From variation to heteroglossia in the study of computer-mediated discourse
Jannis Androutsopoulos (University of Hamburg)

Variation and variation analysis in computer-mediated discourse studies
Despite the novelty of approach that labels such as ‘Internet linguistics’ (Crystal 2006) seem to imply, most language-focused research on computer-mediated communication – or perhaps more accurately, computer-mediated discourse – tends to situate itself within established approaches in socially oriented linguistics. As a consequence, the state of the art in CMD studies is characterized by attempts to adapt established methods of data collection and analysis to new environments of discourse, and new techniques in Internet research tend to supplement rather than replace concepts and methods from linguistic scholarship (Herring 2004, Androutsopoulos/Beißwenger 2008). Variation is one such key concept in CMD studies. As any search across classic publications in the field goes to show, variation is an ubiquitous keyword understood in a number of ways: In terms of linguistic structure, it is usually viewed as variation between “speech” and “writing” or standard and vernacular linguistic forms. In terms of contextual factors, language variation online has been associated with the effect of digital communications technologies on written language, or with the absence of institutional regulation, or with people’s opportunities to realize digital modes and genres in ways that are much more diverse than their apparent technological standardisation.

Within the broader context of "coding-and-counting" approaches in computer-mediated discourse analysis (Herring 2004), the study of linguistic variation is delimited from other quantitative studies of language online (such as Yates 1996, Tagliamonte/Dennis 2008) by using the linguistic variable as structural unit of analysis. This involves the identification of linguistic variables and the calculation of the frequency of their variants in correlation to linguistic or non-linguistic independent variables (Androutsopoulos 2006:424-6). Based on that premise, researchers have examined linguistic variables from various languages in different CMC modes and in relation to different independent variables. Let us briefly review these aspects: First, researchers traditionally focus on modes of synchronous interactive written discourse such as Internet Relay Chat and Instant Messaging (see Androutsopoulos/Ziegler 2004, Christen/Ziegler 2006, Franke 2005, Paolillo 2001, Siebenhaar 2006, Squires 2010), though forums, blogs, newsgroups and personal ads have
been studied as well (Androutsopoulos 2007, Hering/Paolillo 2006, van Compernolle 2008). The traditional variationist focus on phonology is replaced by the analysis of written representations of phonological variation between standard and dialect or formal and casual style (see work by Christen/Ziegler, Androutsopoulos/Ziegler, Franke, Siebenhaar). Cases of grammatical variation have also been studied (van Compernolle, Siebenhaar, Herring/Paolillo), but only a few studies have examined variation that is specific to written language, such as spelling substitutions and apostrophe usage (Paolillo, Squires). Mixed method studies that combine quantitative and qualitative analysis of style shifting have also been conducted (Androutsopoulos/Ziegler, Franke, Androutsopoulos 2007), and Paolillo’s (2001) social network analysis of language variation in an IRC channel is a combined study of five linguistic variables (three spelling substitutions, use of obscenity, and code-switching into Hindi). The relevant dimensions of variation are between standard and dialect as well as spoken and written style, even though concerns are sporadically raised as to what actually constitutes a ‘standard’ in online settings that lack institutional control (Paolillo, Squires). Independent variables examined include gender, region, age, genre (Herring/Paolillo 2006, Androutsopoulos 2007), type of chat channel (Siebenhaar), and network structure (Paolillo).

Despite its fairly limited empirical coverage in terms of languages (especially English, German, and French), this work has established that language variation online is socially and generically patterned. It has thereby contributed to a differentiated and de-exoticised understanding of CMD. However, in the remainder of this chapter I am concerned less with the achievements of variation analysis and more with its ‘blind spots’, its limits in exploring and theorizing language online. I therefore use this work as a backdrop against which to ask: What else? And what next?

**Limits of a language variation approach to language online**

In particular, five ‘limits’ of a language variation approach will be briefly discussed: its unimodal and monolingual focus, its reliance on the linguistic variable, on predefined independent variables, and on quantification. These ‘limits’ echo, in part, a broader critique of variationism within sociolinguistics, which must remain implicit in this chapter (see Auer 2007, Coupland 2007). At the same time, they reflect a gap between the state of art in CMD research and our experience of contemporary popular, yet still poorly researched digital environments such as social networks and media-sharing websites, which are of central importance to my argument. The contrast between these and ‘older’, but much better
researched CMC modes (such as emails, mailing lists, IRC and IM) is central to the following discussion.

