Discourse in Web 2.0: Familiar, Reconfigured, and Emergent

Susan C. Herring
Indiana University, Bloomington

Introduction

From controversial beginnings, the term Web 2.0 has become associated with a fairly well-defined set of popular Web-based platforms characterized by social interaction and user-generated content. Most of the content on such sites is human discourse, via text, audio, video, and/or static images. It is therefore, in principle, of theoretical and practical interest to scholars of computer-mediated discourse. Yet although discourse-focused studies of individual Web 2.0 environments such as Facebook, Flickr, Twitter, and YouTube are starting to appear (see, e.g., the chapters in Thurlow and Mroczek 2011), systematic consideration of the implications of Web 2.0 for computer-mediated discourse analysis as a whole is lacking. Does discourse in these new environments call for new methods of analysis? New classificatory apparatuses? New theoretical understandings? This paper is a preliminary attempt to address these questions.

After defining Web 2.0 and reviewing its development over the past decade, the computer-mediated discourse analysis (CMDA) paradigm developed by the author (Herring 2004) is briefly reviewed, with the ultimate goal of determining whether--and if so, in what ways--it needs to be revised in light of Web 2.0. As a heuristic to address this goal, a three-part classification of Web 2.0 discourse phenomena is introduced: phenomena *familiar* from older computer-mediated discourse (CMD) modes such as email, chat, and discussion forums that appear to carry over into Web 2.0 environments with minimal differences; CMD phenomena that adapt to and are *reconfigured* by Web 2.0 environments; and new or *emergent* phenomena that did not exist--or if they did exist, did not rise to the level of public awareness--prior to the era of Web 2.0. This classification is loosely inspired by Crowston and Williams's (2000) broad classification of web pages into 'reproduced' vs. 'emergent' genres, but with a focus on discourse, rather than genre. For each category of phenomena, the challenges that it raises for CMDA are discussed.

In concluding, I suggest that this three-way classification can provide insight into why particular discourse phenomena persist, adapt, or arise anew in technologically-mediated environments over time. In so doing, I invoke technological factors such as multimodality and media convergence, social factors at both the situational and cultural levels, and inherent differences among linguistic phenomena that make them variably sensitive to technological and social effects. Suggestions are also made of practical ways in which the classification might guide researchers to frame their studies and select certain methods of analysis. While the reconfigured and emergent categories are especially attractive in that they present new phenomena and raise special challenges for analysts of computer-mediated discourse, I argue that researchers should not neglect what appears

familiar in favor of pursuing newness or novelty: All three categories merit research attention, for different reasons.

Background

What is Web 2.0?

The World Wide Web itself is not new. It was pitched as a concept by physicist Tim Berners-Lee to his employers at CERN in 1990, implemented by 1991, and attracted widespread attention after the first graphical browser, Mosaic, was launched in 1993.[1] The early websites of the mid-1990s tended to be single-authored, fairly static documents, and included personal homepages, lists of frequently-asked questions (FAQs), and e-commerce sites. The late 1990s saw a shift towards more dynamic, interactive websites, however, including, notably, blogs (Herring, Scheidt, Bonus, and Wright 2004) and online newssites (Kutz and Herring 2005), the content of which could be--and often was--updated frequently and which allowed users to leave comments on the site. These sites foreshadowed what later came to be called Web 2.0.

The term itself was first used in 2004 when Tim O'Reilly, a web entrepreneur in California, decided to call a conference he was organizing for "leaders of the Internet Economy [to] gather to debate and determine business strategy" the "Web 2.0 Conference" (Battelle and O'Reilly 2010; O'Reilly 2005). At the time, the meaning of the term was vague, more aspirational and inspirational than descriptive. As a business strategy, "Web 2.0" was supposed to involve viral marketing rather than advertising and a focus on services over products. One of O'Reilly's mantras is "Applications get better the more people use them" (Linden 2006). Today the term refers, according to Wikipedia (2011b) and other online sources, to *changing trends in, and new uses of,* web technology and web design, especially involving participatory information sharing; user-generated content; an ethic of collaboration; and use of the web as a social platform. The term may also refer to the *types of sites* that manifest such uses, e.g., blogs, wikis, social network sites, and media-sharing sites.

From the outset, the notion of "Web 2.0" was controversial. Critics claimed that it was just a marketing buzzword, or perhaps a meme--an idea that was passed electronically from one Internet user to another--, rather than a true revolution in web content and use as its proponents claimed. They questioned whether the web was qualitatively different in recent years than it had been before, and whether the applications grouped under the label Web 2.0--including such diverse phenomena as online auction sites, photo-sharing sites, collaboratively-authored encyclopedias, social bookmarking sites, news aggregators, and microblogs--formed a coherent set. Tim Berners-Lee's answers to these questions was "no"--for the inventor of the web, the term suffered from excessive hype and lack of definition (Anderson 2006).

In response to such criticisms, O'Reilly (2005) provided a chart to illustrate the differences between Web 2.0 and what he retroactively labeled "Web 1.0". This is shown

2

in modified and simplified form in Figure 1. The phenomena in the second column are intended to be the Web 2.0 analogs of the phenomena in the first column.

Web 1.0	Web 2.0
Personal websites	Blogging
Publishing	Participation
Britannica online	Wikipedia
Content management systems	Wikis
Stickiness	Syndication
Directories (taxonomies)	Tagging (folksonomies)

Figure 1. Web 2.0 vs. Web 1.0 phenomena (adapted from O'Reilly 2005)

This is no more than a list, however; it does not provide a principled way to determine what should or should not be included as "Web 2.0". An alternative approach to conceptualizing Web 2.0 is from a temporal perspective. Most Web 2.0 phenomena emerged over the last decade; indeed, one criticism of the concept is that it seems to mean nothing more than 'recent websites'—paving the way for 'Web 3.0, 4.0, etc.' to describe ever-newer web-based phenomena in the future. Figure 2 situates phenomena that exhibit characteristics of 'Web 2.0' along a timeline for roughly the first decade of the 21st century. The phenomena included in the figure are not intended to be exhaustive in any sense; however, an attempt has been made to include important exemplars of a range of phenomena. In the figure, applications that run on a platform other than the web but otherwise have charcteristics of "Web 2.0" are indicated in grey font; "Web 2.0" sites that were launched before O'Reilly coined the term in 2004 are indicated in unbolded black fonts; and "Web 2.0" sites launched since 2004 are in bolded black font.[2]
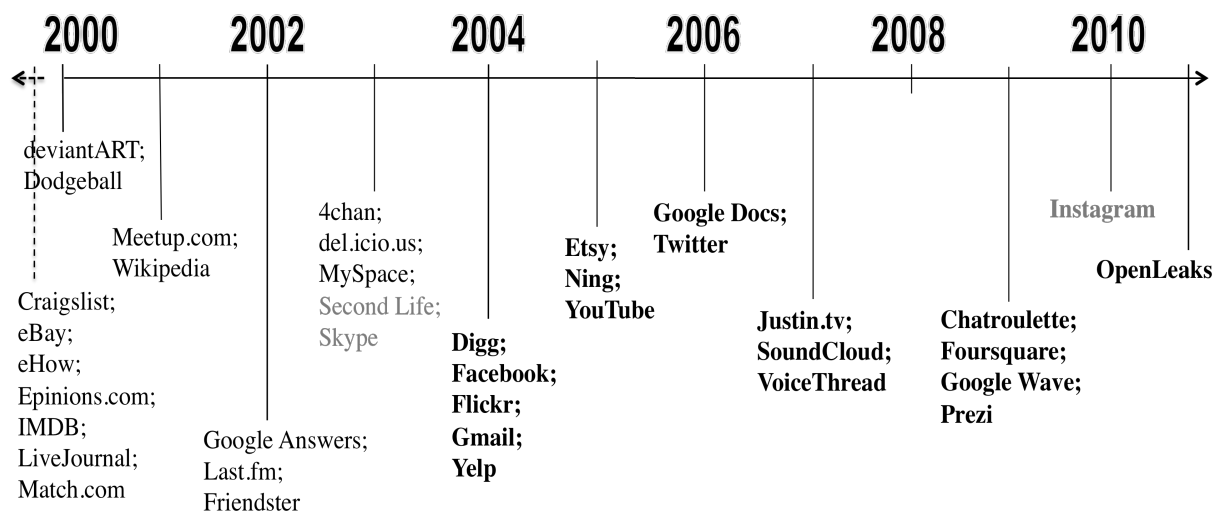

Figure 2. "Web 2.0" timeline (selective)

The most recent decade of collaborative communicative technologies is not a perfect match with "Web 2.0" in the strict sense. Not all of the applications mentioned in Figure 2 are based in the web—for example, Second Life (a three-dimensional graphical world) and Skype (a form of Internet telephony) are applications that run on the internet but not on the Web, and Instagram, a recently-introduced photo-sharing application, runs on mobile phones (although it can share with web-based social network sites). Moreover, a number of sites commonly considered to be "Web 2.0" were launched before 2004, and some—such as eBay, blog hosting sites (such as LiveJournal.com), and online dating sites—date back to the 1990s. Nonetheless, with these exceptions, and together with the characteristics of the sites themselves, the timeline suggests some conceptual coherence that can be attributed to these phenomena. Thus, for the purpose of this paper, the following redefinition of Web 2.0 is proposed:

> Web-based platforms that emerged as popular in the first decade of the 21st century, and that incorporate user-generated content and social interaction, often alongside or in response to structures and/or (multimedia) content provided by the sites themselves.

