistic context (see e.g. Madsen and Stæhr 2014; Stæhr 2014; Stæhr in press). Such research is multi-sited and employs accordingly a number of different methods, adding nuance not only to our understanding of digital communication practices per se, but also their specific functions in people’s communications more generally, as well as broader online-offline dynamics.

Indeed, current research illustrates that there is no need to ‘exocitise’ online data as particularly difficult to analyse or manage (Georgakopoulou 2013), nor to make the study of digital communication an end in itself – especially if we take it as our goal to understand and explain people’s life-worlds and communicative practices comprehensively, and not just taking ‘digital slices’ of them. While early research on technologically mediated

communication focused on the de-contextual log data collected online, in today's ethnographic research some research trajectories involving digital data begin online and others offline, guided by whatever becomes relevant in the field. Current research is thus more 'realistic' in the sense that it is focused on forming accurate pictures of sociolinguistic repertoires and contexts, and digital data is an organic part of this rather than an end in itself, and the internet not a separate sphere of life with no connections to the offline world.

One surprisingly little explored concept in current research on digital communications is the relatively recently introduced notion of *media ideologies*. Ilana Gershon (2010: 3), drawing on Silverstein's (1979) notion of 'language ideology', defines media ideologies as "(...) a set of beliefs about communicative technologies with which users and designers explain perceived media structure and meaning. That is to say, what people think about the media they use will shape the way they use media." Further, Gershon (ibid.: 5) makes use of the notion of 'remediation', referring to the fact that people define each technology in relation to the other technologies available to them. This means that people make communicative decisions based on what they deem the most 'appropriate' medium for the specific communicative task at hand. This can include all kinds of considerations, ranging from aesthetic ones to the perceived effectiveness or quickness of a medium in offering communications for uptake. Finally, Gershon's (ibid.: 6) highly illuminating study on breaking up through (social) media introduces the concept of 'idiom of practice' to highlight the idea that "(...) people have implicit and explicit intuitions about using different technologies that they have developed with their friends, family members, and coworkers." That is, a group of people with a shared understanding of the use of a specific technology will use it to communicate in a particular way. These notions help explain for instance family debates on media use where teenagers can see their parents' phone calls as 'embarrassing' and 'disruptive', and prefer text messages or chat software – which they would use with their peers as the default mode of communication – as a preferable means of interaction. Similarly, issues related to 'formality' or the 'standardness' of language used in communications can partly be explained with these concepts – they have to do with differing understandings as to what is 'appropriate' or 'good' language to be used in which medium.

Thus far, Gershon's useful concepts remain largely without applications, while they seem to possess plenty of explanatory power for understanding people's digital communication practices. Indeed, in order to explain people's linguistic and discursive choices in digital

environments, attending to *both* language ideologies and media ideologies would perhaps provide powerful explanations as to what people do and why they do it. Attending to people's media ideologies will also help making connections to broader socio-cultural issues such as the discussions on privacy referred to above, and hence help explain choice of media for communication and their specific functions.

Digitalisation continues to have profound effects on people's everyday lives and communicative practices, and while digital divides persist – with many lacking either internet or device access, or both – it can also be suggested that it is not only the lives of those who are heavily 'wired' that are shaped by the recent developments; also the lives of people with no or a lesser degree of digital engagement are influenced by the very absence of these tools for communication. Consequently, while appropriations of new communication technologies can provide exciting data on practices and interactions, (ethnographic) research should also not ignore the new types of 'have-nots' appearing as a result of digitalisation. Velghe (2011a, 2011b; also Blommaert & Velghe 2012), for instance, presents an interesting and informative ethnographic case of resource-scarce digital communication in a South African township. Inquiries into digitally deprived contexts also help shed light on digital communications in a broader sense, and studying local appropriations of technologies adds nuance to how we understand specific digital communication platforms and help move away from assumptions of universal digital experience: what 'blogging', 'YouTubing' or 'social networking', for instance, means in each case and particular context is a matter of (ethnographic) investigation, not assumption.