As the literature review above suggests, a language variation approach to language online seems most efficient with CMC modes which enable users to approximate conversational interaction, are available in large and easily extractable volumes, and rely on written language as main resource for the construction of meaning. However, in contemporary web 2.0 environments (on which more below) language comes integrated in visually organised environments, verbal exchanges tend to be more fragmented and dependent on multimodal context, and meaning is constructed through the interplay of language and other semiotic means. Second, a language variation approach seems best suited for the analysis of relations within the same linguistic system, and therefore less well equipped to address multilingual discourse and code-switching, whose importance to CMD is by now well attested. Third, a language variation approach tends to exclude features that are not easily operationalised as linguistic variables. This may affect single semiotic features such as emoticons as well as discourse phenomena such as language or script choice. Fourth, the quantitative premise of variation studies implies that features that are scarce in frequency may be excluded, even though they might be pragmatically and socially meaningful in the data. Finally, the reliance on independent, non-linguistic variables may lead to a preference for data in which such variables can be construed. This in turn may marginalize CMC contexts whose classification in socio-demographic terms is not evident (see discussion in Androutsopoulos 2006: 424-5). The preference for clear-cut independent variables such as gender and age may reflect scholarly convention rather than the categories that are relevant to participants in online communication.

Awareness of such limitations is probably as old as language-centred CMC research, and the insight that CMC data is often more messy and unpredictable than what a language variation approach would be comfortable with, is not new. One 1998 monograph concludes on the observation that CMC is characterised by linguistic elements and fragments from different discourse worlds, which are put together to a specific “style-mix” (Runkehl/Schlobinski/Siever 1998:209). Five years later, Georgakopoulou (2003) suggests that "CMC is by no means a homogeneous and centralized site: in contrast, it encourages hybridity, diversity of voices and ideologies, and expression of difference". Another five years later, Tagliamonte/Dennis (2008:26) view a "quintessential characteristic of IM discourse" in the "consistent juxtaposition of 'forms of a different feather'”. While to some researchers this is no more than a footnote, others have used ideas from pragmatics,
conversation analysis, style and interactional sociolinguistics in order to study CMC modes not primarily as technological containers of speech, but as sites of users’ social activities with language. In such approaches, the classification of language use on dimensions of variation is complemented by an attention to the situated exploitation of linguistic difference, which doesn’t shy away from the importance of singular, unrepeated instances of linguistic difference if and as used in a strategic, yet non-quantifiable way. Likewise, the correlation to pre-defined social categories is replaced by a focus on identities as discursive constructions which participants claim and negotiate by drawing on a variety of semiotic means (e.g. Danet/Ruedenberg-Wright/Rosenbaum-Tamari 1997, Tsiplakou 2009). This and other work offers examples for the code-centred stylistic choices that Georgakopoulou (2003) views as characteristic of informal CMC. It suggests that what characterises language and discourse online is neither a specific new pattern of variation between two predetermined poles nor a new language variety, but rather a heightened attention to all aspects of written language as a key mode of signification.

As far as language use online is concerned, then, variation analysis leaves a number of ‘blind spots’, which have been partially circumvented by focusing on specific CMC modes at the expense of others. Interactional and ethnographic research has acted to some extent as complement and corrective to variation studies, and its insights resonate with the heteroglossia approach outlined below. However, its focus still being on classic CMC modes, neither its research questions not its data take the socio-technological evolution of digital communication into account.

**Web 2.0: participation, convergence, and discourse**

Out of the overwhelming of digital communication modes available today, typical web 2.0 environments such as social networking and media-sharing websites predominate popular practice and imagination. Even though there is no easy and straightforward distinction between an 'old' and a 'new' web, these environments share technological, sociological and structural features that clearly separate them both from earlier stages of the Web as well as from the pre-Web applications linguists are so familiar with (Cormode/Krishnamurthy 2008). Some of these differences are captured by the concepts of convergence and participation, which, even though not directly referring to language, have important, yet not well-understood implications for language online (Androutsopoulos 2010). Simply put, participation relates to the accessibility of localized, bottom-up production and distribution of
online content, while convergence refers to the fusion of formerly distinct technologies and modes communication into integrated digital environments (Jenkins 2006).

An implication of participation from the perspective of CMD studies is that it enables specific sets of role relationships and repertoires of digital media practice. Web 2.0 is radically different to earlier stages of the Web in the modes of production and consumption it enables. Instead of a clear-cut role distribution between professional content production and read-only consumption, web 2.0 sites of social networking and media sharing feature “a co-mingling of commentators and creators, and every visitor has the opportunity to click, comment, create, etc.” (Cormode/Krishnamurthy 2008). In this context, the effect of convergence is a dramatic increase in the range of available semiotic and resources and their combinatory potential.

Drawing on concepts more familiar to sociolinguistics and discourse analysis, we might say that participatory and convergent digital environments are characterised by processes of multimodality and multi-authorship: their content is produced by multiple participants, simultaneously and in part independently of each other; and they host and integrate complex combinations of media and semiotic modes, including written text (and, increasingly, speech), standing image, moving image, colour and graphic design.

