are you still awake…?

Personal Written Communication: From Early Modern English Letters to Electronic Communication of Today

Thesis

presented to the Faculty of Arts
of the University of Zurich

for the degree of Doctor of Philosophy

by
Claudia Aeschbacher
of Rüderswil / Bern

Accepted in the spring semester 2008
on the recommendation of
Professor Dr. Andreas H. Jucker and
Professor Dr. Daniel Schreier

Zurich, 2008

CONTENTS

PART IV: CONCLUSION

LIST OF TABLES, FIGURES, AND GRAPHS

Tables

Figures

Graphs

Acknowledgements

This dissertation, while an individual work, benefitted from the support, insights, and direction of several people. First and foremost, I wish to express my heartfelt thanks to Prof. Dr. Andreas H. Jucker, who encouraged me to pursue a doctoral degree and agreed on being my main advisor. I want to thank him for constantly supporting my work, always making time to read my work at its various stages, and discussing any problems that this research faced. His expertise in the field of pragmatics has never ceased to impress me and he contributed considerably to the quality of this dissertation. In addition, I wish to thank my second advisor, Prof. Dr. Daniel Schreier, for his support, constructive criticism, and the care with which he reviewed the original manuscript. At this point I also wish to express my gratitude to the Research Committee of the University of Zurich for financially supporting my dissertation project. It would not have been possible to dedicate two years fulltime to this project without this financial support, and I very much appreciate that the Research Committee made this possible. Special thanks also go to my proofreaders, Sonia Paget, and my dear friend, Alain J. Escher, who both provided insights that guided and challenged my thinking, substantially improving the finished product.

In addition to the technical and instrumental assistance above, I received equally important assistance from family and friends. In particular my partner, Kaspar H. Hürlimann, and his son, Joshua Li, have provided on-going and loving support in every imaginable way throughout the dissertation process, from the first thought about this project to printing, submitting, and publishing the work. I simply cannot express in words how much this means to me, I thank you both very much. Also, many thanks go to Bruno Seifert and Bianca Gasparoli, my close friends since childhood, who probably know best what an incredible journey and achievement it has been for me to get to the point of becoming a Doctor of Philosophy. Next, I would like to thank my "girl crew", Zahida Huber Vogt, Silvia Pavlik, Siri Fischer, and Eveline Bernasconi, for providing me with "girl power" and the much-needed opportunities to charge my batteries. I am grateful too for the support of Silvia Zoe Germann, who has always cheered me up when things looked tight and did not take "no" for an answer. And the same is true for Sylvie Dolivo, who never grew tired of listening, may I have been enthusiastic or worried about the outcome of this project. Many thanks also go to Barb Breustedt, my dear friend and much appreciated colleague at the English department of the University of Zurich, for motivating me

"to hang on in there" and for always being there for me to talk things through. Also, I happily thank Lukas Bänninger and Regula Wieser, both of them good friends and former fellow students of my second subsidiary subject Sociology, who have made my academic studies so much more fun than they otherwise would have been. Special thanks also go to Susan Stucki, who has been most inspiring in that she kept reminding me to remain focused and be proud of my work.

Further, I would like to thank my parents, Richard and Therese Aeschbacher, without whom I would not be the person I am today. And I am also grateful that my brother, Rico Aeschbacher, has always believed in me, his little sister, and has given me the feeling that I can accomplish anything. Finally, I wish to thank the informants and respondents of my study (who remain anonymous for confidentiality purposes). Their data, comments, and insights helped create an informative and interesting project with opportunities for future research.

Claudia Aeschbacher
November 2008

Abstract

Ever since the broad population has gained access to the Internet and mobile telephony, there has been an ongoing discussion about how new media, such as the computer and the mobile phone, influence the way we communicate. In particular, it is the textual (as opposed to oral) and the personal (as opposed to professional) interaction such as computer-mediated communication (CMC) and short message service (SMS) texts sent via the mobile phone that have received scholarly attention. Although in the case of CMC the attention has been considerably larger than with SMS discourse. In any event, compared to older forms of (hand-)written correspondence, we are nowadays able to communicate in writing with transmission speed reduced to a minimum. Long gone are the days when a letter or a card took days (or weeks) to reach its destination, and to receive a reply would require yet more patience. These days it is possible to send a written message from Wellington to Zurich and receive a reply all in less than a minute, or, even communicate in (near-)synchronous fashion in certain online settings such as Chat rooms. And not only can we communicate faster, we can also communicate with more people at the same time. In the case of anonymous settings (a feature typical of many services provided by the Internet) this includes personal communication with complete strangers, which is new to the field of personal correspondence. The systematic investigation into five natural language corpora (of the text types SMS discourse, e-mail, Web Chat, personal homepage, and 17th century handwritten letter) seeks to shed light onto idiosyncrasies of personal written communication that are assumed to be connected to changes in the communicative context. It is hypothesised that these changes in the communicative context reflect in how authors contextualise their writing.

PART I: PRELIMINARIES

1. Introduction

(1)[1] are you still awake…?
(SMS text / author: male, 35 yrs)

The message that lends itself as part of the title for this dissertation was sent via a mobile phone's short message service (SMS) in 2003, and it is the immediacy of 21st century electronic communication that has made personal written messages like this possible. There have been some radical changes in the field of personal communication with respect to the composition and transmission of messages compared to the era of letter writing. In case of the message shown in example (1), where the term *still* implies that it is probably night time and that there is a possibility that the recipient is already asleep, high transmission speed is needed in order for the message to stand a chance of being read before the recipient falls asleep should he/she be still awake. As to what purpose exactly the sender had in mind by sending this SMS text is unknown. Because even if the recipient was awake at the time the message was received, how was he or she supposed to react? In any event, SMS discourse shows idiosyncrasies that are connected to the communicative context and this is evident in the texts produced on this medium. The same is true of messages composed on the computer. Investigating the contextual idiosyncrasies of modern personal written communication in comparison to a more traditional form of correspondence, the handwritten letter, is the aim of this study.

[1] In examples throughout this dissertation, names framed with angle brackets indicate that they have been made anonymous (in consideration of upper- and lowercase writing in the original messages). Furthermore, misspellings, alternate spellings (in EEC) or seemingly odd word choices in the original data, such as *threat* instead of *treat* (cf. example (2)), will not be annotated by "sic." in order to avoid lengthy wording. Underline will be used throughout this dissertation to highlight particular aspects in examples, should italics be used in addition to underline for means of emphasis, then this will be pointed out. Three dots surrounded by round brackets, (…), will indicate that the text of the message, from which the example was taken, continues on the same line (or is preceded by text on the same line), whereas the absence of those brackets indicates that either a paragraph occurred in the original text at that point, or that it is in fact the end (or the beginning) of the message. Last but not least, all examples will be identified in terms of text type and basic demographics of the author (if known), if these identifications include the term "excerpt", then the text shown is preceded or followed by more text in the original messages, the absence of this term indicates that the example is in fact the complete message (as is the case with example (1)).

Due to its immediacy, computer-mediated communication (CMC) and SMS discourse are often said to be reminiscent of spoken language and to contain features of orality. For example, the use of emoticons that emulate facial expressions (such as the "smiley") is aimed at compensating for the lack of visual paralinguistic cues in both CMC and SMS discourse. Onomatopoeic expressions (as in *boohoo* which means 'to weep noisily' and at the same time imitates that sound), on the other hand, simulate the phonological aspects of spoken language. The differences between the two production modes, spoken and written language, and their connected concepts of orality and literacy will be subject to elaborate discussion at a later point (see section 4.2.). However, even without any theoretical background most would agree that electronic communication is different from older forms of personal correspondence, but find it at the same time difficult to pin down those differences.

Is it just a matter of the immediacy of the new media, or is it the new media as such, or do we communicate altogether differently in the 21st century? Both the advent of the Internet as well as the short message service have enabled us to send and receive written messages at very high transmission speeds, regardless of time and place, and in the case of anonymous settings on the Internet this also includes an unknown audience. This has inevitably changed perceptions about the communicative context, about "how and why" people communicate with each other on a personal level. It can be assumed that technological developments in the media for written communication have influenced the field of personal communication more drastically than other types of formal communication, such as business correspondence, which has been and undoubtedly always will be more constrained by writing conventions than informal communication. Also, while the communicative interaction in anonymous settings with an unknown audience is new to the field of personal correspondence, many types of public discourse are designed to reach a wide and unknown audience (such as newspapers).

It can thus be hypothesised that any media-related changes in the contextuality of written text is more distinctively traceable in personal written correspondence. However, an empirical investigation into the contextuality of personal written communication is coupled with two important decisions before such an investigation can be attempted. First, one needs to select from a wide array of personal communication on offer, above all in the field of Internet correspondence, which types to consider and then collect the data. Second, one needs to decide how to study the collected data from an analytical point of view. In the

following sections, I will discuss the aim and scope of this study and address these issues in more detail.

1.1. Aim and scope

The decision which types of personal written communication to consider for this study was connected to questions of comparability and availability. Also, different types of personal correspondence may pursue different communicative goals. It was thus crucial for the purposes of this study to select types of correspondence that are comparable in the way they address their readership. This is believed to be the case with the following five types of personal written communication: the handwritten letter, e-mail, SMS messages sent via the mobile phone, the (digital) personal homepage, and electronic Chat performed online. Electronic Chat, which will henceforward be referred to as "Web Chat"[2] (the capitalisation indicating communication performed electronically, as opposed to the noun *chat* in lowercase writing that denotes spoken interaction), differs from the other types of correspondence in that several participants are engaged in a communicative exchange at the same time. However, Web Chat is still felt to qualify for this study because the main communicative goal of users, similar to all other text types that involve one author at the time, is to converse on a personal level with the readership.

