

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/95063/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Spilioti, Tereza 2017. Media convergence and publicness: towards a modular and iterative approach to online research ethics. *Applied Linguistics Review* 8 (2-3) , pp. 191-212. 10.1515/applirev-2016-1035

Publishers page: <http://dx.doi.org/10.1515/applirev-2016-1035>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# Media convergence and publicness: Towards a modular and iterative approach to online research ethics

Tereza Spilioti (Cardiff University)

FINAL MANUSCRIPT, accepted September 2016

**Abstract:** The aim of the article is to build a bridge between assumptions about publicness and ethics in traditional (mass) media research and similar issues pertaining to research ethics in so-called new media environments. The article starts off with unpacking ‘publicness’ as defined in authoritative ethical guidelines that regulate research on (and through) media. It points to the challenges media convergence - and, particularly, the increasingly multimodal, multiauthored and multimedial content of websites - have brought to perceptions of publicness, as previously understood in mass media research. With reference to language-focused research on multilingual digital writing in such contexts, I critically engage with ethical tensions related to collecting and analysing internet data, on the one hand, and presenting and publishing data extracts from new media contexts, on the other. Drawing on modularity as a key organising principle of web design and discourse (Androutsopoulos 2010: 208; Pauwels 2012: 251), the article proposes a modular and iterative approach to research ethics that takes into account the complex and fluid configuration of web environments and attends to the conditions of multiple authorship and multiple publics that are increasingly typical of such contexts.

**Keywords:** Media convergence, research ethics, publicness, privacy, consent, copyright, modularity, websites

## 1. Introduction

The aim of this article is to build a bridge between assumptions about publicness and research ethics in traditional (mass) media research (e.g. newspapers, TV and radio) and similar issues pertaining to research in so-called ‘new media’ environments. The reason for doing so is two-fold: first, discussions about online ethics are inevitably informed by perceptions of publicness previously articulated in relation to research on mass media; and, second, convergence of multiple modes and media in everyday communication makes ‘old-new’ dichotomies particularly difficult to keep separate. By moving beyond the ‘old-new’ dichotomy in relation to research ethics in media investigation, the article complements previous work that offers historicised understandings of textual and mediated discourse practices (e.g. Herring 2004: 3, Jones 2011: 336), genres (e.g. Heyd 2016: 89-90), and media ideologies (e.g. Thurlow 2013: 243; Spilioti 2016: 134). More specifically, understanding how publicness has been defined and applied to the process of researching mass media discourse paves the way for contextualising and critically revisiting ethical debates about collecting and analysing internet data, on the one hand, and presenting and publishing data extracts from new media contexts, on the other.

Media convergence, as a key property of new media environments (cf. ‘convergent media computer-mediated communication’ Herring 2013: 4), challenges assumptions about publicness. In particular, the study of websites, which increasingly include multimodal, multiauthored and multimedial content, questions any assumed clear-cut boundaries between public and personal data. With reference to language-focused research on multilingual digital writing in such contexts, I critically engage with the challenges to the process of ethical decision-making and the potential risks for internet users of a priori assumed public-private dichotomies. Drawing on modularity as a key organising principle of web design and discourse (Androutsopoulos 2010: 208; Pauwels 2012: 251), the article proposes a modular and iterative approach to research ethics that takes into account the complex and fluid configuration of web

environments and attends to the conditions of multiple authorship and multiple publics that are increasingly typical of such contexts.

## **2. Mass media and the ‘public domain’: Research ethics**

Any empirical research endeavour involves information management, typically referred to as data collection, analysis and presentation in academic discourse. In the discipline of applied linguistics, in particular, such information can be intimately linked with the persons who provide, produce and consume texts the researcher is interested in. It is this relationship between the researcher and the persons researched that is primarily facilitated and regulated by the key principles of research ethics, such as respect to human dignity, autonomy, and safety, as well as protection from harm (BAAL 2016; AoIR 2012; ESRC 2015). Acknowledging all stages of the research process, from initial design to dissemination (AoIR 2012: 3), we realise that ethical considerations do not only target the process of data collection and analysis (i.e. informed consent to access, record, store and analyse language/texts) but they also concern the process of data presentation and dissemination (i.e. consent or permission to quote and present language/texts).

With respect to the above ethical considerations, media discourse, available in various forms of published press, radio, television and film, has attracted the attention of language and communication scholars, as mass media are deemed to offer ample and convenient access to language data (see Yates 1996: 30, for a similar argument about new media). With media discourse being produced for purposes other than language research and arguably readily accessible, research using such sources seems to kill two birds with one stone: eliminating any influence on the data caused by the presence of the researcher - what Labov (1972) called the ‘observer’s paradox’ - and quickly passing through ethical scrutiny from institutional review panels.

Understanding mass media research as low - or no - risk of harming the persons researched draws on a key premise that distinguishes personal data from material already available in the public domain. To quote the UK’s Economic and Social Research Council framework for research ethics,

‘While data collected and stored as a record at an individual level are considered personal data, material already in the public domain are not. For example, published biographies, newspaper accounts of an individual’s activities and published minutes of a meeting would not be considered “personal data” requiring ethics review, nor would interviews broadcast on radio or television or online, nor diaries or letters in the public domain.’ (ESRC 2015: 12)

Unpacking the notion of ‘already available in the public domain’ as used for press and broadcasting material, it appears that publicness, here, is understood in terms of three key assumptions. First, publicness is associated with a domain, i.e. a space that makes such material available and, thus, accessible to the wider public, including researchers (see also Giaxoglou this issue). Second, publicness is conceived in terms of the material’s purpose and targeted audience: material is produced to be consumed by the public, understood in the context of broadcasting as a mass and unidentified audience (cf. the ‘broadcast audience’ Marwick and boyd 2011: 128-129). Third, publicness is associated with the roles and relevant identities of the authors of such media texts: for example, media professionals, politicians, and sports figures are usually invoked in their professional or public roles in mass media contexts.

