

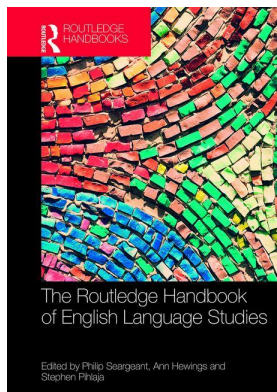
This article was downloaded by: 10.2.97.136

On: 07 Dec 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



The Routledge Handbook of English Language Studies

Philip Seargeant, Ann Hewings, Stephen Pihlaja

The language of social media

Publication details

<https://test.routledgehandbooks.com/doi/10.4324/9781351001724-22>

Tereza Spilioti

Published online on: 20 Jun 2018

How to cite :- Tereza Spilioti. 20 Jun 2018, *The language of social media from:* The Routledge Handbook of English Language Studies Routledge

Accessed on: 07 Dec 2023

<https://test.routledgehandbooks.com/doi/10.4324/9781351001724-22>

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: <https://test.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The language of social media

Tereza Spilioti

What is social media?

The chapter aims to provide a critical review of current research on the language of social media and point to its developing research agendas. After defining what is commonly understood by ‘social media’ and delineating its key distinctive properties, I will first position research on the language of social media vis-à-vis recent developments in studies of digitally mediated language and communication. I will then shed light on current research foci, particularly the study of language and social variation, on the one hand, and language practices, on the other. Issues of self-presentation, identity, and community will be revisited in light of current debates in the field. The chapter concludes by reviewing points of contention regarding methodological issues and developing agendas related to critical and ethical research on social media.

Although there is no consensus about who coined the term ‘social media’, it is almost certain that academics cannot take credit for it: the term appeared and was popularised first among the creators and developers of the very internet sites and services which are today known by the term (Bercovici 2010). Following, to some extent, technologists’ use of the term, academics have also defined social media broadly as referring to a range of internet-based sites and applications that promote, are designed for, and/or are adopted for social interaction between participants (Page et al. 2014; Seargeant and Tagg 2014; Georgakopoulou 2015; Leppanen et al. 2017). The use of such a broad definition has the advantage of bringing together research on various sites and services, ranging from instant messaging apps to social network sites and virtual worlds, and contributes to the development of a critical mass for research on the language of social media. Nevertheless, assuming that technologies for social interaction, such as email, TV, radio, telephone, postcards, and epistolary writing, predate the so-called social media, we need to clarify further some of the distinctive qualities associated with academic and lay understandings of the specific internet-based sites.

In order to clarify what is distinctive about social media, we will focus on three elements that have been hailed as the hallmarks of web 2.0; namely, participation, convergence, and social network architecture. Unlike mass media communication (e.g. newspapers, radio, TV), social media content ‘is as much a product of participation as it is of traditional

creative and publishing/broadcast processes' (Seargeant and Tagg 2014: 4). *Participation* for social media users involves generating and sharing content either publicly or to a selected group of people in a collaborative and often creative manner. Such participation takes place in complex communicative environments that display *convergence* of previously distinct technologies and afford the integration of multiple semiotic modes for user-generated content. For example, Facebook users can move across distinct spaces affording either one-to-many posts (e.g. Facebook wall; News Feed) or one-to-one messaging (e.g. Messenger) and they can also combine a range of modalities by integrating text, image, audio, and video in their status update or chat. At the same time, participation is afforded by a *social network architecture* which

allow[s] individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system. (boyd and Ellison 2008: 211)

The social network architecture may vary from site to site. For example, the *connections* between the members of a social network can range from symmetrical ones (e.g. Facebook connections, where the user has to accept the invite and thus ratify the connection) to asymmetrical ones (e.g. Twitter, where a user can follow another user, without any ratification from the other party). As a result, the extent to which connections can be visible and clearly articulable also varies between these two platforms, with Facebook affording a more explicit articulation of connections within the network. Nevertheless, one of the elements that remain constant across most social media sites and services concerns the creation of a *personal profile* that enables the user to generate content, link with other profiles, and participate in the network. The generation of a personal profile usually involves providing information about one's self in the form of (user)names, pictures, demographic information (such as age, gender, location), and answers to questions about personal interests, beliefs, and likes. Last, but not least, the social network architecture also affords communicative environments where multiple contexts collapse and previously distinct audiences are brought together into single platforms ('*context collapse*' Marwick and boyd 2011: 115). For example, if the average Facebook user has about 338 Facebook friends (Smith 2014), the potential readers and thus audience of a status update include a large amount of people who probably belong to distinct social groups, ranging from professional to family contacts and friends.

Research on digitally mediated language: historical perspectives

In order to understand the theoretical and methodological momentum of current advances in social media research, we need to contextualise contemporary studies within the short, yet rapidly developing, historical trajectory of the field. Early interest in human communication via networked computers (or computer-mediated communication, henceforth CMC) can be traced as far back as the 1980s and 1990s. The interdisciplinary scope of the then emerging field offered fertile ground for linguistic studies to appear and develop in parallel with CMC research in related areas of sociology, computer science, and media and cultural studies. Although marginalised compared to more mainstream areas of linguistic research, language-focused CMC studies continued to proliferate. Despite their fragmentation in terms of analytical foci and methods, they began to cluster together into a relatively identifiable research strand under the umbrella term 'Computer-Mediated Discourse Analysis' (CMDA,

Herring 2004). To use Androutsopoulos's (2006) waves metaphor, the first wave of CMC research is preoccupied with formal and functional properties of text-based communication in primarily English-speaking internet settings, such as email, echat, online forums and multi-player online games. The focus on issues related to the position of CMC vis-à-vis speech and writing, creativity and play, generic norms and textual coherence reveals the field's early attempts to understand the presumed 'new' and 'novel' in juxtaposition with the 'old' norms (Georgakopoulou 2006). In early CMDA studies, such understandings are often constrained by a logic that sees the medium and its technological properties as key shaping factors for language and communication in digital media ('technological determinism' Hutchby 2001).

