Chapter 1. Introduction

1.1 Speech vs. writing vs. conversational writing

Every day millions of Internet users converse in real time by exchanging messages over computer chat systems. This study documents an investigation of conversational writing as carried out in text-based online chat in the early 21st century. More precisely, it presents a lexico-grammatical and functional linguistic analysis of features in a corpus of conversational writing consisting of synchronous and supersynchronous computer-mediated communication (CMC). Synchronous and supersynchronous computer chat differ in that the former is carried out one turn at a time, in for instance chat channels, whereas the latter is carried out in a split window, in which turns are realized keystroke by keystroke, so that completely overlapping turns are possible. The conversational writing corpus, compiled for the present study, is contrasted with existing corpora of various genres of speech and writing to elucidate the relationship between conversational writing and the spoken and written genres. The study is multidimensional in that it applies Biber's (1988) dimensions of linguistic variation to investigate the discourse. Biber's (1988) methodology lies at the heart of the study, as it enables the systematic assessment of lexico-grammatical patterns. In addition, certain textual, interpersonal and modal aspects of the communication are discussed in the light of e.g. Halliday's model of semiotics (see e.g. Halliday 1985a, 2004).

Previous studies have invariably pointed to the dual nature of computer chat, to its oral and written properties (e.g. Ko 1996, Mar 2000, Crystal 2001, Dresner 2005). The present study acknowledges this characterization, discussing the orality and writtenness of conversational writing, but also attempts to rise above the duality. The foremost aim of the study is to position two modes of computer-mediated communication (one synchronous and the other supersynchronous) on Biber's (1988) dimensions of linguistic variation, using Biber's multifeature multidimensional model. A mode is defined as "a genre of CMC that combines messaging protocols and the social and cultural practices that have evolved around their use" (Herring 2002: 112, drawing on Murray 1988). The synchronous mode investigated in the present study is Internet relay chat (IRC), and the supersynchronous mode is split-window ICQ ("I seek you") chat (for descriptions and screenshots of the modes, see section 2.6). Biber's (1988) dimensions are continua along which spoken and written genres vary with respect to more than their oral and written character, for instance with regard to their informational vs. involved focus. To adapt to Biber's framework, "modes" are also termed

"genres" in this study. Positioning the genres of conversational writing on Biber's dimensions is expected not just to provide a clearer picture of conversational writing as a whole, but also to enable the detailed linguistic description of the discourse in the individual conversational writing genres.

From the inception of human computer-mediated communication, the synchronous and supersynchronous modes of computer chat, with their simultaneously oral and written properties, have puzzled linguists and laymen alike. While linguists carefully analyze the oral and written features of the discourse, the chatters themselves conceive of their communication as "talk." Below are ten turns (examples 1 a–j) sampled from various computer chat channels and private chats in which chatters' metalanguage reveals the perceived nature of the communication.

(1)   a.   hey i'll *talk* to ya all later i need to jet for a lil while
        b.   ah been *talking* while ive been away have you ?
        c.   i like it how you *talk* with me
        d.   anyone want to *talk*
        e.   i was trying to *talk* french….. ouoooo fuooooo pou shou
        f.   where *talking* about hat in a chat
        g.   i wasnt *talkin* to you!!
        h.   jim what r u *talking* about?
        i.   you're not *saying* anything to me except hi
        j.   so you wanna *hear* the rest of the v day story?
                        Internet relay chat and split-window ICQ chat (UCOW)

Computer chatters arguably perceive their communication as *talk*, as *saying* things and *hearing* each other's utterances (cf. Giese 1998, Herring 2011b). Just as in an oral situation, their "talk" occurs in real time; it is spontaneous, interactive and immediately revisable. The oral nature of the communication is also reflected in the very denomination of the medium they use: computer chat. *The Oxford English dictionary* (OED) defines the verb "chat" as "to talk in a light and informal manner; to converse familiarly and pleasantly" and the noun "chat" as "familiar and easy talk or conversation." The conversations in computer chat are fluid; topics evolve and evanesce; feedback is immediate; questions are answered (or not) and emotive content abounds (ranging from affective to adversarial). Even so, it is only through writing that the conversation is made possible; it is conveyed by keystrokes of letters and punctuation, and decoded visually by the recipient. The interlocutors depend on the encoding and decoding of graphemes, much like writers and readers in the written media (books, journals, magazines, hypertext, notes, etc.). Demonstrably, real-time text-based computer-mediated communication is a coin with two sides – the oral and the literate. The present study inhabits

this borderland of speech and writing, the field of tension that constitutes the interface between the spoken and the written, endeavoring to map it out.

Conversational writing is by no means a creation of the computer – it presumably appeared long before this invention. During classroom lessons, for instance, when silence is preferred, students sharing a desk may pass notes between themselves carrying out a silent, written conversation. Such a conversation relies on interlocutors' mutual awareness of each other's presence and immediate attention to the message. The students' messages thus constitute conversational writing (as does their act).[1] The computer, however, has made conversational writing (the textual product) amenable to large-scale study, or rather, the Internet and logging software have. Conversational writing texts in this study are chatted[2] texts produced for social interaction in synchronous and supersynchronous computer-mediated communication (SCMC and SSCMC).[3] As mentioned, the two modes of computer-mediated conversational writing differ in that the conversational writing in SCMC is carried out in, for instance, chat channels to which participants submit their entire turn, one turn at a time, whereas the conversational writing in SSCMC is carried out between two or three interlocutors in a split window, in which turns are realized keystroke by keystroke with possible complete overlap. In either mode, the chatters producing the computer-mediated texts, like the students passing notes, rely on the simultaneous presence of a recipient and expect the recipient's immediate feedback. The synchronicity of their communication enables the interlocutors to affect each other's line of thought before or during its formulation into words, thereby, just as in oral interaction, enabling interlocutors to stake out the direction of the conversation. The following four characteristics in combination, then, provide a working definition of conversational writing: it is written communication 1) for social interaction 2) which requires the simultaneous presence

---

1   "Conversational writing" may refer to both a textual product (a noun) and an act (a verbal noun). The present study is primarily concerned with conversational writing in the former sense, i.e. with the texts themselves.