With regard to the availability of the data, an investigation into personal written communication inevitably faces privacy issues. Informants are in the majority of cases hesitant to give away data to which they are emotionally attached. And although personal homepages are published on the Internet and Web Chats take place in public Chat rooms, this does not mean the data is freely available for academic research. The five text corpora collected for this research were compiled in careful consideration of ethical issues as well as meeting the requirements of academic corpus design (issues connected to data collection will be addressed in section 3.1. in more detail). Thus, the empirical part of this investigation will be dealing with five natural language corpora in the field of Early English correspondence (91 letters), SMS discourse (1000 texts), e-mail (140 messages), the personal homepage (60 pages) and Web Chat (30 Chat sessions). In addition, an online survey was designed for the purposes of this study in which a total of 109 informants took part. Both the five text corpora and online survey will be introduced in detail in sections 3.2 – 3.4. While the online survey was aimed at finding out more about the communicative

[2] See section 2.2.3. on the use-related meaning of the term *Web* in "Web Chat".

strategies of 21st century authors, the main purpose of the text corpora was to verify the analytical framework designed for the purposes of this study by means of a thorough descriptive analysis.

This takes us back to the second decision, pointed out at the beginning of this chapter, in connection with empirical research: after the decision about what sort of data is to be considered and once it has been collected, it follows that this data undergoes a systematic analysis. The design of an analytical framework in order to investigate the contextuality of personal written communication is a complex undertaking for two main reasons: first, it needs to accommodate the fact that different types of communication are subject to investigation, and that second, different factors influence the contextualisation of the individual messages. The analytical framework designed for the purposes of this study has been developed in consideration of these circumstances. It draws on various theoretical concepts in relation to language in use. Its main cornerstones are based on the contextualisation of text in connection to media-related features (and constraints), indexicality and deixis, illocutionary force, and the notion of relevance. All theoretical concepts relevant to this study will be introduced in chapter 4. The analytical framework, designed against this theoretical background, will be discussed in more detail in chapter 5.

Thus, the corpora's main function is to show that contextuality in writing can in fact be investigated comprehensively, and only on a secondary basis are those results thought to show trends and tendencies for each of the different text types. Based on the results generated by the empirical analysis as well as the findings gained by the online survey, educated conclusions will be drawn as to how 21st century authors of personal written communication contextualise their messages in comparison to the age of letter writing—and in how far the advent of new media, such as the computer and the mobile phone, can be made responsible for any of these differences. It is in this sense that this study seeks to contribute towards ongoing research in the field of linguistics that focuses on the idiosyncrasies of informal messaging, may it be handwritten or electronically produced, from a contextual point of view.

1.2. Setting of the study: from non-electronic to electronic communication

The invention of the printing press by Johann Gutenberg in 1450 made the mass circulation of the written word possible—and technological progress has since brought about great changes as far as the tools for production and dissemination of written language are concerned. While I agree with Faulstich (2006: 7) that the "history of the media" is not to

be misunderstood as a "history of communication", I do believe that the type of medium (or service) that is chosen to produce and transmit a message, or the setting in which a conversation takes place, is to a great extent reflected in how we communicate. This is true for spoken communication (face-to-face conversation vs. conversation on the mobile phone), but even more so in the case of all types of written communication, which by definition require a certain medium in order for the message to be produced and transmitted. It thus seems important to briefly mention the evolution of media for non-electronic correspondence to the media for electronic communication of the present day.

As far as the personal message is concerned, the handwritten letter (or card) was for a long time one of the only means to exchange intimacies in written fashion. There were, however, other non-electronic technologies to communicate privately before the advent of the computer. For example, the first practical typewriter, the famous Remington No. 1, was marketed in 1874 in the USA (cf. Hörisch 2004: 211). One medium that can be seen as an "inter-medium" between non-electronic and electronic communication is the telegraph, which was invented in the USA by Samuel F. B. Morse and underwent worldwide diffusion from 1848 onwards (cf. Faulstich 2006: 71). The prefix *inter-* in the term "inter-medium" refers to the circumstance that the telegraph as such is operated electronically, but the printed out telegram is not. This is why the telegraph can be seen as a "partly electronic" means of written communication, forming a link between non-electronic and electronic communication (cf. Faulstich 2006: 70).[3]

Coming back to the handwritten letter, it seems important to point out that this means of communication goes back to ancient times. Depending on the people and their cultures, letter writing experienced its heyday at different periods during the history of the literate world[4]. Specifications with regard to the degree of literacy aside, it can be claimed that writing letters came into fashion on a grand scale with the implementation of the postal service, which made it possible to send and receive letters on a more regular and

[3] We have an interesting parallel scenario when one prints out an e-mail message for whatever reasons. Although the message was produced electronically, printing it out means that the document is transformed into non-electronic ink and paper. The difference to the telegram is that a telegram needs to be printed out in order to be read. An e-mail message, on the other hand, can be read without being printed. If one still decides to print it out, then it is estranged from the medium in which it was both produced and meant to be read, and in a hairsplitting way no longer "e-mail"—it has lost all that once was e-(lectronic) about it.

[4] The specification "literate world" is quite important if one considers that "language is so overwhelmingly oral that of all the many thousands of languages—possibly tens of thousands—spoken in the course of history, only around 106 have ever been committed to writing to a degree sufficient to have produced literature, and most have never been written at all" (Ong 2002: 7). This means that when we talk about "written communication" we in fact talk about a fraction of the world's languages.

also faster basis. For example, the history of the British postal service started when King Charles I opened up his royal mail service to the public in 1635.[5] Beforehand, one had to find other means to having one's letters delivered. In any event, within the text type "letter" several "subtypes" can be distinguished, the most common distinction being "private/personal letters" and "non-private/business/official letters" (Bergs 2004: 208). With respect to this main distinction, Bergs (2004: 223) makes two interesting observations: first, the distinction between private vs. business letters should be seen as two extremes on a continuum and second, that even in eighteenth-century England, hardly any writing could be kept private (that is hidden from others) and that "private in this context is not defined by use, but by content."

It seems safe to assume that these observations for the eighteenth century are also true for the seventeenth century, the time period from which the letter corpus for this study originates. Fitzmaurice (2002: 4) points out that "the letter—its writing, reading, keeping, endorsing and sending—apparently permeated every aspect of English life in the seventeenth and eighteenth centuries" and that "it is the kind of document most commonly written by literate adults." Or, in the words of Nevala (2004: 271): "Letters were an intrinsic part of communication between people living in the seventeenth and eighteenth centuries." It becomes clear that the letter had a very important status in past centuries with regard to personal written correspondence and was undoubtedly the main medium for the written exchange of private and personal issues for a long time. However, this changed drastically with the invention of the computer and its various services, such as the Internet, that enabled humans to communicate with each other via electronic data transmission in a very quick fashion regardless of time and place. With regard to the invention of Personal Computers, so-called microcomputers, they were made possible by two technical innovations in the field of microelectronics: the integrated circuit, or IC, which was developed in 1959; and the microprocessor, which first appeared in 1971.

The precursors of the Internet were, inter alia, the "Arpanet" (Advanced Research Projects Agency Network, as early as the 1960's) and "Telnet" (Telecommunication Network, 1970's) in the military field, as well as NSFNet (National Science Foundation Network, 1980's) that was developed in a scientific context (cf. Faulstich 2006: 172). Although the first steps towards the Internet, as we know it today, were taken in the early

[5] Cf. *The British Postal Museum and Archive*, online source 1.

1960's, it has only become widely popular since the invention of the World Wide Web[6] at the beginning of 1990. In the course of the 1990's, the number of websites increased drastically from 90 (1992) to 7.2 million (2000), of which about 70% were located in the USA (Faulstich 2006: 172). It is estimated that to date, more than a billion people have regular access to the Internet.[7] As the world population is currently just above 6.5 billion, this means almost one sixth of the population may go online anytime.[8]

Obviously, the precursor of mobile telephony was landline telephone technology. Although the first landline telephone was invented in the 19th century,[9] it was only in the 1960's that it became a mass medium (cf. Faulstich 2006: 144). About twenty years later, at the beginning of 1980, the so-called "first generation" of wireless phones (using analogue technology) was introduced, to be followed by what is known as the "second generation" digital wireless phone technologies (cf. Kauffman 2005: 4). Recently, "third generation" wireless phone technology has entered the market, providing improved sound quality and multimedia applications such as e-mail at very high speeds (cf. Kauffman 2005: 5). Regarding mobile phone ownership, according to the GSM (global system for mobile communication) Association, there were more than 1.4 billion mobile phone subscribers in 2004, and it was then predicted that the market would expand to reach 2 billion subscribers (around 31% of the world population) by mid-2006 (cf. Kauffman 2005: 1-2). The predictions were accurate: today it is estimated that there are over 2.6 billion mobile phone subscribers worldwide, with some of the more mature markets having more than 100% penetration, which translates to more than one subscribed mobile phone per person.[10]

---

[6] The history of the World Wide Web, henceforward referred to as the *Web*, began in the 1990's with Tim Berners-Lee and Robert Cailliau. After the American military "gave away" the Arpanet technology (in favour of a more decentralised communication system), first to US universities and then to the rest of the world, not only telephone lines were needed (which had existed for more than 100 years at that point) or personal computers (existing for almost 10 years), but above all a universally applicable code was required—which had not existed before 1990 (cf. Hörisch 2004: 387). It was in this year that Berners-Lee and Cailliau from CERN (*Conseil Européen pour la Recherche Nucléaire*, 'European Organisation for Nuclear Research') developed the HTML-Code as universal access language for the Internet (cf. Hörisch 2004: 387).

[7] It has to be taken into consideration that there exist great differences between social classes and geographical regions concerning Internet access.

[8] Cf. *Internet World Stats: Usage and Population Statistics*, online source 2.

[9] Although credit for inventing the electric telephone goes to Alexander Graham Bell, there were many others involved in the process (for details on the history of the telephone see, for example, Coe 1995).

[10] Cf. *Research and Markets*, online source 3.

