Code-switching in computer-mediated communication

Jannis Androutsopoulos

1. Introduction

This chapter discusses how processes known as code-switching and code-mixing manifest in computer-mediated communication (CMC). Code-switching is generally defined as "juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or subsystems" (Gumperz 1982: 59) or as "the use of more than one linguistic variety, by a single speaker in the course of a single conversation" (Heller and Pfaff 1996: 594). Depending on approach, code-mixing is defined as the juxtaposition of linguistic codes within a single sentence or as an alternation of codes that lacks a specific pragmatic function. Code-switching and code-mixing are common linguistic practices among bilingual and multilingual people and therefore easily find their way into communication via digital media. The research literature discussed in this chapter offers examples of code-switching in a wide range of CMC modes and social settings, including texting among South African young people, chatting among second-generation Indians, emails among Egyptian professionals, forum discussions among Persian expatriates, and fan fiction by Finnish bloggers.

Even though code-switching (CS) online attracted the attention of linguists as early as the mid-1990s (Georgakopoulou 1997; Paolillo 1996), it remains less well researched in comparison to other linguistic processes in CMC.[1] The topic is equally under-researched in contact linguistics and multilingualism studies (see Dorlein and Nortier 2009). In a study of written CS, Callahan (2004: 92) claims that "[t]he majority of code-switching in nonfiction […] is found in advertising and journalistic writing". This marginalisation of CS bears no relation to the spread of the practice itself. Given the importance of multilingualism and the pervasiveness of digital media worldwide, it seems safe to assume that digitally-mediated communication (via both networked computers and mobile networked devices) offers opportunities for written CS at an unprecedented scale.

CS in CMC is relevant not only because it is there (and not yet well understood) but also for the insights it can offer to pragmatics, sociolinguistics, and discourse studies. By its very nature, CS calls into question a number of assumptions that have dominated linguistic scholarship, such as the discreteness of linguistic systems and the primacy of monolingualism as the default condition of language in society (Heller and Pfaff 1996). Likewise, the study of CS challenges fundamental assumptions in CMC studies. Early linguistic research on CMC focused on language/technology relations, and some of its key categories are conceived and best operate within a monolingual frame, such as the position of computer-mediated discourse between spoken and written language. However, CS defies easy classifications based on media factors alone, and it requires an emphasis on the interrelation of medium and social/situation factors (Herring 2007). Moreover, CMC as a discourse field challenges the assumption that spoken face-to-face interaction is *the* essential site of code-switching.

Based on research literature on several languages,[2] the aim of this chapter is to organise the available research evidence, identify commonly asked and still untapped questions, and pinpoint limitations of present scholarship. The chapter does not claim to offer a comprehensive overview of forms and functions of CS on the internet; rather, it can only outline how CS has been studied in the languages and sociolinguistic settings I have been able to identify in the literature (section 4). Still, the available literature offers ample evidence that CMC is a site for the meaningful use of language alternation, and a critical synthesis of available research can offer insights into what are promising perspectives for further research, as well as what methods have been mainly used.

The organisation of this chapter is as follows: After a brief outline of code-switching research frameworks (section 2), I distinguish between conversational and non-conversational CS in CMC and delimit both from other patterns of multilingualism online (section 3). The subsequent two sections provide an overview of research (section 4) and discuss the different types of settings in which CS in CMC has been studied (section 5). The next section turns to forms and functions of CS online (sections 6). The chapter concludes with a discussion of CMC as a new domain for CS (section 7) and an outlook for future research (section 8).

2. Code-switching frameworks

The question of what patterns of CS are attested in CMC environments cannot be answered independently of the frameworks within which CS is studied. In general, four points characterize the state of the art:

- Researchers do not use one single framework of CS analysis but rather a number of different approaches.
- A generally accepted methodology that takes the specifics of CMC into account has not yet been developed. Researchers draw on frameworks originally developed for the analysis of spoken discourse, despite criticisms of the adequacy of such frameworks (Hinrichs 2006: 28-30; Leppänen and Peuronen 2011).
- The predominant perspective is pragmatic and sociolinguistic rather than grammatical and linguistic (see also Dorleijn and Nortier 2009: 133). While structural descriptions are not entirely absent from the extant literature, its main aim has been to understand the pragmatic functions, social purposes, and interactional dynamics of CS online.
- The correspondence of online written CS to its offline spoken counterpart is a common concern, but it is also a contested issue, as will become obvious in this chapter.

In terms of frameworks, the literature contains elements of the "three most influential contributions to theory in the sociolinguistic branch of CS studies" (Hinrichs 2006: 28): The first is the markedness model of Carol Myers-Scotton (1993), in particular her concepts of code-switching as a marked (i.e., unexpected, unconventional) or unmarked (expected) choice. Second are concepts introduced by John J. Gumperz (1982), such as the distinction between situational and metaphorical CS, the distinction between '*we*-code' and '*they*-code', his classification of discourse

functions of conversational CS, and the notion of CS as a contextualization cue. Thirdly, researchers also draw on the conversation-analytic framework for the study of bilingual interaction by Peter Auer (1995, 1998b, 1999, 2000), which builds on and develops some of Gumperz' ideas.[3] Other repeatedly used concepts include the syntactic distinction between inter- and intra-sentential CS; concepts from pragmatics such as politeness, face, and interpersonal alignment (e.g., Georgakopoulou 1997); Myers-Scotton's notion of matrix language; and Auer's notion of base language, which refers to the backdrop against which switches to another language or dialect become meaningful.

In terms of discourse functions of CS, the classifications by Gumperz (1982) and Auer are widely used in the literature, e.g., by Androutsopoulos (2006a, 2007a), Sebba (2003), Androutsopoulos and Hinnenkamp (2001), and Paolillo (1996, in press). Both conceive of CS as a contextualization cue, i.e., a resource used by participants to frame their interpretations of what is being said. In a nutshell, Gumperz' categories include: switching for reported speech; addressee specification; clarification, emphasis, expressivity; message qualification (e.g., separating facts from comment); and contrasting personal with objective viewpoints. Auer distinguishes between preference- (or participant-)related and discourse-related CS. The first comprises switches that suit the speaker's or addressee's preference, as well as instances of language negotiation between the interlocutors.[4] Discourse-related switching in a conversational episode "contributes to the organization of discourse in that particular episode" (Auer 1995: 125). Its subtypes partially overlap with those by Gumperz, one important addition being the focus on CS as a device for the internal organization of conversational turns. How these categories have been applied to CS in CMC is discussed in the overview of research findings below (section 6).

The limitations of a conversation-analytic approach with respect to CMC data are well discussed in the CMC literature (e.g., Beißwenger 2008; Herring 1999). CMC technologies rule out one key mechanism of conversational organisation, the turn-taking system; more generally, the lack of visual channels – and, in asynchronous CMC, the temporal gap between contributions – means that important dimensions of the interactional co-construction of meaning are altered or restricted. However, these restrictions do not rule out the sequential organization of computer-mediated discourse,[5] which can be studied with conversation analytic categories. Furthermore, CMC research has also established that users develop creative procedures to cope with these limitations, including the usage of specific turn-taking signals and linguistic innovations such as emoticons and laughter acronyms. Related to this, and specifically relevant to CS, is Georgakopoulou's (1997) suggestion that the lack of ordinary contextualization cues due to the absence of the visual channel "results in an increased reliance on code-centered contextualization cuing, which would be otherwise delegated to different signals" (158). In other words, CMC interlocutors use code-switching, style shifting, and other manipulations of written signs in order to accomplish pragmatic work that would be accomplished by phonological variation, prosody, gaze, posture, and other cues in ordinary spoken conversation. This establishes a productive theoretical link among linguistic choices, communicative practices, and media affordances.

These are elements of a basic theoretical vocabulary, which CMC researchers appropriate, articulate, and apply in different ways. For example, Hinrichs (2006)

combines ideas from all three frameworks with categories from Creole linguistics. Leppänen (2007) draws on the four types of language alternation in Auer's framework (i.e., insertional switching, insertional mixing, alternational switching, and alternational mixing) to examine alternation between Finnish and English in a range of digital genres. Androutsopoulos (2006a, 2007a) defines the discussion thread of web forums as the equivalent of a conversational episode and therefore as the level at which to determine the base language of discussion (Auer 2000), against which the directionality of switches is examined.