However, there are cases where the three aforementioned assumptions may not justify media material as public. One is programmes that are referred to as ‘public or audience participation programmes’ (Livingstone & Lunt 1994; Thornborrow 2015). Radio phone-ins and reality TV shows (like *Big Brother* or the *Jeremy Kyle Show*) involve ordinary people (i.e. non-media or non-expert professionals) who often disclose personal and sensitive information about themselves. In such cases, researchers arguably deal with what is, essentially, mediated personal data, if the latter is defined as ‘information relating to an identifiable living individual’ (ESRC 2015: 23). Thus far, though, such material is also considered low risk, given that ‘the information contained in the personal data has been made public as a result of steps deliberately taken by the data subject’ (UK Data Protection Act 1998). It can be argued that people who willingly come forward, agree to be auditioned, enter into contractual agreements with media organisations and put themselves in front of open microphones and TV cameras have made decisions that involve assessing, as well as confronting, the publicness of such media interactions. Although one can never be sure about the intentions of such individuals or their understanding of the potential (harmful) impact of public sharing of their personal life on air, researchers into media discourse have long considered such interactions as public. Here, publicness also invokes assumptions about the speakers’ or authors’ agency in making such material accessible in the public space and to be consumed by the public.

Against this backdrop of current authoritative ethical guidelines, the media discourse researcher does not need to seek informed consent for collecting and analysing such material. Her ethical stance involves primarily issues related to the process of data presentation and dissemination, i.e. copyright and respect to intellectual property. In the case of press, TV and radio, ownership of published material lies primarily with the media organisation. As a result, appropriate attribution of material quoted or analysed, as well as respect for restrictions regarding the legal reproduction of media content, constitute the main ethical concerns of the media discourse analyst. To what extent, though, does increasing media convergence challenge current perceptions of publicness and, thus, call for revisiting ethical decision-making in language-focused media research? After defining media convergence, the following section questions assumptions of publicness as material that are (i) accessible in a domain, (ii) produced for a mass public, (iii) authored by public or professional figures and/or (iv) the result of deliberate agentic decisions of individuals.

### **3. Media convergence: Questioning ‘old-new’ dichotomies and ‘publicness’**

Media convergence is a key phenomenon in the contemporary mediascape. From a technological perspective, convergence refers to the merging of previously distinct technologies within a single device or media platform. As a sociocultural phenomenon, Jenkins (2006: 2) defines convergence as ‘the flow of content across multiple media platforms, the cooperation between multiple media industries, and the migratory behaviour of media audiences who would go almost everywhere in search of the kinds of entertainment experiences they wanted’. In other words, convergence resists dichotomies between ‘old’ (e.g. press and broadcasting) and ‘new’ (e.g. email, online forums, Twitter, Facebook, etc) media, by foregrounding the increasing integration of multiple media platforms and the mobility of content across media environments. In the context of broadcasting media, for example, there are three key ways in which intersections between ‘old’ and ‘new’ media manifest. Firstly, public participation from TV and radio audiences is no longer limited to telephone calls or to the odd appearances of lay people in talk shows; instead, broadcast programmes increasingly rely on digital technologies for interaction with their audience and often embed emails, tweets,

and other digital texts in the flow of the TV or radio programme (Thornborrow 2015). Secondly, the public engages with ‘old’ media snippets across digital platforms through commenting on, copying-and-pasting, liking, sharing, editing and remixing TV or radio material (e.g. Page 2012, Georgakopoulou 2014). Thirdly, and primarily discussed in this article, media industries work together to augment the inter-medial presence of news corporations, TV and radio programmes through dedicated websites, Twitter and Facebook profiles, used both for (re)broadcasting content and interacting with different audiences.

The challenges media convergence raises for researchers studying language use online have been repeatedly pointed out in the literature (e.g. Androutsopoulos 2011: 281; Herring 2013: 5; Bolander & Locher 2014: 16; Georgakopoulou & Spilioti 2016: 4). Among the key methodological and theoretical issues noted are: (i) the need for a methodological framework appropriate for analysis of ‘media coactivity’ (Herring 2013) or multitasking across and within media; (ii) attending to multimodality, as the range of available semiotic resources and their potential combinations have dramatically increased; and (iii) orienting to multiauthorship processes, as multiple participants are involved in the production of media content and any boundaries between media producers and consumers become increasingly blurred (cf. ‘prosumer’ Ritzer & Jurgenson 2010).

From a research ethics perspective, media convergence – together with the increasingly multimedial, multimodal, and multiauthored environments it affords – resists assumptions about media content as something available in a domain, i.e. given, static, located in a particular space, and, thus, accessible or amenable for collection and recording. Instead, digital content is mobile and moves – or, ‘flows’ in Jenkins’s (2006) words – across multiple sites for media engagement and activity. How can one discern whether this circulating and mobile content has been produced for a mass and unidentified public – if such a public exists in internet contexts? In a similar vein, how can one identify certain individuals as authors or discern the relevance of their professional or public roles to the communicative context, especially in circumstances of multiauthorship production where boundaries between content producers and consumers are no longer fixed or static? Last, but not least, how easy or clear is it for the ethical researcher to assess individual or deliberate agency in making personal content public, especially in environments where, as we will see, social media plug-ins may create conditions of automatic disclosure? These questions unsettle the key assumptions that arguably underpin established distinctions between personal data and material available in the public domain, complicating ethical decision-making, on the one hand, and questioning the scope of applying this distinction to digital contexts, on the other.

With reference to research on multilingual writing in digital environments, this article revisits the above questions in the context of websites that bring together content produced by multiple authors, for multiple publics and through combination of multiple modes and media. As will be shown, assumptions about the publicness of perceived public facing websites can be nuanced through attention to the ways in which individual areas, or modules, of a website are organised, combined, and discursively constructed by – and for – different social actors.

#### **4. Websites as sources for research on multilingual writing**

Unlike social network sites or personal websites as sources for language-focused research, official websites run by institutions and professional organizations arguably represent a low risk digital environment from the point of view of research ethics. With respect to criteria for

publicness, material embedded in such websites can be considered to be ‘already in the public domain’ (ESRC 2015: 12), as the web domain is created and regulated by relevant professionals (i.e. web designers and organization represented), with content produced for public or mass consumption. Nevertheless, demonstrating shortcomings of the public-personal data distinction even in these less controversial areas can be revealing of the need to radically rethink and move beyond received dichotomies.