(B) Computer-mediated discourse analysis, media convergence, and 'Discourse 2.0'
The second key background component of this paper is discourse analysis and, specifically, the approach to analyzing online discourse that the author has developed over the past 17 years, Computer-Mediated Discourse Analysis (CMDA). CMDA is an approach to the analysis of computer-mediated communication (CMC) focused on language and language use; it is also a set of methods (a "toolkit") grounded in linguistic discourse analysis for mining networked communication for patterns of structure and meaning, broadly construed (Herring 2004).[3]

The term 'Computer-Mediated Discourse Analysis' was first used as the name of a GURT workshop organized by the author in 1994. Since then it has evolved into a model organized around four levels of CMDA: structure, meaning, interaction management, and social phenomena (Herring 2004). The organizational principle of the CMDA toolkit is fairly simple: The basic idea is to adapt existing methods, primarily from linguistics (but in principle from any relevant discipline that analyzes discourse), to the properties of digital communication media. The methods and the phenomena, along with broader issues they address, are then loosely mapped onto four level of hierarchy, from the micro-linguistic, more context-independent level of structure, to the macrolevel of contextualized social phenomena, as summarized in Table 1. (A fifth, non-linguistic "level", participation, is sometimes included as well.)

4

Levels	Issues	Phenomena	Methods
Structure	Orality; formality; efficiency; expressivity; complexity; genre characteristics, etc.	Typography, orthography, morphology, syntax, discourse schemata, formatting conventions, etc.	Structural/Descriptive Linguistics, Text Analysis, Corpus Linguistics, Stylistics
Meaning	What is intended What is communicated What is accomplished	Meaning of words, utterances (speech acts), exchanges, etc.	Semantics, Pragmatics
Interaction management	Interactivity; timing; coherence; repair; interaction as co-constructed, etc.	Turns, sequences, exchanges, threads, etc.	Conversation Analysis, Ethnomethodology
Social phenomena	Social dynamics; power; influence; identity; community; cultural differences, etc.	Linguistic expressions of status, conflict, negotiation, face-management, play; discourse styles/lects, etc.	Interactional Sociolinguistics, Critical Discourse Analysis, Ethnography of Communication

Table 1. Four levels of CMDA (adapted from Herring 2004)

However, in the time since the CMDA approach was originally conceptualized, CMC itself has been undergoing a shift, from occurrence in stand-alone clients such as emailers and instant messaging programs to juxtaposition with other content, often of an information or entertainment nature, in converged media platforms, where it is typically secondary, by design, to other information or entertainment-related activities (Herring 2009; Zelenkauskaite and Herring 2008). This phenomenon, which I refer to as convergent media computer-mediated communication or CMCMC, is especially common on Web 2.0 sites. Examples include text comments on photo-sharing sites; text (and video) responses to YouTube videos; text (and voice) chat during multiplayer online games; and text messages from mobile phones posted to interactive TV programs. Less prototypically (because it involves the convergence of text with text, rather than the convergence of text with another mode), CMCMC is also illustrated by reader comments on news stories; 'talk' pages associated with Wikipedia articles; status updates and comments (and for that

matter, chat and Inbox exchanges) on Facebook profiles; and interpersonal and group exchanges on Twitter.[4]

In fact, the overlap between CMCMC and Web 2.0 is considerable. Almost all so-called Web 2.0 sites feature CMCMC, and almost all CMCMC applications are on the Web. An exception to the former is social bookmarking sites such as del.icio.us, which do not contain CMC; an exception to the latter is text chat in multiplayer online games such as World of Warcraft, which are not hosted on the Web. However, the trend is for increasing convergence, and it would not be surprising if these distinctions disappeared in the future. In what follows, it is assumed that Web 2.0 generally involves CMCMC.

The discourse in these new environments—what we might call CMCMD (convergent media computer-mediated discourse) or 'Discourse 2.0'—raises many issues for computer-mediated discourse analysis. There are new types of *content* to be analyzed: status updates, text annotations on video, tags on social bookmarking sites, edits on wikis, etc. New *contexts* must also be considered—for example, social network sites based on geographic location—as well as new (mass media) audiences, including in other languages and cultures. (Facebook, for example, now exists in "localized" versions in well over 100 languages [Lenihan, 2011].) Discourse 2.0 manifests new *usage patterns*, as well, such as media co-activity, or near-simultaneous multiple activities on a single platform (e.g., Herring, Kutz, Paolillo, and Zelenkauskaite 2009) and multi-authorship, or joint discourse production (e.g., Androutsopoulos 2011; Nishimura 2011). The above reflect, in part, new *affordances* made available by new communication technologies: text chat in multiplayer online games (MOGs); collaboratively editable environments such as wikis; friending and the "walled gardens" of Facebook; social tagging/recommending; and so forth. Last but not least, Discourse 2.0 includes *user adaptations* to circumvent the constraints of Web 2.0 environments: e.g., interactive uses of @ and #, as well as retweeting, on Twitter (e.g., boyd, Golder, and Lotan 2010; Honeycutt and Herring 2009) and performed interactivity on what are, in essence, monologic blogs (e.g., Peterson, 2011; Puschmann, in press). Each of these issues deserves attention, and some are starting to be addressed, but on a case-by-case, rather than a systemic, basis.

In the face of so many apparently new phenomena, one might question whether research into Discourse 2.0 should continue to be called computer-mediated discourse analysis. The term 'computer-mediated' originally referenced the term 'computer-mediated communication', which is still preferred among communication scholars. But communication technologies are increasingly moving beyond computers. Mobile phones can be considered honorary computers, but voice calls, for example, challenge that characterization, as does television-mediated conversation (Zelenkauskaite and Herring 2008). Some recent discourse-focused publications use such alternatives as 'digital media' and 'new media' discourse. However, the term 'new media' is lacking in historical perspective, and 'digital media' is too broad, referring as it does to video games as well as communication devices, although 'digital discourse' (e.g., Thurlow and Mroczek 2011)

makes clear that discourse is in focus. It well may be that in the coming years, the dust will settle and a descriptive term that is neither too narrow nor too broad will emerge as the obvious candidate. For now, CMDA still seems a useful term, in that it makes the link to CMC and CMD transparent and is based on established tradition, so it will continue to be used in the remainder of this paper.

Discourse in Web 2.0: An organizational lens

As a conceptual aid to making sense of Discourse 2.0 on a larger scale, in this section I propose a three-part organizational lens, a broad categorization scheme according to which discourse in Web 2.0 environments can be classified. This scheme challenges the prima facie assumption that all phenomena that seem new are actually new, classifying them instead in relation to their antecendents (or lack of antecedents) as *familiar*, *reconfigured*, or *emergent*. At first blush, the scheme may appear simply to reproduce chronology—older phenomena might be expected to be familiar, and new ones to be emergent—but in fact it does not. As the discussion below shows, the reverse pattern sometimes holds.

Web genre classification

The Discourse 2.0 classification scheme proposed here was inspired by Crowston and Williams (2000), who in the mid-1990s (the first version of their paper was written in 1996) were among the first scholars to attempt to classify genres of web pages. They proposed a simple two-way classification—reproduced vs. emergent—but noted that some pages could be classified as adaptations of reproduced genres.

　　　The following are examples of Crowston and Williams's categories applied to traditional web content. Reproduced content (sometimes referred to as "shovelware") includes course syllabi, scholarly articles, and meeting minutes, which in the early years of the web (and sometimes still) were often created with word processing software offline and simply uploaded to the web. Adapted genres include newssites, geneology sites, and e-journals, which have tended to preserve the basic genre conventions of their offline precursors but with adaptations, such as user commenting and hyperlinking, that leverage the affordances of the web. As examples of emergent genres, Crowston and William mention "hotlists" of links and personal homepages; blogs and wikis have also been claimed to be indigenous to the web (e.g., Blood 2002). In the case of blogs, however, several scholars (e.g., Herring et al. 2004; Miller and Shepherd 2004) have pointed out the continuities between paper diary writing, for example, and blogging, such that nowadays most blog researchers would consider blogs to be an adapted genre. As this last case illustrates, there is a tendency to claim 'newness' when historical precedents in fact exist, as part of a general ahistoricity that characterizes much new media research; this is a risk associated with applying Crowston and Williams' classification. The scheme can also be somewhat subjective, depending on what facets of a web genre one chooses to focus on

and the familiarity of the researcher with previous genres. Nonetheless, the classification can lead to interesting insights.

One such insight is that there appears to be a trend over time for web genres to shift along a continuum from reproduced to adapted to emergent, with the seemingly paradoxical effect that as genres age, they move along the continuum in the direction of "emergence". An example is online news sites, which tended to be *reproduced* from print newspapers in the early days but have become increasingly *adapted* to the web with the inclusion of user comments, multimedia, hyperlinks, and other interactivity and navigation features (Eriksen and Ihlström 2000). Similarly, the social network site Facebook initially combined the format of "face books" from Harvard University dormitories with web-based features such as commenting (Wikipedia 2011c) and was thus *adapted*, but over time it has added and combined so many features (including embedded graphics, games, polls, and various modes of CMC) that it can now arguably be considered *emergent*. This is not just because of the number of features, but because the whole is qualitatively different from the sum of its parts, and it has no single offline or online precedent.