The field's growing familiarity with the 'novelties' of the medium, as well as opening up to concurrent methodological and theoretical advances in socially-minded linguistics, paved the way for the so-called second wave of CMC research, coined as such in Androutsopoulos (2006). From formal and functional properties of digitally mediated texts, the focus shifts to users, their language practices and the ways in which identities, relations, and communities are formed in a range of internet contexts, including discussion forums, text and instant messaging, social media profile pages, and online games. At the same time, there is a move from the earlier monomodal and monolingual research on primarily English-speaking texts towards studies that examine the combination of multiple modes (mainly image and text) and multiple languages, varieties, and styles on the internet. Furthermore, although not widely acknowledged, the second wave also critically engages with metalinguistic discourses about digitally mediated language and communication that are circulated in 'old' media (e.g. press, radio and TV), frame folk understandings of the internet, and interplay with wider language and social ideologies.

After the first and second wave, voices now converge towards the realisation that a third wave is currently sweeping through the field (Georgakopoulou and Spilioti 2016). Social media and their communicative exigencies play a key role in shaping the foci and methods of recent research. Following on from the second wave, the focus on users and their practices remains, with a heightened attention to practices that are intimately linked with the design and techno-social environment of social media platforms (e.g. sharing, liking, tagging). Identities, relations, and communities are still central foci of interest but they are revisited in light of the various audiences and global networks afforded by social media and, most importantly, in relation to the interplay and mutual embeddedness of face-to-face and mediated (or offline and online) contexts. Furthermore, there is a proliferation of multimodal and multilingual research that attempts to gauge the increased reliance on the visual mode, on the one hand, and the creative and dynamic manipulation of varieties, styles, and language forms, on the other. As argued in the last section, the critical engagement with wider discourses and ideologies moves away from framings of new technologies in old media and puts under the microscope the discourses and scripts of social media platform templates and algorithms that mediate user-to-user interaction.

The language of social media I: English language and social variation

Drawing on the key premise that language is heterogeneous, research on the language of social media explores variation in the realisation of language forms associated with the English language system. In this line of research, linguistic variation at any level of language description, including spelling, morphology, lexis, syntax, or pragmatics, is often associated with external factors ranging from properties of the technology (e.g. character limit or

platform prompts) to social categories such as participant gender, age, or ethnicity. The latter preoccupation that explores language variation in relation to social aspects of the user's identity or the specific situation is more typical of what we have identified as the second wave of CMC research.

In one of the earliest book-length examinations of social media discourse, Zappavigna (2012) points out a range of alternative spelling forms that vary from standard orthography, such as emoticons, repeated letters or punctuation marks, vowel omission, and phonetic respellings. In her data, it proves difficult to discern specific patterns or to correlate non-standard variants with social factors, as usage is highly variable even within single tweets. As Zappavigna (2012) points out, this can be understood as a sign of fluidity and instability of established norms in the then emerging communicative context of Twitter. Indeed, research on more established genres such as blogs and multi-player online games has fruitfully pointed out how alternation between standard and non-standard spellings correlates with social factors, such as the blogger's gender and residence (Hinrichs 2012) or the player's role status and virtual proximity with another player in the game world (Iorio 2009). Beyond spelling, Herring and Paolillo (2006) explored stylistic variation in blogs and indicated that blog genres (i.e. diary vs. filter entries) co-vary with stylistic features associated with 'female' and 'male' language (regardless of the author gender).

Similarly, more recent research on Twitter offers further insights into sociolinguistic variation in social media. Gender is one of the key social variables that have received due attention. Bamman et al. (2014) focus on lexical variation in order to study gendered language styles in a large corpus of 14,000 Twitter users. Their results reveal that although there are some strong gender orientations (e.g. higher frequency of personal pronouns and emotion words in female authored tweets), there are some 'outliers' that do not fit mainstream language and gender classifications. They suggest that cluster analysis of lexical variables can provide a more nuanced understanding of other potential factors at play. More specifically, the gender make-up of the wider social network emerges as a key social factor, as lexical and stylistic choices are found to co-vary with the gender composition of the network, rather than the writer's assumed gender ('homophily effect').

The role of the network audiences and potential addressees in understanding stylistic variation is further supported by Squires's (2014) study of a mass media novel phrase and its adoption, circulation and diffusion on Twitter. Her findings reveal that non-standard spellings and informality markers were most frequently used in tweets that were not directed at the celebrity associated with the novel phrase. Similarly, Eisenstein (2015) notes that alternation between non-standard and standard spellings depends on whether tweets are addressed to individual users or tweeted to the general Twitter audience.

In addition to metadata about authors' and addressees' gender and status (e.g. celebrity, ordinary) in social media, analysts can also have access to information about the geographical location of the author, if the latter has activated the GPS or location function. The study of geolocated tweets or Facebook messages has the potential of using social media research to explore the interplay between regional and spelling variation. An example of such research is Eisenstein's (2015) study on the demographics of neography (broadly defined as unconventional spelling), whereby non-standard spellings (e.g. g-deletion in <ing>) are mapped against geographical regions in the US, based on metadata from geolocated tweets. His findings reveal that non-standard spellings are more frequent in tweets from areas populated by more individuals who identify as African Americans, whereas standard spellings appear more among tweets from areas where the majority includes individuals who identify as White. This pattern of sociolinguistic variation is strikingly similar with how the equivalent

phonological variables (e.g. g-dropping in <in>) behave in the corresponding geographical regions, suggesting potentially a more profound connection between spelling and phonological variation (Eisenstein 2015).