As mentioned in the previous section, a key factor in challenging current ideas about publicness and ethics is media convergence. In the context of broadcast organizations, convergence between the so-called ‘old and new’ media is manifest, among others, in websites dedicated to popular TV and radio programmes. Such websites typically have the following purposes: to provide the programme’s audience with further opportunities to (re)consume broadcast material, to enable them to engage with a wealth of other material relevant to the show’s topic (e.g. alternative news sources) and to interact with media professionals (e.g. show presenters) and other audience members. These purposes usually shape - and are fulfilled in - two distinct communicative spaces: (i) an edited web space that features content designed and/or selected for (re)broadcasting to TV/radio and web audiences (cf. section 5.1) and (ii) an interactive space that includes messages posted by visitors to the website who often identify themselves as audience members of the particular programme (cf. section 5.2). Content in both areas is highly dynamic and volatile, with material and messages updated (i.e. added or deleted) and reconfigured on a regular basis.

From the perspective of language-focused research, the conditions of mobility, multiauthorship, multimodality, and multimediality afforded in such environments often result in the mobilization and, at times, strategic and reflexive configuration of multiple linguistic resources, such as styles, registers, codes, etc (cf. Deumert 2014; Tagg 2015). For that reason, the developing field of the sociolinguistics of writing acknowledges the central role of research on digitally-mediated communication in reconceptualising writing (Lillis & McKinney 2013: 421-424) and pushing written multilingual discourse from the periphery to the core of sociolinguistic research on multilingualism (Sebba 2013: 98-99).

#### **4.1. ‘Greekophobia’ and multilingual writing online**

To illustrate the challenges to ethical decision-making and the fuzzy boundaries of publicness when studying multilingual writing in digital environments and, particularly, in websites associated with ‘old’ media programmes, I draw on ongoing research on media representations of language variation and hybridity. Previous research has shown that social representations of stylistic variation which people find funny in media performances can be revealing of the socio-cultural values associated with certain language and cultural practices at a given moment (e.g. Coupland 2001). In recent work, I focused on a satirical radio show (Ελληνοφρένεια, lit. translated as ‘Greekophobia’) that is broadcast live on one of the Greek national radio stations (Real FM). The show has a strong political edge, using banter and pranks to play with, ridicule and criticize prominent political figures, institutions, widely held beliefs and stereotypes. In my study of the radio show’s pranks targeting members of the public, I discussed how perceived competence in English is increasingly becoming an index of social identity in the contemporary Greek context, giving rise to stylizations of (Greek-accented) English as a powerful device for subversion and humour (Spilioti forthcoming).

Since 2007 when the show first aired on national radio, ‘Greekophobia’ have expanded their inter-medial presence: there is now a TV programme (with the same title) using satirical

sketches to comment on current political affairs. The show also has a Facebook profile page, a Twitter account followed by 395,000 users (at the time of writing), and an official website ([www.ellinofreneianet.gr](http://www.ellinofreneianet.gr)). The website is updated daily, as each TV and radio show is uploaded after its live broadcast, and provides listeners (and viewers) with an online archive of all aired shows. Typical of the volatile nature of digital environments, the web domain of 'Greekophrenia' has changed its overall design and user interface three times since 2012 when I first turned my attention to the site. Initially, it hosted its archive, together with a wealth of material from other websites, such as video documentaries, news articles, blogs, comments from fans and visitors of the site, links to other web environments and simple interactive games, where the goal was, for instance, to throw tomatoes at prominent political figures. In 2013, the homepage of the website was revised and some of the initial content (e.g. interactive games, links to external websites about recommended books) was not included; visitor comments were nevertheless retained as a key feature in the sub-pages archiving the radio and TV episodes. Since December 2015, the comments section is no longer available and the main navigation hyperlinks have changed, though with content from different media sources still featuring across the website. Similar to other websites dedicated to TV and radio programmes, the Greekophrenia webpages combine spaces for content (re)broadcast and interaction with/among audience members, include material authored by (and for) different people, invite visitors to engage with multiple activities (reading, listening, watching, gaming, commenting), and embed volatile content that is added, and deleted, on a regular basis. In other words, content is far from static resulting in ethical tensions for the researcher who deals with sources whose status of public accessibility shifts through the different stages of the research process.

For the language-focused researcher, the Greekophrenia website displays language hybridity in the form of mixing Greek and English elements. Unlike the pattern of English/national language bilingualism, identified by Androutsopoulos (2012) as 'English-on-top', in German and Greek websites, English is represented through marked visual and graphic (e.g. typographic) resources, such as archaic typefaces and Greek alphabet letters. Parallel to the radio stylizations of (Greek-accented) English, the website documents stylistic variation in written representations of English and Greek evident not only in the edited web content but also in comments by fans. In order to address such representations of written variation and multilingual writing, my research focuses on how these resources are strategically deployed in the different areas of the website and how they are used (reacted to, and interpreted) by the range of content contributors to the website.

In terms of research ethics, when one makes the transition from researching a broadcast radio programme to examining webpages, particularly those containing comments from fans, is one still dealing with material already available in the public domain? Considering that there are usually two distinct communicative spaces in such websites, we might want to address this question separately in each area. The edited web space includes content that is accessible to all internet users and intended for public distribution and mass consumption. It also undergoes processes of heavy editing, selection and artful presentation in line with the professional practices and standards of relevant industries. The interactive space with visitors' comments can be understood as an alternative format enabling audience participation and interaction with the programme's producers/presenters. For example, comments that clearly address the radio producers, in the case of the Greekophrenia website, are reminiscent of telephone calls by radio listeners who often congratulate presenters on their programme.<sup>1</sup>

---

<sup>1</sup> Since comments are no longer accessible or available, as they have disappeared from the site, I am offering, here, constructed examples that echo the content and tone of messages appearing in that section:

Following current guidelines regarding research on media discourse, it appears that both spaces can be considered as belonging to the public domain. Both edited web content and comments are embedded within a wider domain that is created and regulated by professionals, i.e. the show's producers/presenters and web designers, for mass consumption. In addition to the fact that visitors deliberately post a comment, the topic of such comments also orients to the public broadcast of the show, presenting striking similarities with spoken material one could research without consent from the lay people participating in such programmes. So far, it appears that questions about ethics in relation to web content on sites dedicated to TV or radio programmes can be answered in terms of well established procedures for researching media (e.g. broadcast) material.