3. Distinguishing code-switching from multilingualism in computer-mediated discourse

As illustrated by the introductory definitions, CS is typically thought of as a process of (informal or institutional) spoken interaction. This raises the question of how to transfer defining conditions such as "speech exchange" or "conversational episode" to computer-mediated discourse (CMD), where language is typed and the very notion of conversation is contested, as discussed earlier. A rather restrictive reading would be to limit the scope to what has been identified as "interactive written discourse" (Ferrara, Brunner, and Whittemore 1991), i.e., dyadic or multi-party verbal exchanges via technologies of synchronous CMC such as Internet Relay Chat, web-based chat systems, or Instant Messaging, which some researchers identify as the closest approximations of spoken conversation. However, a considerable part of the literature examines CS in asynchronous (dyadic or multi-party) CMD, in particular email, newsgroups, web forums, and texting. Still other researchers have located bilingual discourse in edited genres aimed at a reading audience, such as weblogs and fan fiction. These, in turn, are reminiscent of more traditional written CS, which is a practice known at least since medieval poetry and ranges from fictional representations of conversational dialogue (as in novels or stageplay) to diary writing and newspaper discourse (Androutsopoulos 2007b; Callahan 2004). The approach advocated here is an inclusive one. I suggest that there is no *a priori* reason not to consider blogs or social networking profile pages as sites of bilingual discourse. Nevertheless, a terminological and analytical distinction is required between two main types of CS in CMC: For convenience, I shall call them 'conversational' (dyadic or multi-party, synchronous or asynchronous) and 'non-conversational' (edited and published by a single author), bearing in mind the caveats mentioned earlier.

With these distinctions in mind, CS will now be distinguished from other aspects of multilingual CMD. The aim in doing so is to determine when the coexistence of more than one language in CMD constitutes CS and to identify the discourse units, or equivalents of a "speech exchange", in which CS can be located.[6] I distinguish CS from four other patterns of multilingualism in CMC:

a) The 'multilingual Internet' as a whole
b) The coexistence of different languages on a web page or thread
c) Language choices for emblems
d) Sequential language choices lacking a dialogical interrelation

The first pattern refers to the multilingualism that emerges from the coexistence of different websites, channels, forums, etc. on the web. For instance, Internet Relay

Chat (IRC) as an entire system is massively multilingual, and so are platforms like *flickr* and *blogger.com* in their entirety. While this level is sometimes referred to in discussions of the multilingual internet (see, e.g., some chapters in Danet and Herring 2007a), CS must be located on the more concrete level of individual web pages, discussion threads, or chat channel sessions. At this level, it is necessary to single out the multilingualism that emerges from the coexistence of different discourse units on a single web page. Contemporary websites are composed from textual units of diverse kinds and origins – editorial content, user-generated content, advertisements, graphic-designed banners, user comments – which are adjacent in screen space. Multilingual surfaces emerge when such adjacent units "speak", or are cast in, different languages or dialects as a result of different purposes, audiences, or production processes. Consider *flickr* pages, for example, where headlines, banners tags, and comments are not necessarily in the same language (Lee and Barton 2011). Consider media-sharing sites, where the language of posted items is often different from that of comments, and social network profile pages, where different 'friends' may contribute wall posts in different languages. These are multilingual or indeed heteroglossic discourse spaces (Androutsopoulos 2011), but they do not automatically constitute instances of CS. To the extent that their constituents differ in terms of authorship and production process, they often cannot be conceived of as part of one "episode"; what holds them together is their spatial coexistence in product and reception and not their dialogic orientation to each other. Indeed, the units that make up multilingual web surfaces are often monolingual in themselves. However, nothing prevents some of these units from containing CS. Think of an online newspaper where comments to news items come in different varieties (e.g., standard and dialect) or languages (e.g., the national and a minority language), thus constituting candidates for CS.

Still at the level of single web pages or threads, the third distinction is between CS and language choices for emblems, i.e., textual units that identify and represent individual or institutional actors in CMC. Emblems include website names and screen names, slogans, user signatures, and navigation bars: They are usually graphically designed and placed on website layout. There is evidence that emblematic language choices can extend the multilingual make-up of a website or discussion forum by introducing linguistic resources that are not regularly used in ongoing user discussions or editorial copy (Androutsopoulos 2006a, 2007b). This strategic allocation of languages leads to a sort of emblematic bilingualism*,* which does not challenge the dominant language in terms of informational load but selects another code as relevant to the identity of an institution or individual. Important as these processes may be to multilingual discourse on the web, they are not instances of CS (to the extent that these textual units do not include CS in themselves), even though they may coexist with instances of CS: Think of a blog that uses emblems in order to make a minority language visible, while its entries and their comments include CS between that minority and the respective majority language.

Finally, conversational CMD modes need to be scrutinised in order to assess if the linguistic diversity they host constitutes CS. Consider IRC, a model case of interactive written discourse. Even if IRC is carried out in a language other than English, the system-generated messages that announce users who are entering or leaving a channel are automatically cast in English (Siebenhaar 2005, 2006). While they contribute to the multilingual make-up of channels, these automated messages are not part of participants' code choices and therefore need to be excluded from CS

analysis (as Siebenhaar explicitly does). In a similar vein, we may ask whether the lists of comments that respond to a 'spectacle' or 'prompt' in web 2.0 environments (a blog post, photo, video, song, etc.) qualify as instances of CS. Comments may in principle be posted in any language, and even though the language of the 'prompt' will often make a particular language relevant, it is not uncommon to find comments in different languages. However, whether these individual code choices by different commenters are sequentially related to each other needs to be demonstrated rather than assumed *a priori*. My observations suggest that comments often respond to the 'prompt' rather than to other posts, and while they all contribute to the multilingual character of a chain of comments, they are not forcibly instances of CS (although they can draw on CS, of course, in their own internal organisation).

In conclusion, even though all CS entails a juxtaposition of linguistic codes, as Gumperz' seminal definition puts it (see section 1), there is good reason to adapt a restrictive view of what juxtapositions will constitute CS in CMD. Multilingual CMD environments are shaped at different levels by contrastive language choices which are motivated and meaningful, but for these contrasts to qualify as CS, evidence is required that they are in some way dialogically interrelated by responding to previous, and contextualizing subsequent, contributions.

4. Code-switching across CMC modes

While limited in number, studies of CS in CMC have examined a range of media modes and sociolinguistic settings, using a range of different methods in the process. Table 1 provides an overview of the research summarized in this chapter. Its main categories – mode, participation framework, languages involved, social settings, and methods – orient both to the distinction between medium and situation factors in Herring's (2007) CMD classification scheme and to respective distinctions in studies of code-switching in writing (Callahan 2004). The dimension of synchronicity (synchronous and asynchronous modes) is not listed separately, as it is a stable feature of each mode, but it is discussed in detail in the next section.

As regards CMC modes, most literature considers traditional, pre-web modes of interpersonal CMC (IRC, email, mailing lists, Usenet groups) and only a few web-based modes (discussion forums). This results in a focus on more language-heavy modes, whereas multimodally intensive modes such as media-sharing sites and profile pages are hardly examined. As a consequence, the tantalising question of how to deal with CS when modes other than written text are heavily involved in the production of meaning has not been dealt with systematically.[7] Nonetheless, this coverage still allows us to conclude that code-switching may in principle occur in any CMD mode, be it unidirectional or interactive, synchronous or asynchronous, dyadic or public, private or professional.

The category of participation framework asks what roles of participation are made possible by and conventionalised in the usage of a given CMD mode. Participation structure is a situation factor in Herring's (2007) classification scheme. However, it is in part a medium factor as well, since different CMC modes facilitate different participation frameworks. Instead of the usual labeling in terms of number of participants – one-to-one, one-to-many, many-to-many – I prefer the terms 'private'

and 'public', the former corresponding to dyadic exchanges or those among a limited number of known participants, while the latter refers to communication in a public CMD environment such as a channel or forum, which by definition includes unknown participants. I return to the consequences of this distinction for bilingual discourse below (section 5.2).

Table 1. Classification of selected research publications on code-switching in CMC
(Publications listed by mode and chronologically within each mode. Modes: s=synchronous, a=asynchronous. Methods: QN=quantitative, QL=qualitative. Updated to June 2010.)