The language of social media II: language and social practices

Beyond associations between language and social factors, research on the language of social media has focused on language and social practices that are afforded by particular exigencies in social media environments, such as (i) the design of the specific platform and (ii) the users' agency and engagement with particular aspects of the design. The relation between the two is bidirectional, with examples where not only do the design features shape user practice (e.g. 'like' buttons for positive evaluation on Facebook) but also user practice shapes the development of design interface (e.g. the creation of Twitter 'reply' button as a result of @username practices to respond to a tweet). More so than in other areas, this line of research develops through cross-fertilisations between media/cultural theory and discourse analysis, operationalising and adapting concepts introduced in media and cultural studies to the needs of a more (micro) language-focused approach to communication in social media. In this section, we will focus primarily on the practices of addressing, tagging, sharing, and meming.

Having a conversation on social media: addressing and tagging

Maintaining a sustained conversation between two parties can be a challenge in 'noisy' social media environments (Honeycutt and Herring 2009) populated by millions of users and by large volumes of continuously updated content (e.g. texts, images, videos). Nevertheless, it appears that social media users often target particular addressees, rather than an undifferentiated mass, with their messages. The consideration of specific addressees and their social characteristics (e.g. gender or celebrity status) is implied in findings about spelling and style variation discussed in the previous section. In addition, practices of addressivity (i.e. prefacing a message with the username of the intended addressee) that predate the birth of social media (Herring 1999) also enable users to maintain one-to-one conversations in public multi-party and multi-network environments. For example, Twitter users employ @mentions (that is, @username) to target specific addressees and 'establish [...] the conditions for threading multiple tweets together as a "conversation"' (Squires 2016: 242). Similarly, in one of the few studies of turn management and conversational coherence in YouTube comments, Bou-Franch et al. (2012) demonstrate how cross-turn addressivity and turn entry/exit devices (such as discourse markers like 'listen', 'that's all', question tags, and aphorisms) are used as turn management signals and capitalise on the conversational potential of the site.

Another practice for attracting the attention of another user and marking content as relevant or salient is through tagging. Tagging can be broadly defined as the use of a hyperlinked keyword that attaches a label to user-generated texts, images, audios, and videos. In the context of Facebook, tagging another user in photos or status updates results in notifying the other party about what is shared and opens up the potential for social interaction. In social media platforms like Twitter, Instagram and Flickr, tagging can take the form of hashtags where the label-word/phrase is prefaced by the # symbol. Unlike the Facebook type of user-tagging, hashtags have an organisational function as they are based on topics or ideas. Clicking on a particular hashtag results in threads of uploaded tweets or photos that cohere around a particular label-word/phrase. In other words, as Zappavigna (2012) puts it, hashtags result in 'searchable talk' and contribute to the organisation and flow of social media content.

In one of the earliest studies on Twitter discourse, Page (2012) has broadly identified two types of hashtags: ‘topic-based’ when the tag specifies a topic (e.g. #brexit) and ‘evaluative’ when the tag offers an evaluative comment on the topic discussed (e.g. #brexshit).

Participating in social media: sharing

In his seminal paper about the emergence of ‘sharing’ as a keyword, John (2013) argues about the centrality of sharing as a defining practice in the so-called social media. By tracing the meanings and uses of ‘share’ in social network sites over a decade, he concludes that sharing in social media involves two main activities: (i) distributing content by making texts, links, photos, videos available to others, and (ii) communicating thoughts, emotions, and big and small events of one’s everyday life. Starting with the first understanding of sharing as distributing content, we can discern a number of technologically afforded acts that aim at (re)distributing content. In the communicative dynamics of Twitter, for example, sharing can be achieved through a range of acts: ‘tweeting’ for broadcasting and thus making content available to others, ‘retweeting’ for rebroadcasting another user’s content through one’s own stream, ‘modified tweeting’ for selecting and commenting on parts of another user’s tweet, and ‘embedding’ by rebroadcasting tweets within a different communicative context (e.g. a blog or a newspaper website). All these practices of sharing also indicate a further quality of social media, namely the portability of texts and any elements they consist of. As noted by Squires (2016: 244) in her discussion of tweets, ‘they come with a detachability that makes them particularly amenable to de-and re-contextualization’.

In an attempt to unpack further John’s second understanding of sharing as communicating, Androutopoulos (2014: 17) probes into processes of de/recontextualisation and defines sharing on Facebook ‘as a set of practices by which individuals entextualize significant moments for, and with, their networked audience’. Taking as unit of analysis the Facebook ‘wall event’ that consists of the user’s initiating post (i.e. Facebook status update) followed by ‘likes’ and one or more responses by members of the audience, he outlines three stages in social media sharing: ‘*Selecting* refers to the choice of moments to share; *styling* concerns how to entextualize what is being shared; and *negotiating* refers to the audience engagement that follows up on acts of sharing’ (Androutopoulos 2014: 8). In other words, research on sharing practices involves analysing (i) what is selected to be shared (e.g. routine activities or life events; completed events or moments on the go); (ii) the range of resources used for styling the selected shareables (e.g. styles, language forms, photos, links, etc); and (iii) the audience responses following up the shareable and the ways in which they impact on overall wall event’s style and content. An application of such an approach to sharing is offered by Giaxoglou (2015), who examines mourning practices online and draws attention to the ways in which participants select, style, and negotiate significant moments in the mourning process on public RIP Facebook walls.