This trend is consistent with the observation in media studies that new information and communication technologies (ICTs) are first put mainly to old uses, and uses that take advantage of the full potential of the technologies emerge only later, if at all (e.g., Winston 1998). In the case of web genres, the mechanisms underlying this shift are, on the one hand, the incorporation of new media affordances into familiar text types, and on the other hand, increasing media convergence. When familiar media combine, they often do so in ways that result in qualitatively different, hence what can be considered emergent, phenomena.

In the present classification scheme, the term 'familiar' is used rather than 'reproduced' to suggest continuities in (rather than copies of earlier) discourse phenomena. Familiar patterns are presumed to have manifested more-or-less continuously on the web from the 1990s until the present, albeit not necessarily on the same sites. The term 'reconfigured' is used rather than 'adapted' to highlight the structural reshaping of some discourse phenomena that takes place in Web 2.0 environments. To be sure, users adapt to changes in ICTs, but their adaptations are only relevant to CMDA when they manifest in tangible discourse behaviors. However, the term 'emergent' is preserved from Crowston and Williams (2000), in that new web genres and discourse phenomena can equally well be said to emerge—that is, to become evident where previously there was no general awareness of their existence.

In the following sections, the proposed rubric of familiar-reconfigured-emergent is applied to classify examples of Web 2.0 discourse phenomena, with discussion of some of the challenges each category raises for computer-mediated discourse analysis.

A three-part classification scheme

Familiar aspects of Web 2.0 discourse

Contrary to the impression conveyed by much "new media" research that almost everything on the Web today is new and different, in fact a great deal, and perhaps the majority, of Web 2.0 discourse phenomena are familiar. For one thing, text remains the predominant channel of communication among web users, whether it be in blogs, microblogs, wikis, comments on news sites, or web discussion forums. The latter, in particular, remain very popular, and illustrate many of the same kinds of phenomena as did asynchronous online forums in the 1990s. These include non-standard typography and orthography, code switching, gender differences, flaming, and email hoaxes/scams, to mention just a few.[5]

Contemporary practices, especially when situated in new environments, may not seem familiar at first blush. Vaisman (2011) recently documented a creative typographic practice among Israeli pre-teen girl bloggers that she calls 'Fakatsa' style ('fakatsa' meaning silly, fashion-conscious girl), in which letters of the Hebrew alphabet are replaced by characters of similar appearance, including numbers, symbols, and letters from the roman alphabet. Vaisman gives the following example, the userID of one of the bloggers, which translates as 'Gal (a Hebrew name) the hot babe':

~תי§!¢ה~5ג:°י

In standard Hebrew typography (Arial font), this name would be written: גל הכוסית . The section symbol § in the Fakatsa version replaces the Hebrew letter that resembles (in this font) a small 'o', the cent symbol ¢ replaces the Hebrew letter that looks like a backwards 'c', and the Arabic numeral 5 substitutes for the Hebrew letter *lamed* ל; these substitutions are based loosely on graphical resemblance. Moreover, the Fakatsa version of the word includes extra symbols—such as a tilde ~ in two places and the dots on the right—that are purely decorative.

On the face of it, Fakatsa writing appears novel and exotic. While nonstandard typography and orthography are familiar characteristics of CMC (cf. "Netspeak" expressions such as *ur gr8* 'you're great', which themselves have antecedents in offline writing [Crystal 2008]), Fakatsa seems to take this practice to a new level, including requiring norms of reading that differ from those for netspeak. Yet as Vaisman points out, Fakatsa shows many parallels with Leetspeak, an English subcultural variety that originated with computer hackers in the 1980s. Leet replaces letters from the roman alphabet with non-alphabetic characters and symbols based on visual similarity, similar to Fakatsa, as illustrated in the following example: 1 k4n 7¥p3 £337 'I can type Leet.' The similarities between the two varieties do not appear to be due to any direct contact between them; rather, they share a common generative principle that appears to have been exploited spontaneously in the two different writing systems (for further examples of graphic-based

9

substitutions in computer-mediated Arabic and Greek typography, see Danet and Herring 2007). This example illustrates the first level of CMDA: structure.

Historical continuities in CMD are also evident on the broader level of social interaction. Gender differences in discourse style were documented in public online discussion forums and chat rooms throughout the 1990s that showed males to be more assertive, insulting, sarcastic, and profane and females to be more accommodating, supportive, affectionate, and upbeat (e.g., Cherny 1994; Herring 1993, 2003a). These patterns reproduced gender styles in spoken conversation as described, e.g., by Tannen (1990). Moreover, despite a tendency of scholars and lay people alike to imagine that such stylistic differences reflect outmoded gender role differences that have tended to disappear over time, a recent study of teen chat sites (Kapidzic and Herring 2011a) found similar differences in message tone: In 2011, males were still significantly more aggressive and flirtatious and tended to be more sexual, whereas females were significantly more friendly in their chat messages. Girls also still used more emoticons, especially those representing smiles and laughter, than boys did (Kapidzic and Herring 2011b), as shown in Figure 3. Similar to the findings of earlier research (e.g., Wolf 2000), the only emoticon that boys used more was the winking face, which is associated with both sarcasm and flirtation.
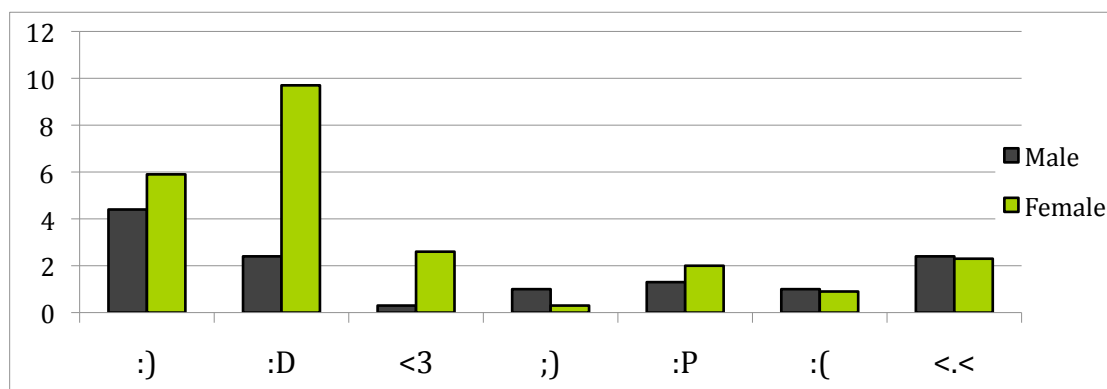

Figure 3. Emoticon use in teen chat per 1000 words by gender (Kapidzic and Herring 2011b)

Familiar Discourse 2.0 phenomena lend themselves readily to CMDA in its current form, since they are the kinds of phenomena the paradigm was designed to address. The challenges they pose mainly concern identification: Familiar phenomena may be mistaken for new (as a researcher unfamiliar with Leet might be tempted to do with Fakatsa) or assumed to be different by virtue of the passage of time (as in the case of online gender styles). There is a need to trace relevant antecedents to gain perspective where familiar online discourse phenomena are concerned, in order to do conscientious research. This, in turn, requires some familiarity with earlier CMDA research. Alternatively, familiar phenomena may simply be passed over by researchers in favor of newer, more "exotic" CMD phenomena.

Reconfigured aspects of Web 2.0 discourse

While Discourse 2.0 may not be as different as popular belief holds, changes have taken place in both technology and in communicative practices. These have resulted in the reconfiguration or reshaping of a number of aspects of CMD. Examples include personal status updates, quoting/retweeting, 'small stories',[6] and customized advertising spam, which might on the surface appear new but have traceable online antecedents, as well as reconfigurations of such familiar phenomena as topical coherence, turn-taking, threading, and intertextuality.

      As a first illustration, let us consider personal status updates of the sort that have been popularized on Facebook and Twitter, in response to a generic prompt such as 'What are you doing?', 'What's on your mind?', or 'What's happening?' The issues raised by the analysis of status updates involve the CMDA levels of structure (syntax), meaning (semantics and pragmatics), and interaction management. Lee (2011) provides the following examples in her study of Hong Kong users of Facebook, which compares the communicative functions of status updates produced before and after Facebook changed the default response format from 'Name is' (with the 3rd person singular form of the verb 'to be') to simply 'Name':

1. Amy is in a good mood.
2. Snow is "I've seen you in the shadow".
3. Kenneth quitting facebook.
4. Ariel thinks that no news is good news.
5. Katy : ?

Example 1 is a grammatical sentence formed from the prompt 'Amy [name of user] is'. Example 2 is also built on the 'name is' prompt, but what follows (*"I've seen you in the shadow"*) is a song lyric; the result is not a grammatical sentence. Status update 3 (*Kenneth quitting facebook*) is also ungrammatical, but for a different reason: It lacks the auxiliary verb 'is', which by this time had been omitted from the default Facebook prompt. The presence, explicit or implicit, of 'is' in the Facebook response format thus results in some status descriptions that are syntactically nonstandard.