2   "Chatted" is recurrently used as an adjective in this study, by analogy with the adjectives "spoken" and "written," to denote the texts of computer-mediated conversational writing (thus, "chatted" texts/corpora/words etc. are contrasted with "spoken" and "written" texts/corpora/words etc.).

3   Other texts may also be produced in synchronous and supersynchronous CMC, for instance in office suites for collaborative writing (e.g. in Google Docs). Documents co-authored in the document window of such collaborative writing software, however, are typically expository prose, spreadsheets and presentations, and not conversational writing. Collaborative writing texts are not considered in this study. The present study is concerned only with conversational writing intended for social interaction.

(physical or virtual) of producer and recipient, 3) in which interlocutors expect immediate feedback (i.e. within seconds) and 4) during which the discourse may be reconfigured by the participants while under construction (e.g. as interlocutors are able to influence each other's line of thought).

Linguists studying computer-mediated discourse have characterized both asynchronous texts (such as e-mail and computer conferencing texts) and synchronous texts (computer chat) as intermediate between speech and writing. Investigating asynchronous CMC (ACMC), Collot & Belmore (1996: 28) call the communication a "hybrid" variety of English, Yates (1996: 46) concludes that it is "neither simply speech-like nor simply written-like," and Davis & Brewer (1997: 2) call it "writing talking." Studying synchronous CMC, Ferrara et al. (1991: 10) call the "interactive written discourse" a "hybrid register that resembles both speech and writing, yet is neither." Similarly, Foertsch (1995: 304) finds the electronic discourse to occupy "the middle ground between oral and written discourse," but also makes clear that the "most compositional formats" (cf. ACMC) fall closer to the written side whereas the "most interactive formats" (cf. SCMC) fall closer to the oral side. A number of empirical investigations of SCMC have shown that its discursive content is intrinsically oral in nature. Werry (1996) exemplifies richly from SCMC to show how the discursive style of the communication simulates face-to-face spoken language. Schulze (1999) points to the inherent interactiveness as the most important characteristic of SCMC, presenting non-verbal and paraverbal properties as well as means for signaling presence cues and status information as features that make SCMC similar to spoken communication. Hård af Segerstad (2002: 246) characterizes SCMC as "a form of conversation, which happens to be written down instead of spoken."

Although most previous studies show that computer-mediated discourse defies simple classification into speech or writing, they point to the important assumption that synchronously mediated texts display more speech-like properties than asynchronous texts (cf. Korsgaard Sorensen 1993, Herring 2001, Sveningsson 2001, Hård af Segerstad 2002, Condon & Čech 2010, Georgakopoulou 2011a). Very few linguists have studied supersynchronously mediated conversational texts, even though several have suggested such studies, e.g. Hård af Segerstad (2002: 269) and Freiermuth (2003: 183). Herring (2004a, 2007) suggests synchronicity as a useful parameter for distinguishing among modes of CMC, seeing that synchronicity is a "robust predictor of structural complexity, as well as many pragmatic and interactional behaviors, in computer-mediated discourse" (Herring 2007: 14). Given its greater interactiveness, supersynchronous conversational CMC might thus display even more speech-like properties than synchronous conversational CMC,

a possibility that makes a contrastive study of the two highly desirable. However, as Herring (2011b) points out, genres of conversation (oral as well as computer-mediated) should be studied not only with regard to the oral vs. written dimension, but could be situated along various other dimensions (cf. Biber's 1988 study). The present investigation, consequently, takes all of Biber's six dimensions into account (e.g. informational vs. involved production and narrative vs. non-narrative concerns, dimensions further described in section 2.3) for the classification of synchronous and supersynchronous conversational CMC.

Herring (2011b) notes that the scholarly assessment of relative degrees of conversationality in different CMC modes is straggling; "no single set of methods is employed, or questions asked, across the collection that would make the results of the individual studies directly comparable with one another" (2011b: 7). Calling for research in the field, she emphasizes that the "systematic consideration of what it means for CMC to be 'conversational' is still lacking" (2011b: 3), as is the systematic comparison of multiple modes of CMC using a "common set of methods" (2011b: 7). The present study is a first step towards remedying these shortcomings; it intends not only to describe conversational writing, but intends to do so using Biber's (1988) systematic multifeature multidimensional (MF/MD) methodology. Positioning the two modes of CMC on Biber's dimensions enables not just the systematic comparison of the modes, but also the systematic comparison of the modes (genres) relative to other genres of writing and speech. Although none of Biber's (1988) dimensions makes a simple, dichotomous distinction between writing and speech, the dimensions differentiate among literate and oral genres in different respects. Among the genres situated by Biber (1988) on the dimensions are face-to-face and telephone conversations. By relating the conversational writing genres to these conversational genres on the different dimensions, lexico-grammatically, situationally and functionally, it is possible to determine the degree of orality in conversational writing. A high degree of orality means that the conversational writing genre displays features with great resemblance to spoken conversations, or even displays features or levels beyond current notions of orality, thus re-defining what it means to converse in real time. In the next section, the hypotheses regarding the relationship between the conversational writing genres and oral conversations are presented, along with the research questions to be addressed in the study.