Authors	CMC Modes	Partici- pation	Languages	Participants, Social setting	Methods
Paolillo 1996, 2001, in press	IRC (s) Usenet (a)	Public	English/Hindi English/Punjabi	Ethnic minority	QN, QL
Androutsopoulos and Hinnenkamp 2001	IRC (s)	Public	German/Greek Turkish/German	Ethnic minority	QL
Tsaliki 2003	IRC (s)	Public	Greek/English	Ethnic minority	QL
Androutsopoulos and Ziegler 2004	IRC (s)	Public	standard German/ dialect	Citychat	QN, QL
Siebenhaar 2005, 2006, 2008	IRC, Webchats (s)	Public	Swiss German dialects/standard German	Youth culture, Flirt communities	QN, QL
Hinnenkamp 2008	IRC (s)	Public	Turkish/German	Ethnic minority	QL
Georgakopoulou 1997, 2004	Email (a)	Private	Greek (L1)/English	Friends group	QL
Hinrichs 2006	Email, Forums (a)	Private Public	Jamaican Creole (L1)/English	Univ students Jam. diaspora	QN, QL
Warschauer et al. 2007	Email (a)	Private	English/Egyptian Arabic (L1)	Young professionals	QN
Lee 2007	Email (a), ICQ (s)	Private	Cantonese (L1)/English	Cantonese University students	QN
Goldbarg 2009	Email (a)	Private	Spanish (L1)/English	Graduate students	QN
Tsiplakou 2009	Email (a)	Private	Greek (L1)/Eng- lish/French/Greek Cypriot dialect	Fellow academics	QN, QL
Deumert and Masinyana 2008	SMS (a)	Private	isiXhosa (L1)/English	Young adults	QN, QL
McClure 2001	Mailing lists (a)	Public	English/Assyrian	Ethnic minority	QL
Sebba 2003	Bulletin board (a)	Public	English (L1)/stylized Creole	Pop culture, Comedy fans	QL
Androutsopoulos 2004	Forums, Guest- books (a)	Public	German (L1)/English	Music youth culture	QL
Sperlich 2005	Forums (a)	Public	Niuean (L1)/English	Local community, Diaspora	QN, QL
Androutsopoulos 2006a, 2006b, 2007b	Forums (a)	Public	German/Greek, Persian, Hindi, Arabic	Ethnic minority	QN, QL
Leppänen 2007; Leppänen et al. 2009	Forums, Blogs (a)	Public	Finnish (L1)/English	Youth cultures, music, fan fiction, sports	QL

The third column lists the languages that are (most) relevant to each study.[8] A number of sociolinguistic constellations are evident here, between majority and

minority (heritage, immigrant, community) language, between national language and English, between a Creole and its lexifier (Hinrichs 2006), standard and dialect, and standard and stylized vernacular speech ("Ali G language", Sebba 2003). Code-switching between varieties of one language is reported from the Greek- and German-speaking areas and studied along the same functional lines as bilingual CS. For the sake of overview, this table downplays the few documented instances of "polylingual languaging" (Hinnenkamp 2008; Jørgensen 2008; Tsiplakou 2009), which involves the (playful) use of bits and pieces from different languages, language varieties, or styles (see sections 6, 7).

The user groups and social settings that are examined in each study are categorised in the fourth column. The label "ethnic minority" brings together situations in which an immigrant or diasporic group uses a minority and a majority language. Other labels foreground online practices by young people, youth culture groups, or local communities, in which relations between national language and English, or standard and dialect, have been examined. Other documented cases involve multilingual academics or professionals, small language communities (Sperlich 2005), and a message board devoted to a Creole-speaking comedian (Sebba 2003). The overview does not mean to suggest rigid constraints between language pairs and social settings. Some language pairs (especially those involving English) are instantiated in different social settings, with distinct CS patterns in each case. For example, English/Creole CS shows different patterns in emails among Jamaican students and on forums by diasporic Jamaicans (Hinrichs 2006). Likewise, CS between national language and English CS is markedly different in youth-culture contexts and among elite bilingual expatriates (see section 5).

The rightmost column features a broad distinction between qualitative and quantitative methods of analysis. The 'qualitative' label comprises methods from conversation, discourse, narrative, or style analysis, which have been used for the study of both conversational and non-conversational CMD. Such research often involves elements of online ethnography, whereby researchers work with data from social networks they themselves belong to (as with Georgakopoulou 1997 and Tsiplakou 2009) and focuses on functions rather than structures of CS. 'Quantitative' encompasses quantifications based on questionnaire data (Goldbarg 2009; Tsiplakou 2009) or coding of textual data (as with Paolillo 2001; Siebenhaar 2008). What is remarkable is the frequency of mixed method approaches, which combine a 'bird's eye view' of the distribution of languages over a large data set with a detailed view of local processes of switching and mixing. Siebenhaar (2008) adds an intermediate level, "windowed or moving average analysis", in which quantification zooms in to slices of time or channel activity as opposed to whole channel comparisons. Leppänen and Peuronen (2011) argue that the transfer of frameworks developed for the study of spoken language and interaction to the study of written multilingual CMC is not adequately problematised. They also point out that when studying edited genres such as blog posts and fan fiction, methods other than conversation-analytic ones, such as narrative analysis or stylistics, are required.

5. Code-switching in technological and social context

5.1. Effects of synchronicity

We may now ask which of the categories reviewed in the previous section have been developed into testable hypotheses. In principle, CS is subject to the entire range of interrelations between medium and situation factors that "have been observed to condition variation in computer-mediated discourse" (Herring 2007: n.p.). In the literature, however, there is only one robust hypothesis that is specifically formulated with a view on CS. It goes back to John Paolillo (in press) who, in a study of English/Punjabi in IRC and Usenet, found that IRC data contain creative conversational CS, whereas Usenet data are limited to formulaic code-switching (such as quoting poetry and using routine phrases). Paolillo's generalisation is that synchronous modes of CMD will contain more conversational CS than asynchronous ones, other things being equal. Besides its quantitative sense, 'more' can be understood in a qualitative sense, meaning a broader range of usage patterns or a richer repertoire of pragmatic functions of CS. Paolillo's prediction is independently confirmed by Lee (2007: 203-204), who finds code mixing to be much more common in ICQ data than in emails by the same users. She attributes this to both synchronicity and formality, because the emails she studied include institutional exchanges, whereas her ICQ data are predominantly social interaction.

The synchronicity hypothesis is no doubt a strong one with respect to medium-specific differences across CMC modes. Synchronous CMC enables exchanges that unfold over several turns, with rapid transitions and relatively short turns, thereby resembling social interaction. In asynchronous modes, individual contributions and transition gaps between each contribution tend to be longer, creating more distance to prototypical interaction. This rationale ties in well with the assumption that synchronous CMC modes are 'closer' to spoken language than synchronous ones (e.g., Dorleijn and Nortier 2009: 130). However, it may be difficult to isolate medium factors from social and situational ones empirically. Lee's (2007) case study, in which synchronous and asynchronous data from the same individuals are compared, is a particularly fortunate one, but in public CMC environments such as chat channels and newsgroups, discourse is shaped by the technological properties of CMC synchronicity as much as by social and pragmatic factors such as individual linguistic repertoires, specific interpersonal relationships, interactional activities, and so on.

Moreover, Paolillo's hypothesis does not rule out CS in non-conversational CMC modes. The articulation of CS and mode described by Paolillo for U.S.-based Hindi and Punjabi communities, such that creative CS holds true for the synchronous mode and fixed CS for the asynchronous mode, does not necessarily hold true for other sociolinguistic settings. Some of the best examples of the creativity and playfulness of code-switching online come from asynchronous (but private), rather than synchronous, public modes (see Georgakopoulou 1997; Sebba 2003; Tsiplakou 2009). Hinrichs (2006) argues that the planned character of (asynchronous) CMC invites rhetorical uses of CS such as double-voicing and stylization. In that sense, asynchronous (public or dyadic) modes are interesting sites of CS online, precisely because they differ from the conditions of interpersonal interaction. Therefore, the synchronicity hypothesis should not lead us to assume that asynchronous modes lag behind synchronous ones in all aspects of CS. It seems more productive to theorise

each mode as affording different opportunities for bilingual discourse. While synchronous modes enable mediated interaction to unfold in sequentially-related turns which can be sustained for considerable time, thereby replicating conversational CS to the largest extent possible, asynchronous modes offer options of planning and quoting, by which distinct patterns of CS are made possible. I return to these issues in the concluding sections of this chapter (see sections 6 and 7).