      The update in 4 (*Ariel thinks that no news is good news*), in contrast, is syntactically well formed, but pragmatically appropriate in the simple present tense only if Ariel's thinking is generally true, which is not the most likely interpretation in this context (it is more likely that Ariel is responding to some specific news, or lack thereof). This use of simple present tense is presumably a carry-over of the simple present tense of the former 'is' in the prompt. Its use in place of, e.g., the present progressive lends the utterance a performative feel, as if Ariel performs the act of 'thinking that no news is good news' through typing it. Finally, if Katy's status update consisting of only a question mark

(presumably to indicate that she is confused or does not know what to say) in 5 is treated as an utterance, it is arguably both syntactically and pragmatically ill-formed.[7]

In fact, these kinds of utterances have characterized CMC since the early days of Internet Relay Chat (Werry 1996) and MUDs and MOOs (text-based virtual reality environments; Cherny 1994, 1995, 1999). Those chat environments made available special commands (called 'action descriptions' on IRC and 'emotes' on MUDs and MOOs) that produced 3rd person present tense descriptions of 1st person actions and states, such as 'Chris is in a bad mood' and 'Lynn waves'. These often had a performative flavor, especially when used to describe actions rather than states (for further discussion of performativity in CMD, see Virtanen, this volume and forthcoming). Utterances that play with the convention itself, analogous to 5 above, were also common in early chat environments (in a perfect parallel, Cherny 1999 gives examples of MOO utterances consisting solely of 'username ?', as well as 'null-emotes', emotes left intentionally blank such that only the username appears).

However, status updates do not simply reproduce these earlier practices. Rather, they have been structurally and functionally reconfigured in comparison to action descriptions and emotes. Syntactically, the inclusion of 'is' in an earlier version of Facebook has led to a greater use of 'is' constructions, even when these are not prescriptively correct. One of the author's Facebook friends has continued to start each of his updates with [Name] is, inserting the 'is' as a stylistic affectation even when another finite verb is present in the utterance, e.g., *David is has a headache*. Functionally, status updates on Twitter and FB serve as prompts that trigger comment threads, unlike the earlier constructions, which are single utterances.

The second example of a reconfigured phenomenon also involves interaction management—specifically, repeating parts of another participant's message in one's own message to create cross-message coherence. 'Retweeting' on Twitter is the inclusion of a previous message ('tweet') in a new message, sometimes with a comment added. Boyd et al. (2010) give the following typical example:

RT @StopAhmadi Bring down Khomeini's website

Here, the (unidentified) user makes use of the abbreviation RT to retweet wholesale a tweet that was originally addressed to StopAhmadi (on the uses of @ as an addressivity marker, see Honeycutt and Herring 2009). Presumably the retweeter does this for the purpose of spreading further the sentiment expressed in the original tweet, *Bring down Khomeini's website*.   Another example of retweeting given by boyd et al. is more complex, in that it contains multiple levels of embedding:

@AndreaJarrell: Via @mStonerblog: RT @zephoria: new blog post "Is Facebook for old people?" socioecon and race are most interesting here http://bit.ly/v0aPS.

In this tweet, the user is addressing to AndreaJarrell a message originally received from mStonerblog, who in turn retweeted a message from zephoria (*new blog post "Is Facebook for old people?" http://bit.ly/v0aPS*). (The latter two instances of @ conventionally signal that what follows is a Twitter username and can effectively be ignored for the purposes of understanding this tweet.) To complicate the message further, the user inserts a comment of his/her own (*socioecon and race are most interesting here*) before zephoria's link.

On the face of it, retweeting might seem a prime example of a discourse phenomenon that is new, that has arisen in response to the novel affordances and constraints of Twitter. In fact, retweeting is a modern form of the older practice in textual CMC of 'quoting' in asynchronous messages (Severinson Eklundh 2010; Severinson Eklundh and Macdonald 1994). Quoted and retweeted segments are both flagged by a conventional symbol (an angle bracket > preceding each quoted line in Usenet, where the convention originated). Both incorporate the words of others in one's own message (intertextuality, e.g., to establish context), and embedded quotations such as those in Twitter example 2 above can occur in both CMC modes. The following schematic from Severinson Eklundh (2010) illustrates multiple embeddings in a traditional asynchronous system (Figure 4).

```
text by B.
greeting by B.
┌──────────────────────────┐
│pointer to A.             │
│> greeting by A.          │
│> text by A.              │
│> greeting by A.          │
│┌────────────────────────┐│
│>│pointer to B            ││
│>│>┌─────────────────────┐│
│>│>│pointer to A.        ││
│>│>│> text by A.         ││
│>│>│text by B.           ││
│>│>│> text by A.         ││
│>│>│text by B.           ││
│>│>│greeting by B        ││
│>│ └─────────────────────┘│
│> └────────────────────────┘
└──────────────────────────┘
```
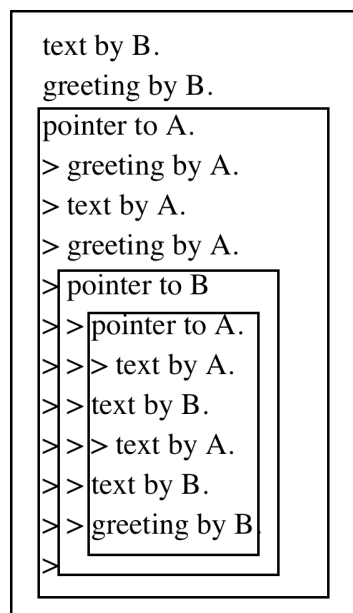
Figure 4. Schematic structure of a message that A received with three levels of quotation (from Severinson Eklundh 2010)

Although some systems include quoted text automatically when the user responds using the 'reply' function, users can also set the default such that quoted text is not included or edit the quoted text selectively. Evidence of the latter can be seen in the message in Figure 4, where B had previously responded to A point by point, interleaving his text (at two levels of embedding) with text by A (at three levels of embedding).

13

As the figure suggests, in quoting the repeated material is set off on a separate line, whereas on Twitter, because of the character limit on tweets, quoted and non-quoted material appears together without line breaks, giving the two practices different surface appearances. Retweeting can thus be considered a form of quoting that is adapted to, and reconfigured by, the Twitter environment. The reconfiguration involves condensing quoted content to fit within 140 characters, including omitting information and reducing URLs to abbreviated bit.ly links, and using new (user-generated) conventions for referring/attributing concisely, resulting in a visually very different format (that includes a small image of the sender, as well), but with a similar underlying function.

The third and last example of a reconfigured phenomenon also involves interaction management, this time with a focus on topical coherence. Conversational exchanges on many Web 2.0 platforms tend to be prompt focused—that is, comments respond to an initial prompt, such as a news story, a photo, or a video, more often than to other users' responses. While this pattern was attested in older CMD, notably in online learning contexts where students would all respond to an instructor's prompt, it was not common in public discourse on the internet. Rather, patterns of topical development tended to involve step-wise digression away from the original topic in multiparty exchanges in chat rooms (Herring 1999) and discussion forums (Lambiase 2010). These differing patterns are illustrated in Figure 5 using the Visual DTA (Dynamic Topic Analysis) tool developed by Herring and Kurtz (2006). The diagram on the left, showing the prompt-focused pattern, represents a YouTube comment thread in response to an Aljazeera English news video (Abdul-Mageed 2009), and the diagram on the right, showing step-wise topical digression, represents a conversation on a recreational IRC channel (Herring 2003b; Herring and Nix 1997). In the diagrams, messages are numbered in chronological order along the y-axis, and the x-axis represents cumulative semantic distance from the initiating message.
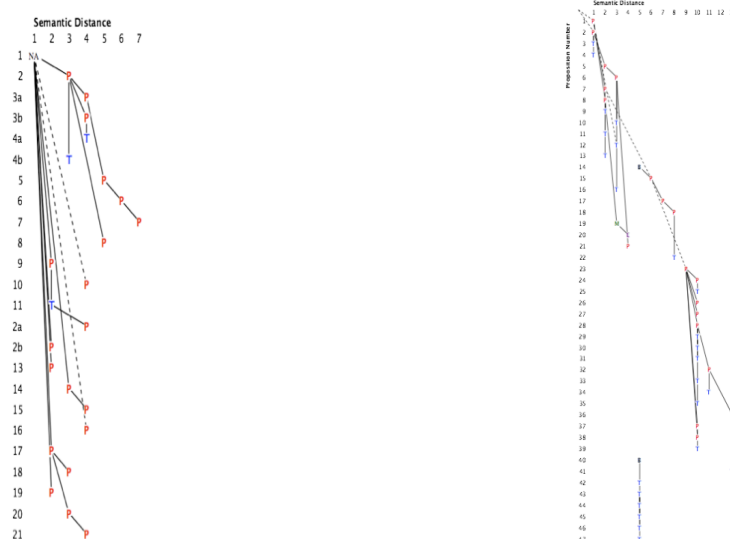

Figure 5. Two patterns of topical development. Left: prompt focused (YouTube comment thread; Abdul-Mageed 2009). Right: step-wise progression (IRC; Herring 2003b).