1.2 Aim and scope of the study

The principal aim of the present study is to position two genres of conversational writing, one of synchronous and the other of supersynchronous CMC, on Biber's

(1988) six dimensions of textual variation; Biber's methodology and dimensions are described in section 2.3, and the positions of conversational writing genres are presented and discussed in chapter 5. The dimensions distinguish spoken and written genres but are not strictly scales of variation between speech and writing; rather, they are "fundamental parameters of linguistic variation among English texts" (Biber 1988: 200).[4] Biber's (1988) *multi*-dimensional model, by definition, substantiates that there is no single dimension of orality vs. literacy (written-ness); rather, texts vary on several dimensions at one and the same time. Three of the dimensions (1, 3 and 5) in themselves can be said to distinguish between spoken and written discourse (despite evincing some overlapping genres of speech and writing), but the other three dimensions (2, 4 and 6) do not correspond to this distinction. Determining the degree to which conversational writing resembles oral conversation therefore imperatively entails consideration of the genres' positions on all six dimensions. It is simply not adequate to equate the two on the basis of one dimension only; rather, in the multidimensional model, "two genres are 'similar' to the extent that they are similarly characterized with respect to all dimensions; they are 'different' to the extent that they are distinguished along all dimensions" (Biber 1988: 168).

It was mentioned in the previous section, and above, that the degree of orality in conversational writing can be determined by considering how conversational writing relates to oral conversations on Biber's (1988) dimensions. Systematic correspondence between the conversational writing genres (from SCMC and SSCMC) and oral conversations (face-to-face and telephone conversations) on all dimensions should then suggest a high degree of orality in conversational writing (although such a correspondence has to be functionally attested to be conclusively established). Biber (1988) mentions face-to-face conversations as a stereotypically oral genre, as "having the characteristic situational features that are most typical of speech" (1988: 162). His analyses characterize discourse as highly oral when it displays characteristics of involved production, situation-dependent reference and non-abstract content, as opposed to highly "literate" discourse, which displays features of informational production, explicit, elaborated reference and abstract content (1988: 162–163), all characterizations based on dimensions 1, 3 and 5. Special attention is thus paid to these dimensions in the investigation of the orality of conversational writing, even though, admittedly, the full picture of the nature of conversational writing emerges only through the

---

4   In variation studies, studying speech means studying transcribed speech, hence spoken "texts."

overall consideration of all dimensions together, and such overall consideration is also forthcoming, in the penultimate chapter of this study.

The assumption underlying the two hypotheses to be tested in this study is that most written discourse is conveyed in one-way communication or in asynchronous exchanges (with delay between production and reception), whereas most oral discourse is conveyed in synchronous exchanges (with no delay between production and reception), and that the degree of orality increases with the degree of synchronicity (cf. Korsgaard Sorensen 1993, Condon & Čech 2010). As mentioned, Herring (2007: 14) suggests synchronicity as "a useful dimension for comparing different types of CMC with spoken and written discourse." The present study consequently acknowledges the importance of the synchronicity of communication in the various genres under study, alongside the analysis of the genres' positions on Biber's dimensions. By virtue of being communicated in real time, the discourse of the conversational writing genres is expected to approximate the discourse of oral conversations, despite its not being spoken. The two hypotheses underlying the present investigation are presented below. The first is derived from previous research on CMC (e.g. Foerstch 1995, Sveningsson 2001, Hård af Segerstad 2002, Herring 2007), and it forms the point of departure for the second hypothesis.

- Synchronous conversational writing displays a higher degree of orality than asynchronous CMC
- Supersynchronous conversational writing displays a higher degree of orality than synchronous conversational writing

None of Biber's dimensions explicitly distinguishes between asynchronous and synchronous discourse; instead, as described, "orality" in the present study is determined by a genre's similarity to oral conversations. A factor concomitant with such similarity, however, is synchronicity; oral conversations are indeed synchronous. Conversational writing, as mentioned, comprises synchronous and supersynchronous communication. The supersynchronous mode surpasses oral conversations in that interlocutors in SSCMC can carry out conversations in complete overlap for an extended period of time (as supersynchronous conversational writing is carried out in a split window into which both interlocutors type at once), even if this opportunity is not taken at all times. Such complete overlap is possible in oral conversations too, but is usually avoided as extended overlap renders the communication incomprehensible (cf. Herring 1999). In SSCMC, by contrast, extended complete overlap does not affect the comprehensibility of the communication. Supersynchronous conversational writing can thus be regarded as exceeding oral conversations in synchronicity (which explains its denotation as *super*synchronous); see also table 1.1 below (to be explained in section 1.3).
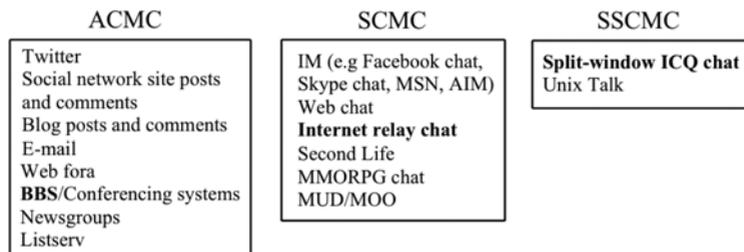
Determining the degree of orality in SSCMC consequently entails taking into account not just the similarity of supersynchronous conversational writing to oral conversations, but also the paradoxical possibility of the former exceeding oral conversations in "orality," as seen from the perspective of synchronicity. This will be borne in mind in the interpretation of the results of the study, especially in the consideration of the positions of the supersynchronous conversational writing genre on Biber's dimensions, more precisely, on dimensions 1, 3 and 5, the dimensions with an "oral" end.