5.2. Effects of public and private CMD

Some of the best examples of bilingual online creativity come from asynchronous, private exchanges, such as emails among friends and colleagues. By contrast, most available research is on public CMC modes such as forums and chat channels, whose participation framework is at odds with the typical situational conditions of private, conversational CS. Public discourse by definition entails an audience that is, at least in part, unknown to the speaker. In media discourse studies, this is captured in the notion of overhearers (Hutchby 2006), which corresponds to the category of 'lurkers' in Internet culture. Overhearers are legitimate, ratified participants, who are neither known to the speaker nor actively involved in an exchange, but whose presence may nonetheless have an effect on a speaker's audience design (Bell 1984) – that is, the way language choices are tailored to parts of the audience. All public conversational CMD includes overhearers by definition, and even though contributions in public CMC environments are typically directed to specific addressees who have a primary impact on language style, they are always co-directed to overhearers. At the same time, CS presupposes a bi- or multilingual audience that is able to understand the codes at hand and to draw inferences from the way speakers juxtapose and alternate between these codes. This inferential capacity of participants is precisely a base condition of much conversational CS.

However, such a bottom line cannot be taken for granted in public CMD. There is evidence to suggest that public CMD spaces create favourable conditions for the "functional marginalization" (Paolillo 1996) of minority and heritage languages on the internet. Digital networks that explicitly focus on a shared ethnolinguistic identity may in fact be used by ethnolinguistically heterogeneous individuals, leading to an orientation to the majority language as a common denominator. The reasons for such heterogeneity may be diverse: The ethnic communities that use CMD spaces are often undergoing intergenerational language shift, some of their members may not be fluent in the heritage language anymore, or users may be ethnically mixed, including members of the ethnic majority group. In such cases, the communicative aim of reaching as many audience members as possible may override the preference for the heritage language, and at the same time the wish to index ethnolinguistic identity may lead to patterns of formulaic and emblematic code-switching from the dominant, majority language into the heritage language, as described by Paolillo and Androutsopoulos, among others. However, it seems important not to generalise this finding, as chat channels with a predominant use of migrant/ethnic languages have been studied as well (e.g., Androutsopoulos and Hinnenkamp 2001; Hinnenkamp 2008; Tsaliki 2003).

By contrast, private CMD provides different conditions for recipient design (to use a term from conversation analysis), as participants can rely on a much greater inferential potential; i.e., they can count on their code-choices and switches being

understood by virtue of common background knowledge and common practices. Speakers have more leeway to explore playfully the associative potential of language, dialects, and styles in their shared repertoire (Tsiplakou 2009).

5.3. Research on different social settings

Common sites for CS online are public CMD spaces by and for immigrant, diaspora, and ethnic minority groups, in which CS between the minority (migrant, community) and the respective majority language has been identified as the main pattern of bilingual discourse. Studies of this type of CS are often linked to an interest in language maintenance, and CMC has been associated with hopes and expectations of maintenance as much as with anxieties of loss and acceleration of ongoing processes of linguistic and cultural shift (Sperlich 2005).

Another user population that has attracted research attention is young people and their local or 'glocal' language practices associated with music and media culture. In the literature reviewed here, such groups are located in Egypt, Finland, Germany, and South Africa, and one of their conspicuous language practices in various CMC modes is CS between the respective national language and English. Diverse as they may be, these studies suggest patterns of "minimal bilingualism" (Androutsopoulos 2007b), in which sets of English chunks and formulaic routines (including greetings and farewells, interjections and discourse organisers, requests, slogans, etc.) are inserted into the base national or majority language. Their choice is often indexical to the groups' lifestyle orientations, including stylized representations of vernacular "Englishes" (Androutsopoulos 2004, 2007b).

Research from Finland offers evidence for an even broader range of online bilingual practices involving English in Finnish youth cultures (Leppänen 2007; Leppänen et al. 2009). One such practice is alternational CS in fan fiction (i.e., stories written by fans about popular fiction characters or settings), whereby Finnish and English are used for narrative and reported speech, respectively, so that neither constitutes the predominant language of discourse. In another digital genre, diary weblogs, English is the matrix language into which Finnish cultural keywords and expressions are inserted. In all these cases, there is no evidence that this usage mirrors corresponding bilingual styles in face-to-face communication. Rather, it is CMD that enables these bilingual practices in the first place. In addition to these Finnish cases, there is some evidence of CMD from what one might call elite bilinguals (i.e., academics, white-collar professionals) with an L1 other than English, whose private email communication shows complex patterns of switching and mixing (Georgakopoulou 1997; Tsiplakou 2009, both on L1 speakers of Greek).

The social conditions of CS online may of course be more complex that that. For instance, the members of fan communities within one nation-state are often of ethnically diverse origin, and the cultural practices they orient to may originate beyond the English-speaking world. One anecdotal example is the web forum of a local network of salsa fans in northern Germany (based on unpublished material), which includes native speakers of German, Spanish, and various other migrant languages. Regardless of their origin, forum members orient to Spanish as a symbol of their common cultural practice, and the usage of Spanish in this forum is often

formulaic, resembling the patterns of minimal bilingualism that are often identified with English. Spanish is also used among Spanish-speaking participants, and so are migrant languages (Turkish, Polish) among their respective speakers. The outcome is a complex polylingual space, in which the majority language, the language linked to the particular cultural practice, and immigrant languages all find their place.

5.4. Effect of social variables

The qualitative orientation of much research reviewed here means that the situated expression of social identities through CS (see section 6) has attracted more attention than the testing of hypotheses based on large corpora. It is therefore not surprising that the effect of different social variables on CS online has not been systematically examined. However, a few sources of evidence are available. In his work on Swiss-German chats, Siebenhaar (2005, 2006) identifies the combined effect of age group and channel range on dialect/standard switching. Swiss teenagers prefer local IRC rooms and use their dialect as the default code, whereas middle-aged users prefer a cross-regional flirt chat, in which they use a considerable amount of standard German as well. Paolillo (2001) and Ziegler (2005) provide evidence for the impact of institutional roles within Internet culture on CS. Paolillo finds that operators of the IRC channel #india switch less into Hindi than peripheral users. As Paolillo points out, an identity-based hypothesis would predict the opposite, i.e., that core members of the IRC network use the heritage language to a higher degree. Ziegler (2005) finds that in the local IRC channel of a German city, #mannheim, channel operators use higher amounts of the local city dialect, thus CS from and to standard German, than ordinary chatters. Paolillo offers a pragmatic explanation of his finding based on an attention-based hypothesis: U.S.-born IRC users of Indian descent display ethnic affiliation by other means than heritage language, and it is peripheral (rather than core) members of the chat community who code-switch in order to attract attention. Tsiplakou (2009) presents findings from a questionnaire study that aimed at identifying variables predicting the degree of CS in email (between Greek and English, among native Greeks), one such variable being the use of English at home.

6. Code-switching patterns

CS in CMC is not confined to just a few typical patterns of usage. In terms of structure, reported patterns of usage range from a few formulaic switches to dense, multilingual code-mixing and 'polylingual languaging' (sections 6, 7). A restriction to formulaic CS has been reported for public CMD in some ethnic communities (see work by Paolillo and Androutsopoulos), for English/Jamaican Creole in personal emails (Hinrichs 2006), and for national language/English in youth-cultural contexts (see section 5.3). Bilingual code-mixing has been reported for a number of languages and modes, including personal emails among fellow academics in Cyprus (Tsiplakou 2009), chatting among youth of Turkish descent in Germany (Hinnenkamp 2008), and "mixed messages characterized by English-isiXhosa code-mixing and code-switching" in texting among South African youths (Deumert and Masinyana 2008: 137). Deumert and Masinyana (2008) claim that these linguistically mixed messages mirror code-mixing as the unmarked choice in the spoken vernacular usage of these youths. The percentage of mixed text messages in their data is 23% (in a corpus of 312 messages), and they are further differentiated by communicative purpose. Code-

mixing occurs in text messages on social arrangements and information exchange, but less so in romantic messages.