This means that more than double the people own a mobile phone than have access to the Internet. Inevitably, this has to do with acquisition costs, as buying the equipment to go online via personal computer (PC) is more expensive than buying a mobile phone. This also suggests that mobile phone ownership disperses across more social classes than is the case with computer equipment. However, both types of media have become immensely popular during the last two decades.

Most interesting for this research is a service that started out as a "by-product" of mobile telephony: the sending of text messages via the short message service. Around ten years after the first mobile phone models were marketed in the early 1980's, it was discovered that it is also possible to send short text messages via the mobile phone. The first SMS text was sent in Great Britain in 1992, and as such it is a relatively young medium for written interaction (cf. Schlobinski et al. 2001: 4). Undoubtedly, the sending of SMS texts ceased to be a by-product long ago: the worldwide total of text messages sent in 2001 was estimated to be around 360 billion (cf. Curwen 2002: 16), for the year 2004, this estimate went up to 750 billion and one can assume that for the year 2008, the volume of SMS messages will have increased considerably. This shows that textual interaction via the short message service has become an established and widespread form of communication since its advent in the early nineties.

After this brief detour into the history of the media, which provides the larger (historical) setting for this research project, it is apparent that the 1990's are a landmark with regard to the spread of electronic written correspondence, the proliferation of which has been welcomed by many and despised by some.[11] Although handwritten messages are still present in the 21st century, communication without the services provided by the Internet and mobile telephony are to most people a thing of the past.

1.2.1. What is CMC?

It seems important to point out early on that the term CMC subsumes many discourse types of which only a small selection (for reasons connected to comparability) are addressed in this study. Next to e-mail, Web Chat, and personal homepages there are various other online settings where people come together and converse with each other: Multi User Dungeons (virtual role-playing games via e-text and instant messaging, usually referred to

[11] Not everyone appreciates the e-mail revolution. For example Terry Waite used the opportunity of a launch of the new Collins Dictionary to complain about the effect e-mail language has on English: "This e-mail English, bashed out without capitals, paragraphs and any idea of composition, is . . . irritating, tiring to read, and often simply unreadable" (Waite, quoted in Shortis 2001: 81).

as MUDs), Newsgroups (users can engage in discussions on all sorts of topics by gaining access to a Newsgroup and make their contributions via e-mail, also referred to as *pull-principle*), Mailing-Lists (information on all sorts of topics are sent automatically to subscribed members, also referred to as *push-principle*), Internet Telephony (similar to traditional telephony, people talk with each other via microphones installed in or attached to the computer), Video Conferencing (where people can talk and see each other via webcam and microphone over great physical distances), Online-Radio, Online Newspapers and Web-TV—to name the most important ones. A good overview of the different types of languages and communications in the Internet is given in Crystal (2001), Döring (2003), and Schlobinski et al. (1998). However, for reasons of interest and comparability I have decided to concentrate on *personal* written communication, preferably of a one-to-one or one-to-many type of discourse. Yet MUDs, Newsgroups, and Mailing-Lists are varieties of many-to-many discourse, Internet Telephony and Video Conferencing are types of oral (in the case of Video Conferencing also visual) discourse, and Online Radio, Newspapers and TV are variants of public/professional discourse. This is the reason why this study focuses on e-mail, Web Chat, and personal homepages as three variants of personal communication on the Internet.

Before turning to preliminary remarks on text (types), context, contextuality, and contextual effects, I would like to avert potential misunderstandings with regard to terminology as used to refer to the "substance" investigated in this study: although there are slight differences in meaning and scope of reference between the terms *communication*, *correspondence*, and *discourse*, they will be used interchangeably to refer to written language of dialogical character, produced with an (implied) reader in mind. Should any of these terms be used to refer to spoken language (except for *correspondence*, which by definition refers to written text only) then this will be pointed out specifically.

1.3. Text (types), context, contextuality, and contextual effects: preliminary remarks and working definitions

Before providing a brief overview of the chapters, it is appropriate to first address some key concepts as used in this dissertation. The main focus of this research is placed on how changes in the communicative context have affected personal written communication over time. Thus, for the purposes of this study, the notions of "text (types)", "context" and "contextuality", as well as "contextual effects" are of particular importance in connection

with the theoretical background (see chapter 4), as well as the methodology of the empirical investigation (see chapter 5).

1.3.1. The notion of text (types)

The investigation of five different text corpora is aimed at giving insight into the contextuality of written communication and its changes over time. However, referring to the data as different text corpora, i.e. compilations of different types of texts, raises an important question: what is "text"? And, following from that: what are "text types"? It is beyond dispute that the notion of text is, on a general level, referring to the production of language in either spoken or written fashion. However, on closer inspection, it is somewhat unclear what exactly a text is, or rather, which features can be said to be "strictly textual" (such as the spoken or written word) and whether they can be separated from features that can be classified as non-textual (such as paralinguistic cues). For example, can the sending of an SMS message, also referred to as *texting*, that consists solely of an emoticon, for example a smiley :-), be referred to as an SMS *text*? It becomes even more complex in written online settings that combine text with pictures, symbols, and links to other documents. Although commonly referred to as "hypertext", it still does not come closer to a concrete definition of "text".

I thus decided to adopt Grunder's (2001: 86) definition of "text" that views all systems of signs and symbols as text, consisting of "alphanumeric characters, spoken language, music, still pictures or moving pictures to mention only a few examples." Hence, any written or graphic transmission of information qualifies as text. Hence, an SMS message that consists solely of an emoticon qualifies as an SMS text, as does a personal homepage that contains pictures and symbols. Still, to distinguish between the notions of "hypertext" and "text" is useful, as hypertext denotes a different text structure (non-linearly linked text segments), which can occur in both off- and online written settings (see hereto also 2.2.4. and 4.2.1.). In addition, the notion of "text" is also understood to differ from "message" in that it is seen as a closed system, whereas a message, of which text is a part, is seen as an open system with less concrete boundaries. I will come back to the concept of message, and possible ways of defining what a message comprises, later in this study (section 4.1.).

The notion of "text" has been defined as a system encompassing all sorts of symbols and graphic representations. Hence, in order to be able to differentiate between

different kinds of text, a classification into different text types is needed—but according to which criteria? The answer to this question appears to be somewhat unclear:

> The concept of text types is fairly recent addition to the instrumentarium of synchronic and historical linguistics. True, certain categories of texts have always been used in their everyday meaning, for instance by compilers of anthologies. However, even in our times, terms such as "letters", "hymn", "obituary" or "joke" have normally not been analysed with sufficient precision or with a view to comprehensiveness. In consequence we have never got close to understanding which or how many text types there are in a particular culture nor with what distinctive features they can be delimited from each other.
>
> (Görlach 2004: 102)

It seems as if the notion of "text types" suggests quite fuzzy categories. However, Görlach (2004) proceeds to offer a definition of "text type" which has been adopted for the purposes of this study. According to Görlach (2004: 105), who includes both spoken and written text in his definition, "a text type is a specific linguistic pattern in which formal/structural characteristics have been conventionalised in a specific culture for certain well-defined and standardised uses of language so that a speaker/hearer or writer/reader can judge:"

> a) the correct use of linguistic features obligatory or expected in a specific text type (including the choice of appropriate language and register);
> b) the adequate use of the formula with regard to topic, situation, addressee, medium, register, etc.;
> c) the identification of intentionally or inadvertently mixed types, or their misuse;
> d) the designation of the text type (speakers not only know what features characterise a *telegram* but also know the name).

Görlach then continues to explain each of the points made above in more detail. A discussion thereof would, however, go beyond the scope of this study. Most important for this study is Görlach's (2004) general statement that a definition of "text type" is culture-dependent and characterised by both form and function of a particular text variant. Detailed descriptions of both form and function of the five different text types analysed in this study will be presented at a later stage.[12]

The notion of "text type" is, of course, also related to "genre" as "texts can be sorted into generic categories in almost any imaginable way" (Unger 2006: 4). Unger (cf. 2006: 5, emphasis original) also points out that individual genre labels are constructed

[12] A good overview of pragmatic aspects (i.e. the function) of each of the five text types can be found in chapter 5, Table 5.1., structural aspects (i.e. the form) of the five text types are discussed in chapter 6.

on the basis of quite diverse criteria, for example, "the label *detective story* refers to aspects of what the story is about", whereas "*dialogue* relates to the external properties of the talk exchange such as number of participants." However, Unger (cf. 2006: 5) also highlights the fact that such generic classifications have been given different labels in literature: the most commonly used terms being *genre*, *discourse type*, or, as discussed above, *text type*. Unger (2006: 5, emphasis original) continues to explain that "different terms are used in different disciplines" where "*genre* is mostly used in literary theory" and "*discourse type* and *text type* are more often used in linguistic studies of discourse." Most importantly, though, is Unger's (cf. 2006: 5) observation that the notion of "genre" on the one hand, and "discourse type" and "text type" on the other, are not crucially different concepts, "at least not one which could be easily captured in pre-theoretical terms." It is in this sense that these terms will be used interchangeably throughout this study.[13]

1.3.2. The notions of context, contextuality, and contextual effects

Having defined what the notions of "text" and "text types" encompass for the purposes of this study, I would now like to turn to the three concepts of "context", "contextuality", and "contextual effects" that are inextricably interlinked with each other, yet denote different linguistic phenomena. All three concepts are also complex in character; in particular the notion of "context" as its definition involves both inside- and outside-of-text features. With respect to "context", Asher (1994: 731, quoted in Fetzer 2004: 1) even claims that "context is one of those linguistic terms which is constantly used in all kinds of context but never explained." While the different approaches to context will be discussed in the theoretical part of this study (chapter 4), I will at this point introduce my working definitions for both "context" and "contextuality":

[13] Bergs (2004: 208) argues in his paper on the typology of letters that "the 'terminological maze' (Moessner 2001) of text types, genres, styles, and registers is still a matter of dispute and controversial debates." Bergs (2004: 208), however, does not intend to be "yet another voice in this sometimes very dissonant chorus" and makes no claims as to whether the distinctions introduced in his paper "are a matter of text type, genre, register, or style," but treats these notions as interrelated concepts and settles for the terminology "text type" for the purposes of his paper. On a more general note, see Görlach (2004: 23ff.) for an impressive list of hundreds of different English text types. For a diachronic perspective on English text types, see, for example, Diller and Görlach (2001).