The corpus of conversational writing to be investigated in the present study, the "Uppsala Conversational Writing Corpus" (UCOW), was recorded and annotated in a research project culminating in this book, and will be described in detail in sections 3.1–3.3. The corpus consists of conversational writing from SCMC, as instantiated in Internet relay chat (IRC) chat channels, and from SSCMC, as instantiated in private split-window ICQ chats. The two genres to be positioned on Biber's dimensions are therefore labeled "Internet relay chat" and "split-window ICQ chat," respectively. The genres are exemplars of SCMC and SSCMC, much as face-to-face conversations and a range of other genres exemplify speech, and as e.g. academic prose and a host of other genres exemplify writing. The categories speech, writing, ACMC, SCMC and SSCMC each have the working label of "medium" in the present study. In a slightly opportunistic account of Internet language, Crystal (2001) subsumes the various modes of textual CMC under one linguistic variety labeled "Netspeak." "Netspeak," according to Crystal, "is something completely new […] something fundamentally different from both writing and speech […] in short, a fourth medium," the first three being speech, writing and sign language (Crystal 2001: 238). The present study recognizes the relative novelty of textual CMC, but stresses the heterogeneity of this communication, above all hesitating to draw conclusions as to the linguistic nature of all CMC. Rather, conclusions are drawn only with respect to conversational writing, as instantiated through IRC and split-window ICQ chat, or with respect to one of these modes.

To understand the diversity of CMC modes for social interaction, see figure 1.1. Out of all these modes (further explained in sections 2.5 and 2.6), the present study covers only one mode of SCMC (Internet relay chat) and one mode of SSCMC (split-window ICQ chat). With regard to ACMC, the study comments on various previous research, especially that of Collot (1991) and Collot & Belmore (1996) on bulletin board system (BBS) communication, as well as Yates (1993, 1996). ACMC is brought in mostly as a quantitative point of reference in this study, as the ACMC data to be compared to conversational writing consists of linguistic frequency counts, derived from previous research, more than actual corpus texts (since few texts have been made available from the comparable studies).

The hands-on empirical investigations in this study thus focus primarily on conversational writing, that is, on the conversational discourse in SCMC and SSCMC.

Figure 1.1: *Examples of asynchronous, synchronous and supersynchronous modes of written CMC.*[5]

ACMC	SCMC	SSCMC
Twitter Social network site posts and comments Blog posts and comments E-mail Web fora **BBS**/Conferencing systems Newsgroups Listserv	IM (e.g Facebook chat, Skype chat, MSN, AIM) Web chat **Internet relay chat** Second Life MMORPG chat MUD/MOO	**Split-window ICQ chat** Unix Talk

The research questions to be answered in the course of the present study are the following four (parentheses indicating chapters in which the questions are addressed):

---

5 Some modes deserve explanation as they may be unfamiliar to present-day readers. A BBS is a bulletin board system run on a server to which users log in to post and exchange messages (Collot 1991, Collot & Belmore 1996). BBSs peaked around 1996 but were rapidly replaced by web fora upon the popularization of web browsers (the hypertext protocol). Computer conferencing systems have come in an abundance of modes (besides BBSs), e.g. CoSy, VAX Notes, Confer, First Class; those investigated by linguists include CoSy (Yates 1993, 1996) and VAX Notes (Davis & Brewer 1997). Newsgroups are hosted by Usenet servers and accessed in client programs via users' selective subscription (Herring 2002, Paolillo 2011). Listservs are electronic mailing list software applications that allow users access to global interest groups, also by subscription (Herring 1996b). Like BBSs, newsgroups and listservs have largely been superseded by web fora and other web-based applications. IM is an umbrella term for instant messaging software, of which early applications include Microsoft Messenger (MSN) and America Online instant messenger (AIM). Linguistic studies of IM include Baron (2004, 2010) and Tagliamonte & Denis (2008). Second Life is a graphic online virtual world in which users interact as avatars. Also graphic, MMORPGs are massive multiplayer online role-playing games, e.g. World of Warcraft. MMORPGs have largely superseded MUDs, multi-user dungeon games, and MOOs, i.e. MUD object-oriented applications, which are text-only virtual worlds (Reid 1994, Cherny 1994, 1999, Herring et al. 2009). Second Life, MMORPGs, MUDs and MOOs all allow synchronous chat among participants in the virtual worlds. A predecessor of split-window ICQ chat is Unix Talk, also carried out in a split window and realized character by character with possible complete overlap.

- What is the linguistic nature of conversational writing and the genres studied here, IRC and split-window ICQ chat? (Chapters 4, 5 and 6)
- How does conversational writing carried out in SCMC and SSCMC, respectively, relate to writing and speech? (Chapters 4, 5 and 6)
- How do the genres of SCMC, SSCMC and ACMC relate to oral conversations on Biber's (1988) dimensions? (Chapters 5 and 6)
- Does conversational writing carried out in SCMC and SSCMC constitute a modality of its own? (Chapter 6)

The first two research questions are treated extensively in chapter 4, which contrasts the conversational writing genres with the media of speech, writing and ACMC (the latter represented by Collot's 1991 BBS conferencing genre), in chapter 5 with regard to Biber's dimensions and in chapter 6, summarizing and discussing the results. The third question is partly addressed in chapter 5, in which the positions of Internet relay chat (SCMC) and split-window ICQ chat (SSCMC) on Biber's dimensions are presented, as well as those of BBS conferencing (ACMC), and partly in chapter 6, which discusses the CMC genres' similarity to oral conversations on Biber's dimensions. By relating the CMC genres to oral conversations on Biber's dimensions it is also possible to address the two hypotheses posed at the beginning of this section, which suggest different degrees of orality in texts from the three CMC media.