MAIN RESEARCH METHODS

As mentioned above, this chapter understands ethnography not as a set of field methods, but as an approach, and hence not reducible to specific techniques. Even if ethnography was reduced to techniques, when it comes to studying 'the internet' or digital environments, it would be extremely difficult to outline a simple set of techniques to follow: there is no 'one size fits all' solution, not least due to the myriad of different communicative environments that digital communication encompasses. Also, ethnography as an approach in any case is methodologically flexible and adaptive: it does not confine itself to following specific procedures, but rather remains open to issues arising from the field. With digitalisation and

the attendant new types of communicative environments, debates have emerged as to how exactly study these new forms of interaction, and whether ‘pre-digital’ methodologies and approaches can be successfully applied in research on them. However, rather than discussing which ‘offline’ methodologies could be successfully applied to researching online environments and how, the questions raised by the study of digital environments could be used to reflect more broadly on methodologies and their epistemological nature. As Hine (2013 [2005]: 9) points out, in line with the ethnographic commitment to reflexivity, “The question is much more interesting, potentially, than whether old methods can be adapted to fit new technologies. New technologies might, rather, provide an opportunity for interrogating and understanding our methodological commitments. In the moments of innovation and anxiety which surround the research methods there are opportunities for reflexivity.”

Participant observation, traditionally one of the ethnographic staples for understanding local practices and meaning-making, is one example of a technique which has often featured in such ‘moments of innovation and anxiety’. The reasons for this are manifold, yet all have to do with the fact that the study of digital communications always involves the screen in one way or another. One of the issues has to do with the fact that researchers can now lurk – ‘participate’ invisibly and unbeknownst to the people whose activities are being observed – while being entirely immersed in the environment and activities in question; it is as if the ethnographic ‘fly on the wall’ was now wearing an invisibility cloak. Arguments around the invisibility issue range from suggesting that such lurking is not ethnographic observation in the traditional sense (hence ‘participate’ in scare quotes above) and hence not ‘proper’ ethnography, to idealising the situation by claiming that the invisibility guaranteed by the screen between the researcher and the researched presents a unique opportunity for collecting ‘natural’ data, as the informants are not aware of their informant status and hence do not modify their behaviour accordingly. The latter point of course raises all kinds of ethical questions, a broader discussion on which is beyond the scope of this chapter (see ‘Critical issues and topics’ above).

As regards the issue of ‘participation’, a further point of debate is the fact that online communication is easily collectable, printable and screenshotable – entire histories of activity can be made into ‘data’ with a couple of clicks without ever having witnessed the interactions while they actually unfolded. This raises the question as to whether the ethnographer should always be ‘there’ to observe interactions as they take place, so as to be ‘immersed’ in the

situation and directly experience and witness the interaction as it unfolds – with the lags in communication, the editing and deletion of posts and messages, the floods of commentary in discussion forums and blogs etc. – *in real time*. This would mean in many cases that the poor ethnographer would be able to get very little sleep: with translocal digital communications, one's archive of materials grows potentially non-stop, 24 hours per day. Investigator triangulation is of course an option, with shifts in observation, but this is naturally not always feasible. This means that in many cases the ethnographer will have to do with *products* rather than *processes* and, modifiable and editable as digital artifacts are, what remains visible is the end result of possibly countless edits, changes and deletions. Some platforms give ethnographers useful research assistance in this respect – for instance Facebook shows which posts have been edited and which have not (and also gives a separate stamp for mobile posts), and time stamps on different kinds of online platforms give indications as to how interactions have unfolded in time, giving possible cues also for further inquiries. Such digital traces potentially give valuable information and leads to follow for the ethnographer who will also be interested in the processes of semiotisation and meaning-making, not only the final product.

However, whatever methods will be applied, one thing remains constant in studying digital environments – there is always the issue of the screen. Being able to read and watch on the screen can of course be seen as one of the advantages of digital ethnography: field sites are accessible and data available for the ethnographer potentially all the time, even on the go, if smartphones or other mobile devices are used. What can be viewed as downsides include that we indeed only see what is on the screen. In the case of multi-functional platforms such as Facebook where multiple channels of communication are available, what is observable on the screen can be misleading, or at least only provide a partial image. For instance, while somebody may seem like an inactive or infrequent Facebook user based on observations of their profile, they may at the same time be actively sending private messages and chatting with their connections; or, what may seem like status updates without any reaction or commentary from other users may in fact be the object of heated chatting or private messaging. While a full ethnography of such a multi-channelled site can be admittedly difficult to accomplish, in any case the (semi-)public profile is only *part* of the whole experience of using such a site. What is visible on the screen is thus only part of the story.