These dimensions of media and semiotic complexity, I argue, exhaust the potential of variation analysis and call for analytical and theoretical attention to current web environments: while most research on language variation online is dedicated to pre-web and web 1.0 modes, contemporary web surfaces are semiotic patchworks in which authors, media and voices co-exist in ways that make processes that were sporadically pinpointed by earlier research empirically ubiquitous and theoretically pressing.

**Introducing heteroglossia**

Heteroglossia sets in at this point, as one concept which, I argue, might offer an alternative perspective of linguistic heterogeneity online. However, heteroglossia is an elusive and slippery concept, and in this section I indicate its complexities besides discussing some of its definitions and analytical applications.

Originally introduced by Mikhail Bakhtin in his 1934 essay "Discourse in the Novel", the notion of heteroglossia has been embraced by literary studies and sociolinguistics (see Vice 1997:18-21; Lähteenmäki 2010). Bakhtin originally speaks of “the co-existence of socio-ideological contradictions between the present and the past, between differing epochs of the past, between different socio-ideological groups in the present, between tendencies,
schools, circles and so forth” (1981:291). From the range of available scholarly definitions, two I find particularly insightful are by Ivanov (2001:95) who defines heteroglossia as “the simultaneous use of different kinds of speech or other signs, the tension between them, and their conflicting relationship within one text”; and Bailey (2007:257) who suggests that “Heteroglossia addresses (a) the simultaneous use of different kinds of forms or signs, and (b) the tension and conflicts among those signs, based on the sociohistorical associations they carry with them”.3

Compared to sociolinguistic concepts like code-switching (Bailey 2007) and variation, the notion of heteroglossia entails a number of differences. Bakhtin locates it at a double level: in the novel, which uses language to articulate tensions between voices of different social origin and standing, and at the same time in the social reality a novel reflects or draws on (see also Lähteenmäki 2010:23). Therefore heteroglossia does not just occur, as one might say with regard to language variation, but is made: it is fabricated by social actors who woven voices of society into their discourses, contrasting these voices and the social viewpoints they stand for. Such an understanding of heteroglossia as the outcome of purposeful, and often artful, semiotic activity ties in well with the concept’s use to study relationships between linguistic diversity, social difference and power in media discourse.4 At the same time, it does not prevent its application to social interaction. As Bailey points out in his study of bilingual interaction, heteroglossia is more inclusive than variation in that it can address both mono- and multilingual discourse (2007: 258). Indeed, any kind of linguistic difference at different levels of linguistic structure can be potentially viewed as heteroglossic, to the extent it is drawn into meaningful oppositions by participants. Because heteroglossia “takes as its starting point the social and pragmatic functioning of language” (Bailey 2007:262) rather than linguistic form as such, it enables analysts to contextualize locally contrastive patterns of usage within larger social, historical, and ideological processes. Thus heteroglossia goes beyond the co-occurrence of and shift between languages or language varieties, and focuses on “the coexistence of different competing ideological points of view” (Lähteenmäki 2010:25) that are indexed by language in specific communicative situations.

However, the literature reviewed here has not produced an agreed methodology on how to identify instances of heteroglossia in discourse, which units of analysis to work with, and which levels of linguistic structure to consider. To the contrary, researchers have located heteroglossia in widely different discourse domains and processes, and while one might argue that such width of application is made possible by the fluidity of the concept, it arguably makes its application more complex and entails risks of subjectivity and lack of transparency.
Especially when contrasted to the clear-cut operationalisations of variation analysis, heteroglossia appears very difficult to operationalise (Bailey 2007:263). However, bearing in mind the critique of variationism above, this openness and flexibility should be thought of as an advantage. They allow us to “think big”, offer space to envisage heteroglossic relations between signs of various kinds and structural properties, whose coexistence and dialogue may be established at different levels of discourse.

**Locating heteroglossia in digital discourse environments**

Summing up the previous discussion, heteroglossia invites us to examine contemporary environments of computer-mediated discourse as sites of tension and contrast between linguistic resources, social identities and ideologies. However, the discussion also makes clear that heteroglossia is not a ready-made concept; it needs to be tailored to the conditions of discourse under investigation. A crucial issue is therefore how to locate its spaces of articulation: Is digital heteroglossia, one might heuristically ask, text- or practice-based? In other words, does it emerge in the eye of the beholder, or can it also be pinpointed as a constellation of textual/semiotic resources? What is its level of articulation? Is it the entire web? An entire website or just one full webpage? A thread of posts or just a single post within a thread? A single text within a web page or just some constituents of that text?