As mentioned, the working label for speech, writing, ACMC, SCMC and SSCMC is "media." Speech and writing are known from previous research to be separate modalities, as is sign language; see figure 1.2.[6] The present study leaves sign language out of account, but attempts to answer the fourth research question, as to whether conversational writing constitutes a fourth modality, as it is conceptualized along the dashed line in figure 1.2. This fourth research question will not be addressed until chapter 6, when all results have been presented, as the answer must be backed up by substantial evidence. In the meantime, the genres are subsumed under their media categories when compared to speech and writing. This means that, in chapter 4, the media of ACMC, SCMC and SSCMC are compared to the media of speech and writing whereas, in chapter 5, the genres

---

6   A modality is a "means of production/reception" (Herring 2007: 5), i.e. a means of materializing a linguistic message. Three modalities are regularly recognized in linguistics: speech, writing and sign language (cf. Baron 1981). Crystal (2001) calls each of the three a medium. Crystal (2008b: 300) notes that speech is regarded as "the 'primary medium'" and "writing the 'secondary' or 'derived' medium," and that various branches of linguistics may denominate these modalities instead of media.

of IRC (representing SCMC) and split-window ICQ chat (representing SSCMC) are compared to various genres of speech and writing. ACMC is subsumed under the written modality throughout this study, as ACMC is not defined as conversational writing.[7] Owing to the lack of corpus texts (from Collot's 1991 ACMC study), however, the investigation of ACMC is limited; instead, discussions mainly revolve around conversational writing as compared to traditional writing and speech (cf. the genres studied in Biber 1988). Figure 1.2 illustrates the working relationship between modalities, media and genres/modes in the present study, in which genres are subcategories of the media writing and speech, and modes are subcategories of the three CMC media.

Figure 1.2: Working relationship between modalities, media and genres/modes in the present study.
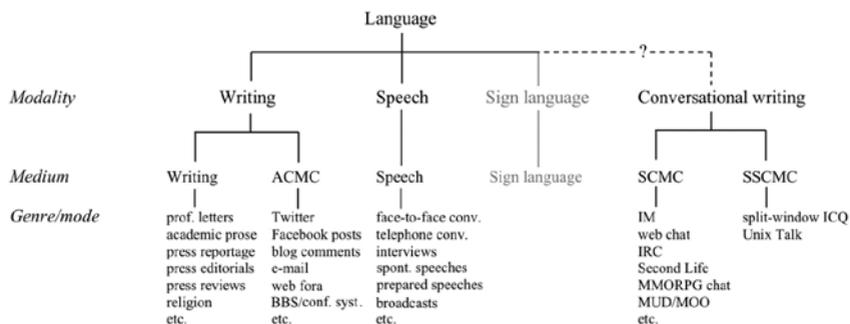
The primary purpose of the first results chapter, chapter 4, is to document the features that are salient in conversational writing when the genres of SCMC and SSCMC studied are compared to speech and writing at the level of medium. The quantitative findings presented in the chapter utilize the mean frequencies and standard deviations of Biber's (1988) linguistic features, and interpretations draw on the fact that the features in conversational writing that deviate most from the mean of all spoken and written genres (considered in Biber 1988) are those that

---

7   Drawing on Morrisett (1996), Mann & Stewart (2000: 182) note that ACMC "has been associated with […] characteristics found in traditional writing forms," including "having the time to study, analyse and reflect on incoming messages and being able to compose responses carefully" (ibid.), but also acknowledge that other analyses have characterized ACMC as a hybrid variety between writing and speech (as noted in section 1.1 above).

most distinctively characterize conversational writing. In addition, the chapter takes up salient features in the conversational writing corpus that are not included among the features studied in Biber's (1988) multidimensional methodology, among them paralinguistic features. The analysis of conversational writing thus sets out broadly, relating computer-mediated communication to speech and writing, and proceeds to more fine-grained scrutiny, comparing the genres of conversational writing to the multiple genres of speech and writing, all in order to adequately answer the research questions posed above.

1.3 Synchronicity of communication

The present study recognizes synchronicity as a useful construct for classifying text-based computer-mediated communication (as suggested in Herring 2004a, 2007). Accordingly, figure 1.1 illustrated the synchronicity of communication in a number of CMC modes. As mentioned, the analysis of conversational writing in this study is based on the UCOW components Internet relay chat, representing SCMC, and split-window ICQ chat, representing SSCMC. The UCOW findings are related to the findings in previous research on ACMC, especially on BBS conferencing (Collot 1991, Collot & Belmore 1996), but more importantly, to the genres studied by Biber in his account of textual variation in English (Biber 1988). Concomitant to the classification of synchronicity in the CMC genres/modes (figure 1.1), therefore, is the consideration of the synchronicity of communication in Biber's (1988) genres. Biber studied six genres of speech (face-to-face conversations, telephone conversations, interviews, broadcasts, spontaneous speeches and prepared speeches) and 17 genres of writing (including professional letters, academic prose, press reportage, press editorials, popular lore, general fiction and official documents). The spoken texts derived from the London-Lund Corpus, LLC (Svartvik 1990), and the written texts from the Lancaster-Oslo/Bergen Corpus, LOB (Johansson et al. 1978), and two collections of letters. A complete list of the texts included in Biber's (1988) study is given in Appendix I. To begin to relate the CMC genres/modes to Biber's genres with regard to synchronicity of communication, they are here conflated into one list, table 1.1, along with a number of other existing and hypothetical genres (not yet classified as such).