In terms of discourse functions, the literature drawing on widely accepted classifications of conversational CS (see section 2) has produced evidence for a number of discourse functions of CS in CMC.[9] These include:

a) switching for formulaic discourse purposes, including greetings, farewells, and good wishes;
b) switching in order to perform culturally-specific genres such as poetry or joke-telling;
c) switching to convey reported speech (as opposed to the writer's own speech);
d) switching with repetition of an utterance for emphatic purposes;
e) switching to index one particular addressee, to respond to language choices by preceding contributions, or to challenge other participants' language choices;
f) switching to contextualize a shift of topic or perspective, to distinguish between facts and opinion, information and affect, and so on;
g) switching to mark what is being said as jocular or serious, and to mitigate potential face-threatening acts, for example through humorous CS in a dispreferred response or a request;
h) switching to or from the interlocutor's code to index consent or dissent, agreement and conflict, alignment and distancing, and so on.

Some of these functions – especially (a) and (b) – have been found to favor a sustained use of minority or migrant languages, while for others, the pragmatic effect is created through the situated contrast between the codes involved (Hinrichs 2006 provides an extensive discussion of this distinction).

While the comparability of CS in CMC with general discourse functions of conversational CS is thus in principle firmly established across languages, modes, and social settings, individual manifestations of CS in CMD data may be difficult to categorize, and switching and mixing may co-occur in the discourse of one user or community. A few examples from my own research on predominantly German-speaking diasporic forums are presented below to illustrate these points (Androutsopoulos 2006a, 2006b, 2007a). The following three examples come from an Indian, Persian, and Greek web discussion forum, respectively, and exemplify three variants of CS within a post:[10]

(1) **Excerpt from the Indian forum, *theinder.net* (base language is German, English italicized)**

im westen ist es auch *tradition* jungfräulich in die ehe zu treten!!! das tun auch einige (bsp. spanierinnen, italienerinnen, etc..)! wieso wird immer der westen für alles verantwortlich gemacht??? *is there no gravity in indian brains?*

['in the west, no sex before marriage is also a tradition!!! And some stick to it as well (e.g., Spaniards, Italians etc.)! Why is the West always being blamed for everything??? *is there no gravity in indian brains*?']

(2) **Excerpt from the Iranian forum, *iran-now.de* (base language is Persian, German italicized)**

Bare B. juuuuuuuuuun, kamelan hagh dari. in marda ham khasisan, ham bimazan wa ham gedan!!! *Genau das dachte ich mir auch*, chon iran-klick waghat klick mikhado *sonst nix, aber neeeeeeeeeeeee* A. agha *kann uns ja net* ghalbesh az ma beporseh. ghorbuuuuuuuuuuuuune harfat khanumi. P. e *immer noch* ghamginiani

['That's true, dear B, you're absolutely right. These men are stingy, boring, and quite the poor creatures!!! *That's exactly what I thought*, because on iran-klick you just need to click, *nothing else, but nooo* Mister A. *just can't* ask in advance. I love you for your words, my dear. P is *still* sad']

(3) **Excerpt from the Greek forum, *greex.net* (base language is Greek, German italicized)**

edo iparxi pollous ellines apo tin makedonia epidis i *wirtschaftliche lage* tous den einai kali...palia i makedoni itane plousioi.. eftiaxnan gounes ktlp ala tora pige i *wirtschaft* me tis gounes *den bach runter*

['there are many greeks from macedonia here but their *financial situation* is not good… macedonians were rich in the past, they were trading with furs, but now the fur *business* is going *down the drain*']

In these and other forums examined in that study, ethnic minority languages mainly occur in isolated, insertional switches. In Excerpt 1, taken from a discussion thread on premarital sex, the concluding switch into English serves to accentuate the writer's critical conclusion and sets it off from the preceding argumentation. This type of bilingual discourse is quite common in these forums, and several types of discourse functions listed above are instantiated in this insertional manner. In Excerpt 2, the first instance of German can be analysed as a switch that marks a change of perspective from the previous evaluation of "these men" to the writer's own views – "That's exactly what I thought". However, this line of interpretation becomes increasingly difficult as the post unfolds, and no contextualizing function can be identified for the penultimate adverbial phrase in German, which is more typical for code-mixing. In Excerpt (3), whose base language is Greek, the three German phrases (*financial situation, business, down the drain*) have neither any obvious discourse function nor do they serve a referential necessity. In both cases, a Greek equivalent would presumably have been readily available, and the writer seems to select German lexical items and idioms for reasons of habit or convenience. The last instance of German in this excerpt is an idiom (*den Bach runtergehen* 'go down the drain'), and the writer expresses the finite verb in Greek, and the idiomatic phrase in German. In Auer's discourse-functional framework (see section 2), such cases are also typically classified as code-mixing.

Working with classifications of discourse functions provides an initial overview of patterns of CS in a CMC environment and a useful point of entry for exploratory research. Cumulatively, analyses along these lines offer valuable evidence for the regularity and conventionality of CS online as well as for its functional similarity to CS in other discourse environments, and thereby contribute to its normalization.

However, too heavy a reliance on classifications also entails the risk of reducing analysis to a simple "category check" that disconnects CS from the conversational activity in which it is embedded and may result into a decontextualized listing of CS instances. That danger can be reduced by ethnographic knowledge of the population and digital platform under study and by detailed sequential analyses that take into account "the place within the interactional episode in which languages alternate" (Auer 1998b: 3), the way switches align to previous code choices of other speakers, and the way they index participants' background knowledge.

More specifically, the functional analysis of CS in CMC needs to transcend the level of single turns or posts and examine the sequential organisation of code-switching within threads of dialogically related posts or messages. CS is embedded in the "polylogues" (Marcoccia 2004) that unfold on spaces of public, asynchronous, thematically focused discourse, and makes use of the specific affordances provided by these spaces (see also discussion in Georgakopoulou 1997; Siebenhaar 2006). In the case of the Germany-based diaspora forums, switches within a post are a resource for responding to different addressees and engaging with different strands of a discussion thread. Users sometimes start in one language, quote from previous posts which are cast in a different language, respond to that quote in its language choice, and then return to their original code choice (Androutsopoulos 2006a, 2006b, 2007a).

In addition, a pragmatically informed micro-analysis of CS in CMC will aim to examine "how, within frameworks of generic assumptions and expectations, speech communities draw upon their linguistic resources in order to maximize the effectiveness and functionality of their communication" (Georgakopoulou (1997: 160). In such an analysis it is possible to identify how different codes in a group's usage take on pragmatic functions and identity values, which cannot be assumed *a priori* based on the wider cultural associations of these languages. The use of linguistic heterogeneity to index social identities is a key issue in much of the work reviewed here. Some researchers draw on Gumperz' distinction between 'we-code' and 'they-code', which was originally equated with minority and mainstream language, respectively. This works well when a few salient instances of a code are used as a means to signal the ethnic identity associated with that code, and such 'we code' signals reported repeatedly include greetings, openings, closing, slogans, and the like (Androutsopoulos 2006a; Hinrichs 2006; Paolillo 1996). This approach is complemented by a more fluid and dynamic understanding of language/identity relations, in which 'we'/'they' contrasts are locally constructed in discourse. In diaspora forum discussions, for example, the 'we-code' is not always the minority language, and the 'they-code' is not always the mainstream language. Writers may switch into their heritage language to index traditionalist views they distance themselves from, or, in another case, switch into the majority language to emphasize their own multicultural outlook (see Androutsopoulos 2006a; Hinrichs 2006).

Working within a constructivist language and identities framework, Tsiplakou (2009) examines how email writers draw on Greek, English, and further linguistic resources (French, Cypriot Greek dialect, stylized sociolects of Greek) in order to act out "localized performativities", i.e., contextually constructed social identities, which participants playfully claim for themselves, stylize, or parody. Hinrichs (2006) identifies three types of identity-related alternation between Standard English and Creole in private emails among Jamaican students: use of Creole for self-

identification and message framing (e.g., greetings, farewells, terms of addressing); use of CS to organize different narrative activities and to set apart 'we' and 'they' perspectives; and double-voicing, with Creole being employed in the creation of stereotypical local speech styles. In the last pattern, writers exploit the "potential of Patois to make salient certain cultural values and personae" (2006: 134) such as the "country bumpkin" personae, which writers playfully or ironically associate themselves with in specific thematic contexts and speech acts, for instance to mitigate boasting. As Hinrichs (2006: 134) suggests, these latter expressions of identity-related CS are "especially at home with the written medium", because they "involve the highest degree of planned, rhetorical use" of Creole.