Another issue related to the screen is that we do not, for instance, know whether the identities people establish or what they present about themselves are biographically or demographically ‘accurate’ information in the sense that they correspond with what is on the other side of the screen. The ‘authenticity’ of information is of course a broader issue regarding online materials: it is sometimes difficult, if not impossible, to establish whether information is accurate or not, and the argument could be made that, while there is no way of knowing whether something is ‘real’ or not, such material would not be useful as data. The willingness to deem something as not worthy of research simply because it does not correspond to (assigned) ‘real’, offline identities is of course in itself an interesting phenomenon, and there often seems to be an ideological understanding of the ‘offline’ somehow being primary over what is ‘online’ in terms of ‘real’ selves. An important point here is that identities and self-representation are *contextual*: they appear with a specific function and uptake in mind. This also goes for ‘false’ profiles – these serve a purpose for their creators, and should not be automatically dismissed as uninteresting. However, from the perspective of research which does not go beyond the screen this of course poses a problem if demographic or biographic accuracy (e.g. in terms of physical location, age or gender) is taken for granted in drawing for instance sociolinguistic conclusions which rely on demographic correspondence between ‘online’ and ‘offline’. In such cases, it would take an ethnographer to go beyond the screen to find out how and for what purposes specific meanings are made; data mining, or the collection of log data, will not get us beyond the ‘on the screen’ understanding.

However, in going beyond the screen it should also be remembered that people are not ‘cultural or linguistic catalogues’, as Blommaert and Dong (2009) put it: people do not have an opinion on or a straightforward explanation for everything (they do), nor is every aspect of our behaviour easily verbalised (see also Briggs 1986; Blommaert and van de Vijver 2013 on a discussion on ‘methodological loops’). Interviews are thus not necessarily the magic fix. In any case, while with digital culture things have changed – we do have new kinds of socio-cultural activity, new types of environments, and this may require us to be methodologically creative – at the same time, the principles of ethnography stay the same, and ethnography has been through innovations before: there was the time when ethnography first went to school contexts, for instance. As cultures and societies develop, so does ethnography, and what it does is use its inherent adaptivity and flexibility in trying to find out what exactly is going on. This is particularly useful considering the speed and scope of change thanks to digitalisation.

FUTURE DIRECTIONS

Broader changes in (digital) culture and the internet will continue to shape the nature of ethnography. For instance, the emergence of ‘nonymous’ (e.g. Zhao et al. 2008) spaces such as Facebook – i.e. private companies providing social network platforms on which people are encouraged to present themselves with their ‘real’ names – has clearly introduced a change in the functions of online environments in everyday life as well as the study of online environments. It would of course be naïve to assume that simply because nonymity is encouraged, everybody presents themselves with their ‘real names’ – as many of us know from our own contact lists on for instance Facebook. However, nonymity has clearly been a tendency which has to do not only with the changing face of (online) sociality, but also with the interests of private companies offering such services as well as other parties (e.g. governments) with a need to get to people’s ‘real’ identities, and their networks. As to the changing face of sociality,

(...) all kinds of sociality are currently moving from public to corporate space; even as little as ten years ago, the coding of social actions into proprietary algorithms, let alone the branding and patenting of these processes, would have been unthinkable. Today, Facebook, Google, Amazon and Twitter all own algorithms that increasingly determine what we like, want, know, or find. (van Dijck 2013: 37)

The changes can thus be rapid, and changes in communicative environments continue to introduce changes in the shape and functions of people’s communications. Ethnography will have to follow suit, and it only remains to be seen what kind of transformations lie ahead of us. Technological innovation thus always forms a context informing our inquiries on digital communications.

A further context that should be evoked here is the ongoing debates on ‘Big Data’, propelled by the perceived ease of collecting huge amounts of (‘natural’) data from people’s online activities. The amount of potential data entered online every day on social network sites alone is astronomical, and a lot of it is easily available for researchers. However, to assume that big data research gives access to *all* the data is misleading – only social media companies

themselves have access to ‘truly big’ data; for academics, a lot of it is beyond reach. This has prompted some company researchers to even suggest that academics should not take the trouble of studying social media, as they themselves are the only ones to have privileged access to all of the data, for instance all the tweets featuring a specific hashtag (boyd & Crawford 2011).