> For the purposes of this study, the notion of communicative CONTEXT is defined as a dynamic construct that is being shaped in any given personal written communicative exchange by the following outside-of-text parameters: media competence and communication patterns of addressor and addressee(s), mutual and encyclopaedic knowledge of addressor and addressee(s), relationship between addressor and addressee(s), and the medium (medium-specific features). Further, the notion of communicative context is also shaped by inside-of-text parameters such as (non-)contextual effects.
>
> For the purposes of this study, the notion of CONTEXTUALITY is understood to be a by-product of the overall communicative context that is traceable in any given personal written communicative exchange. Contextuality is thus defined as the physical evidence of (non-)contextual features that can be traced in any given written communicative exchange. Features of textual contextuality include adopted communication channels (e.g. the written word, pictures, means of emphasis), inter- and intra-textuality of a written message, different types of deixis (personal, spatial, temporal), grammatical mood and syntactic structures, the evaluation of textual units in terms of (non-)contextual effects, length as well as organisation of the text.

Thus the notion of "context" concerns the general circumstances and environment that the different types of communication are embedded in, i.e. the communicative setting (see also Figure 4.2), whereas the term "contextuality" denotes specific context-related features that can be traced in the textual form of the different types of personal written communication (see also Table 5.1).

The concept of "contextual effects" was developed by Sperber and Wilson within the larger framework of Relevance Theory (RT), which was first formulated in 1986. According to Sperber and Wilson (cf. 2004: 250), the central claim of RT is that the expectations of relevance raised by an utterance are precise enough, and predictable enough, to guide the hearer towards the speaker's meaning. If adapted to written correspondence, this translates to the parallel-scenario of raising expectations of relevance by a written contribution that are precise enough, and predictable enough, to guide the reader towards the author's intended meaning. Whether applied to spoken or written language, contextual effects are highly dependent on the communicative context and, as mentioned above in the working definition for "context", are important inside-of-text parameters that shape the dynamic context of any ongoing communication. Sperber and Wilson (1986, 1995) defined four different types of contextual effects as processed by the hearer (or reader) in the comprehension of new information, to which an additional fifth one was added for the purposes of this study. There is no need to go into too much detail at this point, since the main tenets of RT, and the concept of contextual effects in particular, are treated more elaborately in chapters 4 (*Theoretical background*), 5 (*Methodology*), and

12 (*Contextual effects in personal written communication*). What should be kept in mind, however, is that Sperber and Wilson's theoretical concept of contextual effects has, to my knowledge, not yet been empirically applied to naturally produced language data. This means that the empirical approach of adapting the theoretical concept of contextual effects, and applying it to natural language data, is rather experimental (the nature of these adaptations will be discussed in more detail in 12.2.1.). However, this also means that the concept as proposed by Sperber and Wilson has been adapted to the point where it could be said to be estranged from the larger context of RT.

1.4. Structure of the study and overview of chapters

This dissertation is structured into four main parts: (I) PRELIMINARIES, (II) THEORETICAL SCOPE AND METHODOLOGY, (III) DISCUSSION, and (IV) CONCLUSION. The following brief overview of the chapters will illustrate the structure and contents of this study.

The first part (PRELIMINARIES) comprises three chapters (1-3) of which chapter 1 (*Introduction*) discusses the aim and scope of this study, illustrates the setting of this study and provides and defines key terminology as used in this study. Chapter 2 (*Previous work*) introduces selected previous work published in the field of both electronic as well as non-electronic correspondence. The last chapter of Part I, chapter 3 (*Data collection and composition of corpora*), gives an overview of the process of data collection and designing the online survey, as well as providing details on the composition of the different text corpora compiled for this study.

The second part (THEORETICAL SCOPE AND METHODOLOGY) consist of chapters 4 and 5. Chapter 4 (*Theoretical background*) introduces the various theoretical works that have been consulted. The chapter is divided into six sections that each deal with different contextual features in- and outside of the written text. Chapter 5 (*Methodology*) focuses on the analytical approach to the empirical investigation of the natural language corpora. It also introduces in detail the main research categories, and most importantly for the discussion in Part III, the analytical framework (code system) according to which the empirical investigation was conducted.

The third part (DISCUSSION) can be seen as the centrepiece of this dissertation: after having systematically applied the analytical framework to different types of empirical data, the results that this investigation generated are discussed in seven chapters (6-12). On a general note, all chapters in Part III feature manifold examples from each of the five natural language corpora. They are thought to both illustrate and enrich the discussion of

the findings. Chapter 6 (*Written communication: media-related features and constraints*) investigates the structure of the different text types and looks into uses of the communication channels that are available in the different media. Chapter 7 (*Personal reference*) deals with forms of address, whereas chapter 8 (*Textual reference: text as networks*) looks into how far personal written messages can be seen as isolated text entities or textual networks. How text is contextualised in terms of space, is analysed in chapter 9 (*Spatiality in text*), and chapter 10 (*Text is time-bound*) focuses on how personal written communication is embedded in a timeframe. The focus of chapter 11 (*Grammatical mood and syntactic structures as indicators for authorial intention*) shifts to a more grammatical point of view, looking at personal written correspondence in terms of indicative vs. hypothetical mood and comparing the distribution and function of different syntactic structures across the five text corpora. Chapter 12 (*Contextual effects in personal written communication*) is of a more experimental approach. It concludes Part III by investigating what sort of contextual effects are prevalent in the processing of the different types of correspondence.

The fourth and final part of this dissertation (CONCLUSION) is comprised of chapter 13 (*Conclusion and outlook*), which provides a concluding summary that re-addresses the main findings of this study (13.1.) and presents a critical evaluation of the analytical framework (13.2.). The chapter closes with reflections on what the future holds for personal written communication (13.3.).

2. Previous work

This study has a twofold focus in that it introduces an analytical framework of quite complex theoretical background, as well as an empirical investigation into five different types of personal written correspondence based on that analytical framework. While this chapter will concentrate on previous work in the fields of the personal written correspondence, the various theoretical works will be addressed in chapter 4 (*Theoretical background*).

There have been many contributions in each of the fields of personal communication analysed in this study. This chapter attempts to summarise the key tenets of academic research in each area and give an overview of different approaches. With regard to previous work in the field of the handwritten letter, the focus will be restricted to the Early English letter from the Early English Period (in particular the 17th century, from which the letter sample for this study originates). On a more general note, the overview given in this chapter is focused on works that bear a direct connection with the issues treated in this study. However, a substantial amount of previous work has been published in the area of personal written communication. Above all CMC has received a lot of scholarly attention since its advent at the beginning of 1990. Due to reasons of scope, the findings of the individual studies will be presented in a brief and summarised fashion.

2.1. Non-electronic communication

The handwritten letter is the oldest form of personal written communication and next to communicating content, letters may also have an emotional value:

> Beyond the words in their texts, books and letters serve as physical embodiments of experience and repositories of memory . . . . Letters convey richly nuanced messages in the handwriting of the writer, the type of writing instrument, the size, quality, colour, shape and texture of the paper used, and so on.
>
> (Danet 1997: OD)[14]

[14] The following differentiation will be made when quoting from or referring to works published online: documents in PDF (Portable Document Format) contain page numbers, which will be indicated when quoting a text passage. Documents in HTML (Hyper Text Mark-up Language), however, do not contain page numbers and when quoting a text passage from a HTML source, this will be indicated by OD (meaning 'online document without page numbers'). On a general note, all online references were re-checked on 28.10.2008 (shortly prior to publication of this dissertation) with respect to their functionality and, in case of changes in their URL (Uniform Resource Location), updated.

The above quote captures what a letter may convey beyond words and how it can be symbolised rather than merely be viewed as a piece of paper. The most telling case for the importance of letter-as-object is, of course, the love letter, with millions of people preserving old piles of handwritten love letters for decades (cf. Danet 1997: OD). Although this study focuses on the textual side of letters (the words), it seems important to be aware of such paralinguistic aspects of letters, since they have considerably contributed towards the letter persisting in being a means of communication to the present day.

2.1.1. Early English correspondence (EEC)

Being one of the oldest forms of personal written communication, academic research on the idiosyncrasies of epistolary correspondence is a long established tradition. The decision to look at a sample of Early English correspondence from the 17th century (as opposed to letters from a later century) was based on its origin from a time when the letter was the only means to personally communicate in written form. The handwritten letter thus serves as a prime contrast in terms of production compared to electronically produced text types. The letter sample originates from the Early Modern English Period, covering the time span of ca. 1500 – ca. 1800 (cf. Fennell 2001: 1).[15]

A lot of empirical work on EEC, this study being no exception (see section 3.2.), is based on the Corpus of Early English Correspondence (CEEC)[16], compiled by the Sociolinguistics and Language History project team at the Department of English, University of Helsinki.[17] The CEEC letter collection covers the time period 1420 – 1681 and contains 6039 letters of about 2.7 million words (cf. Nevalainen 1996: 3). The main aim of the project, initiated in 1993, was to test the applicability of modern sociolinguistic methods to historical language data. Terttu Nevalainen and Helena Raumolin-Brunberg, two researchers involved with both the collection and the analysis of the CEEC, contributed seminal works (see for example, Nevalainen & Raumolin-Brunberg 1996, 2003) in the field of historical sociolinguistics based on the CEEC. Nevalainen and Raumolin-Brunberg (1996, 2003) place their main focus on dialectal variation and change within the CEEC, in

[15] It cannot be the purpose of this study to give an overview of the history of English, but see Barber (1997) and Fennel (2001), who approaches the issue from a socio-linguistic point of view, for insightful discussion on the origins of English and how it evolved into the Modern English of today.