7. Computer-mediated discourse as a new domain of multilingual code-switching

Based on the preceding discussion, we are now in a position to assess the relation of CS in CMC to spoken conversational CS and written CS. The reviewed literature does not offer a generally agreed position on this, and its suggestions mirror to some extent the wider discussion on spoken and written aspects of language on the internet. While CS in CMC obviously qualifies as written in terms of the written representation of linguistic signs, it also bears resemblance to spoken conversational CS, most obviously in terms of its dialogic context and its discourse functions (see section 6). From the perspective of contemporary language-focused CMC scholarship, however, both a dichotomy of that sort and the homogenisation of various CMD genres and practices as "language on the Internet" seem rather crude. CMC is generally viewed as a heterogeneous domain of discourse, in which traditional dichotomies between written and spoken, private and public, immediate and mediated discourse, are blurred.

Theorizing CS in CMC needs to take the specific pragmatic and social conditions of written language use in digital media more systematically into account. Writing in networked digital media is different from other types of written discourse in a number of ways: It is dialogical, i.e., oriented to particular addressees, and often embedded in multi-party conversational sequences; it also is often vernacular, i.e., located outside of educational, professional, and other institutions; and it is often simultaneously used together with other semiotic resources. Taking these properties of digital writing into account when studying CS in CMC will contribute to a deconstruction of spoken/written dichotomies and to a move beyond the assumption that only spoken conversational CS constitutes "authentic" CS, and therefore sets the benchmark against which CS in CMC ought to be assessed. Such expectations are familiar from other domains of written CS, notably in fiction, where CS "authenticity" is assessed in terms of fictional settings, characters, and the thematic content of a novel (Callahan 2004: 99-111). These criteria no doubt operate, albeit implicitly, when researchers assess the ethnic or diasporic "authenticity" of a web forum or chat channel. Notions of authenticity are also reproduced within CMC, for example in the assumption that only synchronous CMC modes will host "real" CS (see sections 3, 5). An alternative and, I would suggest, more productive approach would be to ask how CS is used as a pragmatic resource under the specific conditions of CMD, and how specific conditions of written online discourse can give rise to distinct CS practices. Two relevant points are touched upon here: planning and the semiotics of writing itself.

Planning focuses on the difference between the immediate production of speech in a spoken conversational setting and the gaps between production and reception that are inherent in CMC, especially its asynchronous modes. The availability of planning time and the lack of visual cues are at the core of the assumption that CMC is characterised by code-centered contextualization cues (Georgakopoulou 1997, 2003). However, the assumption that a "higher level of consciousness [...] seems inevitable in producing written CS" (Dorleijn and Nortier 2009: 131) should not be generalised prematurely. It is well known that CMD is sometimes produced quickly and spontaneously, while in other cases it may involve extensive drafting and rewording. It is therefore more useful to expect different levels or degrees of conscious organisation of bilingual discourse, depending on mode and circumstances.

Moreover, it seems necessary to distinguish between various potential consequences of planning for CS. One of these, as already noted (section 6), relates to the composition of contributions within multi-party conversational exchanges. Asynchronous modes, especially, enable writers to deploy strategies of participation in which a single post may compile responses to a number of previous posts; in multilingual contexts, writers may thereby use post-internal CS as a resource for distinguishing between addressees or perspectives. What is particular to CS in this regard is its strategic deployment in a context of discourse organisation that is uniquely digital.

A second and less well understood impact of planning is on the potential avoidance of or preference for certain types of switching or mixing. While examples like (3) may reflect the mixed language style that has been reported as typical for the speech of immigrant background youth (e.g., Hinnenkamp 2008), language mixing is relatively rare across the forums of that study. I suggested (Androutsopoulos 2007a: 347-348) that this may be due to the combined impact of the conditions of publicness and asynchronicity: The metalinguistic awareness involved in planning and editing posts may inhibit the spontaneous, unconscious process of code-mixing, and the public character (and ethnolinguistic heterogeneity) of these forums is at odds with the situational conditions that favour the occurrence of code-mixing. In this case, the conscious production of discourse seems to result to a higher degree of formality, in that language contact patterns that are at home in vernacular, intimate settings are avoided.

However, the opposite tendency is also documented. That is, the planning opportunities afforded by asynchronous CMC may enable participants to use language mixing in creative and sometimes masterful ways that might not have occurred in speech. A few case studies offer evidence for this tendency, in part drawing on the notion of 'polylingual languaging' (a playful use of all linguistic resources available to speakers/writers in a given context, regardless of degrees of linguistic competence or ethnolinguistic affiliation; Jørgensen 2008).

One example comes from a discussion thread on language and ethnicity from a German-based Greek forum (Androutsopoulos 2006b), where a poster argues that Germany-born Greeks would permanently mix languages in both countries. A short response to this – *Korrekt. richtisch* stin teleia *gebracht* ('Correct. Got right *to the point*') – starts in German, switches intrasententially to Greek for one phrase ('to the

point') and concludes in German. The phrase switched into Greek is idiomatic in German, although not in Greek, and the resulting mix sounds pragmatically odd when taken literally. However, this mix encapsulates the "essence" of the preceding discussion, i.e., the group's mixture of identities and languages, and even though the phrase itself would be a rather odd and unexpected instance of mixing in speech, its symbolic meaning can readily be understood by its audience and ties in well with everyday and academic interpretations of mixed talk as symbol of hybrid identities (Hinnenkamp 2008).

This and other research (notably Hinnenkamp 2008; Tsiplakou 2009) suggests that even though their pragmatic force depends on shared knowledge and pragmatic conventions, some instances of code-mixing in CMD go beyond a simple reflection of spoken conversational patterns (Tsiplakou offers evidence from spoken conversation to support that point). What the language mixing instances reported in these studies share are their implicitly or explicitly metalinguistic character and their occurrence in discourse that focuses on key identity issues of a group or community. Viewed this way, some code-mixing in CMC is pragmatically effective precisely because it diverges from colloquial bilingual usage to "iconise" beliefs or values shared by the participants. However, it is very difficult to say whether the distinctive feel of these mixing instances is due to their particular syntactic structure or rather to their rhetorical tailoring of code-mixing to the gist of a given discourse episode.

A second point that merits more attention in future research is the relevance of the mode-specific resources of writing (i.e., orthography, spelling variation, and even script choice) in the production of CS in CMC. Again, different aspects need to be distinguished preliminarily. One is the impact of standardised orthographies on the selection of base language in CMD. Hinrichs (2006) reports that unlike everyday communication in Jamaica, where Creole is the default medium of communication, its lack of a standardised orthography makes it less suitable to that purpose in CMD, so that the Jamaican students and expatriates in his study subjects basically draw on English, occasionally switching to Creole. However, in other settings, writers may stick to their vernacular in CMD despite its lack of a standardised orthography, for example in German-speaking Switzerland (Siebenhaar 2005, 2006), where the online spelling of local dialects tends to reflect vernacular pronunciation more accurately than traditional dialect spelling does.

A further issue, again following up on the notion of code-centered contextualisation cues, is the role of spelling in signalling CS. It is obvious that CS in CMD is produced by representing in writing another language (or dialect), but what is less obvious is that this can be done based on different orthographic conventions. The switch from one language to another usually co-occurs with, and is indexed by, a switch between the respective orthographies, albeit not forcibly so. Representing one language in the orthography of another may sometimes be a matter of necessity, but it may also be a more or less conscious choice and thereby a source of pragmatic meaning in its own right. There is also evidence that spelling may be exploited in its own right, with CMD writers drawing on the contrast with normative orthography to create pragmatic meaning. A striking example is the deliberate "mixing of alphabetic conventions" reported for German-Turkish chatters who create mixed-language conversations as "German words and even phrases get a kind of Turkish wrapping", which consists of Turkish orthography (Hinnenkamp 2008: 262, 266). Spelling

Deutsch (in correct German orthography) as *Doyc* (based on Turkish orthography) is, as the context makes clear, not a typo or spelling mistake but a conscious blend of language and orthography designed to elicit pragmatic meaning, which ties in with the debate on language and ethnic boundaries that dominates that chat session. Likewise, although at a different level of written structure, there is evidence of purposeful "script switching" between native and Roman script in the Romanised transliteration of different languages on the Internet (see chapters in Danet and Herring 2007a).

Such evidence, I argue, suggests that the study of CS in CMC primarily in terms of its apparent "authenticity" or correspondence to spoken conversational CS may be limiting, and that important insights will be gained by theorising the written digital mode not as a limitation, but as a new set of conditions for the deployment of multilingual resources in discourse.