Table 1.1: Principal synchronicity and direction of communication in various genres. Genres studied appear in bold script, other genres in normal script[8]

Type of communication	corpus		asynchronous one-way	asynchronous two-way	synchronous one-way	synchronous two-way	supersyn-chronous two-way
Speech (conversation)							
face-to-face	LLC, SBC	A				●	(●)
audiovisual telephone		A				●	(●)
audiovisual Skype telephone		A				●	(●)
telephone	LLC	A				●	(●)
Skype telephone		A				●	(●)
Ventrilo		A				●	(●)
Conversational writing							
split-window ICQ chat	UCOW	V				●	●
Internet relay chat	UCOW	V			(●)	●	
Second Life		V			(●)	●	
MMORPG chat		V			(●)	●	
notes passed face-to-face		V			(●)	●	
IM, e.g. Facebook chat		V	(●)	(●)	(●)	●	
Speech – continued							
interv., publ. conv., debates	LLC	A				●	
spontaneous speeches	LLC	A			●	(●)	
prepared speeches	LLC	A			●		
TV/web broadcasts[8]		A	●		●		
audio broadcasts	LLC	A	●		●		
voicemail, Heytell		A	●	(●)			
Asynchronous CMC							
SMS		V	(●)	●	●	(●)	
Twitter		V	●	●	(●)	(●)	
e-mail		V	●	●	(●)	(●)	
Facebook posts, comments		V	●	●			
blog posts and comments		V	●	●			
newsgroups, **BBS**, web fora	ELC other	V	●	●			
Writing							
personal letters	Grabe	V	(●)	●			
professional letters	Biber	V	●	(●)			
FAQs		V	●	(●)			
Wikis		V	●	(●)			
posted personal notes		V	●	(●)			
academic prose	LOB	V	●	(●)			

---

8   TV/web and audio broadcasts, of course, may contain e.g. conversations – the categorization here indicates only the communicative purpose of the broadcast to the public and the conventional response from the public.

31

Type of communication	corpus	asynchronous one-way	asynchronous two-way	synchronous one-way	synchronous two-way	supersyn-chronous two-way
commercial web pages		V	●			
direct mail letters		V	●			
press reportage	LOB	V	●			
press editorials	LOB	V	●			
press reviews	LOB	V	●			
religion	LOB	V	●			
hobbies	LOB	V	●			
popular lore	LOB	V	●			
humor	LOB	V	●			
biographies	LOB	V	●			
general fiction	LOB	V	●			
mystery fiction	LOB	V	●			
science fiction	LOB	V	●			
adventure fiction	LOB	V	●			
romantic fiction	LOB	V	●			
official documents	LOB	V	●			

A – mainly auditory reception
V – mainly visual reception

Table 1.1 outlines the distribution of genres according to their principal synchro-nicity and direction of communication.[9] Communication may be asynchronous, synchronous or supersynchronous. Asynchronous and synchronous communi-cation can vary in direction, i.e. it can be one-way or two-way. Professional letters that elicit no response from the recipient, for instance, are communicated one way, presumably the default direction of such letters (indicated by a bullet in table 1.1). A professional letter that is responded to, however, becomes part of a two-way asynchronous transaction (indicated by a parenthesized bullet). Super-synchronous communication, however, is by default two-way.[10] In individual genres, texts may be communicated with different synchronicity and direction.

---

9   Table 1.1 includes SMS among asynchronous CMC. Although similar to ACMC modes (cf. figure 1.1), SMS is, strictly speaking, not CMC, but rather telecommunication, and therefore not included among the ACMC modes in figure 1.1.

10 The delineation of two-way communication here differs from Herring's (2001) defini-tion of two-way transmission, as two-way communication here includes asynchronous writing, whereas Herring (2001) regards asynchronous writing as one-way transmis-sion. Herring's definition of two-way transmission includes only modes (genres) in which speaker and addressee perceive the message as it is produced, such as oral conversations and supersynchronous conversational writing.

The bullets in table 1.1 indicate the main types of communication carried out in the genres, i.e. the types of communication used to fulfill the communicative purposes of the genre. By inference, face-to-face conversations are usually synchronous (as indicated by a bullet), but may be supersynchronous for very limited periods of time (as indicated by a parenthesized bullet). Split-window ICQ chat, by contrast, is both synchronous *and* supersynchronous, as its keystroke-by-keystroke means of transmission enables the communication to fluctuate between being realized in consecutive, synchronous, turns up to extensively overlapping, supersynchronous, turns.

Internet relay chat, on the other hand, is only carried out synchronously, i.e. turn by turn, and one-way messages, i.e. turns not responded to, are more likely in the Internet relay chat channels than in the split-window ICQ mode (adding a parenthesized bullet for the former, but not for the latter, in table 1.1). The genres to be contrasted in the present study are marked in bold in table 1.1, and the corpora from which they derive are indicated in the second column. (Other existing and hypothetical genres are interspersed among these, inter alia to illustrate a number of linguistically understudied genres.)

The genres in table 1.1 are ordered from top to bottom by their principal degree of synchronicity. Genres of similar synchronicity are only tentatively ordered relative to each other (the written genres in one-way communication, of course, defy ranking altogether). The table nevertheless serves to illustrate an important point. Before conversational writing, synchronous communication relied almost exclusively on the acoustic channel, i.e. on auditory reception (indicated as A in table 1.1). Auditory (A) and visual (V) reception were then largely on a par with the distinction between speech and writing. Table 1.1 illustrates how conversational writing challenges this division; not only is conversational writing synchronous, like spoken conversation, but it also challenges conversation with a more synchronous genre, one amenable to extended supersynchronous communication: split-window ICQ chat. As will be seen in this study, Biber's (1988) multidimensional methodology is a highly useful tool for distinguishing speech and writing on dimensions other than oral (cf. A) vs. literate (cf. V). Yet, split-window ICQ chat was not around at the time of the 1988 methodology's conception, nor was written synchronous communication included in Biber's study, cf. Internet relay chat. Taking synchronicity, especially supersynchronicity, into account is therefore imperative for an adequate description of the orality of conversational writing, in the interpretation of the positions of conversational writing on Biber's dimensions (chapters 5 and 6), as well as in the preceding comparison of conversational writing to speech and writing (chapter 4). Table 1.1 serves to conceptualize the parameters to bring into those considerations.