[16] For more detailed information on both informants and the material they provided for the CEEC see Nevalainen and Raumolin-Brunberg (2003: 43ff.).

[17] Cf. *Research Unit for Variation, Contacts, and Change in English*, online source 4.

consideration of how social factors such as the background of informants (for example, family ties or social status) as well as gender are variables in the diffusion of linguistic changes in the periods investigated (cf. Nevalainen & Raumolin-Brunberg 2003: 110).

Although my work is less concerned with dialectal variation and change, and more concerned with the contextuality of text, the work by Nevalainen and Raumolin-Brunberg gave me an insight into the history of sociolinguistics, as well as a corpus to work with. This is a very relevant point as the aim to empirically investigate personal correspondence involves the collection of natural language data. Data collection of personal, and often intimate, correspondence involves informants that are willing to part with something they would otherwise classify as private. This aspect complicates data collection in the field of personal correspondence considerably, and I was thus grateful to be able to use an already existing letter corpus for the purposes of this study (see also chapter 3 on data collection and corpus design).

With regard to contextual issues as treated in this study, it emerged that the *Journal of Historical Pragmatics* was a useful source concerning Early English letter writing and its pragmatic aspects, in particular the special issue *Letter Writing* (2004, 5:2). Beginning with an introduction by Nevalainen (2004) on the history of letter writing, the special issue then looks into contextual issues such as Wood's (2004) reflections on text in context, and Tanskanen (2004), who investigates the intertextual networks in the correspondence of a female author. In the same issue and also insightful are Bergs's (2004: 207) considerations on the letter and "its subdivisions into smaller groups of texts (i.e. subtypes such as 'requests', 'orders', or 'reports') on the basis of socio-psychological and pragmatic dimensions" as well as Nevala's (2004) paper on address formulae in 17th and 18th-century letters. In line with the subject matter of personal address is Tieken-Boon van Ostade's (2006) paper *"'Disrespectful and too familiar?'* Abbreviations as an Index of Politeness in 18th-Century Letters" that sheds light onto the dynamics behind abbreviated names in greetings and farewells, and thus addresses a particular type of personal deictic reference in Early English correspondence.

Textual reference, another type of deixis, is particularly interesting in the view of the fact that personal written communication does in most cases not consist of isolated messages. They are usually embedded in larger discourse structures in the sense that they are preceded and/or followed by other messages. This is not only interesting from a discourse structure point of view, but also in particular from a contextual perspective: in how far do authors contextualise their writings by means of reference to other texts? Not

much material was available dealing with this subject matter in older types of correspondence. However, Claridge's (2001) paper on discourse deixis in Early Modern English texts (based on the *Lampeter Corpus of Early Modern English Tracts*)[18] was helpful as a point of reference for the analysis of this study into the textual deixis of EEC.

Some typical examples of personal correspondence written in English between 1500 and 1700, the Early Modern English period, can be found in Cusack (1998), who investigates the English language of the 16th and 17th centuries. Addressing different features of EEC in her reader, illustrated with a collection of texts from different decades within the 200 years covered, she provides an overview of the language and letter writing strategies "of the real-life people of the time" (Cusack 1998: vi). Questions dealing with issues of context and the linguistic construction of epistolary worlds, as well as the pragmatic forces of epistolary communication in general, are addressed in Fitzmaurice's (2002), *The Familiar Letter In Early Modern English: A Pragmatic Approach*. Fitzmaurice (2002: 1) points out early on in her book that the familiar letter is "a pragmatic act that is embodied in a text that responds to a previous text, whether spoken or written, and at the same time anticipates new texts." It is in this sense that a letter can be seen as a link in an inter-textual chain.

Having briefly introduced the works consulted to learn more about other researchers' approaches to EEC, I will now turn to previous work in the field of modern electronic communication. The next section is divided into four parts that each provide an overview of previous work for each of the four text types of modern written communication under investigation.

2.2. Electronic communication

The possibilities for personal written communication in the 21st century are ample: from analogue to digital, various media can be used to establish personal contact. New technologies and changing communicative patterns have considerably impacted the nature of interpersonal exchange in written correspondence. Interestingly, Hiltz and Turoff's (cf. 1993: 509) observation that for most forms of communication, societies have had decades or even centuries to develop cultural guidelines on appropriate use, whereas for CMC there has not been time to develop an adequate ethical and legal framework to emerge to regulate its use, still holds—even though it was made over a decade ago. The difficulty in

[18] The *Lampeter Corpus of Modern English Tracts* contains 120 texts published between 1640 and 1740, 12 for each decade, amounting to around 1.1 million words (cf. Claridge 2001: 56).

regulating the use of CMC, and the Internet in general, is definitely connected to the fast growing number of (anonymous) users that gain access to the Internet every day. In any event, it is undisputed that the "lack of outside control" over the material published on the Internet makes this material immensely interesting for the producer in terms of creativity as well as for the researcher that aims at investigating natural language use.

This study concentrates on electronic discourse via the short message service (SMS discourse), e-mail, Web Chat, and personal homepages which are set in contrast with the personal handwritten letter of the 17th century. Apart from SMS texts, which can be sent from a computer but are mainly sent via the mobile phone, all types of electronic discourse relevant to this investigation are computer-mediated. And although the computer is often regarded as being one medium (or machine), its multitude of features allow discourses of a very diverse nature. This is probably one of the reasons why CMC in general, and e-mail discourse in particular, have received a lot of academic attention from different disciplines including linguistics, communication studies, sociology, psychology, information science and the newly formed departments of Cyber- or Internet studies (cf. Preece et al. 2003: 8). One particularly useful source to find out more about the nature of CMC is the *Journal of Computer-mediated Communication* (cf. online source 5), an online platform that has published a vast number of academic papers looking into CMC from different perspectives.

2.2.1. SMS discourse

Communication via SMS texts, a service based on GSM technology, permits an interpersonal exchange of electronic text messages with a maximum number of 160 alphanumeric characters (at the time of corpus collection, see 3.3.1.) via the mobile phone or the Internet. In contrast to CMC, SMS discourse via the mobile phone is faced with software-limitations that decrease the possibilities of text-configuration considerably. This is important in terms of contextuality insofar that users have limited options on a visual basis—words are the main carrier of the message. There is no possibility, for example, to highlight something deemed important with colour or bold font.

Although millions of people use the short message service to communicate with each other every day, scientific research on this subject is still astonishingly scarce. In fact, only few studies place their main focus on SMS discourse. Also, most of the existing research deals with SMS texts composed in German (for example, Androutsopoulos & Schmidt 2001, Höflich 2001, Schlobinski et al. 2001, Döring 2002a/b, Schmidt & Androutsopoulos 2004), Norwegian (Ling 2005), or Finnish (Kasesniemi & Rautiainen

2003), rather than SMS discourse performed in English. Furthermore, linguistic investigation on German SMS discourse generally places its emphasis on the medium's impact on language use, above all the small number of characters per text entity, whereas the dialogic structures and interactional features of SMS discourse are neglected (cf. Schmidt & Androutsopoulos 2004: 52).

As has been pointed out above, there is little previous work dealing with English SMS texts available. To my knowledge, the only empirical investigation into English SMS was carried out by Mori and Lycos (2000). The main focus of their study, however, is how the British public use SMS texts in flirtatious circumstances. There is no investigation into the textual features, structure, and style of the text messages in their paper. Another publication dealing with English text messages is by Bryant et al. (2006) who investigate how, amongst other types of communication, text messaging via the short message service is used by adolescents in social networking. The study's empirical investigation is based on questionnaire data collected from 7th grade college students at a middle school in the United States (cf. Bryant 2006: OD), coming to the conclusion that while the teenagers in their study used SMS discourse to maintain social network ties, they did not use it to create new ones. However, Bryant et al. (2006) also highlight the fact that there was a low adoption rate for mobile telephony in their sample, meaning that many of the students interviewed did not own a mobile phone, and that further research "among older adolescents who are more likely to have their own mobile phones, should provide better data in these areas" (Bryant et al. 2006: OD).

Another research branch focuses on the phenomenon of mobile telephony as such, and looks into its effects on social and individual life (as investigated by Plant 2000, Fox 2001, Brown et al. 2002, Geser 2004). For example, Plant's (2000) report is the result of a global enquiry[19] into the social impact of the mobile phone. Part of this investigation was to interview "a wide range of individuals about their use and perception of the mobile phone, their attitudes to other mobile users, and their sense of the mobile's social and cultural effects" (Plant 2000: 24). Since mobile phone calls can come at any time, at any place, and in the company of any number of other people, Plant (2000: 30) argues that "the etiquette of handling mobile phone interventions has become a matter of some debate" and

[19] Plant's (2000) primary research was conducted in eight major locations: Tokyo (Japan), Beijing and Hong Kong (China), Bangkok (Thailand), Peshawar (Pakistan), Dubai (United Arab Emirates), London and Birmingham (GB), as well as Chicago (USA). Plant (2000: 24) points out that "while these cities are by no means representative of their regions, or of the wider world, the international scope of this research has made it possible to identify significant ways in which local economic, technological, political and cultural conditions shape the use and perception of the mobile."

that "in many parts of the world, the ability to handle them—on the part of the recipient and all those within earshot—has become an important social skill." Plant (2000: 37) also states that both the nature of the social group one is part of is important to mobile phone behaviour, but that the location, in which the call is taken or made, is equally significant. Generally, the mobile phone is less likely to be seen or heard in more formal contexts than in less formal settings.

However, the amount of academic attention given to SMS discourse is insufficient, and this study aims at contributing insight into the idiosyncrasies of SMS discourse from a linguistic perspective.