Although there is a growing body of research on sharing as entextualising significant moments in social network sites, there is less attention to related phenomena of virality and mimicry where particular signs are shared and spread at an astonishing speed by – and to – large numbers of people. From a media and cultural studies perspective, Shifman (2012) makes a distinction between virality and mimicry and sees the degree of sign modification as a key difference between the two. For example, a viral YouTube video is spread without significant change, whereas a memetic video ‘lures extensive creative user engagement in the form of parody, pastiche, mash-ups or other derivative work’ (Shifman 2012: 190). Varis and Blommaert (2014) show how a socio-discursive perspective that approaches such phenomena

as social semiotic activities offers a more nuanced perspective, blurring the boundaries between virality and memicity. Beyond (re)contextualisation, the study of sharing in memes invokes the concept of ‘resemiotization’ which

refers to the process by means of which every ‘repetition’ of a sign involves an entirely new set of contextualization conditions and thus results in an entirely ‘new’ semiotic process, allowing new semiotic modes and resources to be involved in the repetition process.

(Varis and Blommaert 2014: 8)

Memes illustrate this process by establishing some generic recognisability through repetition of certain sign elements, on the one hand, and producing different communicative effects through modification and adjustment to specific situations and for different audiences, on the other. As a result, a language-focused approach to memes can investigate the range of resources that are mobilised in the complementary processes of establishing recognisability while producing different effects. Such resources include the visual architecture of the meme (e.g. design, layout, colour), its speech act format (e.g. ‘keep calm and [*directive*]’), its textual-stylistic features (e.g. lolspeak), or, in the case of mashup memes, a combination of the above originating in different established memes (Varis and Blommaert 2014).

The language of social media III: identity & community

Social identities and the study of language as indexical of belongingness to wider social groupings (such as gender, geographical location, and ethnicity) have preoccupied research on language and social variation on social media. At the same time, though, there is a wealth of studies that have undertaken a more discourse-oriented approach and explored identity not as a stable or fixed property but as multiple, fluid, and constructed through discourse and in interaction with others (Bucholtz and Hall 2005). Shying away from pre-existing social categories and paying attention to the ongoing social construction and negotiation of identities are also appropriate for environments that have often been hailed as anonymous and, more importantly, display a social network architecture that capitalises on the reflexive design of personal profiles for multiple audiences. What research on the language of social media can contribute to the study of reflexive performances of identity for networked audiences is attention to the ways in which verbal and other semiotic resources are mobilised to achieve such performances.

The mobilisation of multiple languages and scripts – or, to be more precise, of multiple forms associated with particular languages and scripts as sociocultural constructs (e.g. English, French, Latin, or Greek alphabet) – have attracted the attention of language scholars. From earlier studies of code-switching and language choice to more recent research on trans-languaging and trans-scripting, English-related forms have been noted to appear among the linguistic repertoires of social media users in very diverse contexts (e.g. Chinese and Spanish Flickr users, Lee and Barton 2011; Nepalese undergraduate students on Facebook, Sharma 2012; members of Finland-based football web forums and Finnish footballers on Twitter, Kytola 2016). Although English-related forms are used with varying and ambivalent indexical values and meanings, this line of research converges in foregrounding the multiple and hybrid identities performed in social media: from the negotiation of ‘glocal’ identities on Flickr (Lee and Barton 2011) to the varying ‘translocal’ positions projected by users with real or aspirational transnational trajectories (Sharma 2012; Androutsopoulos 2014; Kytola 2016).

In addition to linguistic codes and scripts, previous literature has paid attention to time and space, especially explicit references to and markings of time and physical or geographical space as resources for identity formation on social media (Georgakopoulou 2015). For example, spatio-temporal references contribute to the construction of expert identities in online reviews (Vasquez 2015), to the authentication of diasporic identities (Heyd and Honkanen 2015), and to the maintenance and affirmation of friendship (Cohen 2015). Platform-specific discursive features, such as hashtags on Twitter, can also be mobilised as resources for self-presentation. For example, Page (2012) and Lorenzo-Dus and Di Cristofaro (2016) demonstrate how the varying types of hashtags (e.g. topic-based vs. evaluative) and their thematic co-occurrence contribute to different ways of self-branding among celebrities, corporations, ordinary people, and influential citizens on Twitter. The mobilisation of visual resources for self-presentation remains relatively under-explored in language-focused research. A notable exception is Georgakopoulou's (2016: 300) study of pictures and selfies by adolescent women on Facebook where she concludes that 'far from being narcissistic expressions of "ideal selves", selfies emerge as contextualized and co-constructed presentations of self'.

With social media platforms increasingly marketing new services as tools for sharing stories (e.g. from Storify to Snapchat or Instagram Stories app), narrative is emerging as one of the key modes for participating and relating with others. Of course, such narratives seldom have the form of prototypical, long, and monologic stories; instead, they are often fragmented, non-linear, co-constructed and open-ended (cf. small stories approach Georgakopoulou 2007). It is during such acts of story-telling that social media users create particular roles for themselves in interaction with their networked audiences (Dayter and Muhleisen 2016). For example, various forms of story-telling enable users to construct a counselling or advisor persona in online health contexts (Thurnherr et al. 2016), a credible and authentic persona in online dating ads (Muhleisen 2016), or a figure of authority in the reading and interpretation of the Bible among Evangelical Christian YouTube users (Pihlaja 2016).