Tracing trajectories of practices such as audience participation and interaction with media professionals across media is important because it offers a more contextualised understanding of communicative topics, purposes and roles. This is particularly useful at the start of a project, because it helps the researcher to avoid being trapped within the confines of the argument of equating access with publicness and to resist being caught in the popular and a-historicised hype of 'newness', whereby 'new' media involves 'new' procedures. But, as will be shown in the following section, there is value in revising such initial claims by paying attention to the individual building blocks of websites and developing gradually a modular and iterative approach to ethics.

## **5. Opening Pandora's box of 'convergence': Revisiting publicness in the research process**

What counts as data in a multi/bilingual website, together with the ethical issues raised by such methodological decisions, is a reflexive process shaped and driven by one's research question and wider epistemology. Recent research on multilingualism and writing (Sebba 2013; Lee 2016: 121; Tagg & Seargeant 2016: 341) has stressed the multimodal aspect of such phenomena, driven by a shift in sociolinguistic research towards writing as an (ideological) semiotic process (Lillis & McKinney 2013; Maybin 2013: 554). This multimodal approach involves taking into account the 'visual contexts' of multilingual discourse, 'as a text surrounded by other texts, potentially with differing font sizes, colours and styles all potentially providing context for interpreting the content of the [multilingual] text' (Sebba 2013: 102). This theoretical shift in the approach of the written word as a (multi)semiotic object has methodological consequences: from the decontextualized written word as unit for analysis, the focus shifts to the study of writing in multimodal ensembles, i.e. multimodal representations or interactions that are 'seen as a material outcome or trace of the social context, available modes and modal affordances, the technology available and the agency of an individual' (Jewitt 2013: 254-255). Following this approach, attention to users' – and, at the same time, web designers' – strategic use of visual means to organise content on a webpage helps the researcher to delineate multimodal analytical units within which multilingual writing can be explored in relation to the surrounding linguistic, graphemic, and visual resources. Revisiting ethics iteratively within and across such multimodal ensembles is paramount as the varying social and material constraints at play in each unit impact upon perceived boundaries between public and personal data.

---

'Congratulations on your show – the best satire on the Greek radio!'; 'When will you upload the prank with the old lady asking for Oprah Winfrey's show?'; 'Well done for taking the mick of this government – corruption!'.

## 5.1. Ethics and edited web content: towards a modular approach

Edited content on official websites of media institutions and professional organisations is generally thought to belong to the producers and owners of the web domain. This approach to web domain as an undifferentiated whole is challenged by research on web discourse that foregrounds ‘modularity’ as a key organisational concept (e.g. Pauwels 2012: 251): content of individual webpages is organised into building blocks or modules that have been selected and combined, either advertently or inadvertently, into a coherent whole. As a result, in addition to the aforementioned methodological issues pertaining to the study of writing, the identification of modules that make up a website is useful for processes of ethical decision-making, as each building block may have different conditions of authorship and, thus, perceived ownership of creative content.

In order to identify the relevant modules on the multimodal space of a web page, visual elements can be combined with functional criteria. Visual elements (as suggested by Kress & van Leeuwen 2006: 214-218), particularly lines that frame and separate spatially contiguous texts, as well as colour, font and shape discontinuities that distinguish visually different areas of the web domain, have been particularly useful for identifying areas likely to operate as separate modules on edited web content on the *Greekophrenia* website. Drawing on visual resources, the identification of relevant building blocks also takes into account the extent to which the visually and spatially distinct modules fulfil various functions. With respect to websites dedicated to TV / radio programmes, such functions orient to (i) information structure or organization (menus, internal links, navigational tools), (ii) (re)broadcasting of media material (external links and other embedded textual, audio and visual material), and (iii) interactivity with site visitors (internal links and prompts to post comments).

The application of these criteria to websites reveals that different modular parts entail varying attention to copyright issues. For example, elements designed by the site developers, such as the ‘*Greekophrenia on the net*’ logo, site menus and internal links, represent creative content belonging to the site. Proper attribution and permission by the creators and owners of the website to reproduce such content in published material is necessary, similar to practices used for reproducing mass media material. In contrast, decisions about copyright permission become less straightforward when considering other areas that embed and (re)broadcast, for example, articles or news bites from journalists or bloggers who are not directly involved in the site design and who probably produced such texts for different audiences. The multiauthored composition of edited web content, together with the multimedial make up of the web area embedding the show’s radio or TV episodes, present cases where web discourse involves texts produced for different publics and exhibiting different conditions of media production (e.g. broadcast vs. internet). The challenges arising from convergence, here, concern the need to clarify copyright issues related to the reproduction of creative content for presentation and publication purposes. Contacting the site developers and owners and seeking permission for reproduction in academic research is the starting point.

Although it is advisable to contact web developers and owners of individual sites, digital affordances for copying, pasting, embedding and circulating content often obscure ownership and complicate copyright permission. In such cases, researchers can think of alternative ways for (re)presentation of web material. In fact, what seems like an obstacle at first sight can be a useful research tool. Publishing research on any type of discourse data involves a process of selective presentation of data in order to illustrate and support the arguments made. Rather than approaching potential issues with reproduction of multimodal and multiauthored discourse as

mere challenges, the time is ripe for developing appropriate methods for presenting such sources where copyright clearance may not be adequate or logistically possible. Similar to transcribing and presenting (multimodal) spoken discourse, arbitrariness of symbols and selectivity in discourse (re)presentation are inherent properties of this part of the research process. Research on multimodal and multilingual writing can draw on publishing practices of discourse analytic research on artwork, where strict copyright restrictions do not allow the reproduction of such material for publication purposes. Nevertheless, this does not hinder researchers from examining and analysing such material as evident in Jaworski's (2014) study of multilingualism and heteroglossia in artwork by multimedia and performance artists. In such cases, visual material can be (and are) represented through the use of diagrams (Jaworski 2014: 154). The following section makes use of a diagrammatic representation of web discourse in order to discuss ethical and privacy issues raised by research on the interactive areas of websites dedicated to TV and radio shows.