2.2.2. E-mail

Sometime in 1971, the American computer engineer Ray Tomlinson managed to send the first e-mail message (cf. Preece et al. 2003: 2). It took another ten years before e-mail software hit the market and there has been no stopping it ever since. Today, more than thirty years after the first e-mail message was sent, people living in technologically developed countries can hardly imagine a life without computers and e-mail anymore, both from a professional as well as recreational point of view. Although e-mails are, technically speaking, unlimited with regard to text-length, they are generally seen as a tool for immediate communication and thus tend to be on the short side. CMC was initially designed for conveying short, goal-directed information which resulted in messages of a few lines. Only in the 1990's, as the handling of e-mail became more user-friendly and of a more social nature, did message-length cease to be predictable (cf. Baron 2000: 241).

Being one of the most popular means of correspondence in the field of CMC, e-mail discourse has received a lot of academic attention since it became available to the wider public in the early 1990's. According to Crystal (cf. 2001: 128), communication via e-mail portrays a wide range of stylistic expressiveness, from formal to informal. This is confirmed by Gains' (1998) analysis into the textual features of commercial vs. academic e-mail messages. Gains (1998) comes to the conclusion that commercial e-mails in general follow the established conventions for standard written English, whereas the data from academic sources indicates that some users view the medium as a pseudo-conversational form of communication. In any case, I agree with Baron (cf. 2000: 248) who states that most of the research to date on electronic communication has looked at the one-to-many conversation, rather than the one-to-one conversation that characterises most e-mail correspondence. She also claims that "large-scale studies of one-to-one electronic dialogue

are more challenging" (Baron 2000: 148), as many people are not willing to part with the content of their inbox and/or sent items folder for academic purposes. This is one of the main reasons why less is known about private e-mail exchanges between individuals than about business e-mail correspondence.

Personal e-mail correspondence is also interesting from a social perspective. For example, how the medium is used to establish and maintain relationships. Suler (2003: OD) states in his paper "E-Mail Communication and Relationships" that "E-mail may be the most important, most unique method for communicating and developing friendships since the telephone." According to Suler (cf. 2003: OD), avid e-mail users see the computer as a major feature of their interpersonal and/or professional life, including dyad relationships and group membership. In addition, the quality of the relationship between e-mail correspondents rests on their writing skills: "the better people can express themselves through writing, the more the relationship can develop and deepen" (Suler 2003: OD). Next to these far reaching effects of e-mail on social relationships, Suler (2003: OD) also looks at the individual sections of a typical e-mail message (the sender's e-mail username, subject line, greeting, body, sign-off line/name, and the signature block of an e-mail message) from a social point of view. Other papers that deal with e-mail and its interrelations and effects on social network ties, include, amongst others, Wellman (1997), Riva and Galimberti (1998), and, more recently, Bergs (2006).

Early research on CMC and particularly e-mail discourse (in the period between 1975 and 1990) had primarily focused on its instrumental aspects (cf. Danet 1997: OD), for example, how it affected communication among staff in teleconferencing and occupational networking (for example, Short et al. 1976, Garton et al. 1997). One of the seminal early works is titled *The Network Nation: Human Communication via Computer*, written by Hiltz and Turoff and published in 1993. Hiltz and Turroff (1993) look into the nature, potential applications and impacts of computerised conferencing, and they also make projections about the future use of teleconferencing technology. Essentially, Hiltz and Turoff (1993: 486) argue that their systematic investigation of computerised conferencing, used at the workplace or for personal purposes, has resulted in "the recognition that the technology of CMC has provided an opportunity to improve our fundamental understandings of human communication." Hiltz and Turoff (cf. 1993: 400) also raise awareness that computerised conferencing is connected to questions of policy, law, and regulation, and that future application of computerised conferencing "will depend critically on the resolution of the information rights issues" (Hiltz & Turoff 1993: 414). Looking at

these last arguments more than a decade later, it is obvious that Hiltz and Turoff were accurate in their predictions as there is still no legislation operative that regulates legal matters connected to the use of e-mail and, in particular, the Internet (see hereto also 3.1.).

The focus then shifted towards investigating whether CMC is a distinctive variety (for example, December 1993, Haase et al. 1997, Gains 1998, Runkehl et al. 1998, Rheingold 2000 [1993]), coming to the conclusion that generally, there are different varieties of language used for different services that the Internet offers. For example, the varieties used in personal compared to commercial e-mail (among others, Danet 1997 and Gains 1998). More recently, the investigation of features of orality in CMC has come into focus (for example, Wenz 1998, Crystal 2001, Danet 2002, Jacobson 2002, Dürscheid 2002a, Döring 2003), as well as how authors of e-mail compensate for reduced cue-bandwidth in their writing (for example, Bertacco & Deponte 2005: OD). One observation shared by all researchers is that personal e-mail typically contains considerably more features of orality than e-mail used for work purposes. Another area of interest is in how far the advent of computer technologies challenges our attitude towards the spoken and the written, electronic or printed, word (see Crystal 2004, Baron 2005). It is this branch of study in particular, that aims at shedding light onto why, amongst other online text types, personal e-mail is believed to be reminiscent of spoken conversation.

2.2.3. Web Chat[20]

> In an exchange whose primary purpose is not to facilitate the serious exchange of information but to be polite, entertaining, discreet or diplomatic, the participants may behave as though they share common, contextual meanings even though they know they do not.
>
> (Fitzmaurice 2002: 179)

Before the invention of the Internet, group communication was restricted to face-to-face contact, or in other words, required the physical presence of the participants. With the inception of the Internet, this situation changed drastically. As we have seen above, communication via e-mail or short message discourse can be used to bridge large distances, but they are both mainly one-to-one types of personal correspondence as well as

[20] From a technical point of view, Chat performed online is a form of computer-mediated communication that is based on the client/server principle. There are different formats of electronic Chat available, such as Internet Relay Chat (IRC) or Web Chat. While IRC requires the download and installation of a program (mIRC software), Web Chat is easily accessible via Microsoft Internet-Explorer (cf. Schlobinski et al. 1998: 84, Storrer 2001: 441). I decided to collect and analyse Web Chat for reasons of accessibility.

being asynchronous in nature. Chat differs from e-mail and SMS discourse in that it offers the option of near-*synchronous* group conversation (one-to-many, many-to-many) and private conversation (one-to-one) as participants need to be online at the same time.

Early academic research into Chat conversations, of which the vast majority takes place anonymously, has mainly focused on identity issues as well as its impact on social network ties (for example, Bechar 1995, Bays 1998, Paolillo 1999, Rheingold 2000, Herring 2001, Döring 2003), with the general conclusion that participants may become confused about the idea of being able to take on any "online-identity" they choose. Yet participants sticking to the same nickname[21] and meeting up regularly in the same Chat room are very well capable of establishing strong social network ties, even though they do not know more of each other than participants are willing to reveal (see, for example, Goutsos 2005). This is usually the case in Chat rooms (or discussion forums) with specific discussion topics, where participants come together to support each other or discuss certain themes. However, the majority of Chatters chat in several rooms, sometimes simultaneously, and also tend to change their nicknames according to the communicative mood they are in, and the name/topic of the Chat room (cf. Dittmann 2001: 47ff.).

More recently, there have been analyses into the idiosyncratic uses of language in Chat with regard to topic organisation as well as discourse structure in general (for example, Hentschel 1998, Baron 2000, Crystal 2001, Dittmann 2001, Storrer 2001). Storrer (cf. 2001: 452ff.) observes two distinctive features in connection with the discourse structure of Web Chat: first, text production is hidden and second, the sequencing of contributions follows what Storrer refers to as the "first come first serve"-principle: the typed comments appear on screen as uploaded by the server. According to Storrer (cf. 2001: 453), both the hidden text production as well as the "first come first serve"-principle make it difficult to follow conventional turn-taking rules in Web Chat; partly because one cannot see who is currently typing a contribution, and partly because the sequencing of the contributions is affected by typing speed and the capacity of the server. With regard to features of orality, Storrer (cf. 2001: 462) points out that Chat, albeit typed, resembles spoken discourse in that it is situational, interactive, and simultaneous. Furthermore, as the individual turns are bound to the ongoing conversation and directly refer to each other, they follow the typical patterns of spoken discourse (cf. Storrer 2001: 462, see also Kilian 2001).

[21] It is possible in certain Chat rooms to register a nickname. The use of a particular "nick" (as referred to in Chat jargon) is then restricted to one user.

Androutsopoulos and Ziegler (2003) analyse language variation on the Internet, in particular regionalisms in Web Chat. They (2003: 251) argue from a pragmatic point of view, claiming that "regional features in chat interaction can work as contextualisation cues of various sorts, e.g. indicating shifts in the structure of modality of interaction or contributing towards the construction of social stereotypes." According to Androutsopoulos and Ziegler (2003: 251), participants in Chat rooms draw on "regional speech" (dialect) in particular contexts of Chat interaction. For example, the comparison of four different city-related German Chat rooms displayed "differences in the distribution of regional features, which map the north-south divide of German dialects" (Androutsopoulos & Ziegler 2003: 251). On a general level, Androutsopoulos and Ziegler (cf. 2003: 275) argue that the specific use of regional expressions is seen as a social symbol by Chatters.

Another area of academic interest concerns the circumstance that anonymous Chat fosters the communicative interaction between strangers which inevitably raises questions on how Chatters contextualise their writings on the level of personal reference. Wirth (cf. 2005: 75) observes that Chatters show tendencies to adapt their personal references depending on the nicknames of their interlocutors. For example, they tend to address a Chatter with a nickname that contains the term *girl* (as in *Cybergirl*) as female—even though they are aware that these pseudonyms may have little to do with the "real person" behind the nickname. This is in fact one of the essential aspects of Chat discourse: reality is not of main interest but more how Chatters portray themselves in the virtual world.