The quick answer is – all of the above. Heteroglossic relations can manifest at different levels of linguistic and textual organisation, and such relations can co-exist, owing to the convergent, multimodal and multi-authored structure of web 2.0 environments. However, any text-based analysis of heteroglossia requires an ‘anchor’, a pivotal point in discourse structure. With regard to social network and media-sharing websites, the anchor I propose in this research is the web page, including personal profile pages and pages hosting media items. Two caveats seem important here. First, we focus here specifically on dynamically generated web 2.0 pages (Cormode/Krishnamurthy 2008) with all their textual components featuring a variety of authors, genres, modes, media, and so on. Second, an emic justification for that choice is the observation that web 2.0 pages are functional units tied to social actors or media items – in other words, they “belong” to someone (such as a profile page or blog) or host a particular semiotic artefact (a video, photo or other media item). Web 2.0 pages constitute a mesolevel of discourse structure, providing rich context for the analysis of microlevel elements (such as posts) and constraining the overwhelming mass of content that is experienced at the level of entire web sites. Using web 2.0 pages as units of analysis is not a
handicap to an ethnographic approach to CMD, because the sense of ownership related to blogs and profile pages may motivate contact to and communication with content creators.

Against that backdrop, I now identify potential locations of heteroglossic contrasts, moving between the level of an entire web page and its textual constituents. It is especially at a microlevel of analysis, zooming on single utterances by the same speaker/author, that my observations resonate with previous findings on synchronous or asynchronous CMC. Consider an example from a prototypical web 2.0 environment:

**Excerpt 1: YouTube comment**

```
xXxCroatiaStylexXx (2 months ago)
des is doch echt so geil zefix oida^^
so sama hoit mia bayern :D^^
```

This comment refers to an amateur video, which appropriates a global pop song to stylize local identity in Bavaria, Southern Germany (discussed in Androutsopoulos 2010). Like most comments in this thread, it gives praise to the video (the first line reads: ‘this is really great, mate’) and asserts a collective local identity (second line: ‘that’s how we are, we Bavarians’). Both sentences are cast in Bavarian dialect, as evidenced in the orthographic representation of dialect features and the use of dialect lexis (such as the interjection zefix). What strikes me here is less the comment’s alignment to the language style of the commented video than the contrast between the comment proper and the screen name. While the former uses dialect to praise the video and claim (jocularly, perhaps) local identity, the latter is cast in English and signifies at the propositional level a different identity (namely, Croatian). I interpret this as heteroglossic: two different languages explicitly indexing two different identities, moulded together into one post yet at the same time differentiated in terms of its functional components. Here, as elsewhere, the elements participating in a heteroglossic contrast belong to different parts of a genre, and previous research suggests that the difference between a post proper and accompanying emblematic elements, such as names and mottos, is consequential to style and language choices. At the same time, the example makes transparent the inevitable, it seems, subjectivity of such analysis, and the tension between a screen-based and a user-based approach to CMD: That contrast is salient to me as analyst on sociolinguistic grounds, but not necessarily to other posters or viewers of that thread or indeed its author, at least as far as we can tell without asking them. I return to this point in the conclusion.
Such tensions of language and identity, viewpoint or participation role have of course been reported from older CMC modes as well, such as chat channels, forum discussions or email exchanges among friends or colleagues (e.g. Hinrichs 2006 and Tsiplakou 2009, in part with disparate labels). Moving one step further, they may manifest not just within single utterances but in the dialogic and sequential relation among utterances, such as threads of comments or posts, in which different participants draw on linguistic heterogeneity to contextualize specific readings of their contributions or to position themselves towards others (Androutsopoulos 2007).

Further instances of heteroglossic articulation can be identified in and around what I call ‘spectacles’: participatory multimedia content such as videos and photos, which is uploaded, displayed, commented and rated on media-sharing websites (Androutsopoulos 2010). The notion of spectacle operates at two levels: that of text that is made available to a web audience, and that of dialogue between the text and the responses contributed by that audience. Even though spectacles are not heteroglossic by default, their composition and commenting may entail elements of heteroglossia. This is especially true for spectacles which appropriate and recompose popular texts, e.g. by adding new footage, voice tracks, subtitles and so on. To the extent such modifications aim to stage tensions between life worlds or worldviews, or to express critical, oppositional or subversive views, for instance in the context of political discourse, heteroglossia emerges as key element in the grassroots participatory culture of media production (Jenkins 2006, Androutsopoulos 2010). The style choices of spectacles are framed by the emergent and interactively shaped choices of audience responses, potentially giving rise to additional tensions which may reflect different aesthetic and ideological responses to a spectacle. In the Bavarian context investigated in Androutsopoulos (2010), one instance of heteroglossic contrast between spectacle and comments indexes long-standing animosities between two neighbouring regions, while another instance expresses tensions between globally circulating media content and its local, parodist appropriation.