The availability of data mining techniques has also given rise to new innovative approaches such as ‘ethnomining’ (Aipperspach et al. 2006) which attempts to bring together ethnographic insight and data mining. Both big data and ethnography have their applications and are suited for different tasks. From an ethnographic perspective, big data research is not necessarily without problems, though. For instance, sampling Twitter accounts to probe into people’s communicative practices can be highly problematic if generalisations are made based on the assumption that ‘the bigger the better’ – the more representative, the more accurate (see boyd & Crawford 2011 for a broader discussion on the challenges related to big data). Instead of a ‘the bigger the better’ attitude, ethnographers would be more inclined to endorse the idea that ‘small is beautiful’. Thorough ethnographic investigations provide in-depth understandings of the particularities of the cases studied, and provide ecologically valid information. danah boyd’s most recent book (2014), for instance – providing an account of American teens’ mediatised lives – is a perfect example of the kind of rich and detailed account only possible as a result of longitudinal ethnographic engagement in the field, and in the future we will hopefully see similar studies focusing specifically on language. Long-term ethnographic engagement is of course time- and resource-consuming, but such investments are necessary to provide detailed and situated accounts of communicative practices. Indeed, with technological advances, “We’ve entered an era where data is cheap, but making sense of it is not.” (boyd 2010: np).

In some respects, we have become full circle from what we started out with in the introduction and the context-poor, or contextless, early research on technologically mediated communication. Data mining is now widely seen as *the* approach to studying online communication: it has the aura of being easy, ‘comprehensive’ and ‘objective’. Boellstorff et al. (2012) recall what could have been an unfortunate example of automated analysis of chat logs from a virtual world gone wrong: in the machine-collected data they discuss, the word ‘bunnies’ appeared as thematically significant, and left the quantitative researcher puzzled as to the apparent heavy interest in rabbits in the data. Had it not been for a colleague with a

contextual understanding of what exactly had been the object of discussion in the data set, it would have been difficult to establish that the people in the chat were in fact not discussing rabbits at all, but ‘bunny slippers’ – a type of shoe that increases jump height. This is the kind of insight that ethnographers can produce, and the contributions of ethnographers are very much needed in providing realistic, in-depth explanations of digital communications.

Digitalisation has in many respects changed, even complicated, things for ethnographers. However, what could easily be presented as ‘problems’ in digital ethnography should perhaps not be seen as problems at all. It is not simply that we should keep thinking about these issues in the euphemistic frame of ‘challenges’, but rather see these issues as indicative of the kind of cultural and societal change societies are undergoing thanks to digital technologies. Blommaert and van de Vijver (2013; see also Arnaut 2012, Blommaert 2013) identify complexity, mobility and dynamics as key features of today’s world of rapid social change. The world is increasingly complex, and the changes are in many respects dramatic, and not only for those who have the means and skills – digital and media literacies – to participate, more or less fully, in these developments. These changes affect all of us: those without the resources and the means to participate digitally are affected by the very fact that they stay outside of these new means of communication. For charting these socio-cultural developments, (digital) ethnography is exceptionally well-equipped: as Blommaert (2007: 682, emphasis original) puts it, “One rather uncontroversial feature of ethnography is that it addresses complexity. It does not, unlike many other approaches, try to reduce the complexity of social events by focusing *a priori* on a selected range of relevant features, but it tries to describe and analyze the complexity of social events *comprehensively*.”

REFERENCES

Aipperspach, Ryan, Tye Lawrence Rattenbury, Allison Woodruff, Ken Anderson, John F. Canny & Paul Aoki 2006. Ethno-mining: Integrating numbers and words from the ground up. Technical Report, Electrical Engineering and Computer Sciences, University of California at Berkeley. <http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-125.pdf> (Last viewed 12 July, 2014)

Androutsopoulos, Jannis 2008. Potentials and limitations of discourse-centred online ethnography. *Language@Internet* 5, article 9.

<http://www.languageatinternet.org/articles/2008/1610/androutsopoulos.pdf> (Last viewed 12 July, 2014)

Arnaut, Karel 2012. Super-diversity: Elements of an emerging perspective. *Diversities* 14 (2), 1-16.