8. Outlook and directions for future research

Research has only just started tackling the massive bilingualism and multilingualism that occurs as global multilingual populations increasingly gain access to digital communications media. Much remains to be done in documenting different sites and types of CS online, and systematic comparisons among modes, language, and settings are needed. Most lacking are, first, studies of private, dyadic data; second, cross-media and cross-mode comparisons of CS usage based on the same writer(s); third, multimodal data from social networking and media-sharing websites; and fourth, case studies of multilingual CMD in transnational work teams. In terms of method, research is moving away from static classifications and towards ethnographically and pragmatically informed analyses of the local interactional purposes that CS serves in its generic and sequential context. However, mixed-methods combinations of qualitative and quantitative techniques, e.g., by means of questionnaires or language choice analyses, are bound to remain productive and insightful. One limitation of current research is the restriction to single modes, which are analytically examined in isolation. Motivated by practical necessities as that may be, it creates an isolationism that runs counter to actual computer-mediated practices, which are spread across modes and platforms in combinations and routines that are not yet well understood. There is reason to assume that CS patterns will often cut across modes, and understanding such code/mode repertoires will deepen the understanding of the specific properties of CS online.

Two issues that cut across the sections of this chapter are how to theorise the relationship between medium and social/situation factors, on the one hand, and between online written and offline spoken CS, on the other. Even though it might seem customary to think of CS on the internet in terms of its "authenticity" or correspondence to an assumed spoken conversational blueprint in the usage of an individual or community, I have argued that CMD is unscripted, dynamically unfolding communication in its own right. Taking this into account would cast doubt on both the necessity and the means of establishing such authenticity. This chapter has presented a number of cases where bilingual practices are not verbatim reproductions of face-to-face interaction patterns but, judging from ethnographic and linguistic evidence, specific to CMD. I therefore suggest that rather than examining CS online in terms of its authenticity or equivalence to offline speech, a more productive

question to pursue is how CS is used as a resource, under the specific conditions of communication offered by digital media.

Notes

1. Only two volumes to date include more than one contribution dealing, at least in part, with CS (Danet and Herring 2007a and a 2006 special issue of the *Journal of Sociolinguistics* 10[4]), while in other recent books the topic is either entirely absent or marginal (e.g., Baron 2008; Crystal 2006; Rowe and Wyss 2009).
2. The languages of publication covered are English and German. Literature in other languages can be traced through the reference list.
3. The latter is used by Androutsopoulos and Hinnenkamp (2001), Androutsopoulos and Ziegler (2004), Androutsopoulos (2006a, 2007a), Hinnenkamp (2008), Hinrichs (2006), Leppänen (2007), Siebenhaar (2005, 2006), and Tsiplakou (2009).
4. Auer points out that "a speaker may simply want to avoid the language in which he or she feels insecure and speak the one in which he or she has greater competence. Yet preference-related switching may also be due to a deliberate decision based on political considerations" (1995: 125).
5. Following Herring (2004), I use computer-mediated discourse (CMD) to indicate a narrower focus on the use of semiotic resources, whereas CMC denotes a broader view on communication processes facilitated by digital technologies.
6. In earlier work (Androutsopoulos 2006a) I located CS within a quadripartite matrix of multilingualism on ethnic portals. Its main site is regular user-contributed discourse, especially in forums and occasionally in journalists' edited content as well. This is distinguished from user and site-specific emblems, such as names, mottos, and slogans. These distinctions are taken up and recontextualised in the present discussion.
7. This mirrors the limitations of research on multilingualism in CMC in general (see Danet and Herring 2007b: 24).
8. The label 'L1' (first language) was only added if clearly stated in the research, usually based on controlled socio-demographic data. Other studies either do not control for this factor or they study public CMD settings where assigning an L1 is not straightforward.
9. All of the following studies include discussions of these: Androutsopoulos (2006a, 2007a), Androutsopoulos and Hinnenkamp (2001), Dorleijn and Nortier (2009), Georgakopoulou (1997), Hinrichs (2006), Kadende-Kaiser (2000), McClure (2001), Paolillo (1996, in press), Sebba (2003), Sperlich (2005), and Tsiplakou (2009).
10. The examples are anonymised, and presentation follows Androutsopoulos (2006a, 2006b).

References

Androutsopoulos, Jannis   2004   Non-native English and sub-cultural identities in media discourse. In: Helge Sandøy (ed.), *Den fleirspråklege utfordringa / The Multilingual Challenge*, 83–98. Oslo: Novus.

Androutsopoulos, Jannis   2006a   Multilingualism, diaspora, and the

Internet: Codes and identities on German-based diaspora websites. *Journal of Sociolinguistics* 10(4): 524–551.

Androutsopoulos, Jannis   2006b   Mehrsprachigkeit im deutschen Internet: Sprachwahl und Sprachwechsel in Ethno-Portalen. [Multilingualism on the German Internet: Language choice and code-switching on ethnic portals.] In: Peter Schlobinski (ed.), *Von *hdl* bis *cul8r*. S*prache und Kommunikation in den Neuen Medien* [From *hdl* to *cul8r*. Language and communication in New Media]*, 172–196. (Thema Deutsch, Band 7.) Mannheim: Dudenverlag.

Androutsopoulos, Jannis   2007a   Language choice and code-switching in German-based diasporic web forums. In: Brenda Danet and Susan C. Herring (eds.), 340–361.

Androutsopoulos, Jannis   2007b   Bilingualism in the mass media and on the Internet. In: Monica Heller (ed.), *Bilingualism: A Social Approach*, 207–230. New York: Palgrave Macmillan.

Androutsopoulos, Jannis   2011   From variation to heteroglossia in the study of computer-mediated discourse. In: Crispin Thurlow and Kristine Mroczek (eds.), *Digital Discourse: Language in the New Media*, 277–298. New York: Oxford University Press.

Androutsopoulos, Jannis and Volker Hinnenkamp   2001   Code-Switching in der bilingualen Chat-Kommunikation: ein explorativer Blick auf #hellas und #turks. [Code-switching in bilingual chat communication: an exploratory analysis of #hellas and #turks.] In: Michael Beißwenger (ed.), *Chat-Kommunikation,* 367–402. Stuttgart: Ibidem.

Androutsopoulos, Jannis and Evelyn Ziegler   2004   Exploring language variation on the Internet: Regional speech in a chat community. In: Britt-Louise Gunnarsson, Lena Bergström, Gerd Eklund, Staffan Fridell, Lise H. Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundgren, and Mats Thelander (eds.), *Language Variation in Europe*, 99–111. Uppsala: Uppsala University Press.

Auer, Peter 1995   The pragmatics of code-switching: A sequential approach. In: Lesley Milroy and Pieter Muysken (eds.), *One Speaker, Two Languages*, 115–135. Cambridge, UK: Cambridge University Press.

Auer, Peter, ed.   1998a   *Code-Switching in Conversation*. London: Routledge.

Auer, Peter 1998b   Introduction: Bilingual conversation revisited. In: Peter Auer (ed.), 1–24.

Auer, Peter   1999   From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3(4): 309–332.

Auer, Peter   2000   Why should we and how can we determine the 'base language' of a bilingual conversation? *Estudios de Sociolingüística* 1(1): 129–144.

Baron, Naomi S.   2008   *Always On: Language in an Online and Mobile World.* New York: Oxford University Press.

Beißwenger, Michael   2008   Situated chat analysis as a window to the user's

perspective: Aspects of temporal and sequential organization. *Language@Internet* 5, article 6. http://www.languageatinternet.org/articles/2008/1532/

Bell, Allan 1984 Language style as audience design. *Language in Society* 13: 145–204.

Callahan, Laura 2004 *Spanish/English Codeswitching in a Written Corpus.* Amsterdam/Philadelphia: John Benjamins.

Crystal, David 2006 *Language and the Internet*, 2nd Ed. Cambridge, UK: Cambridge University Press

Danet, Brenda and Susan C. Herring, eds. 2007a *The Multilingual Internet: Language, Culture, and Communication Online*. New York: Oxford University Press.

Danet, Brenda and Susan C. Herring 2007b Introduction: Welcome to the Multilingual Internet. In: Brenda Danet and Susan C. Herring (eds.), 3–38.

Deumert, Ana and Oscar Masinyana Sibabalwe 2008 Mobile language choices – The use of English and isiXhosa in text messages (SMS): Evidence from a bilingual South African sample. *English World-Wide* 29(2): 117–147.