1.4 Notes on terminology

A few notes are in order with regard to the terminology applied in the present study. Table 1.1 in the previous section lists in bold the "genres" to be considered in the study. The term "genre" is used here in analogy with Biber (1988) referring to "categorizations assigned on the basis of external criteria" (1988: 70), that is, criteria related to the author's or speaker's communicative purpose. The genres in Biber (1988) are largely adopted from the categories distinguished in the corpora from which they derive (see Appendix I) and constitute "text categories readily distinguished by mature speakers of a language" (Biber 1989: 5). The present study agrees with Biber's (1988, 1989) definition of genre, finding the CMC modes studied, especially IRC and split-window ICQ chat, equally distinguishable on external criteria as Biber's genres, i.e. on the basis of their external format and the distinct situational setting of their production, and thus worthy of the designation of genre (although they may also be referred to as modes).

In addition to genre, however, the term "register" is also regularly used in linguistic studies to refer to situationally defined varieties of speech and writing. While some studies exclusively use the term "genre" (e.g. Biber 1988, Biber & Finegan 1989, Swales 1990, Love 2002), others use the term "register" (e.g. Atkinson & Biber 1994, Biber 1995, Biber et al. 1999, Conrad 2001). Some have attempted to draw theoretical distinctions between genres and registers, e.g. Ferguson (1994) who regards "genre" as "[a] message type that recurs regularly in a community" and "register" as "[a] communication situation that recurs regularly in a society" (1994: 20–21), while others have used the terms rather interchangeably, e.g. Biber (1993: 244) who uses both to refer to "situationally defined text categories." In 1995, Biber notes that there is "no general consensus within sociolinguistics concerning the use of *register* and related terms such as *genre* and *style*" (1995: 8, original italics), reviewing attempted distinctions as "quite abstract and vague" (1995: 9). In his 1995 study, Biber opts for "register" as a general cover term for all aspects of variation in use, admitting that it corresponds closely to his earlier use of "genre" (1995: 10).

Discussing genre categorizations in corpora, Lee (2001) points out that several genres in corpora really denote sub-genres (e.g. the five fiction genres in LOB), rather than situationally defined varieties, but calls for calm among linguists; "we need not be unduly worried about whether we are working with genres, sub-genres, domains, and so forth, as long as we roughly know what categories we are working with and find them useful" (2001: 52). Advocating consistency in any approach, Lee proposes the usefulness of seeing the terms "genre" and "register" as two different angles, or points of view; "genre" being used to talk about

"membership of culturally-recognisable categories" (2001: 46) and "register" being used to talk about "lexico-grammatical and discoursal-semantic patterns associated with situations" (ibid.). Genres, says Lee, are instantiations of registers (as a genre may invoke more than one register), and "so will have the lexico-grammatical and discoursal-semantic configurations of their constitutive registers, in addition to specific generic socio-cultural expectations built in" (2001: 46–47). As if in line with Lee's reasoning, Biber & Conrad (2009), although they draw distinctions between "register," "genre" and "style," opt to focus mostly on the register perspective, as it is seen as valid for the description of all text varieties. Whereas the genre perspective focuses on the "conventional structures used to construct a complete text within the variety" (2009: 2), say Biber & Conrad, the register perspective can be used to analyze "any text sample of any type" (ibid.).

The present work studies conversational writing from both the genre and the register perspective (leaving individual "styles" largely out of account) but employs the term "genre" from both perspectives in the multidimensional analysis. The conversational writing genres have been identified on the basis of external criteria, and the study aims to identify the genres' lexico-grammatical patterns, i.e. the "conventional structures" that characterize them as varieties. This is done by contrasting the texts of conversational writing with texts from spoken and written genres, for which the conventional structures have been pre-defined (as sets of co-occurring linguistic features) by Biber (1988). Moreover, as the conversational writing genres here in effect represent only one register each, the genre/register distinction is not of central concern. Accordingly, instead of using "register," like Biber (1995), in a way that is similar to Biber's earlier use of "genre," the present author simply opts for using the original term "genre," as defined in Biber's early work (1988, 1989, Biber & Finegan 1989), in the first place.

In Halliday's model of semiotics (e.g. Halliday 1978, Halliday & Hasan 1989) and systemic-functional linguistics (e.g. Halliday & Hasan 1989, Martin 1992, Halliday 2004), on the other hand, the term "register" is the mainstay construct, whereas "genre" is peripheral. A register is a functional variety of language (Halliday 1978, 2004). It is defined on the basis of three variables of context taken together: field, tenor and mode,[11] which essentially represent what is going on in the course of the language exchange, who is taking part, and what role the

---

11 Halliday's variable of mode is paradigmatically distinct from mode defined as a genre of CMC. The semiotic notions field, tenor and mode will be explained further in section 2.4. Mode in the Hallidayan sense will be referred to as "semiotic mode," whenever discussed in non-Hallidayan contexts, to set it apart from "mode" used to denote genres of CMC.

language is playing, respectively. ("Register" in the Hallidayan sense will be further explained in section 2.4.) In the systemic-functional framework, "genre" and "register" are said to represent different semiotic planes. "[A] genre is a staged, goal-oriented, purposeful activity in which speakers act as members of [a] culture" (Martin 2001b: 155), that is, in order to fulfill certain communicative purposes. "Register," by contrast, has more to do with the particular linguistic choices communicators make in a certain genre. The choices result from the contextual variables field, tenor and mode. Genre thus corresponds roughly to "context of culture" and register to "context of situation" (Martin 2001b: 155). In the present study, the systemic-functional concept of "genre" will not be employed. Rather, in the analysis of conversational writing, only the Hallidayan concept of "register" will be used, to discuss linguistic features associated with the field, tenor and mode of the discourse, i.e. the distinct situational setting of the discourse. The configuration of the field, tenor and mode is realized in any given text:

> Any piece of text, long or short, spoken or written, will carry with it indications of its context. We only have to hear or read a section of it to know where it comes from. This means that we reconstruct from the text certain aspects of the situation, certain features of the field, the tenor, and the mode. Given the text, we construct the situation from it. (Halliday & Hasan 1989: 38)

In sum, in the present study, whenever "genre" is used, it is with regard to Biber's framework, and to the extent that "register" is used, it is employed in one of two ways: 1) as interchangeable with Biber's notion of genre, since, as mentioned, several authors have used "register" and "genre" interchangeably (as will be seen, for instance, in section 2.2 surveying previous research into speech and writing) or 2) in connection with semiotic, systemic-functional interpretations. The context of each discussion will clarify in which meaning the term "register" is employed.