It goes without saying that any identity projections or claims invoke wider groupings and collectivities. In other words, it is impossible to study identity without referring to the concept of 'community'. Similarly to identity, assumptions about community as static, fixed, and delineated by specific (and pre-existing) geographical or temporal boundaries have been challenged by social media research. Instead, identity projections and claims online contribute to the participants' alignment with strangers on the basis of shared interests, topics, and experiences (e.g. photography, Lee and Barton 2011; football fandom, Kytola 2016; interpretation of religious texts, Pihlaja 2016). In such affinity spaces, language researchers can explore the development and negotiation of shared repertoires of rituals, routines, vocabulary, and styles that contribute to the emergence and formation of local communities of practice. For example, Potts (2015) demonstrates how the introduction of homosocial meaning and homosexual innuendo by prominent gamers contributes to the formation of a self-policing community of practice that is more tolerant and rejects bigotry in YouTube comments by heterosexual male fans. Furthermore, increased attention has been given to the so-called 'hashtag communities' of Twitter and social media 'where people have an "ambient affiliation" (Zappavigna 2012) to a topic or issue around which they interact' (Seargeant and Tagg 2014: 12). In such communities, connections between people and content are felt to exist but they are not necessarily defined or visible in a clear manner. Hashtags play a key role in making such connections findable and searchable but it is the feeling of ambient sociability rather than any clear articulation of such connections that is of the essence in such environments. For Varis and Blommaert (2014), memes as signs almost empty of semantic

meaning represent prime resources for achieving ambient sociability and generating temporary collectives. If we accept that social media groups display loose bonds around transient and superficial interests, then memes and, more importantly, the practice of meme sharing offer ‘brief moments of focusing on perceived recognizable and shareable features’. The exact meaning of such features in a given context is highly indeterminate but the recognisable act of innocuous sharing creates a ‘structural level of conviviality’ and generates transiently focused collectives (Varis and Blommaert 2014: 18–19).

Researching the language of social media: mixed methodologies

The rapid development of the field has left it rather fragmented in terms of the range of methodologies used for researching and analysing the language of social media. Nevertheless, one could argue that the multiplicity of perspectives brought about by such fragmentation is one of the field’s strengths: the lack of convergence towards a single dominant paradigm has enabled research to develop simultaneously across a range of linguistic research traditions, on the one hand, and facilitated the use and application of mixed methodologies to the study of social media discourse, on the other. In fact, a commitment towards ‘mixed methods’ research arguably represents a current point of convergence in the field (Bolander and Locher 2014: 20–22), although the term is used for various stages of the research process and in relation to various combinations of existing methodologies.

In terms of data collection, a mixed methods approach is often equated with the complementary use of log file data (i.e. ‘the stored, static records of message sequences that have been put into their particular order by a server feature and that are displayed as a message protocol on the users’ screens’ Beisswenger 2008) and interviews with the people participating in the social media platforms from which the log file data are extracted. In Androutsopoulos’s (2013) terms, it combines ‘screen-based’ with ‘user-based’ data and it develops in line with the wider shift in the field’s preoccupations from formal, functional and textual properties of social media to users’ practices for self-presentation, (dis)identification and community building online. Some notable examples of such work include Lee and Barton’s (2011) study of multilingual writing practices on the photo-sharing site Flickr, Tagg and Seargeant’s (2016) research on language choice as audience design on Facebook, and Georgalou’s (2015) study of time and age identities on Facebook.

There is also a wealth of research where the process of collecting log file data and interviews is paired with systematic observation and immersion in the communities and cultures of social media users (either online only or both online and offline). Some recent examples of this type of ‘virtual fieldwork’ (Androutsopoulos 2013) or, more broadly, ‘digital ethnography’ (Varis 2016) can be found in research on player communities in virtual games (e.g. Newon 2011), as well as on Twitter communities of practice based around shared interests (e.g. ballet in Dayter 2016).

Beyond the combination of various methods for data collection, mixed methods research has also been used to refer to ‘multi-sited’ approaches to the study of social media. The multi-sited element is evident in research that attends to the multiple sites for user participation and engagement afforded by social media platforms. For example, YouTube research may involve the study of language resources and practices within and across three inter-related but distinct sites: (i) the video clip uploaded by the user, (ii) the ‘comments’ space populated by other users’ responses to the video and to one another, and (iii) the ‘hosting space’, that is ‘the individual webpage on which each video is framed with additional information’ (Androutsopoulos and Tereick 2016). Such multi-sited and mixed methods studies also attend

to the multiple semiotic resources mobilised in the respective spaces of social media, undertaking a multimodal perspective or, at least, committing to the study of the interplay between the verbal and the visual. Furthermore, the term multi-sited is also increasingly used to refer to research that spans across multiple web sites and social media platforms; for example, the study of a particular hashtag, its meanings and potential for affiliation and community building, can take the researcher to a journey across social media, from Twitter to Facebook to its further embeddings in online news streams. Although still in its infancy, the mobility in the research process of exploring links and moving around in the networks has potential as it replicates the ways in which users engage with the openness, diversity, and connectivity of social media (cf. ‘guerrilla ethnography’ Yang 2003; and Deumert 2014).

In terms of data analysis, a mixed methods perspective entails not only the combination of qualitative and quantitative methodologies but also the development of eclectic approaches that draw on more than one framework for analysing language. For example, attention to both visual and verbal aspects in social media environments has given rise to studies that combine social semiotics with (critical) discourse analysis. The relative ease of access to large bodies of text-based data, generated for social interaction but still persistent on screen, has paved the way for an increased use of language corpora and corpus linguistics tools. Either alone or in combination with other frameworks from sociolinguistics or discourse analysis, corpus analysis is often used to address quantitatively-oriented research questions about language use on social media. Nevertheless, the field is still trying to grapple with appropriate ways for combining the smallness of qualitative micro-linguistic research with big data and their potential for quantitative analysis of not only user-generated texts but also abundant meta-contextual information (metadata) generated with such texts (Page 2016).