Both of the above are thematically "coherent" patterns of interaction involving multiple participants that begin with an initial prompt, but in the YouTube comment thread (and much other Discourse 2.0) responses are mostly directed to the prompt throughout, resulting in limited topical development, whereas in the IRC example the topic shifts progressively from the upper left to the lower right through interaction among participants. The former pattern can be considered an adaptation of topical coherence to convergent media (CMCMC) environments in which entertainment and news content is presented for users to consume and comment on. As a consequence of this adaptation, patterns of interaction are reconfigured not just at the individual level but at the group level.

Because they may appear quite different on the surface, reconfigured phenomena are at an even greater risk than familiar phenomena of being mistaken for "emergent" CMD, with consequent loss of comparative insight. An additional challenge posed by reconfigured phenomena for CMDA is the need to abstract common structures, functions, and/or social dynamics across different media affordances in order to identify what they are reconfigured from and the reasons for the reconfigurations.

Emergent aspects of Web 2.0 discourse

Given that much of what has been claimed to be unprecedented on the web has been found, upon deeper examination, to have online and/or offline antecedents, caution must be exercised in asserting that any phenomenon is entirely new. This section tentatively identifies several Web 2.0 discourse phenomena that appear to be emergent and unprecedented,[8] at least as common practices. These include the dynamic collaborative discourse that takes place on wikis, along with conversational video exchanges, conversational exchanges via 'image texts', and multimodal conversation more generally. In the domain of non-bona fide discourse, the phenomena of "sockpuppets" and "link building" could also be mentioned.

Collaborative text production of the sort that takes place on Wikipedia represents a new kind of online discourse. It is democratic and anarchic: There is no central organization, and anyone can contribute to any part of a text. It is massively multi-authored by Internet users who usually do not know one another. It leaves a manifest trace both 'front stage', in Article pages, and 'backstage' (cf. Goffman 1959), e.g., in Talk pages. Moreover, every addition, deletion, or alteration of the text is preserved in History pages, which in themselves constitute a new kind of text. It is process focused rather than product focused; even its most stable content, articles, are dynamic documents subject to frequent updating. To be sure, texts are sometimes constructed collaboratively in offline contexts, such as group reports, but typically each person is responsible for one part; the number of contributors is limited; their talk about the task is ephemeral; the tasks are centrally coordinated; and at some point, the product is deemed complete and no further editing takes place. The anarchic nature of contribution to Wikipedia, in combination with the

platform's ease of updating and technical affordances that make process visible, results in a discourse context that seems qualitatively unprecedented.

Interestingly, despite being arrived at by entirely different means, the text of Wikipedia articles can be strikingly similar to that of traditional print encyclopedias, both in quality of content (Giles 2005) and in style (Emigh and Herring 2005). In a corpus-based analysis of structural markers of formality and informality in 15 articles in four corpora—Wikipedia, Everything2 (an online knowledge repository where only individual authors can modify their articles), a traditional print encyclopedia, and Wikipedia Talk pages—Emigh and Herring found that the degree of formality in Wikipedia and the traditional encyclopedia was statistically identical, whereas Everything2 and the Talk pages were significantly less formal. As an explanation for this counterintuitive result, Emigh and Herring suggested that Wikipedia contributors have internalized cultural norms of encyclopedic style, which includes formality. However, this does not explain how the actual editing process takes place over time, or how hundreds of strangers come to an apparent consensus. A tool for visualizing the creation of Wikipedia articles over time has been developed by Viégas, Wattenberg, and Dave (2004), but it has yet to be applied by discourse analysts.

Another emergent Discourse 2.0 phenomenon is the use of channels other than text, and semiotic systems other than verbal language, to carry on "conversational" exchanges. Exchanges in which the turn unit is a video created by an individual and uploaded to a website have been analyzed by Pihlaja (2011) for YouTube, with a focus on verbal metaphor development, and by Kendall (2007) for a user-created animation sharing site, with a focus on the role of visual themes in creating cross-video coherence. McDonald (2007) analyzed conversational exchanges of still webcam images on a graphical community blog, describing four strategies used to create coherence across images: positional play (e.g., showing a picture in which a person is pointing to another picture on the site that is outside the picture frame), animation, text-in-image, and image quotes. In image quotes, a picture or part of a picture posted by a previous contributor is used, sometimes with modification, in a response, as illustrated in the sequence in Figure 6 (from McDonald 2007).

In the first image in this sequence, user Butah P. wishes luck to Valrik, who is leaving a web community to form his own. Valrik ('Val') thanks Butah in the following image, showing himself holding up a screen with a cartoonish drawing of himself over the URL of his new website. Image (c) from user Dana wishes Valrik luck and says, via text, that they will miss him. The latter idea is taken up and exaggerated in image (d) from Runz with Sizzers, who "chopped" the cartoon image of Valrik's head from image (b) and photoshopped it over a tombstone, implying that Valrik has died. In the last image of the sequence, Val denies this by photoshopping in Runz with Sizzers' entire image and superimposing the red 'no' circle over it; he also adds the text: 'I'm not dead! I've just moved to [new URL]'. In this series of exchanges, the cartoon image of Valrik serves as a
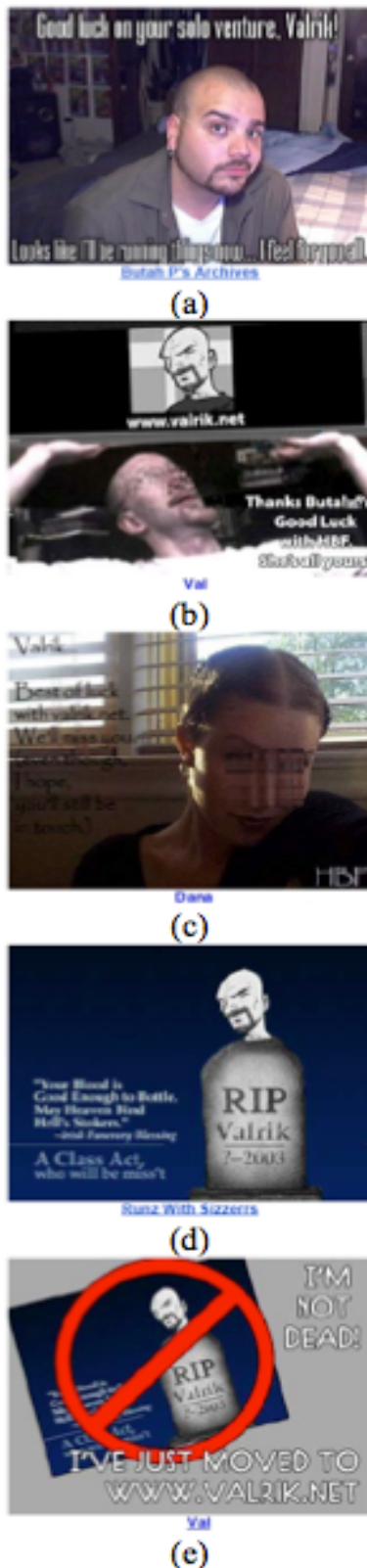
16

Figure 6. A conversation with image quotes from a community blog (McDonald 2007)

reference object that is repeated and transformed dynamically through group interaction. Image quotes are also a key feature of discourse in 'image boards' (image-based online discussion boards) such as 4chan (Bernstein et al. 2011) where they often 'meme', or spread virally.[9]

While in the pre-Internet era videos (or films) were sometimes made that responded to other videos (or films), such practice was restricted mostly to artistic contexts due to the cost and special equipment required; video is now cheaper and easier to create, enabling qualitatively different kinds of communication to take place. And while images with text have been around since the first illuminated manuscripts, and include such familiar genres as comics and children's books, dynamic image texts that develop collaboratively, as in Figure 6, appear to be a recent phenomenon, one enabled by popular access to drawing and photo modification software.

Media convergence is also resulting in a trend for multiple modes of verbal communication to co-occur in the same platform. For example, an individual may respond (asynchronously) to a YouTube video either via text comment or video; may chat (synchronously) while playing World of Warcraft via text or voice; and may send text messages (either synchronous or asynchronous, depending on whether the addressee is logged on) and/or speak (synchronously) to an interlocutor over Skype. On these platforms the mode options are accessed somewhat differently and tend to be used for different communicative purposes (e.g., Newon 2011; Pihlaja 2011; Sindoni 2011), although they may be used co-temporaneously. On the multimodal discussion site VoiceThread.com, in contrast, asynchronous conversations take place in which comments via text, audio, and video are fully integrated within a single interface display, as shown in Figure 7. In the figure, the three modes of commenting are displayed as they appear when each type of message is played back in a

VoiceThread. The case in point involves discussion of a video (in the center of the screen) about the dangers of speeding as part of a high school driver education class (from Herring and Demarest, under review). The contents of the three comments are as follows.

Audio [teacher]: "Hey ya' guys, welcome to .. VoiceThread. I want you to watch this video clip, and give your feedback, either through a video recording, or a voice recording. It's a short clip on speeding, and I'm interested to hear your thoughts."