Moving “upwards” to the level of an entire web page, we find that user-contributed content may contrast to the linguistic design of the surrounding web interface. More specifically, heteroglossic effects can be located at the relation between components of a web page which are designed and authored by different individuals and institutions: user interface, institutional and participatory content, advertisement space. These components can all come in one single language, but often don’t, thereby reflecting relations of inequality between languages as much as user-specific conditions of participation. In my own experience with
social network and content-sharing sites, which I’d like to draw on here by means of example, users’ language choices often part ways from those of the website frame and those of advertisements. For example, as a speaker of Greek, my own YouTube experience is always framed by a language other than Greek, as that language is not (yet) included into the site’s language localization options. These are restricted to a handful of languages, including fairly small ones in terms of speaker numbers (such as Suomi and Svenska), but not Ukrainian, Punjabi or Arabic. As with other corporative websites (Kelly-Holmes 2006), decisions on linguistic localization are based on the power of languages in the digital economy rather than their mere number of speakers. Another example is facebook, where the linguistic activities on my own profile ‘wall’ are usually trilingual (in English, German and Greek), while entries on my ‘home page’ involve a larger number of languages as selected by my ‘friends’. In addition, there is by now quite an extensive choice of ‘primary languages’ for the website’s user interface. However, setting my interface to English or Greek will not stop Google ads targeting me in German, based on the automatic recognition of my Internet Protocol address. Similar tensions may emerge wherever contributions of different social origin or participation role coexist: “consumers” reviews and professionally authored copy in online shops, user reviews and promotional copy on web mapping applications, and so on. In all these cases, the composition of digital spaces can be thought of as heteroglossic, as the linguistic choices of various components reflect and reinforce boundaries between actual user activity and its multiple corporate framings.

**An example: heteroglossic relations on a profile page**
The previous section identified potential locations of heteroglossic contrasts in web 2.0 environments, from the micro level of a single utterance to the configuration of entire web pages, and from processes already familiar from earlier CMC modes to those that seem specific to the participatory and convergent structure of web 2.0. Against that backdrop, this section illustrates how an analysis of heteroglossia on the web might proceed. The example comes from MySpace, one of the most popular social networking websites worldwide (boyd 2008). It is the profile page of Mantoinette, a German musician in her twenties, which has been online since 2006. The version I analysed (by now modified) is from spring and early summer of 2009. Based on procedures of discourse-centred online ethnography (Androutsopoulos 2008), my analysis is based on repeated views of the profile page, selection of textual data for linguistic analysis, and an interview with the producer, all of which contributes to the interpretation offered in what follows.⁵
The analysis reconstructs Mantoinette's profile page as a multi-purpose discursive space where different types of relationships and discourse activities are acted out. Following the profile's linear order, it shows how M's own contributions and those by her 'friends' come together to form a semiotic collage that is reminiscent of 1990s personal homepages, albeit with a higher density of multimodality and multi-authorship. That collage is framed in a twofold way that lies beyond M’s control: by the viewer’s choice of the language of interface, which determines the automatically generated headings of various text boxes (e.g. X's Friends Space), and by the advert banners at the top of the page whose language follows the viewer’s location. Against that backdrop, heteroglossic relations will be identified in the ways different representations, activities and relationships constructed on the profile page are realized through distinct, and possibly conflictual, choices of semiotic resources.

Figure 1: Schematic representation of Mantoinette’s MySpace page in June 2009.
Not to scale. Profile's own descriptors are set in *italics*.
MySpace offers its users different opportunities to select and arrange page elements in individual ways (boyd 2008). Mantoinette’s profile is customised in some respects, such as colour design and embedded media, but rather standardised in others, especially sequential structure. A schematic representation of its components and their sequential order is given in Figure 1. One aspect this scheme does not capture is M’s customised use of background colour: the page background is a red floral wallpaper, and the various text boxes are given a dark red background, framed by a golden hairline. M’s ‘contact box’, an automatically generated list of contact options, is given the background photo of a stylus, thereby indexing, as M explains in the interview, her commitment to music. Another standardised text box is the 'calling card' with information about the profile owner, including her name, a photo, a tag line with her music style, her location (Germany), and a slogan (*thevoiceisthesisterofthesoul*).

M’s music player, to the right of the calling card, is a standard feature of musician profiles on MySpace. It offers at the time of observation four songs sung by M, all in English.

The longest pieces of verbal text on the page are M's self-presentation blurb (General Info) and the testimonial placed to its right (About). The former is quite short and fully in English (Excerpt 2), the latter is longer and comes in a German and an English version (Excerpts 3 and 4 with two sentences from each version).

**Excerpt 2: General Info**

**Contemporary soultunes move and influence me the most, I think.** In some ways I feel kinda „at home“ within rap, but soul defines it wonderfully concrete for me according to the attribute „feeling“.(shouldn’t mean rap has no feeling for me =))

**Excerpt 3: About, German version:**

Marie-Antoinette ist eine Künstlerin, die schon sehr früh ihren Lebensweg im Singen, in der Performance, in der Musik gesehen hat. [...] Die Jahre, die ich mit ihr arbeiten durfte, waren eine große Bereicherung, denn sie sucht unermüdlich den Kern in allen Dingen und bringt ihn dann stimmlich ausdrucksstark ans Licht. [...] 