Beaulieu, Anne 2004. Mediating ethnography: Objectivity and the making of ethnographies of the internet. *Social Epistemology* 18, 139-163.

Blommaert, Jan 2007. On scope and depth in linguistic ethnography. *Journal of Sociolinguistics* 11 (5), 682-688.

Blommaert, Jan 2013. *Ethnography, superdiversity and linguistic landscapes. Chronicles of complexity*. Bristol: Multilingual Matters.

Blommaert, Jan & Jie Dong 2009. *Ethnographic fieldwork. A beginner's guide*. Bristol: Multilingual Matters.

Blommaert, Jan & Fie Velghe 2012. Learning a supervernacular: Textspeak in a South African township. *Tilburg Papers in Culture Studies*, Paper 22. https://www.tilburguniversity.edu/upload/f3224909-1d96-4a80-9a36-bba9c83239eb_tpcs%20paper22.pdf (Last viewed 12 July, 2014)

Blommaert, Jan & Fons van de Vijver 2013. Good is not good enough: Combining surveys and ethnographies in the study of rapid social change. *Tilburg Papers in Culture Studies* Paper 65.

Boellstorff, Tom, Bonnie Nardi, Celia Pearce & T. L. Nardi 2012. *Ethnography and virtual worlds: A handbook of method*. Princeton: Princeton University Press.

boyd, danah 2008. How can qualitative internet researchers define the boundaries of their projects: A response to Christine Hine. In Markham, Annette N. & Nancy K. Baum (eds.) *Internet inquiry: Conversations about method*. Los Angeles: Sage, 26-32.

boyd, danah 2010. Privacy and publicity in the context of Big Data. WWW 2010, Raleigh, North Carolina, April 29, 2010. <http://www.danah.org/papers/talks/2010/WWW2010.html> (Last viewed 12 July, 2014)

boyd, danah 2014. *It's complicated. The social lives of networked teens*. New Haven: Yale University Press.

boyd, danah & Kate Crawford 2011. Six provocations for Big Data. Paper presented at *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, Oxford Internet Institute, September 21, 2011. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431 (Last viewed 12 July, 2014)

Briggs, Charles L. 1986. *Learning how to ask. A sociolinguistic appraisal of the role of the interview in social science research*. Cambridge: Cambridge University Press.

Burrell, Jenna 2009. The field site as a network: A strategy for locating ethnographic research. *Field Methods* 21, 181-199.

Coleman, Gabriella E. 2010. Ethnographic approaches to digital media. *Annual Review of Anthropology* 39, 487-505.

Dijk, José van 2013. *The culture of connectivity. A critical history of social media*. Oxford: Oxford University Press.

Garcia, Angela Cora, Alecea I. Standlee, Jennifer Bechhoff & Yan Cui 2009. Ethnographic approaches to the internet and computer-mediated communication. *Journal of Contemporary Ethnography* 38, 52-84.

Geertz, Clifford 1973. Thick description: Toward an interpretive theory of culture. In *The interpretation of cultures: Selected essays*. New York: Basic Books, 3-30.

Gershon, Ilana 2010. *Breakup 2.0. Disconnecting over new media*. Ithaca: Cornell University Press.

Georgakopoulou, Alexandra 2013. Small stories research & social media: The role of narrative stance-taking in the circulation of a Greek news story. *Working Papers in Urban Language & Literacies*, Paper 100. <http://www.kcl.ac.uk/sspp/departments/education/research/ldc/publications/workingpapers/th-e-papers/WP100.pdf> (Last viewed 12 July, 2014)

Herring, Susan C. 2004. Computer-mediated discourse analysis: An approach to researching online behavior. In Barab, Sasha, Rob Kling & James H. Gray (eds.) *Designing for virtual communities in the service of learning*. New York: Cambridge University Press, 338-376.

Hine, Christine 2000. *Virtual ethnography*. London: Sage.

Hine, Christine 2013 [2005]. Virtual methods and the sociology of cyber-social-scientific knowledge. In Hine, Christine (ed.) *Virtual methods: Issues in social research on the internet*. London: Bloomsbury, 1-13.

Hymes, Dell 1996. *Ethnography, linguistics, narrative inequality. Towards an understanding of voice*. London: Taylor & Francis.