Video [teacher]: "Hey hey! It's not actually Steve Perry, it's just me … checking this thing out. Uh, this video clip … is pretty neat, I think … I especially enjoy his Australian accent, and I know it's Australian bec- for a few reasons, uh .. the biggest one of them obviously being that the car's steering wheel … is on the right hand side of the car! So. Another cool thing that I … liked about this video, other than the information in it, was the special effects they used, he- he's sorta walked out in front of the car, uh which .. I wouldn't think anybody would normally do, but he did it! And, uh, Steve Perry – come back for us all, please!"

Text [student]: "People may think that they have a quick reaction time but in fact it doesn't matter whether you can break fast or you have good breaks on the car its what speed you are going before you notice something that either has run out on the road or that you are about to hit something...be aware on the suroundings [sic] and do the speed limit or this could be you!!!"

Figure 7. An audio comment (top), a video comment (middle), and a text comment (bottom) in a VoiceThread (Herring and Demarest, under review)

Herring and Demarest argue that VoiceThread provides an unprecedented, authentic environment for research into mode choice, and they profit from it to analyze the effects of mode on participation by gender, expression of positive or negative attitude, and use of metadiscourse to index social awareness, all of which are found to vary according to mode. Audio and video comments in VoiceThreads are made more often by males than by females and contain more self-reference than text, whereas the text comments use relatively more negative language.

Some Web 2.0 sites take the trend towards convergence further by allowing users to embed text directly into video. An example is collaborative video annotations on
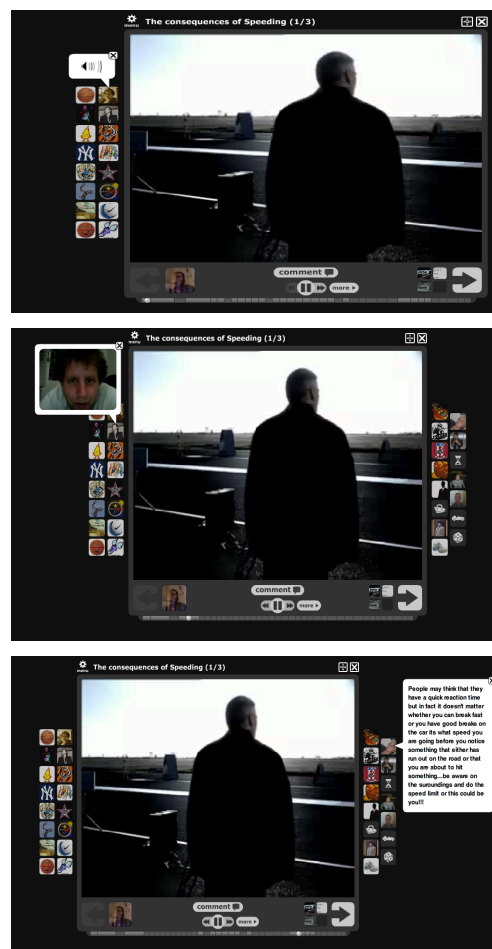
YouTube, an innovation that came about when some users discovered that it is possible to share with other users the URL to the interface that allows the person who uploads a video to add textual annonations to it.[10] Howard (2011) analyzed the referents, tenor, and functional moves performed through such textual annotations posted by education students on a video of a teacher teaching a class (see Figure 8), with the aim of determining the effectiveness of the environment for fostering critical discussion. His findings were mixed; critical moves were evident but there was little reference to specific comments among contributors, hence the "discussions" were not very interactive.

Figure 8. Collaborative text annotations on a YouTube video (Howard 2011)

Finally, as regards non-bona fide communication, artificial intelligence (AI) agents (or 'bots') have been present in online chat spaces since the early days of the public Internet, and they have occasionally succeeded in fooling some people into thinking they were human (Turkle 1995). But recent developments have largely bypassed sophisticated AI, creating fraudulent participants in online forums through the manipulation, computational or otherwise, of simple notions such as frequency and timing of participation. For example, multiple fake online personas, or "sockpuppets", are generated automatically and, posing as real people, made to join online forums for various purposes, including posting comments to create the illusion of consensus on controversial issues (Romm 2011). Another practice, link building, involves unobtrusively inserting links to a company in web forums in order to boost the company's search engine ranking. One

strategy for doing this is to infiltrate a forum and post messages, and after the discussion has moved on to another topic (and the forum moderator is presumably no longer paying attention), edit one's posts to include the links (Haltom 2011). These practices raise the question of how the discourse of non-bona fide participants (including and in addition to their participation patterns) differs from that of real participants, a question that computer-mediated discourse analysts are well qualified to address, and that has practical implications for fraud detection.

With the exception of the articles cited above, however, little analysis of these new discourse phenomena has yet been conducted. A general challenge for emergent media environments is that they need to be analyzed descriptively first before more sophisticated, theoretically-informed analyses can be carried out; this often results in a lag between the emergence of new environments and scholarly understanding of them. The discourse that takes place in emergent environment also raises numerous challenges for the CMDA paradigm. Non-bona fide discourse participants call for new ways of understanding 'participation'. Wikis call for new conceptual understandings of what constitutes a text, what can be data for analysis, and what 'authorship' means. Finally, multimodal discourse requires the analyst to devise new analytical methods and to draw from theoretical frameworks outside linguistics (e.g., visual semiotics). This last point leads to the question of whether multimodal discourse requires a new level of analysis in the CMDA "toolkit", or whether it can be accommodated within the existing paradigm, e.g., with the addition of semiotic methods to address each level. If the latter, at what level does integrative analysis of the meanings and functions of the complex whole take place, and what theories and methods exist to guide it? Table 2 suggests one possible format for incorporating multimodal discourse into the existing CMDA paradigm, as a fifth level.

Level	Issues	Phenomena	Methods
Multimodal communication	Mode effects; cross-mode coherence; reference and address management; generation and spread of graphical meaning units; media co-activity; etc.	Mode choice; text-in-image; image quotes; spatial and temporal positionality and deixis; animation; etc.	Social semiotics; visual content analysis; film studies (?)

Table 2. A possible fifth level of the CMDA paradigm

Practically speaking, there is also a need to devise parallel transcription and visualization displays for textual and non-textual communication, which differ in a number of respects, including temporally. Such displays should include the representation of silences, and the new norms and meanings surrounding them, in multimodal environments.

Finally, it is likely that some apparently emergent phenomena will have antecedents that one is not aware of. Given the inherent difficulty of supporting a claim (i.e., of newness) that requires negative evidence, researchers need to be circumspect in making such claims.

Implications

The tripartite familiar-reconfigured-emergent classification presented above raises questions that have the potential to lead to new theoretical insight. Why, for example, do some discourse phenomena persist, others adapt, and others arise anew in technologically-mediated environments?

I have previously (e.g., Herring 2007) emphasized the importance of taking into account facets of the technological medium and the social context in analyzing computer-mediated discourse. In the case of Discourse 2.0, technological facets that are especially relevant are media convergence and multimodality, including use of images and channel choice. Social facets that continue to be especially relevant include number of participants, anonymity, communicative setting, and cultural context. To these two faceted dimensions, I propose adding a linguistic dimension, based on the observation that different linguistic phenomena appear to be variably sensitive to technological and social effects. One tentative generalization that follows from the examples presented above is that social phenomena such as gender styles (level 4 of CMDA) seem most resistant to technological reshaping, perhaps because they exist at a higher level of abstraction and their expression is not bound to a specific communicative modality. In contrast, interactional phenomena (level 2 of CMDA) seem most likely to be reconfigured. This is understandable in that changes in web design often affect turn-taking,[11] yet "conversational" exchanges persist even in platforms not primarily designed to support them (e.g., Abdul-Mageed 2009; Kendall 2007; Zelenkauskaite and Herring 2008), necessitating reconfigured strategies of interaction. However, emergent phenomena do not appear to be associated with any particular linguistic level; the analyses cited above involve all four CMDA levels: structure (e.g., formality markers), meaning (e.g., functional moves), interactional coherence (e.g., cross-turn reference), and social phenomena (e.g., expressions of sociability and negativity), as well as the fifth "level" of participation. Even if a new level is posited for CMDA to capture the semiotics of multimodal communication, as in Table 2 above, discourse constructed collaboratively through online multimodal systems is only 'emergent' at the present time; in the future it will be 'familiar' or perhaps 'reconfigured' by newer multimodal systems. Thus it is difficult to link emergence with any specific aspect of language use; rather, in the examples presented here, it is tied to technological developments.

Classifying online discourse phenomena in terms of their novelty can make explicit what phenomena are new and in need of basic descriptive research; these "low-hanging fruit" can be attractive objects of study. Reconfigured phenomena can suggest interesting comparative studies that shed light on the effects of technological change on online

discourse. Familiar phenomena can be rewarding to analyze, too, especially when familiar patterns are thought to no longer exist or when they are incorrectly labeled as new. Recognizing what is familiar or reconfigured as such is an important antidote to the tendency towards ahistoricity in new media studies. The three-way classification scheme has other practical benefits, as well. Once a topic for research has been selected, the scheme can be used to frame a research study, select literature for review, determine appropriate methods of analysis, and make interpretive comparisons. However, the scheme should not be applied too strictly; it is coarse grained and intended as a first-pass classification of discourse phenomena, subject to refinement from further investigation.