Dittmann (cf. 2001: 51) looks at Chat room discourse from yet another angle and draws attention to "freedom of speech", and refers to the circumstance that participants may optionally contribute to the conversation at any time. However, one needs to bear in mind that if a contribution in a Chat room (a so-called "turn") remains unattended by the other participants, it might be classified as "freely spoken" but at the same time also as "unheard" (in the sense of 'not attended to'). In the terminology of RT, this would then be classified as a contribution of new information that has seemingly not produced a contextual effect large enough for any of the other participants to follow up on it. However, concerning discourse strategies as employed by Chatters, Baron (cf. 2000: 234) points out an interesting aspect in connection with the non-physicalness of Chat room interaction: because no one monitors whether the "disclosures" in a Chat conversation are accurate or not, authors have far more control in managing their side of the discursive exchange, as opposed to more physically revelatory circumstances such as face-to-face

encounters. This being precisely what this study is drawing on: contextuality that is consciously constructed by the author.

2.2.4. Personal homepage and the Internet

The homepage is the entrance page to a website[22] that is published and maintained by an individual person or, in some cases, a family or an informal, small group (cf. Döring 2002c). The ownership status of personal homepages can usually be determined from page titles or headings, allowing those websites maintained by organisations, institutions or formal groups to be distinguished from personal homepages. Of interest for this study are personal homepages created by individuals (as opposed to groups), and it is assumed that a particular personal homepage is the creation by the individual to whom it refers—the main purpose being that of asynchronous self-presentation with a considerably larger potential audience than in traditional modes (cf. Chandler & Roberts-Young 1998: OD). Notably, although aimed at self-presentation and directed at an unknown readership, personal homepages can be, and according to Döring (cf. 2002c: OD) often are, very diverse with regard to topic, text type, and language style. Also, "the volume of personal homepages varies between one document and hundreds of files" and "the number of external links ranges from zero to more than a thousand" (Döring 2002: OD) which makes a classification of homepages with regard to size and degree of hypertextuality very difficult.

Since the advent of the Internet, many of the services it offers have been subject to ample academic research. However, similar to SMS discourse, the field of personal homepages has been largely overlooked: "in many psychology, sociology, linguistics, or communication studies books about the Internet . . . personal homepages are, apart from a few exceptions, not discussed" (Döring 2002c: OD). Most of the academic discussion on personal homepages can be found on the Internet. Books on personal homepages, on the other hand, are few (for example, Boardman 2005). One reason for this may be that they are a highly fluid medium and constantly subject to change—by the time a book is published, probably none of the websites discussed therein have remained unchanged, if they still exist at all.

[22] The pages of a website are accessed from a common root URL (Uniform Resource Location), which is called the homepage of the website. The URLs of the pages organise them into a hierarchy, although the hyperlinks between them control how the reader perceives the overall structure and how the traffic flows between the different parts of the sites. Both Runkehl et al. (1998) and Döring (2003) give a good overview on the technical details of the Internet and the services it provides.

The majority of early as well as current academic research on personal homepages concentrates on identity issues, Internet presence in connection to gender issues being the most widely discussed aspect[23], and other psycho-sociological themes connected with self-presentation on the Internet (among others, Erickson 1996, Chandler 1998, Chandler & Roberts-Young 1998, Ishii 2000, Wellman 2001, Arnold & Miller 2003, Döring 2003, Ridings & Gefen 2004). For example, Arnold and Miller (2003) introduce a framework for understanding web page identity in their paper "Self in Web Home Pages: Gender, Identity and Power in Cyberspace." In attempting to provide such a framework, Arnold and Miller (cf. 2003: 79, emphasis original) argue that people take active decisions about the ways they organise and classify their and others' *actions* and are not puppets of predetermined character or stereotypical role behaviour. One of Arnold and Miller's (2003: 80) essential claims is the following:

> If we believe that the self is constructed out of the doing of things, then the new thing to be done, for instance, constructing a personal Website, will give the possibility of new aspects of the self. The opportunity to make a complex, multi-layered, but controlled presentation—the hypertext self—does raise new possibilities for how people can think about themselves, and get others to think about them.

This means, for example, that people make conscious decisions about what kind of self-image they want to publish on their personal homepages. Arnold and Miller (2003) collected a random sample of 35 homepages created by males and 35 authored by females and analysed the main features of the first page that loaded on each site. In connection with photographs, Arnold and Miller (2003: 81) identified four categories for the portrayal of online self-image by means of photographs:

1. straight: an image which purports to be a straightforward likeness;
2. joke: a distorted or caricatured or unrepresentative image: cartoon, baby photo, author just falling off bike, author caricatured as frog, etc.;
3. symbolic: an image which represents a human being, but not the actual person who posted the page. This if often a piece of clip art, like a cherub or a generic silhouette;
4. none: no images of humans.

[23] While looking at gender aspects themselves, Arnold and Miller (2003: 80) make an interesting claim concerning the focus on gender differences in the self portrayal on the Web: "Looking for differences between the genders has long been criticised as being sexist in itself. Why establish differences, unless it's for the sake of validating discrimination?" However, Arnold and Miller (2003) are careful and successful in not discriminating anyone in their paper. Also, see Wilkinson and Kitzinger (1995) for a discussion of feminism and discourse from a psychological perspective.

Arnold and Miller (cf. 2003: 81) observed that several (15 out of 35 pages for both groups) did not contain any images at all and that men's pages had more "straight images" than pages authored by females (10 compared to 6). However, the big difference was in the other two categories: "joke images" only featured on a total of 4 men's homepages, whereas "symbolic images" featured only on a total of 10 women's homepages (and were thus the most common form of image on women's pages). It seems as though females are more reluctant to publish "straight" pictures of themselves than males.

Another aspect of personal homepages (and websites in general) that has generated a wider interest is the notion of "hypertext". The possibility that text published on the Internet can be linked to other text online in a non-linear fashion has raised questions with regard to online discourse structure (among others, Mitra 1999, Karlsson 2002, Mejías 2004a/b, Askehave & Ellerup Nielsen 2005, Jakobs & Lehnen 2005, Stein 2006, Mancini et al. 2007). For example, Mejías (2004a: 6) identifies online discourse as "more than just a synthesis of its predecessors" in that it goes "beyond the elements borrowed from oral and written speech."[24] Mejías (2004a: 6, emphasis original) then goes on to identify two characteristics that make online discourse unique: "*hypertextuality* and *distributed discursivity*", where "hypertextuality refers to the ability of one online text to link to any other text, thus incorporating it into a discursive thread." "Distributed discursivity," on the other hand, "refers to the social aspects of online discourse that are implemented across time and space in unique, unprecedented ways by online technologies" (Mejías 2004a: 7). Mejías (2004a: 7) argues that online discourse, as found on personal homepages and other sites in the Internet, "unfolds in complex social patterns" and that it is "highly collaborative and communal." This in turn calls for new online discourse models and tools in order to be able to better understand the structures and functions of online discourse. In Mejías (2004b), such a new approach to online discourse is introduced and referred to as "distributed textual discourse". Essentially, distributed textual discourse (DTD) "can be thought of as a supplement to other online communication as well as a stand-alone software application" (Mejías 2004b: 2). In other words, Mejías (cf. 2004b: 2) is analysing how current technologies could be applied in new and better ways to support online discourse.

Next to idiosyncrasies in the structure of online discourse, the notion of hypertext located in the Internet has also raised questions connected to deixis (as discussed in

[24] See also the discussion on the notion of "cyberdiscursivity" in 4.2.1.

De Saint-Georges 1998, Krikorian et al. 2000, Loehr 2002, Herrring et al. 2003, Mancini et al. 2007). On a general note, the non-linearity of hypertext challenges the traditional deictic reference systems of personal, spatial, temporal, and textual deixis (see hereto also 4.3.1.). For example, De Saint-Georges (1998) analysed a sample of 38 homepages authored by students with respect to person, time, and place deixis. In terms of place deixis, De Saint-Georges (1998: 71) observes that "the use of spatial deictic expressions can be roughly divided between: (i) deictic expressions anchored to the text itself" and "(ii) deictic expressions anchored to the text pointing to other documents exophorically." With regard to person deixis, De Saint-Georges (1998: 73, emphasis original) notes that personal homepages can be seen as "'stories' in which (i) in most cases, the 1st person pronoun dominates the text with various degree of presence of the author," and that "(ii) the pages are often headed by a title which is the *name* of the individual to which the personal homepage belongs." Concerning time deixis, De Saint-Geroges (1993: 75) observes that the writers of the homepages analysed largely favoured present tense, often in the progressive form, over any other tense and that this, together with temporal adverbs and complex time adverbials, "tend to combine to create a sense of immediacy, of activity in progress."

Loehr (2002: 29), on the other hand, investigates some linguistic aspects of hypertext, "using as data Web pages from college students and commercial enterprises" and looks in particular into the syntactic and deictic forms of hypertext anchors (hyperlinks) and the quantitative relationship between hypertext and spatial deictics. Loehr (2002: 37) concludes that hyperlinks tend to be noun phrases, "with a lesser concentration of complete sentences," and "although one might intuit a relationship between place deixis and hypertext, none could be substantiated." However, Loehr (cf. 2002: 37) also points out that hyperlinks do contain a greater concentration of the spatial deictic *here*, owing to the popularity of the phrase *click here*. Interestingly, as to the types of deixis found inside hyperlinks, "the majority were person deictics referring to the speaker, probably due to the fact the pages studied were homepages, in which authors present themselves to their audience" (Loehr 2002: 37-38). It will be interesting to compare Loehr's (2002) findings with the results this empirical investigation into the use of deictics in homepages has generated.

Still, little is known about the idiosyncrasies of the language used on websites, even less so about therewith connected contextual issues. Yet, "the real publishing revolution has come from the fact that anyone can put up their own site without, initially, having it

vetted by anyone" (Boardman 2005: 37). This means that there is a vast amount of unrevised material on the Internet. At first glance, this is a very fortunate situation for linguists looking for genuine data in the field of personal communication. First, it can be assumed that people are willing to part with whatever intimate information they publish on the Internet, and second, this information will in most cases be composed by the person who runs the website, unrevised by any third parties. However, as will be discussed in more detail in the next chapter, the collection of material published on the Internet is constrained by issues of copyright comparable to print works.