## 5.2. Multimedia convergence and interactivity: from copyright to privacy issues

One of the most distinctive features of language-focused research on multimedia and multilingual texts, such as those found on websites, concerns the meaning-making dynamics associated with hybrid, heteroglossic and multivoiced environments (Deumert 2014: 120). Such environments are often afforded through social media plug-ins activated in traditional web (1.0) environments to enhance participation and interactivity. On the Greekophrenia site, Facebook plug-ins enable visitors and fans to post comments under embedded broadcast material, such as past radio and TV episodes.

Typical features of multimedia web environments can include (i) verbal and visual elements (i.e. texts and pictures) designed by the site's developers and owners; (ii) verbal and visual elements designed by the social media platform developers and automatically embedded in the website through plug-in services; and (iii) verbal and visual elements designed by members of the public and posted on the host website. To illustrate how such elements are laid out on the Greekophrenia website (from 2012 to 2015), the following diagram annotates material originating in different domains, produced by different actors (site's producers in **bold**; social media web developers in *italics*; and members of the public in normal font) and, as will be argued below, for different audiences. At the same time, information about the varying linguistic and typographic resources mobilised in each area is provided.

[insert diagram 1 here]

Similar to edited web content, the issue of how to deal with creative content designed by different actors but integrated within the same digital space is also relevant here. The comments section, though, foregrounds an additional complication: unlike the articles, extracts from the broadcast radio/TV shows, and other creative content chosen for inclusion by the site's designers/owners, the comments section includes personal information that is automatically generated from the integration of social media plug-ins. Whose permission does one need in order to use such material for analysis, presentation or publication? In other words, to what extent can we safely assume that automatically disclosed metadata, as well as comments generated by users, in such interactive web areas are produced for a mass public, authored by public/professional figures and/or the result of users' deliberative acts (cf. section 2)?

Attention to the modular organisation of web spaces and the multiple semiotic resources activated in such digital environments provides a more nuanced picture of interactive comment areas on such sites. The diagram reveals the individual comments authored and posted by multiple users, as well as the range of accompanying metadata that are automatically embedded in that area of the website (see elements in italics, on the diagram). For example, the Facebook plug-in automatically displays the profile picture of the commenter, their Facebook user name and selected information from their personal profile (i.e. where they work, where they live, school/university they went to). Ethical decision-making regarding analysis and dissemination of such information raises key issues regarding the publicness of such material. With respect to who such material is produced for, metadata information (profile picture, username, etc) is arguably produced, first and foremost, for the Facebook audiences, rather than visitors to the radio show's website. Whether such information is made available to the site's visitors as a result of the commenter's deliberate act is debatable as this type of personal metadata is automatically transferred through the technological specificities of such plug-ins. More specifically, based on my own experience with using the plug-in on the particular site, users do not have control over the inclusion or exclusion of information, and they cannot post comments as anonymous users. In addition, there is no alert informing commenters about the type of metadata/information that will eventually appear in the comments section. Although one could argue that names and tone of voice also give and give off (in Goffman's 1959 terms) personal identification information about listeners who call radio stations to speak to their favourite presenters, the lack of control by social media users over such transferable personal information is a factor researchers need to take into account when making ethical decisions.

According to current ethical and legal frameworks, the use of such personal identifiable information without consent can be problematic as there is limited (or no) evidence that the metadata became available as a result of steps deliberately taken by commenters to disclose their profile picture or place of residence. In fact, the discussion of metadata in the context of online ethics also demonstrates the limits of current ethical guidelines. The internet researcher is often placed in the unenviable position of second-guessing participants' intentions or reconstructing the steps participants might have deliberately taken in order to post a comment. Participant observation (as I did in my study) and interviews with users can be useful tools for gaining further insights about the process of participating in an online environment. Nevertheless, such tools are still limited as users can over/underreport their intentions and platform interfaces are frequently updated, altering the steps users take to participate and upload comments. One wonders whether ethical recommendations should shift their focus from participants' intentions to researchers' responsibility over information inadvertently embedded in the building blocks of a website. While the process of determining whether data is public becomes increasingly elusive, we (as researchers) can be more sensitive to information we make available to a range of publics through our research. Alertness and sensitivity to the handling of metadata information is important, especially since, unlike tone of voice or accent in recorded media talk, online information can be searched and linked to a person's online (and offline) profile (cf. Zimmer 2010: 319).

In addition to metadata, the interactivity patterns observed in the comments exchanged between visitors of institutional websites call for revisiting claims about publicness that draw on assumptions about discourse produced for public consumption. The interactive affordances of social media plug-ins (e.g. 'reply' or 'like' buttons) often result in exchanges of comments between individual users. In the case of the Greekophrenia site, although the majority of comments are directed towards the show and the site producers, there are cases where users interact between themselves. Such one-to-one publicly accessible exchanges often take the

form of overt disagreements as to what counts as good satire or appropriate humour. Approaching such material as targeting a mass and unidentified public is challenged by users' strategic use of addressivity (referring to the other person's username) as a means for identifying intended addressee(s). These means for users' negotiation of multiple audiences in digital communication indicate varying orientation to different participation frameworks and active engagement in discursively constructing spaces as more or less public (Tagg & Seargeant 2016: 349; Marwick and boyd 2011). In such cases, public-private can be understood as a continuum that is discursively constructed and negotiated by users, rather than a checklist, as put forward current regulatory approaches to ethics (cf. Giaxoglou; and Mackenzie this issue).

Previous research on internet data (e.g. Bolander & Locher 2014) has also pointed out the sensitivity of topics as an important factor in assessing the publicness of a communicative space. Nevertheless, even topics that appear non-sensitive could still raise ethical concerns if they are embedded in discourse activities that are face threatening. On the Greekophrenia website, public debates about what counts as good satire tend to devolve into personal attacks at another user's political beliefs or ethnicity, through comments on another user's choice of profile picture or writing style. In other words, what starts as a public (or safe/ non-sensitive) topic can often take the form of personal disagreement, foregrounding sensitive aspects related to a person's identity and resulting in highly-charged interactions which neither parties might want exposed to external scrutiny. Given the fuzziness of 'topic', alternatives might include the notion of discourse activity that focuses on what is done and achieved through discourse, rather than what is said.