Kozinets, Robert V. 2009. *Netnography: Doing ethnographic research online*. London: Sage.

Leppänen, Sirpa, Samu Kytölä, Henna Jousmäki, Saija Peuronen & Elina Westinen 2013. Entextualization and resemiotization as resources for (dis)identification in social media. *Tilburg Papers in Culture Studies*, Paper 57. https://www.tilburguniversity.edu/upload/a3d5524e-4413-4772-9f96-9fe0ee714c6f_TPCS_57_Leppanen-et-al.pdf (Last viewed 12 July, 2014)

Madsen, Lian Malai & Andreas Stæhr 2014. Standard language in urban rap: Social media, linguistic practice and ethnographic context. *Tilburg Papers in Culture Studies*, Paper 94. <https://www.tilburguniversity.edu/research/institutes-and-research-groups/babylon/tpcs/item-paper-94-standard-language-in-urban-rap/> (Last viewed 12 July, 2014)

Marwick, Alice E. & danah boyd 2010. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society* 13 (1), 114-133.

Murthy, Dhiraj 2008. Digital ethnography: An examination of the use of new technologies for social research. *Sociology* 42 (5), 837-855.

Postill, John & Sarah Pink in press 2012. Social media ethnography: The digital researcher in a messy web. *Media International Australia*. <http://blogs.bournemouth.ac.uk/research/files/2013/04/Postill-Pink-socialmedia-ethnography.pdf> (Last viewed 12 July, 2014)

Robinson, Laura & Jeremy Schulz 2009. New avenues for sociological inquiry: Evolving forms of ethnographic practice. *Sociology* 43, 685-698.

Rymes, Betsy 2012. Recontextualizing YouTube: From macro-micro to mass-mediated communicative repertoires. *Anthropology & Education Quarterly* 43 (2), 214-227.

Sade-Beck, Liav 2004. Internet ethnography: Online and offline. *International Journal of Qualitative Methods* 3 (2), article 4. http://www.ualberta.ca/~iiqm/backissues/3_2/html/sadebeck.html (Last viewed 12 July, 2014)

Shifman, Limor 2011. An anatomy of a YouTube meme. *New Media & Society* 14 (2), 187-203.

Silverstein, Michael 1979. Language structure and linguistic ideology. In Clyne, Paul, William F. Hanks & Carol L. Hofbauer (eds.) *The elements: A parasession on linguistic units and levels*. Chicago: Chicago Linguistic Society, 193-247.

Silverstein, Michael 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23 (3-4), 193-229.

Stæhr, Andreas 2014. Metapragmatic activities on Facebook: Enregisterment across written and spoken language practices. *Working Papers in Urban Language & Literacies* 124. <http://www.kcl.ac.uk/sspp/departments/education/research/ldc/publications/workingpapers/the-papers/WP124.pdf> (Last viewed 12 July, 2014)

Stæhr, Andreas (in press). Appropriation of transcultural flows among Copenhagen youth: The case of Illuminati. In Androutsopoulos, Jannis & Kasper Juffermans (eds.) 'Superdiversity and digital language practices'. Special issue, *Discourse, Context & Media*.

Varis, Piia & Xuan Wang 2011. Superdiversity on the internet: A case from China. *Diversities* 13 (2), 71-83.

Velghe, Fie 2011a. Lessons in textspeak from Sexy Chick: Supervernacular literacy in South African instant and text messaging. *Tilburg Papers in Culture Studies*, Paper 1. https://www.tilburguniversity.edu/upload/5328eb9d-43bb-47d3-9dbc-f9bc38727c6e_tpcs%20paper1.pdf (Last viewed 12 July, 2014)

Velghe, Fie 2011b. Deprivation, distance and connectivity: The adaptation of mobile phone use to a life in Wesbank, a post-Apartheid township in South Africa. *Tilburg Papers in Culture Studies*, Paper 6. https://www.tilburguniversity.edu/upload/cdcdd501-1a7b-4af2-803d-dd51dd6623cd_tpcs%20paper6.pdf (Last viewed 12 July, 2014)

Zhao, Shanyang, Sherri Grasmuck & Jason Martin 2008. Identity construction on Facebook: Digital empowerment in anchored relationships. *Computers in Human Behavior* 24, 1816-1836.