Future directions

Reflecting upon the future of research on language and social media brings forward the realisation that the age of innocence has long passed both for the field and for social media. It is by now a truism that social media is not only a means for communication but also a means for surveillance. Taking the practice of sharing as an example, it is not always clear what we share, with whom, and under what circumstances, when we ‘check in’ on our favourite hangouts or when we agree with the Terms and Conditions of the various platforms and apps. What I suggest as a loss of innocence or, perhaps, a sign of maturity for the field is manifest in the call for critical approaches to the discourse of social media users and, primarily, to the discourse of media platforms and design (e.g. Georgakopoulou and Spilioti 2016: 5).

Research on language and surveillance practices is still in its infancy but studies are already pointing to the role of discourse analysts and sociolinguists in enriching surveillance studies with insights from language theory (e.g. inferential pragmatics Jones 2016; Goffmanian interaction order, as well as contextualisation and inferencing processes Rampton 2016). More specifically, Jones (2016: 409–410) proposes an approach to social media texts as ‘information gathering devices’ that not only collect information about people but also construct discourse positions and social identities for them. This type of surveillance work is achieved by means of ‘subtexts’ (e.g. algorithms, default templates), ‘pretexts’ (e.g. rhetorical strategies that precondition users to opt for a preferred option), and new ‘contexts’ where the circumstances for participation, monitoring, and control over information may be obscure or subject to change (for example, privacy default settings are often updated as social media platforms develop or change corporate hands). Out of the three, we find some preliminary work on subtexts and pretexts: for example, Eisenlauer’s (2014) study on Facebook software

and text automation processes as the ‘third author’ intervening between profile owner and profile recipients, and Blommaert and Omoniyi’s (2006) research on ‘phishing’ strategies of fraud emails. There is still ample scope for further research in this area, particularly in the role of users’ agency in interacting with platform sub/pretexts and in the wider and rather complex contextualisation processes on social media. Although social media platforms deserve a more detailed and critical study, the loss of innocence does not need to shift the field towards an a-critical dystopic view of technologies, especially since, as Seargeant and Tagg (2016) remind us, ‘the filter bubble isn’t just Facebook’s fault – it’s yours’ as well.

Recent developments that have foregrounded the need for language-focused research into surveillance and privacy have also been coupled with a renewed interest towards ethics. Once restricted to the realm of philosophers and ethicists, questions about selfhood, privacy and responsibility now feature in public debate and popular discourse: for example, they are present in users’ everyday practices while choosing the privacy settings of their personal profile, as well as in web designers’ development of such choices for them. Issues of privacy and responsibility, though, are also central to the academic community and debates about research ethics in a changing world.

Although research on social media is developing fast within and across disciplines, research ethics procedures seem to be lagging behind, still driven by a dichotomous understanding of private versus public sphere and by a checklist rationale that orients primarily to prevention of harm. The complex, dynamic, and, at times, unexpected participation formats afforded in social media environments challenge prior understandings of public and private, informed consent and its limitations in online environments, as well as the researcher’s responsibility towards the various stakeholders (users as producers/consumers, platform owners and, of course, the wider public). For issues related to research ethics, the Association of Internet Researchers offers the most up-to-date, complete, and authoritative recommendations for general research on social media (Markham and Buchanan 2012). Similar recommendations for language-focused research can be found in the relevant professional associations, such as the BAAL (British Association for Applied Linguistics) Recommendations on Good Practice (BAAL 2016). Dominated by ethicists and philosophers, issues related to the field of internet research ethics are marginally addressed in applied linguistic research. Nevertheless, linguists and discourse analysts can offer important insights into how privacy and/or publicness are constructed, (re)negotiated and oriented to by participants in – and through – social media discourse. The special issue on the ethics of online research in applied linguistics (Spilioti and Tagg 2016) provides a first step and brings together a range of primarily qualitative studies that have addressed some of these challenges through an approach to ethics as a contextualised, reflexive, and iterative process of decision-making at critical junctures. There is still, though, ample scope for language researchers to engage with internet ethics debates by further unpacking social media discourses and bringing in disciplinary-specific sensitivities to issues of power and control in researcher–researched relationships and networks.

Further reading

- Georgakopoulou, A. and T. Spilioti (eds) 2016 *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge. This 29-chapter edited volume showcases critical reviews of established literature in language-focused research on digitally mediated communication, while engaging with cutting edge research and pointing to new directions for study.
- Page, R., D. Barton, J. W. Unger and M. Zappavigna (2014) *Researching Language and Social Media: A Student Guide*. London and New York: Routledge. A highly accessibly book that offers useful

guidance to undergraduate and postgraduate students on how to design, undertake, and engage with language-focused research on social media.

Sargeant, P. and C. Tagg (eds) (2014) *The Language of Social Media: Identity and Community on the Internet*. Basingstoke: Palgrave Macmillan. This 11-chapter volume represents one of the first comprehensive collections of original research on the language of social media, with a focus on identity and community issues in such environments.

Related topics

- English and social identity
- Media, power and representation
- Sociolinguistics: studying English and its social relations
- Multimodal English.