Furthermore, participant observation in such sites provides researchers with valuable information about control and regulation of information, informing claims about the extent to which site owners may intervene in their professional or public role. In the case of Greekophrenia, the scarce evidence of moderation from the show or site producers resists an interpretation of the comments area as a space owned and regulated by the site developers and owners. Although the comments section falls under the spatial boundaries of the website as a whole, close observation of the aforementioned disagreements and personal attacks reveals no apparent intervention or moderation from the site developers, in the form of either deleting or responding to posts. Assumptions about how content flows in highly convergent media environments (afforded by social media plug-ins) should pay attention to each separate building block, as control over content by website owners varies across the different modules that make up a website.

The modular approach undertaken in this section reveals that assumptions about publicness are challenged by the different elements that are combined on a website. The use of social media plug-ins to afford interactivity and participation blurs perceived boundaries between public and private. More importantly, content travels across platforms, with personal and private information, such as the user's picture or location on their Facebook profile, becoming publicly available as metadata that frame individual comments. Similarly, what starts as a public debate on non-sensitive topics can soon take the form of personal attacks to specific individuals, unsettling any clear-cut public-private distinctions and inviting the researcher to revisit such decisions by focusing, for example, on unfolding discourse activities, rather than topics.

## **6. Towards a modular and iterative approach to ethics in applied linguistics**

Challenges related to media convergence arise from the interplay of multiple modes, authors and media material. Approaching a website as a single domain or space runs counter to the modular make up of such communicative environments which, in turn, resists any blanket ethics statement about the publicness of institutional websites. Instead, unpacking media convergences and any associated ethical challenges foregrounds the need for a more contextualised and process-based approach that addresses ethical issues at different junctures of a research project and opens a dialogue between regulatory-driven ethical guidelines and context-sensitive understandings of ethical dilemmas (AoIR 2012: 4-5).

In the case of dealing with intersections of ‘old’ and ‘new’ media (as in the example of radio/TV programmes’ websites), tracing trajectories of media practices enables the researcher to enrich and contextualise understandings of publicness. It moves beyond equating publicness with accessibility and adds further dimensions to understandings of the public domain, by taking into account discussion topics (cf. Bolander & Locher 2014: 17; Buchanan 2011: 94; Sveningsson 2004) and professional roles of media producers (cf. similar distinction between amateur and commercial web environments in Androutsopoulos 2008: 9). At the same time, a process-based approach involves revising ethical decisions during different research stages. The research questions and methodological premises shape the process of ethical decision-making. For instance, methodological decisions about the study of multilingual writing online, ranging from the written word as individual, decontextualized units to writing as ideological (multi)semiotic process, have an impact on how the ethical researcher deals with informed consent and copyright issues (cf. Page et al 2014: 60).

In terms of consent for using web material for research purposes, a modular perspective to ethical decision-making contributes to the development of a process-based approach to research ethics (AoIR 2012). As shown in the article, websites consist of several, functionally and spatially distinct, areas where conditions of multiauthorship and multimodality are (re)configured in various ways. Different configurations involve varying copyright restrictions and understandings of publicness across and within the pages of a website. Attention to the design, multiple audiences, discourse activities, and interactivity patterns in each individual area is important both for methodological (i.e. digital writing as hybrid, multisemiotic and multivocal) and ethical reasons, as digital content is generated through various means. The article also foregrounds the need for careful ethical consideration of metadata that might get inadvertently collected by the internet researcher. Metadata, such as users’ profile picture and other personal information, are often automatically embedded in websites that use social media plugins. In this case, human agency and publicness as framed in ethical guidelines (i.e. as steps deliberately taken by human subjects) need to be revisited in order to account for technologically mediated agency that involves instances of automatic disclosure of personal information.

Together with a modular perspective, the process of ethical decision-making online is iterative. Understandings of publicness have been revisited across the different stages of the research and across the different modules that make up the main web pages of the site. Concerns about publicness extend to the current moment of writing the article and, thus, the stage of publishing and disseminating research, particularly because visitors’ comments are no longer available on the site. What sort of ethical decisions should one make when certain material no longer exists in the public domain but only in the private record of the researcher? It seems that the process of academic research is also caught up in the ethical conundrums of online scalability and spreadability (boyd 2011: 48). It is a truism that the speed and scale at which online content is circulated are unprecedented but academic research and publishing have always played their

own role in spreading and scaling up (certain) types of discourse. My response to this ethical conundrum was to represent the different areas and scripts in a diagram, rather than including screenshots or verbatim messages from the comments section. Although this would be a questionable practice for publishing results and findings from a language-focused study, such an instrumental approach to research presentation (i.e. moving beyond the publishing of verbatim quotes if justified by the aims of the study) can be useful in digital contexts where public-private discussions are muddying the waters of research ethics.

The process of transforming recorded material in line with ethical commitments is well rooted in research practices of transcribing and analysing spoken discourse (Ochs 1979). Compared to collecting and transcribing speech, written discourse (e.g. written documents, newspaper articles) lends itself conveniently to academic research. In other words, it deceives researchers into thinking that writing is already available for scrutiny, i.e. given material or, to use the Latin term, data ('given') ready to be studied and analysed (Markham 2013). The element of deceit lies in the fact that this reasoning collapses together understandings about the materiality of the sources (persistence of writing on a page vs. volatile speech) and epistemological assumptions about texts and people as objects and 'data subjects', respectively (UK Data Protection Act 1998). Perceiving writing as a static object produced by data subjects shapes our stance towards research on texts and people (cf. Cameron et al 1993: 87). Research on digital discourse carries over these assumptions about written discourse data. Perhaps we need to stop thinking about discourse data in terms of materiality and move beyond objectifying texts and people; then we can embrace richness and complexity as inherent properties of doing research (in any setting, face-to-face or mediated). This can take us a step closer to solutions that orient to research as process and shed light on, rather than obscure, the nuances of communicative encounters, online and offline.