Summary and conclusion

The main arguments of this paper can be summarized as follows. Discourse in Web 2.0 environments is mostly CMCMD; that is, it occurs on converged media platforms. CMCMD still has a strong textual component, and many CMD phenomena carry over into CMCMD. At the same time, CMCMD increasingly co-occurs with information in other semiotic, especially graphical, modes, leading to the emergence of new converged CMD phenomena. These differences notwithstanding, and although multimodal online discourse presents special challenges, CMDA remains a useful lens through which to analyze new social media.

This last point becomes further evident when discourse-focused approaches to contemporary web phenomena are contrasted with Web 2.0 approaches. Each is a lens with a central focus and a periphery, and certain phenomena are outside the scope of each. Web 2.0 as a lens has centrally in its sights phenomena such as user-created content, user control of content, collaborative processes, and folksonomies. Social interaction and communication are on the periphery, and it "does not see" CMD in contexts other than Web 2.0 sites. In contrast, CMCMD as a lens focuses on language, communication, conversation, social interaction, and media co-activity as they occur anywhere online (including via mobile phones), with collaboration on the periphery; non-interactive content is outside its scope. Each lens is valuable, but CMCMD focuses more squarely on phenomena of interest to discourse analysts. At the same time, it overlaps considerably with Web 2.0 and thus remains a timely lens through which to examine online communicative content.

If this paper demonstrates only one thing, it should be that Discourse 2.0 offers a rich field of investigation for discourse analysts. Especially urgently needed in future research is integrated multimodal analysis. Longitudinal analysis is also needed, and it is increasingly feasible, given the preservation of digital records and computer-assisted corpus analysis methods. Large-scale automated analysis can and should co-exist, in the author's view, with ethnographic case studies. Indeed it seems likely that in the short term, integrated analyses of multimodal communication will be qualitative and ethnographic in nature, at least until the interplay of different semiotic systems in online environments is

understood well enough to distill out parameters that can be subjected to systematic, quantitative analysis.

Notes

1. On the creation of the Web, see Wikipedia (2011a).
2. With the possible exception of WikiLeaks, which is the most recent, all of these sites (and many more) can be found in lists of "Web 2.0" sites available online (see, e.g., http://www.sacredcowdung.com/archives/2006/03/all_things_web.html [accessed 2/28/11] and http://edudemic.com/2011/11/best-web-tools/ [accessed 12/16/11]).
3. For some proposed expansions to CMDA methodology, see the articles in Androutsopoulos and Bießwenger, eds. (2008).
4. These are considered to be CMCMC environments because in each case, the site's original purpose was other than conversational exchange.
5. Herring (2001) provides an overview of many of these practices in the CMD of the 1990s. For discussion of the offline historical predecessors of English-language CMC, see Baron (2000).
6. 'Small stories' are nontraditional narratives, often personal rather than fictionalized. In Web 2.0 environments they are reconfigured from their offline antecedents in that they tend to be interactive, hypertextual, and collaborative (see Georgakopoulou, in press).
7. It is syntactically ill-formed, because *Katy: ?* is not a grammatical sentence, and pragmatically ill-formed, because the default expectation of a Facebook status update is that it provide information about what the user is doing or thinking.
8. These examples are advanced with the caveat that more astute researchers than the present author may find plausible antecedents to them. Some could be argued to have antecedents in specialized offline contexts; however, to the best of the author's knowledge, none have previously existed as common communicative practices online or offline.
9. An example of this is 'lolcat' images—photographs of cats with text superimposed (in misspelled and ungrammatical English, e.g., 'I can haz cheezburger')—which originated on 4chan and spread subsequently to other online environments (Bernstein et al. 2011).
10. Craig Howard, email communication, January 8, 2012.
11. Numerous examples of this can be found in the CMD literature (e.g., Cherny 1999; Herring 1999). For an extreme example of reconfigured turn-taking in a synchronous chat system, see Anderson, Beard, and Walther (2010).

References

Abdul-Mageed, Muhammad. 2009. Comments on the Al Jazeera English YouTube Channel. Panel on Convergent Media Computer-Mediated Communication, *Internet Research 10.0*, Milwaukee, WI, October 9.

Anderson, Nate. 2006. Tim Berners-Lee on Web 2.0: "[N]obody even knows what it means". *Ars Technica* [website]. http://arstechnica.com/business/news/2006/09/7650.ars [accessed December 3, 2011]

Anderson, Jeffrey F., Fred K. Beard, and Joseph B. Walther. 2010. Turn-taking and the local management of conversation in a highly simultaneous computer-mediated communication system. *Language@Internet* 7, article 7. http://www.languageatinternet.org/articles/2010/2804

Androutsopoulos, Jannis. 2011. From variation to heteroglossia in the study of computer-mediated discourse. In *Digital discourse: Language in the new media*, ed. Crispin Thurlow and Kristine Mroczek, 277-298. New York: Oxford University Press.

---------- and Michael Beißwenger, eds. 2008. Data and methods in computer-mediated discourse analysis. Special issue of *Language@Internet* 5. http://www.languageatinternet.org/articles/2008

Battelle, John, and Tim O'Reilly. 2010. Web 2.0 Summit is a wrap. *Web 2.0 Summit* [website]. http://www.web2summit.com/web2010 [accessed December 3, 2011]

Baron, Naomi S. 2000. *Alphabet to email: How written English evolved and where it's heading*. London and New York: Routledge.

Bernstein, Michael S., Andrés Monroy–Hernández, Drew Harry, Paul André, Katrina Panovich, and Greg Vargas. 2011. 4chan and /b/: An analysis of anonymity and ephemerality in a large online community. *Proceedings of the Fifth International Conference on Weblogs and Social Media*, 50–57. Menlo Park, CA: AAAI Press.

Blood, Rebecca. 2002. *The weblog handbook: Practical advice on creating and maintaining your blog*. Cambridge, MA: Perseus Publishing.

boyd, danah, Scott Golder, and Gilad Lotan. 2010. Tweet tweet retweet: Conversational aspects of retweeting on Twitter. *Proceedings of the Forty-Third Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Press. http://csdl.computer.org/dl/proceedings/hicss/2010/3869/00/03-06-04.pdf

Cherny, Lynn. 1994. Gender differences in text-based virtual reality. In *Cultural Performances: Proceedings of the Third Berkeley Women and Language Conference*, ed. Mary Bucholtz, Anita C. Liang, Laurel A. Sutton, and Caitlin Liang, 102-115. Berkeley, CA: Berkeley Women and Language Group.

----------. 1995. The modal complexity of speech events in a social mud. *Electronic Journal of Communication* 5. http://fragment.nl/mirror/Cherny/The_modal_complexity.txt

----------. 1999. *Conversation and community: Chat in a virtual world*. Stanford, CA: CSLI.

Crowston, Kevin, and Michelle Williams. 2000. Reproduced and emergent genres of communication on the World-Wide Web. *The Information Society* 16(3):201–216.

Crystal, David. 2008. *Txtng: The Gr8 Db8*. New York: Oxford University Press.

Danet, Brenda, and Susan C. Herring, eds. 2007. *The multilingual Internet: Language, culture, and communication online*. New York: Oxford University Press.

Emigh, William, and Susan C. Herring. 2005. Collaborative authoring on the Web: A genre analysis of online encyclopedias. *Proceedings of the Thirty-Eighth Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Press. http://ella.slis.indiana.edu/~herring/wiki.pdf

Eriksen, Lars B. and Carina Ihlström. 2000. Evolution of the web news genre: The slow move beyond the print metaphor. *Proceedings of the Thirty-Third Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE. http://csdl.computer.org/dl/proceedings/hicss/2000/0493/03/04933014.pdf

Georgakopoulou, Alexandra. In press. Narrative and CMC. In *The handbook of pragmatics of computer-mediated communication*, ed. Susan C. Herring, Dieter Stein, and Tuija Virtanen. Berlin: Mouton de Gruyter.

Giles, Jim. 2005. Internet encyclopaedias go head to head. *Nature* 438: 900–901. http://bert.lib.indiana.edu:2146/nature/ journal/v438/n7070/full/438900a.html

Goffman, Erving. 1959. *The presentation of self in everyday life*. New York: Doubleday.

Haltom, Aaron. 2011. How to spot a link builder. Unpublished paper, Indiana University, Bloomington.

Herring, Susan C. 1993. Gender and democracy in computer-mediated communication. *Electronic Journal of Communication* 3(2). http://ella.slis.indiana.edu/~herring/ejc.doc

----------. 1999. Interactional coherence in CMC. *Journal of Computer-Mediated Communication* 4(4). http://jcmc.indiana.edu/vol4/issue4/herring.html

----------. 2001. Computer-mediated discourse. In *The handbook of discourse analysis*, ed. Deborah Schiffrin, Deborah Tannen, and Heidi Hamilton, 612-634. Oxford, UK: Blackwell Publishers.

----------. 2003a. Gender and power in online communication. In *The handbook of language and gender*, ed. Janet Holmes and Miriam Meyerhoff, 202-228. Oxford, UK: Blackwell Publishers.