A fair number of abbreviations will be used throughout the study. Several of these were encountered in the sections above and, as seen, they are usually explained upon first encounter – if not, and for repeated reference, readers may consult the list of Abbreviations (front matter). Corpus citations throughout the study contain full names of genres and abbreviations for corpora. The corpus citation convention applied for the conversational writing genres is "corpus genre + text number + (corpus name)," e.g. "Internet relay chat text 3b (UCOW)" and "Split-window ICQ chat text 11 (UCOW)," except for when several short samples are conflated into one example, in which case only the genre(s) and corpus name are given, e.g. as in example (1) above, citing "Internet relay chat and split-window ICQ chat (UCOW)." As indicated among conversations in table 1.1, face-to-face conversations from the Santa Barbara Corpus of Spoken American English

36

("SBC," for short), from the 1990s (Du Bois et al. 2000), are studied alongside other genres in the investigation here. A subset of the corpus was adapted and annotated for the present study to supplement the British LLC spoken genres (studied in Biber 1988), for updated reference and a more global approach (see section 3.4 for a description of the procedure). To distinguish among face-to-face conversations from the two corpora, they are named "face-to-face conversations LLC" and "face-to-face conversations SBC" in the discussions. The two are treated as separate genres, although, admittedly, they constitute regional more than situational varieties. Their denomination as individual genres is merely applied for convenience and for consistency with the use of the genre perspective in Biber's (1988) methodology.

1.5 Outline of the study

Following this introductory chapter, the present study is organized into chapters of background (chapter 2), material and method (chapter 3), results (chapters 4 and 5), discussion (chapter 6) and conclusion (chapter 7).

The background, chapter 2, starts out by surveying previous literature on speech and writing, to introduce some of the linguistic features that distinguish between texts from the two modalities. Several of the studies mentioned in the survey are important because they are reflected in Biber's (1988) selection of features, those that he used to map out spoken and written genres. The survey also serves as a theoretical backdrop to the discussions in the results chapters. Next, the chapter introduces Biber's and Halliday's frameworks in separate sections. Biber's approach to linguistic variation is quantitative at its outset, but enables paramount qualitative, functional interpretation, whereas Halliday's approach to linguistic variation is essentially qualitative. The choice of Biber's and Halliday's approaches is partly drawn from Yates' (1993) study of ACMC, as some passages of the present study attempt to parallel Yates' study with analogous analyses of SCMC and SSCMC. Biber's (1988) MF/MD methodology is broadly outlined in chapter 2, as is Halliday's theory of metafunctions in language. The chapter then surveys the literature on computer-mediated communication, among other things to present how previous studies have treated conversational writing. Chapter 2 ends with a description of the interfaces for conversational writing, so as to anticipate the UCOW corpus description in chapter 3.

Chapter 3 is the "Material and method" chapter. It describes the compilation and annotation of the UCOW corpus, the sampling and annotation of SBC, and the application of Biber's (1988) MF/MD methodology to the material. The chapter explains the data retrieval procedure and the calculation of the results.

The quantitative investigation of conversational writing takes all of Biber's 67 linguistic features into account, and the results are presented in two ways: in relation to writing, ACMC and speech, in chapter 4, and in relation to all of Biber's spoken and written genres, in chapter 5.

Chapters 4 and 5 are the results chapters. Chapter 4 focuses on the salient features in conversational writing, e.g. those taken up in previous studies of CMC, such as modal auxiliaries and paralinguistic features, but also features rarely accounted for in quantitative studies of conversational writing, such as inserts (an umbrella term for e.g. interjections and discourse markers, typically found in conversations) and "emotives." "Emotives" is an umbrella term invented in the present study for emoticons (e.g. *:)*, *;)*, :() and sentiment initialisms (e.g. *lol*, meaning "laughing out loud"), both of which add an emotional zest to chatters' utterances. The thrust of chapter 4, however, is to present qualitative analyses of salient quantitative results from the feature counts in the application of Biber's methodology and to contrast measures of lexical diversity (such as type/token ratio, TTR, and lexical density) in the annotated corpora. The most salient linguistic features in conversational writing are those that deviate from the mean of Biber's spoken and written genres by more than two standard deviations, and these, together with other features presented in chapter 4, epitomize the character of conversational writing. Chapter 4 thus constitutes a major step in the description of conversational writing.

Another major step towards the description of conversational writing in relation to speech and writing is taken in chapter 5, which presents the positions of the conversational writing genres (as well as SBC) on Biber's dimensions of linguistic variation. Like chapter 4, the chapter discusses numerous examples from the corpora, to elucidate the nature of conversational writing, but whereas chapter 4 adduces abundant theoretical anchorage to previous research, chapter 5 essentially breaks new ground as regards conversational writing, with fewer references to previous research. Both results chapters, however, contain analyses and discussions, not just the results. Much of the character of conversational writing thus emerges already in the results chapters, even though the penultimate chapter, chapter 6, is dedicated to a crucial, summarizing discussion of all results.

Chapter 6 revisits the hypotheses and research questions posed in section 1.2. The chapter narrows down what answers to these were provided in the study, discusses the findings, and points out what it means for chatted texts to be conversational. Chapter 7, finally, provides a concluding summary of the study and some suggestions for further research.