References

- Androutsopoulos, J. (2006) 'Introduction: Sociolinguistics and computer-mediated communication', *Journal of Sociolinguistics* 10 (4): 419–438.
- Androutsopoulos, J. (2013) 'Online data collection', in C. Mallinson, B. Childs and G. Van Herk (eds), *Data Collection in Sociolinguistics: Methods and Applications*. London and New York: Routledge, 236–250.
- Androutsopoulos, J. (2014) 'Moments of sharing: Entextualization and linguistic repertoires in social networking', *Journal of Pragmatics* 73: 4–18.
- Androutsopoulos, J. and J. Tereick (2016) 'YouTube: Language and discourse practices in participatory culture', in A. Georgakopoulou and T. Spilioti (eds), *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge, 354–370.
- BAAL. (2016) 'Recommendations on good practice in applied linguistics'. Available at www.baal.org.uk/goodpractice_full_2016.pdf.
- Bamman, D., J. Eisenstein and T. Schnoebelen (2014) 'Gender identity and lexical variation in social media', *Journal of Sociolinguistics* 18 (2): 135–160.
- Beisswenger, M. (2008) 'Situating chat analysis as a window to the user's perspective', *Language@Internet* 5, article 6. Available at www.languageatinternet.org/articles/2008/1532.
- Bercovici, J. (2010) Who Coined Social Media? Web Pioneers Compete for Credit, *Forbes*, 9 December, Available at www.forbes.com/sites/jeffbercovici/2010/12/09/who-coinedsocial-media-web-pioneers-compete-for-credit/2/.
- Blommaert, J. and Omoniyi, T. (2006) 'Email fraud: Language, technology, and the indexicals of globalisation', *Social Semiotics* 16 (4): 573–605.
- Bolander, B. and M. Locher (2014) 'Doing sociolinguistic research on computer-mediated data: A review of four methodological issues', *Discourse, Context & Media* 3: 14–26.
- Bou-Franch, P., N. Lorenzo-Dus and P. Garces-Conejos Blitvich (2012) 'Social interaction in YouTube text-based polylogues: A study of coherence', *Journal of Computer-Mediated Communication* 17 (4): 501–521.
- boyd, d. and N. Ellison (2008) 'Social network sites: Definition, history, and scholarship', *Journal of Computer-Mediated Communication* 13: 210–230.
- Bucholtz, M. and K. Hall (2005) 'Identity and interaction: A sociocultural linguistic approach', *Discourse Studies* 7 (4–5): 585–614.
- Cohen, L. (2015) 'World attending in interaction: Multitasking, spatializing, narrativizing with mobile devices and Tinder', *Discourse, Context & Media* 9: 46–54.
- Dayter, D. (2016) *Discursive Self in Microblogging: Speech Acts, Stories and Self-Praise*. Amsterdam: John Benjamins.
- Dayter, D. and S. Muhleisen (2016) 'Telling stories about self in digital contexts: Same, same, but different?' *Open Linguistics* 2: 572–576.
- Deumert, A. (2014) *Sociolinguistics and Mobile Communication*. Edinburgh: Edinburgh University Press.
- Eisenlauer, V. (2014) 'Facebook as a third author – (semi-)automated participation framework in social network sites', *Journal of Pragmatics* 72: 73–85.

- Eisenstein, J. (2015) 'Systematic patterning in phonologically-motivated orthographic variation', *Journal of Sociolinguistics* 19 (2): 161–188.
- Georgakopoulou, A. (2006) 'Postscript: Computer-mediated communication in sociolinguistics', *Journal of Sociolinguistics* 10 (4): 548–557.
- Georgakopoulou, A. (2007) *Small Stories, Interaction and Identities*. Amsterdam: Benjamins.
- Georgakopoulou, A. (2015) 'Introduction: Communicating time and place on digital media – multi-layered temporalities & (re)localizations', *Discourse, Context & Media* 9: 1–4.
- Georgakopoulou, A. (2016) 'From narrating the self to posting self(ies): A small stories approach to selfies', *Open Linguistics* 2 (1): 300–317.
- Georgakopoulou, A. and T. Spilioti (eds) (2016) 'Introduction: Language and digital communication', in A. Georgakopoulou and T. Spilioti (eds), *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge, 1–16.
- Georgalou, M. (2015) 'Beyond the timeline: Constructing time and age identities on Facebook', *Discourse, Context & Media* 9: 24–33.
- Giaxoglou, K. (2015) '“Everywhere I go, you're going with me”: Time and space deixis as affective positioning resources in shared moments of digital mourning', *Discourse, Context & Media* 9: 55–63.
- Herring, S. (1999) 'Interactional coherence in CMC', *Journal of Computer-Mediated Communication*, 4: 0. doi: 10.1111/j.1083–6101.1999.tb00106.x
- Herring, S. (2004) 'Computer-mediated discourse analysis: An approach to researching online behaviour', in S. A. Barab, R. Kling and J. Gray (eds), *Designing for Virtual Communities in the Service of Learning*. Cambridge: Cambridge University Press, 338–376.
- Herring, S. and J. Paolillo (2006) 'Gender and genre variation in weblogs', *Journal of Sociolinguistics* 10 (4): 439–459.
- Heyd, T. and M. Honkanen (2015) 'From *Naija* to *Chitown*: The New African Diaspora and digital representations of place', *Discourse, Context & Media* 9: 14–23.
- Hinrichs, L. (2012) 'How to spell the vernacular: A multivariate study of Jamaican e-mails and blogs', in A. Jaffe, J. Androutsopoulos, M. Sebba and S. Johnson (eds), *Orthography as Social Action: Scripts, Spelling, Identity and Power*. Boston/Berlin: De Gruyter Mouton, 325–358.
- Honeycutt, C. and S. Herring (2009) 'Beyond microblogging: Conversation and collaboration via Twitter', *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences (HICSS-42)*. Los Alamitos, CA: IEEE Press. Available at www.ella.slis.indiana.edu/~herring/honeycutt.herring.2009.pdf.
- Hutchby, I. (2001) 'Technologies, texts and affordances', *Sociology* 35 (2): 441–456.
- John, N. (2013) 'Sharing and web 2.0: The emergence of a keyword', *New Media & Society* 15 (2): 167–182.
- Jones, R. (2016) 'Surveillance', in A. Georgakopoulou and T. Spilioti (ed.), *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge, 408–411.
- Iorio, J. (2009) 'Effects of audience on orthographic variation', *Studies in the Linguistic Sciences: Illinois Working Papers 2009*: 127–140. Available at hdl.handle.net/2142/14815.
- Kytola, S. (2016) 'Translocality', in A. Georgakopoulou and T. Spilioti (eds), *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge, 371–388.
- Lee, C. and D. Barton (2011) 'Constructing glocal identities through multilingual writing practices on Flickr.com[®]', *International Multilingual Research Journal* 5 (1): 39–59.
- Leppanen, S., S. Kytola, E. Westinen and S. Peuronen (2017) 'Introduction: Social media discourse, (dis)identifications and diversities', in S. Leppanen, E. Westinen and S. Kytola (eds) *Social Media Discourse, (Dis)identifications and Diversities*. New York and London: Routledge, 1–36.
- Lorenzo-Dus, N. and M. Di Cristofaro (2016) '#Living/minimum wage: Influential citizen talk in Twitter', *Discourse, Context & Media* 13 (Part A): 40–50.
- Markham, A. and E. Buchanan (2012) 'Ethical decision-making and internet research: Recommendations from the AoIP ethics working committee', Available at www.aoir.org/reports/ethics2.pdf.
- Marwick, A. and d. boyd (2011) 'I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience', *New Media & Society* 13 (1): 114–133.
- Muhleisen, S. (2016) '“More about me” – Self-presentation and narrative strategies in Caribbean online dating ads', *Open Linguistics* 2 (1): 437–449.
- Newon, L. (2011) 'Multimodal creativity and identities of expertise in the digital ecology of a *World of Warcraft* guild', in C. Thurlow and K. Mroczek (eds), *Digital Discourse: Language in the New Media*. Oxford: Oxford University Press, 203–231.