----------. 2003b. Dynamic topic analysis of synchronous chat. In *New Research for New Media: Innovative Research Methodologies Symposium Working Papers and Readings*. Minneapolis, MN: University of Minnesota School of Journalism and Mass Communication. http://ella.slis.indiana.edu/~herring/dta.2003.pdf

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

----------. 2007. A faceted classification scheme for computer-mediated discourse. *Language@Internet* 4, article 1. http://www.languageatinternet.org/articles/2007/761

----------. 2009. Convergent media computer-mediated communication: Introduction and theory. Panel on Convergent Media Computer-Mediated Communication, *Internet Research 10.0*, Milwaukee, WI, October 9.

---------- and Bradford Demarest. Under review. Mode choice in multimodal comment threads: Effects on participation, sociability, and attitude.

---------- and Andrew Kurtz. 2006. Visualizing Dynamic Topic Analysis. *Proceedings of CHI'06*. New York: ACM. http://ella.slis.indiana.edu/~herring/chi06.pdf

----------, Daniel O. Kutz, John C. Paolillo, and Asta Zelenkauskaite. 2009. Fast talking, fast shooting: Text chat in an online first-person game. *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Press. http://ella.slis.indiana.edu/~herring/hicss.bzflag.pdf

---------- and Carole G. Nix. 1997. Is "serious chat" an oxymoron? Pedagogical vs. social uses of Internet Relay Chat. Paper presented at the American Association of Applied Linguistics Annual Conference, Orlando, FL, March 11. http://ella.slis.indiana.edu/~herring/aaal.1997.pdf

---------- , Lois Ann Scheidt, Sabrina Bonus, and Elijah Wright. 2004. Bridging the gap: A genre analysis of weblogs. *Proceedings of the Thirty-Seventh Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Press. http://ella.slis.indiana.edu/~herring/herring.scheidt.2004.pdf

Honeycutt, Courtenay and Susan C. Herring. 2009. Beyond microblogging: Conversation and collaboration via Twitter. *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Press. http://ella.slis.indiana.edu/~herring/honeycutt.herring.2009.pdf

Howard, Craig. 2011. Collaborative video annotation as critical discourse. Paper presented at the Georgetown University Roundtable on Languages and Linguistics, Washington, DC, March 12

Kapidzic, Sanja, and Susan C. Herring. 2011a. Gender, communication, and self-presentation in teen chatrooms revisited: Have patterns changed? *Journal of Computer-Mediated Communication* 17(1):39-59.

Kapidzic, Sanja, and Susan C. Herring. 2011b. Gender, innovation, and non-standardness in teen chat language. Paper presented at the Georgetown University Roundtable on Languages and Linguistics, Washington, DC, March 12.

Kendall, Lori. 2007. Colin Mochrie vs. Jesus H. Christ: Messages about masculinities and fame in online video conversations. *Proceedings of the Fortieth Hawaii International Conference on Systems Sciences*. Los Alamitos, CA: IEEE Press. http://www.ideals.illinois.edu/bitstream/handle/2142/705/co?sequence=2

Kutz, Daniel O., and Susan C. Herring. 2005. Micro-longitudinal analysis of Web news updates. *Proceedings of the Thirty-Eighth Hawai'i International Conference on*

System Sciences. Los Alamitos, CA: IEEE Press.
http://ella.slis.indiana.edu/~herring/news.pdf

Lambiase, Jacqueline J. 2010. Hanging by a thread: Topic development and death in an online discussion of breaking news. *Language@Internet* 7, article 9. http://www.languageatinternet.org/articles/2010/2814

Lee, Carmen. 2011. Texts and practices of micro-blogging: Status updates on *Facebook*. In *Digital discourse: Language in the new media*, ed. Crispin Thurlow and Kristine Mroczek, 110-128. New York: Oxford University Press.

Lenihan, Aoife. 2011. "Join our community of translators": Language ideologies and Facebook. In *Digital discourse: Language in the new media*, ed. Crispin Thurlow and Kristine Mroczek, 48-64. New York: Oxford University Press.

Linden, Greg. 2006. Tim O'Reilly and defining Web 2.0. *Geeking with Greg* [blog], May 14. http://glinden.blogspot.com/2006/05/tim-oreilly-and-defining-web-20.html

McDonald, David. 2007. Visual conversation styles in web communities. *Proceedings of the Fortieth Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE Press. http://www.pensivepuffin.com/dwmcphd/papers/McDonald.HICSS-40.preprint.pdf

Miller, Caroline R., and Dawn Shepherd. 2004. Blogging as social action: A genre analysis of the weblog. In *Into the blogosphere: Rhetoric, community, and culture of weblogs,* ed. Laura J. Gurak, Smiljana Antonijevic, Laurie Johnson, Clancy Ratliff, and Jessica Reyman. Minneapolis, MN: University of Minnesota. http://blog.lib.umn.edu/blogosphere/blogging_as_social_action_a_genre_analysis_of _the_weblog.html

Newon, Lisa. 2011. Multimodal creativity and identities of expertise in the digital ecology of a World of Warcraft guild. In *Digital discourse: Language in the new media*, ed. Crispin Thurlow and Kristine Mroczek, 131-153. New York: Oxford University Press.

Nishimura, Yukiko. 2011. Japanese *keitai* novels and ideologies of literacy. In *Digital discourse: Language in the new media*, ed. Crispin Thurlow and Kristine Mroczek, 86-109. New York: Oxford University Press.

O'Reilly, Tim. 2005. What is Web 2.0? Design patterns and business models for the next generation of software. *O'Reilly Network* [website], September 30. http://oreilly.com/web2/archive/what-is-web-20.html

Peterson, Eric. 2011. How conversational are weblogs? *Language@Internet* 8, article 8. http://www.languageatinternet.org/articles/2011/Peterson

Pihlaja, Stephen. 2011. Cops, popes, and garbage collectors: Metaphor and antagonism in an atheist/Christian YouTube video thread. *Language@Internet, 8*, article 1. http://www.languageatinternet.org/articles/2011/3044

Puschmann, Cornelius. In press. Blogging. In *The handbook of pragmatics of computer-mediated communication*, ed. Susan C. Herring, Dieter Stein, and Tuija Virtanen.

Berlin: Mouton de Gruyter.

Romm, Joe. 2011. Denier-bots live! Why are online comments' sections over-run by the anti-science, pro-pollution crowd? *Think Progress*, February 20. http://thinkprogress.org/romm/2011/02/20/207554/denier-bots-live-why-are-online-comments-sections-over-run-by-the-anti-science-pro-pollution-crowd/

Tannen, Deborah. 1990. *You just don't understand: Women and men in conversation*. New York: Morrow.

Thurlow, Crispin, and Kristine Mroczek, eds. 2011. *Digital discourse: Language in the new media*. New York: Oxford University Press.

Turkle, Sherry. 1995. *Life on the screen: Identity in the age of the Internet*. NY: Simon and Schuster.

Severinson Eklundh, Kerstin. 2010. To quote or not to quote: Setting the context for computer-mediated dialogues. *Language@Internet* 7, article 5. http://www.languageatinternet.org/articles/2010/2665

---------- and Clare Macdonald. 1994. The use of quoting to preserve context in electronic mail dialogues. *IEEE Transactions on Professional Communication* 37(4):97-202.

Sindoni, Maria Grazia. 2011. "Mode-switching": Speech and writing in videochats. Paper presented at the Georgetown University Roundtable on Languages and Linguistics, Washington, DC, March 11.

Vaisman, Carmel. 2011. Performing girlhood through typographic play in Hebrew blogs. In *Digital discourse: Language in the new media*, ed. Crispin Thurlow and Kristine Mroczek, 177-196. New York: Oxford University Press.

Viégas, Fernanda B., Martin Wattenberg, and Kushal Dave. 2004. Studying cooperation and conflict between authors with history flow visualizations. *Proceedings of the 2004 Conference on Human Factors in Computing Systems*, 575-582. New York: ACM.

Virtanen, Tuija. In press. Performativity in CMC. In *The handbook of pragmatics of computer-mediated communication*, ed. Susan C. Herring, Dieter Stein, and Tuija Virtanen. Berlin: Mouton de Gruyter.

Werry, Christopher C. 1996. Linguistic and interactional features of Internet Relay Chat. In *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*, ed. Susan C. Herring, 47-63. Amsterdam/Philadelphia: John Benjamins.

Wikipedia. 2011a. *History of the World Wide Web*. http://en.wikipedia.org/wiki/History_of_the_World_Wide_Web [accessed 4:05 pm EST, December 3, 2011]

----------. 2011b. *Web 2.0*. http://en.wikipedia.org/wiki/Web_2.0 [accessed 4:07 pm EST, December 3, 2011]

----------. 2011c. *Facebook*. http://en.wikipedia.org/wiki/Facebook [accessed 9:54 pm EST, December 3, 2011]

Winston, Brian. 1998. *Media, technology, and society: A history--from the printing press to the superhighway*. London: Routledge.

Wolf, Alecia. 2000. Emotional expression online: Gender differences in emoticon use. *CyberPsychology and Behavior* 3:827-833.

Zelenkauskaite, Asta, and Susan C. Herring. 2008. Television-mediated conversation: Coherence in Italian iTV SMS chat. *Proceedings of the Forty-First Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Press. http://ella.slis.indiana.edu/~herring/hicss08.pdf