**Acknowledgement:** I would like to thank the anonymous reviewers for their insightful comments and feedback, as well as Alison Wray and Caroline Tagg for their continuous support and constructive critique of earlier drafts.

## References

- Androutsopoulos, J. (2008) Potentials and limitations of discourse-centred online ethnography. *Language @ Internet* 5, article 8. Available online at: <http://www.languageatinternet.org/articles/2008/1610> (Last Accessed: 22 February 2016)
- Androutsopoulos, J. (2010) Localizing the global on the participatory web. In N. Coupland (eds.) *The Handbook of Language and Globalization*. Oxford UK: Wiley-Blackwell. 201-231.
- Androutsopoulos, J. (2011) From variation to heteroglossia in the study of computer-mediated discourse. In C. Thurlow and K. Morczek (eds) *Digital Discourse: Language in the New Media*. Oxford/New York: Oxford University Press. 277-298.
- Androutsopoulos, J. (2012) English 'on top': Discourse functions of English resources in the German mediascape. *Sociolinguistic Studies* 6 (2): 209-238.
- AoIR (2012) *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)*. Available online at: <http://aoir.org/reports/ethics2.pdf> (Last Accessed: 22 February 2016)

BAAL (2016) *The British Association for Applied Linguistics: Recommendations on Good Practice*. Available online at: [http://www.baal.org.uk/goodpractice\\_full\\_2016.pdf](http://www.baal.org.uk/goodpractice_full_2016.pdf) (Last Accessed: 20 July 2016)

Bolander, B. and Locher, M. (2014) Doing sociolinguistic research on computer-mediated data: A review of four methodological issues. *Discourse, Context and Media* 3: 14-26.

boyd, d. (2011) Social network sites as networked publics: Affordances, dynamics and implications. In Z. Papacharissi (ed.) *A Networked Self: Identity, Community and Culture on Social Network Sites*. New York/Oxon: Routledge. 39-58.

Buchanan, K. (2011) Internet research ethics: Past, present and future. In M. Consalvo and C. Ess (eds) *The Handbook of Internet Studies*. West Sussex: Blackwell. 83-108.

Cameron, D., Frazer, E., Harvey, P., Rampton, B. and Richardson, K. (1993) Ethics, Advocacy and Empowerment: Issues of method in researching language. *Language & Communication* 13 (2): 81-94.

Coupland, N. (2001) Dialect stylization in radio talk. *Language in Society* 30 (3): 345-375.

Deumert, A. (2014) *Sociolinguistics and Mobile Communication*. Edinburgh: Edinburgh University Press.

ESRC (2015) *ESRC Framework for Research Ethics*. Available online at: <http://www.esrc.ac.uk/files/funding/guidance-for-applicants/esrc-framework-for-research-ethics-2015/.pdf> (Last Accessed: 22 February 2016)

Georgakopoulou, A. (2014) Small stories transposition & social media: A micro-perspective on the 'Greek crisis'. *Discourse & Society* 25: 519-539.

Georgakopoulou, A. and Spilioti, T. (2016) Introduction. In A. Georgakopoulou and T. Spilioti (eds.) *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge. 1-15.

Goffman, E. (1959) *The Presentation of Self in Everyday Life*. London: Penguin.

Herring, S. (2004) Slouching toward the ordinary: Current trends in computer-mediated communication. *New Media & Society* 6 (1): 26-36.

Herring, S. (2013) Discourse in web 2.0: Familiar, reconfigured, and emergent. In D. Tannen and A. M. Trester (eds.) *Discourse 2.0: Language and New Media*. Georgetown University Press. 1-25.

Heyd, T. (2016) Digital genres and processes of remediation. In A. Georgakopoulou and T. Spilioti (eds.) *The Routledge Handbook of Language and Digital Communication*. London and New York: Routledge. 87-102.

Jaworski, A. (2014) Metrolingual art: Multilingualism and heteroglossia. *The International Journal of Bilingualism* 18 (2): 134-158.

- Jenkins, H. (2006) *Convergence Culture: Where Old and New Media Collide*. New York and London: New York University Press.
- Jewitt, C. 2013 Multimodal methods for researching digital technologies. In S. Price, C. Jewitt, and B. Brown (eds) *The SAGE Handbook of Digital Technology Research*. London: Sage. 250-265.
- Jones, R. (2011) C me Sk8: Discourse, technology, and “bodies without organs”. In C. Thurlow and K. Mroczek (eds.) *Digital Discourse: Language in the New Media*. Oxford: Oxford University Press. 321-339.
- Kress, G. and van Leeuwen, T. (2006) *Reading Images: The Grammar of Visual Design*. London and New York: Routledge.
- Labov, W. (1972) *Sociolinguistic Patterns*. Philadelphia: University of Philadelphia.
- Lee, C. (2016) Multilingual resources and practices in digital communication. In A. Georgakopoulou and T. Spilioti (eds.) *The Routledge Handbook of Language and Digital Communication*. London and New York. 118-132.
- Lillis, T. and McKinney, C. (2013) The sociolinguistics of writing in a global context: Objects, lenses, consequences. *Journal of Sociolinguistics* 17 (4): 415-439.
- Livignstone, S. and Lunt, P. (1994) *Talk on Television: Audience Participation and Public Debate*. London: Routledge.
- Markham, A. (2013) Undermining ‘data’: A critical examination of a core term in scientific enquiry. *First Monday* 18 (10). Available online: <http://firstmonday.org/article/view/4868/3749> (Last Accessed: 22 February 2016)
- Marwick, A. & boyd, d. (2011) ‘I tweet honestly, I tweets passionately: Twitter users, context collapse and the imagined audience. *New Media & Society* 13 (1): 114-133.
- Maybin, J. (2013) Working towards a more complete sociolinguistics. *Journal of Sociolinguistics* 17 (4): 547-555
- Ochs, E. (1979) Transcription as theory. In E. Ochs and B. Schieffelin (eds.) *Developmentall Pragmatics*. New York: Academic Press. 251-268.
- Page, R. (2012) *Stories and Social Media: Identities and Interaction*. New York and London: Routledge.
- Page, R., Barton, D., Unger, J. and Zappavigna, M. (2014) *Researching Language and Social Media: A Student Guide*. London and New York: Routledge.
- Pauwels, L. (2012) A multimodal framework for analysing websites as cultural expressions. *Journal of Computer-Mediated Communication* 17 (3): 147-265.