**Excerpt 4: About, English version:**

Marie-Antoinette is an artist, who saw it’s way in singing, performing and music very early in her life. [...] The years I worked with her were a great pleasure because she’s straight looking for the source in everything and brings it to life with her voice very expressive. [...] 

Written from a first person perspective throughout, M’s English usage indexes familiarity with music and computer-mediated discourse (cf. soultunes, kinda, ellipses, emoticons and the decorative use of asterisks). The bilingual testimonial is signed off with
first name + last name + professional title of its author, a voice therapist and M's former singing teacher. Its style approximates that of a reference, with long sentences, complex syntax, characterising and evaluating modifiers, learned vocabulary. The English version is equivalent to the German one in terms of propositional content, but rich in non-native features in English such as missing adverbial suffixes and errors in gender congruence. The two authors use not just different languages, between different genres and language styles in order to convey different, but spatially adjacent perspectives on one individual.

Just below these texts, the embedded photo frames show M in singing poses, and the embedded videos show her perform her songs in English, some preceded by German addressing the audience. The music player and the embedded videos can be operated sequentially or simultaneously, offering a 'sound carpet’ that turns the profile page consumption into a rich multimodal experience. Just below the videos we see a small box with three “Top Friends” and her total number of “friends” (1,020 as per June 22, 2009), followed by a comments list, which extends over two thirds of the visible page. Each comment consists of screen name, photo, date, and the comment proper, sometimes accompanied by visual material; some comments are longer and more colourful than others.

Two types of comments that are common in the context of MySpace are advertising entries (Excerpt 5) and "thanks for adding" messages (Excerpt 6). The former are generally informal, sometimes intentionally heteroglossic (see Excerpt 10 below), and make frequent use of English, which is quite expected in youth-cultural contexts, even though German is clearly the base language in M’s comments space; but they usually lack a personal term of address, indexing their orientation to a public audience. "Thanks for adding" messages are posted after a request to befriend has been accepted by the page owner. They are conventionally brief, consisting of greeting + thanksgiving + politeness formula, again indexing the absence of an established personal relationship. By contrast, other comments on M’s page (Excerpts 7-9) rely on and evoke such personal relationships. Their discourse functions and stylistic choices index a longer communication history between writer and profile host that is not restricted to MySpace, for instance catching up or arranging meetings, referring to past or future events, and so on. In Excerpt 7, note how the ‘thanks for befriending me’ wording is joined by an emoticon to contextualize irony, since this online befriending follows up on an existing offline friendship. In Excerpts 7 and 8, the choice of leave-taking formula (liebe grüße and its acronym, Lg, 'loving greetings') comes from the familiar end of the greetings repertoire in current German CMC. In Excerpts 7 and 9, reference is made to places presumably known to both interlocutors, and Excerpt 9 simulates
a dialogue whose interpretation by the analyst is supported by ethnographic knowledge (e.g. gospel is one style of M’s singing). Typical spoken-like and 'netspeak' features such as informal greetings, contractions and assimilations, lack of normative capitalization and emoticons, occur regularly in this type of comments. Put differently, we see here a relation between the discourse activity carried out in a comment, the social relation between commenter and page host, and the selection of stylistic resources, leading to considerable heterogeneity as one scrolls down the list.

Excerpt 5
User A on 10 Jun 2009 12:54
neuer Track online! "Folge dem Licht" mit Manic und RasRic ab jetzt im player!!!!
Über ein Feedback würde ich mich sehr freuen

‘New track online! “Follow the light” with Manic und RasRic now in the player. I’d greatly appreciate any feedback’

Excerpt 6
User B on 9 May 2009 14:13
Hallo!
Grüße und
Danke für die Freundschaft!
Der IndividualPerformer

‘Hello!
Greetings and
Thanks for the friendship!
The IndividualPerformer’

Excerpt 7
User C on 6 May 2009 16:55
hey danke fürs freunde sein! :) 
hab mich eben fürs fieber angemeldet! *jihaaa*
bis dann also hoffentlich im riff!
Lg E

‘hey thanks for being friends! :) 
I just registered for fever
Well then hopefully see you at riff
Love E’

Excerpt 8
User D on 1 Apr 2009 16:14
hi [M], bist du über Ostern vielleicht in Berlin? liebe grüße!

‘Hi [M], are you over Easter maybe in Berlin? Love!’
Excerpt 9
User E 18 Nov 2008 16:24


Other instances of heteroglossia are located within single entries and identified by salient style contrasts. One example are promotional entries such as Excerpt 10, whose wording articulates two distinct discourses of music and culture: the African-American tradition of soul/funk music and that of German 'contemporary adult' music called Schlager. Both the propositional content and the wording of this entry reinforce that hybrid meeting, and the phrase weird stuff aus deutschen Landen – the nominal phrase indexing music youth culture, the prepositional phrase being typical for agricultural produce – is iconic to that articulation.