Dorleijn, Margreet and Jacomine Nortier 2009 Code-switching and the internet. In: Barbara E. Bullock and Almeida J. Toribio (eds.), *The Cambridge Handbook of Linguistic Code-Switching*, 127–141. Cambridge, UK: Cambridge University Press.

Ferrara, Kathleen, Hans Brunner, and Greg Whittemore 1991 Interactive written discourse as an emergent register. *Written Communication* 8(1): 8–34.

Georgakopoulou, Alexandra 1997 Self-presentation and interactional alignments in e-mail discourse: The style- and code switches of Greek messages. *International Journal of Applied Linguistics* 7(2): 141–164.

Georgakopoulou, Alexandra 2003 Computer-mediated communication. In: Jef Verschueren, Jan-Ola Östman, Jan Blommaert, and Chris Bulcaen (eds.), *Handbook of Pragmatics* (2001 Installment), 1–20. Amsterdam/Philadelphia: John Benjamins.

Georgakopoulou, Alexandra 2004 To tell or not to tell? Email stories between on- and off-line interactions. *Language@Internet* 1, article 1. http://www.languageatinternet.org/articles/2004/36

Goldbarg, Rosalyn Negrón 2009 Spanish-English codeswitching in email communication. *Language@Internet* 6, article 3. http://www.languageatinternet.org/articles/2009/2139

Gumperz, John J. 1982 *Discourse Strategies*. Cambridge, UK: Cambridge University Press.

Heller, Monica and Carol W. Pfaff 1996 Code-switching. In: Hans Goebl, Peter H. Nelde, Zdenek Stary, Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics,* Vol. 1, 594–609. Berlin/New York: de Gruyter.

Herring, Susan C., ed. 1996 *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*. Amsterdam/Philadelphia: John Benjamins.

Herring, Susan C., ed. 1999 Interactional coherence in CMC. *Journal of*

Computer-Mediated Communication, 4(4).
http://jcmc.indiana.edu/vol4/issue4/herring.htm.

Herring, Susan C.   2001   Computer-mediated discourse. In: Deborah Schiffrin, Heidi Hamilton, and Deborah Tannen (eds.), *The Handbook of Discourse Analysis*, 612–634. Malden, MA: Blackwell.

Herring, Susan C.   2004   Computer-mediated discourse analysis: An approach to researching online communities. In: Sasha A. Barab, Rob Kling, and James H. Gray (eds.), *Designing for Virtual Communities in the Service of Learning,* 338–376. Cambridge/New York: Cambridge University Press.

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

Hinnenkamp, Volker   2008   Deutsch, Doyc or Doitsch? Chatters as languagers – The case of a German–Turkish chat room. *International Journal of Multilingualism* 5(3): 253–275.

Hinrichs, Lars   2006   *Codeswitching on the Web.* Amsterdam/Philadelphia: John Benjamins.

Hutchby, Ian   2006   *Media Talk: Conversation Analysis and the Study of Broadcasting*. Maidenhead, UK: Open University Press.

Jørgensen, Normann J. 2008   Polylingual languaging around and among children and adolescents. *International Journal of Multilingualism* 5(3): 161–176.

Kadende-Kaiser, Rose M.   2000   Interpreting language and cultural discourse: Internet communication among Burundians in the diaspora. *Africa Today* 47(2): 121–148.

Lee, Carmen K. M.   2007   Linguistic features of email and ICQ instant messaging in Hong Kong. In: Brenda Danet and Susan C. Herring (eds.), 184–208.

Lee, Carmen and David Barton   2011   Constructing glocal identities through multilingual writing practices on flickr.com. *International Multilingualism Research Journal* 5(1): 39-59.

Leppänen Sirpa   2007   Youth language in media contexts: insights into the functions of English in Finland. *World Englishes* 26(2): 149–169.

Leppänen, Sirpa and Saija Peuronen   2011   Multilingualism on the Internet. In: Marilyn Martin-Jones, Adrian Blackledge, and Angela Creese (eds.), *The Routledge Handbook of Multilingualism*. London: Routledge [in press].

Leppänen, Sirpa, Anne Pitkänen-Huhta, Arja Piirainen-Marsh, Tarja Nikula, and Saija Peuronen   2009   Young people's translocal new media uses: A multiperspective analysis of language choice and heteroglossia. *Journal of Computer-Mediated Communication* 14(4): 1080–1107.

Marcoccia, Michel   2004   On-line polylogue: Conversation structure and participation framework in Internet newsgroups. *Journal of Pragmatics* 36: 115–145.

McClure, Erica   2001   Oral and written Assyrian-English codeswitching. In: Rodolfo Jacobson (ed.), *Codeswitching Worldwide II*, 157–191. Berlin: Mouton de Gruyter.

Myers-Scotton, Carol   1993   *Social Motivation for Code-Switching: Evidence from Africa*. Oxford, UK: Clarendon Press.

Myers-Scotton, Carol   1998   A theoretical introduction to the markedness model. In: Carol Myers-Scotton (ed.), *Codes and Consequences: Choosing Linguistic Varieties*, 18–38. New York: Oxford University Press.

Paolillo, John C.   1996   Language choice on soc.culture.punjab. *Electronic Journal of Communication* 6(3).
http://ella.slis.indiana.edu/~paolillo/research/paolillo.publish.txt

Paolillo, John C.   2001   Language variation on Internet Relay Chat: A social network approach. *Journal of Sociolinguistics* 5(2): 180–213.

Paolillo, John C.   2007   How much multilingualism? Language diversity on the Internet. In: Brenda Danet and Susan C. Herring (eds.), 408–430.

 Paolillo, John C.   In press   'Conversational' codeswitching on Usenet and Internet Relay Chat. *Language@Internet* 8.

Rowe, Charley and Eva L. Wyss (eds.)   2009   *Language and New Media: Linguistic, Cultural, and Technological Evolutions*. Cresskill/NJ: Hampton Press.

Sebba, Mark   2003   'Will the real impersonator please stand up?' Language and identity in the Ali G websites. *Arbeiten aus Anglistik und Amerikanistik* 28(2): 279–304.

Sebba, Mark and Tony Wootton   1998   We, they and identity: Sequential vs. identity-related explanation in code-switching. In: Peter Auer (ed.), *Code-Switching in Conversation*, 262–289. London: Routledge.

Siebenhaar, Beat   2005   Varietätenwahl und Code Switching in Deutschschweizer Chatkanälen. [Variety choice and code-switching in Swiss German chat channels.] *Networx* 43.
http://www.mediensprache.net/de/websprache/networx/docs/index.asp?id=43

Siebenhaar, Beat   2006   Code choice and code-switching in Swiss-German Internet Relay Chat rooms. *Journal of Sociolinguistics* 10(4): 481–509.

Siebenhaar, Beat   2008   Quantitative approaches to linguistic variation in IRC: Implications for qualitative research. *Language@Internet* 5, article 4.
http://www.languageatinternet.org/articles/2008/1615/index_html/

Sperlich, Wolfgang B. 2005   Will cyberforums save endangered languages? A Niuean case study. *International Journal of the Sociology of Language* 172: 51–77.

Tsaliki, Liza   2003   Globalization and hybridity: The construction of Greekness on the Internet. In: Karim H. Karim (ed.), *The Media of diaspora. Mapping the Globe*, 162–176. London: Routledge.

Tsiplakou, Stavroula   2009   Doing (bi)lingualism: Language alternation as performative construction of online identities. *Pragmatics* 19(3): 361–391.

Warschauer, Mark, Ghada R. El Said, and Ayman Zohry   2007   *Language Choice Online: Globalization and Identity in Egypt*. In: Brenda Danet and Susan C. Herring (eds.), 303–318.

Wright, Sue, ed.   2004   *Multilingualism on the Internet*. Special issue, *International Journal on Multicultural Societies* 6(1). UNESCO Social and Human

Sciences. http://www.unesco.org/shs/ijms

Ziegler, Evelyn   2005   Die Bedeutung von Interaktionsstatus und Interaktionsmodus für die Dialekt-Standard-Variation in der Chatkommunikation. [Relevance of interactional status and interactional mode for dialect/standard variation in chat communication.] In: Eckhardt Eggers, Jürgen Erich Schmidt, and Dieter Stellmacher (eds.), *Neue Dialekte - moderne Dialektologie* [New dialects – modern dialectology]*,* 719–745. Stuttgart: Steiner.