Ritzer, G. and Jurgenson, N. (2010) Production, consumption, prosumption: The nature of capitalism in the age of the digital 'prosumer'. *Journal of Consumer Culture* 10 (1): 13-36.

Sebba, M. (2013) Multilingualism in written discourse: An approach to the analysis of multilingual texts. *International Journal of Bilingualism* 17 (1): 97-118

Spilioti, T. (2016) Digital discourses: a critical perspective. In A. Georgakopoulou and T. Spilioti (eds.) *The Routledge Handbook of Language and Digital Communication*. London and New York. 133-148.

Spilioti, T. (forthcoming) Radio talk, pranks and multilingualism: Styling Greek identities at a time of crisis. In J. Mortensen, N. Coupland & J. Thøgersen (eds) *Style, Identity and Mediation: Sociolinguistic Perspectives on Talking Media*. Oxford: Oxford University Press.

Tagg, C. (2015) *Exploring Digital Communication: Language in Action*. London and New York: Routledge.

Tagg, C. and Seargeant, P. (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.

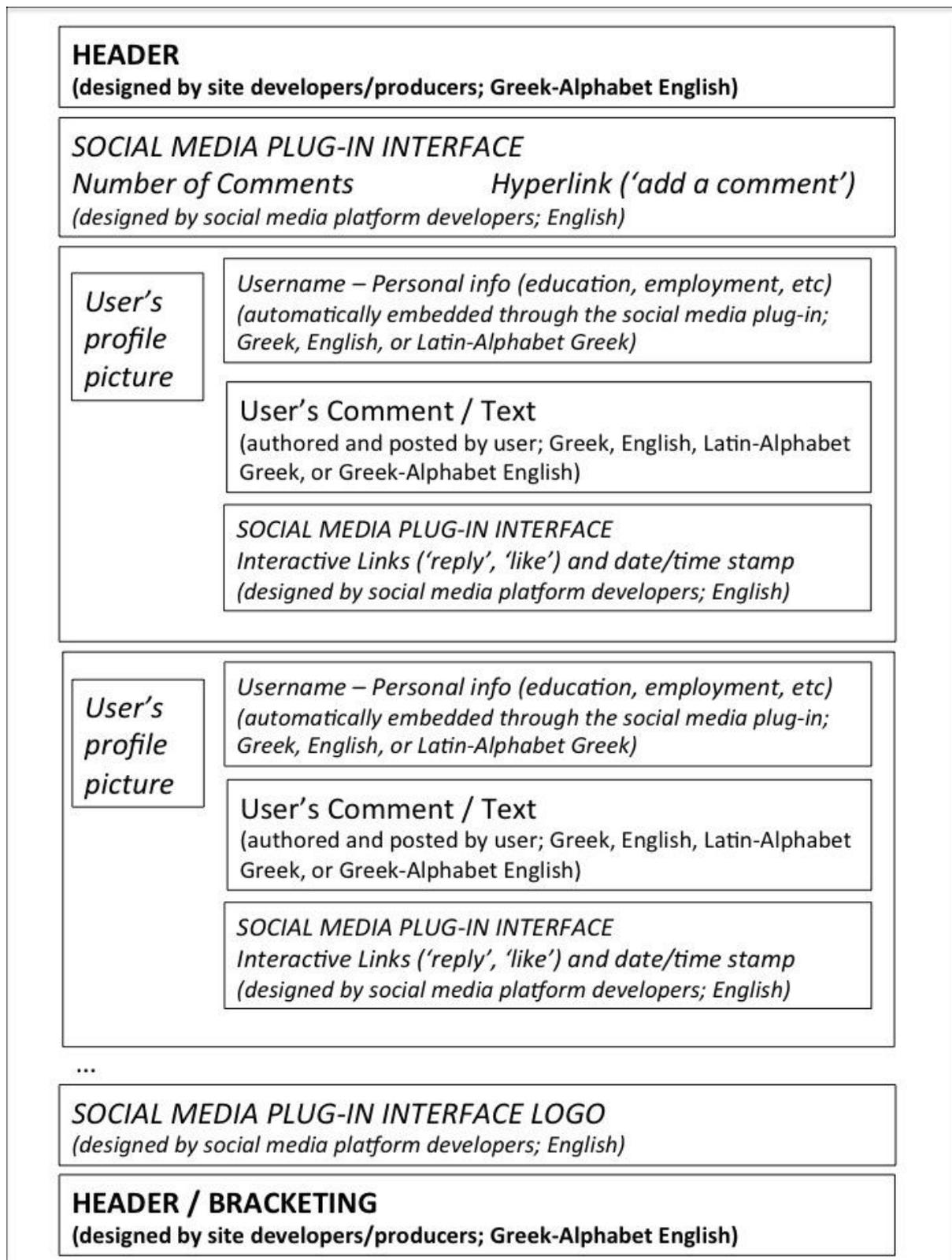
Thornborrow, J. (2015) *The Discourse of Public Participation Media: From Talk Show to Twitter*. London and New York: Routledge.

Thurlow, C. (2013) Fakebook: Synthetic media, pseudo-sociality, and the rhetorics of web 2.0. In D. Tannen and A. M. Trester (eds.) *Discourse 2.0: Language and New Media*. Georgetown University Press. 225-249.

(1998) *Data Protection Act 1998*. Available online at: <http://www.legislation.gov.uk/ukpga/1998/29/contents> (Last Accessed: 22 February 2016)

Yates, S. (1996) Oral and written linguistic aspects of computer conferencing. In S. Herring (ed.) *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*. Amsterdam/Philadelphia: John Benjamins. 29-46.

Zimmer, M (2010) "But the data is already public": On the ethics of research in Facebook. *Ethics, Information and Technology* 12: 313-325.



**Diagram 1.** Modular representation of comments section ([www.ellinofreneianet.gr](http://www.ellinofreneianet.gr) 2012-2015)