Excerpt 10
Dynamite Soul on 17 Dec 2008 19:04
Are you ready for some Super Dynamite Soul?
Plug und Mr. Edd droppen auf Krauty Schlager und German Freak Beats im wesentlichen Weird Stuff aus deutschen Landen. Musik zu der unsere Eltern gerne getanzt hätten, wenn sie ihnen so gekonnt servciert worden ware, wie es die Leipziger Funksoulbrothers auf diesem Mixtape tun.

Damenwahl!

'Are you ready for some Super Dynamite Soul?
Plug and Mr. Edd drop on Krauty Schlager and German Freak Beats essential weird stuff from German soil. Music our parents would have danced to if it only was as well served as Leipzig's own Funksoulbrothers do now on this mixtape.
Ladies' choice!'

Excerpt 11
User F on 31 Jan 2009 16:47
und so saß er da... auf den monitor starrend, der ihn vor kurzem noch mit freude erfüllte. was war mit ihr geschehen, warum antwortete sie nicht?
zu gerne hätte er diese frage beantwortet bekommen, doch der raum verblieb stumm...

wassup girl, why you ain..t answering my mail?? ;-

and so he sat there... staring at the computer screen that filled him with joy just a while ago. what had happened to her, why didn't she respond?
all too much did he wish an answer to this question, but the room remained silent...
wassup girl, why you ain..t answering my mail?? ;-)'
A second type of micro-heteroglossia is in comments that initiate social contact to the page host. While these are similar to promotional entries in their poetic use of language, their pragmatics is oriented to interpersonal relationship rather than commercial marketing. My example is a stretch of comments by one writer, which were posted in a very short period of time, three messages within twelve hours and a fourth one (reproduced as Extract 11) ten days later. Their common pattern of composition is that the concluding line comes in English – the informal, internet savvy style of English M herself uses in her blurb (Excerpt 2). The last message in this series, displayed here, comes after an apparent communication breakdown of several days. What is remarkable is the way it seems together a rather "literary" style of German with a concluding line of stylised vernacular English. This entire post is shaped as a multiple contrast between languages (German/English), styles (literate/vernacular), perspectives (third/second person) and speech acts (wondering/inquiring), and the outcome is arguably a resource for politeness as much as a display of communicative skills on the part of the writer. Here, as elsewhere in CMD, working with linguistic heterogeneity is an option that allows participants to achieve a higher level of quick-wittedness, humour, and skilful playfulness, all of which may increase their communicative attractiveness to others.

*Mantoinette's* profile is not exceptional. Quite the contrary, just about every *MySpace* profile page has a certain heteroglossic potential. Even though “owned” by a particular user, its participative and convergent properties give rise to a variety of coexisting discourse activities and textual components, each with their specific linguistic choices. A variation analysis of such profile pages in terms of spoken/written style, vernacular/standard varieties or German/English language choice is, of course, possible. However, it would probably miss the subtle style differences and indexical tensions that heteroglossia makes us aware of. This echoes Bailey’s (2007) suggestion that heteroglossia, with its social and historic grounding, offers a breadth of perspective that formally defined concepts (in his case, code-switching) seem to lack. Metaphorically speaking, the visual association of a profile page prompted by heteroglossia is less that of a point in variational space and more that of a patchwork of speech styles related to discourse activities, interpersonal relations, ideological positions, and identity claims.

This analysis attempted to present Mantoinette’s profile as a space of heteroglossic relations articulated across and within its distinct components through the juxtaposition of linguistic resources. Their most obvious aspect is probably the contrast between German and English, which can be identified across a number of genres. M chooses English as her
emblematic language in her “calling card” and as her language of self-presentation in her “General Info” box, but the testimonial (“About”) is both in German and English. She sings in English, but addresses her audience in German, most friends’ comments also being in German. This bilingual contrast is re-instantiated within some textual units, as in the embedded videos (performance in English, intro words in German) and the bilingual testimonial. However, caution is needed in order not to decontextualise the two languages from their situated usage. If heteroglossic contrasts may be identified here, these are not between German and English as a whole, but between more specifically contextualized instantiations thereof: between spoken and written material (the former in English throughout, the latter in both English and German); between the language of M’s performance (English) and that of her audience (predominantly German); and between resources selected for distinct modes of identity construction, such as musical performance (sung, in English) as opposed to self-narrative (General Info authored by M, in English) as opposed to others’ utterances on the profile page (written, predominantly in German). Moreover, as we saw with Excerpts 2-3, this is not a contrast of languages but rather language styles selected for different kinds of identity work. We see here once again how the notion of heteroglossia enables us to move beyond binary oppositions of "whole" languages and towards situated uses of linguistic resources.