- Page, R. (2012) 'The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags', *Discourse & Communication* 6 (2): 181–201.
- Page, R. (2016) 'Moving between the big and the small: Identities and interaction in digital contexts', in A. Georgakopoulou and T. Spilioti (eds), *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge, 403–407.
- Pihlaja, S. (2016) "'What about the wolves?'" The use of scripture in YouTube arguments', *Language and Literature* 25 (3): 226–238.
- Potts, A. (2015) "'Love you guys (no homo)": How gamers and fans play with sexuality, gender, and Minecraft on YouTube', *Critical Discourse Studies* 12 (2): 163–186.
- Rampton, B. (2016) 'Foucault, Gumperz and governmentality: Interaction, power and subjectivity in the 21st century', in N. Coupland (ed.), *Sociolinguistics: Theoretical Debates*. Cambridge: Cambridge University Press, 303–328.
- Seargeant, P. and C. Tagg (2014) 'Introduction: The language of social media', in P. Seargeant and C. Tagg (eds), *The Language of Social Media: Identity and Community on the Internet*. Basingstoke: Palgrave Macmillan, 1–21.
- Seargeant, P. and C. Tagg (2016) The Filter Bubble Isn't Just Facebook's Fault – It's Yours, *The Conversation*, 5 December 2016, Available at www.theconversation.com/the-filter-bubble-isnt-just-facebooks-fault-its-yours-69664.
- Sharma, B. K. (2012) 'Beyond social networking: Performing global Englishes in Facebook by college youth in Nepal', *Journal of Sociolinguistics* 16 (4): 483–509.
- Shifman, L. (2012) 'An anatomy of a YouTube meme', *New Media & Society* 14 (2): 187–203.
- Smith, A. (2014) 6 New Facts About Facebook, *Pew Research Center*, Available at www.pewresearch.org/fact-tank/2014/02/03/6-new-facts-about-facebook/.
- Squires, L. (2014) 'From TV personality to fans and beyond: Indexical bleaching and the diffusion of a media innovation', *Journal of Linguistic Anthropology* 24 (1): 42–62.
- Squires, L. (2016) 'Twitter: Design, discourse, and the implications of the public text', in A. Georgakopoulou and T. Spilioti (eds), *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge, 239–256.
- Spilioti, T. and C. Tagg (2016) 'The ethics of online research methods: Challenges, opportunities, and directions in ethical decision-making', *Applied Linguistics Review* 8 (2–3): 163–167.
- Tagg, C. and P. Seargeant (2016) 'Facebook and the discursive construction of the social network', in A. Georgakopoulou and T. Spilioti (eds), *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge, 339–353.
- Thurnherr, F., M.-T. von Rohr and M. Locher (2016) 'The functions of narrative passages in three written online health contexts', *Open Linguistics* 2 (1): 450–470.
- Varis, P. (2016) 'Digital ethnography', in A. Georgakopoulou and T. Spilioti (eds), *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge, 55–68.
- Varis, P. and J. Blommaert (2014) 'Conviviality and collectives on social media: Virality, memes and new social structures', *Tilburg Papers in Culture Studies*, Paper 108: 1–21. Available at www.tilburguniversity.edu/upload/83490ca9-659d-49a0-97db-ff1f8978062b_TPCS_108_Varis-Blommaert.pdf.
- Vasquez, C. (2015) 'Right now versus back then: Recency and remoteness as discursive resources in online reviews', *Discourse, Context & Media* 9: 5–13.
- Yang, G. (2003) 'The internet and the rise of a transnational Chinese cultural sphere', *Media, Culture & Society* 25: 469–490.
- Zappavigna, M. (2012) *Discourse of Twitter and Social Media: How we Use Language to Create Affiliation on the Web*. London: Continuum.

Part 3

Analysing English