Conclusions
Returning to the main argument, this chapter suggests that the notion of heteroglossia offers an alternative to the study of processes of linguistic heterogeneity which are widely observed in CMD yet not adequately captured by the concept of language variation and the methods of variationist analysis. I argued that some of these processes are already well known from earlier modes of CMC, whereas others seem specific to web 2.0, and I proposed a hierarchy of levels of analysis in order to identify them. This includes (a) the composition of web surfaces out of user interface, advertising, institutional and participatory content, (b) spectacles and other participatory content, and (c) relations between content and audience responses.

I conclude on four points, which attempt to capture some specifics of “web 2.0 heteroglossia” as analysed in this chapter. First, heteroglossia seems particularly useful with respect to the layered composition of web pages out of different elements, because it allows us to interrelate phenomena on different levels of analysis, from single semiotic forms to larger textual units, such as posts or spectacles, to sequences of such units on a web page.
Second, one aspect of web 2.0 heteroglossia is its dual nature as both intended and emergent process. In line with its traditional understanding, heteroglossic discourse is an outcome of intentional semiotic action, in which linguistic resources are juxtaposed in ways that index social, historical, ideological tensions and conflicts. However, web 2.0 heteroglossia also appears as a by-product, or side effect, of the composite structure of contemporary dynamic web pages, which are not composed in their entirety by one single author.

Such use of the concept needs to be wary of the pitfalls of technological determinism. I argued that the convergent and participatory composition of web 2.0 sites gives rise to social and ideological contrasts, which, depending on the linguistic resources they are given, can find heteroglossic expressions. However, it is not technology that creates heteroglossia. Web 2.0 may well be experienced in monoglot terms – but it often isn’t, and, one might add, it almost certainly isn’t to speakers of smaller or “weaker” languages. Web 2.0 environments open a range of possibilities for heteroglossic “hot spots”, but their exploitation ultimately depends on institutional and situational context and discourse dynamics. Finally, heteroglossia in this chapter goes beyond the type of class-based social tension which shapes its original understanding and some of its applications to media discourse (e.g. by Georgakopoulou 2000 on film). While classic Bakhtinian heteroglossia revolves around the expression of social conflict and carnivalesque subversion ‘from below’ (White 1993), this chapter also focuses on global/national and global/local relations as new domains of heteroglossic tension. We also saw a stylized usage of linguistic heterogeneity (Extracts 9-10) which ties in well with the insight that in the absence of visual cues or a face-to-face encounter, a playful and creative use of language becomes a key resource for interpersonal communication and identity management in digital media.

Future work on and with the concept will need to refine its operationalisation and diagnostic criteria, and to spell out its conceptual relation to sociolinguistic notions such as style and style-shifting. It will also need to engage with issues of interpretation and perspective, in particular with questions such as: Heteroglossic to whom? And is heteroglossia on the web a property of textual surfaces or rather their situated reception by users? Stronger online ethnography would probably rebalance the dominance of a text-based over a user-based perspective in this chapter, and that, in turn, would help to enhance the robustness of the interpretations offered here. At the same time, we need to acknowledge that heteroglossia is a notion that is sensitive to perspective and time, and therefore to researchers’ social positions and their understandings of social change. The issue therefore will not be to construct an “objectively valid” interpretation of heteroglossia, which would seem an illusory
undertaking, but to lay bare the perspectives from and criteria on which given linguistic practices on the web may be viewed as heteroglossic.

Acknowledgments

First presented to the 5th International Conference on Language Variation in Europe (ICLaVE#5), June 25-27, 2009, Copenhagen. Thanks are due to Mantoinette for permission to use data from her profile page.

References


van Compernolle, R. A. (2008) Language variation in online personal ads from Quebec: The case of ne. Language@Internet, 5, article 1. URL: http://www.languageatinternet.de/articles/2008/


Notes

1 A rare quantitative study of code-switching online is Paolillo (2001).

2 In the original German wording, “daß sprachliche Elemente und Versatzstücke aus diversen Diskurswelten zu einem spezifischen Stilmix zusammengebastelt werden”.

3 Other useful definitions are “the different kinds and uses of speech that struggle within a speech community” (Stivale 1997: 134) and “the diversity of social meaning making practices in a particular community” (Thibault 2004: 41). The definition by Leppänen et al. (2009) – “the coexistence, combination, alternation and juxtaposition of ways of using the communicative and expressive resources language/s offer us” – downplays the aspect of socio-ideological tension that is central to my use of the concept here.

4 Heteroglossia has been used to study advertising discourse (Cook 2001: 187-8), multilingual media in minority contexts (Busch 2006), popular film (Georgakopoulou 2000), and CMC (Stivale 1997, Leppänen et al. 2009).

5 Permission has been obtained by the page owner to publish her screen name and extracts from her profile page.