3. Data collection and composition of corpora

This chapter on data collection and corpus design is prefaced by section 3.1. that outlines considerations on research ethics when working with empirical data, as well as constraints that arise in the course of academic data collection. Ethical issues in linguistic research have always been a matter of concern, but with the advent of the publicly accessible Internet, containing vast amounts of (in some cases anonymously) published data, the issue has become more delicate. Another debated subject matter in corpus linguistics in connection to data collection is the notion of representativeness of corpus design. This will also be addressed briefly in section 3.1. After attending to issues that need to be considered before data may be collected, the focus will then shift to the data that was collected for the purposes of this study. Sections 3.2. and 3.3., respectively, illustrate the proceedings of collecting the non-electronically and electronically produced text corpora. The last section (3.4.) of this chapter provides further information on the online survey, with a particular focus on the composition and type of questions as presented in the questionnaire.

3.1. Data collection: research ethics, copyright, and constraints

Herring (1996) points out two opposing proposals in her article on ethical considerations with regard to the use of online data that, on the one hand, is (often anonymously) published on the Internet for public access, yet on the other hand, raises questions of copyright. The first proposal comes from legal scholars, postulating that all messages posted via computer networks are published works and thus fall under the protection of the copyright law, which would mean giving full references for all quoted material (cf. Herring 1996: 153). But how is one to proceed if the material has been published anonymously? And on another note, what if an author gives consent to using the material, but does not want the source revealed for reasons of privacy? The other proposal concerns the circumstance that all published material on the Internet is potentially private and in order to protect the privacy of the authors, no identifying characteristics of the data should be reproduced in scholarly work, and one should paraphrase messages rather than quote verbatim (cf. Herring 1996: 154.). Not only do these two views contradict each other, but neither of them accommodates a linguistic approach to empirical data collection and analysis.

Another viewpoint is brought forward by Berry (2004: 53), who argues that "privacy is a misleading and confusing concept to apply to the Internet, and that the concept of non-alienation is more resourceful in addressing the many ethical issues surrounding Internet research," suggesting that "ethical considerations should use principles of an open source ethics approach to guide research into the online world." Nevertheless, Berry (2004: 66) comes to the conclusion that open source research "places the researcher in paradoxical situations regarding copyright and fair-use," and that it also raises questions about the use of texts (and other online artefacts) and the rights of the community to control them. The open source approach is further complicated by "digital right technologies" (Berry 2004: 66) and increased copyright protection that are being implemented by private companies for online digital works. Although this does not (yet) concern personal homepages, which are in general neither password protected nor do the pages state that the material is under copyright, it is assumed to become a matter of concern in the future.

However, in order to be able to investigate natural language phenomena, it is necessary to work with naturally produced language samples, i.e. data that was produced in conditions uninfluenced by any academic research purposes. Yet before using any material for academic purposes, the consent of the author(s) to do so is required. While many authors are willing to contribute their data for academic research, most of them prefer to remain anonymous. In the case of private correspondence, the matter is even more complicated. The content of the messages is of intimate nature, and authors often feel strongly about correspondence that contains matters close to their hearts. Intimate information anonymously published and publicly accessible on the Internet then makes the situation quite paradoxical. Since two types of correspondence analysed in this study are performed on the Internet (Web Chat and personal homepages), I was forced to make considerations addressing this paradox prior to data collection.

Information published on a website, with or without password restriction, is the intellectual property of the author, and whoever maintains the website (i.e. the webmaster) is, or can be held, responsible for its content. The authors and webmasters of personal homepages are in most cases the same person and the content (if not otherwise stated) can be assumed the "intellectual property" of the author. A similar scenario applies regarding the copyright of printed vs. online material if one considers that it is not the physical book that is owned, but that it is the thought process and its output that authors wish to protect:

> The mark looks so simple: a "c" enclosed within a circle [©]. Modern authors and readers all but take the mark for granted as part of the obligatory material found in the front of printed works. It signals, of course, who holds the copyright to the work that follows—who owns it (typically the author or publisher). But what is owned? Surely not the physical book itself . . . . The property denoted by the copyright is more abstract. It's the words—the sequence in which they appear and the original ideas they express.
>
> (Baron 2000: 58)

From this point of view, there is no difference concerning copyright[25] between printed and online material and thus the same rules apply. Hence, any material published on a website is the intellectual property of the author. I thus needed permission to analyse the contents of the homepages.[26] This meant contacting the homepage authors and asking for their consent, which in turn constrained my homepage sample to pages that contained either an e-mail address or a link to another page where messages could be deposited. I composed an e-mail explaining the subject matter and sent it to potential informants. It took several months before I had the target number of 60 consents in my inbox. My impression that material published on the Internet comes with a tag of "you may read my homepage, but it is mine" was confirmed because quite a few authors wrote back saying that they would rather I did not analyse their homepage for academic purposes. Although the potential informants were assured that their data would be made anonymous, all of the authors (except one) that consented to the analysis of their homepages pointed out specifically that they neither wanted the URL of their homepages published nor their identity (as portrayed online) revealed. In order to be able to refer to the individual homepages throughout this study, an alphabetical reference system was adopted (see Table 3.2 below).

I was faced with a similar situation with respect to the Chat room data: all of the six Chat room operators that I contacted consented to me using data collected in their Chat rooms for academic purposes, but two out of the six requested that I neither publish the URL of the Chat room nor the nicknames of their Chatters, and four out of six agreed with having their URL published but did not want the nicknames of their Chatters to be revealed. What is interesting (if not puzzling) about this is that nicknames are in most cases

[25] For a brief historical discussion on the notion of "copyright" see Haarmann (1994).

[26] Opinions on this subject matter may differ. There seems to be a continuum of which one pole argues that material published on the Internet becomes "public property" (comparable to newspapers), whereas the other end of the continuum sees the material as copyright protected. The circumstance that the execution of copyright laws varies from country to country (cf. Haarmann 1994: 898) further complicates the issue of published material on the "global Internet". To my knowledge, no generally accepted regulation has so far been established as to how to proceed with material published on the Internet with regard to copyright issues. It was thus decided to be most careful in this respect for the purposes of this study.

anonymous pseudonyms and thus already conceal the "real identity" of the Chatters. Nevertheless, the majority of Chat room operators felt strongly about not publishing them. Their wish for complete anonymity was, of course, honoured and I will be referring to the six Chat rooms as Room I-VI, respectively. Nicknames, where required, were replaced by fictitious ones. Speaking of anonymity, I would like to point out that all corpora were made anonymous prior to analysis and in consideration of publication. The only exception is the letter corpus from the 17th century which included all original names when I obtained it. As there is a time span of more than 300 years between the composition and the analysis of these letters, it was decided to work with the original data.

3.1.2. The notion of representativeness in linguistic corpus design

> The extent to which a corpus can ever be considered to represent a language in general is currently a matter of some contention. In practice, whether a finite sample of a language could ever "represent" the vast amount of a language produced in even a single day is always likely to be, in the final analysis, an act of faith.
>
> (Kennedy 1998: 21)

When is a sample big enough? And in what proportion should different types of informants be "allowed" to contribute to a text corpus that is aimed at representativeness? Working with empirical data involves addressing these questions at some point. The quote above captures quite aptly, and perhaps somewhat discouragingly, that linguistic corpus design faces difficulties in connection with representativeness. According to Francis (1982: 7)[27], a corpus is "a collection of texts assumed to be representative of a given language, dialect, or other subset of a language, to be used for linguistic analysis"—yet what exactly does "representative" mean? Is it aimed at the size of a corpus, or the types of informants (in consideration of age, gender, social status and so forth) that contribute data, or a mixture of both? Due to reasons of scope, it will not be possible to discuss the question of representativeness for empirical linguistic research in general, but I would like to add some considerations concerning this corpus-based study.

[27] This is, in fact, quite an old source. More notably still, this statement by Francis (1982: 7) was made when it was still early days with respect to computer data bases. Nowadays, the computer technology is more advanced and this has changed notions of representativeness with respect to sample size. More recent publications that deal with corpus linguistics (for example, Baker 2006, Mc Enery 2006, Hundt et al. 2007) include considerations in connection with electronic data collection (on- or offline) and storage, as well as the potentially large-scale analysis of electronic data with the help of specifically designed software. Nonetheless, Francis' statement provides a good starting point for a discussion of representativeness in linguistic corpus design.

Given the large number of English speakers (native and non-native) that produce correspondence on a personal level every day, empirical research into this area is faced with an abundance of potential data. There are two main problems connected to this "abundance of potential material": first, the volume of the data is overwhelming, and second, due to personal correspondence being *personal*, it is more difficult to actually collect that data in comparison to, say, public discourse as found in magazines and newspapers. Furthermore, since informants have to be informed before linguistic analysis can be carried out, the linguist is faced with the circumstance that informants contribute pre-selected data. In the case of personal correspondence, this means that informants are likely not to give away their most intimate pieces which in turn has an influence on the composition, and thus the representativeness, of the corpus. This is an aspect of corpus linguistics that most researchers are confronted with. Yet in consideration of the ethical issues outlined above, there is no other solution than to accept that a sample contains only the data informants were willing to part with. However, I do agree with Nevalainen and Raumolin-Brunberg (2003: 29) that "personal correspondence stands out as first-class primary data" for studies taking the social aspects of correspondence into account.

More problematic with regard to corpus compilation is Nevalainen and Raumolin-Brunberg's (2003: 29) claim that without identifying the individual letter writers with regard to their social backgrounds, personal histories and relationships to the recipients of the correspondence, "sociolinguistic analyses and interpretations could not be made." While I acknowledge that information on the social background of the writers and readers allows for a more detailed picture regarding interpersonal correspondence, the full particulars of the correspondents are, in my opinion, not necessary if the text itself is the main focus of analysis, even if this includes sociolinguistic aspects. All text types investigated in this study are of a personal nature, and from a sociolinguistic point of view, it is assumed that the correspondents either know each other personally (as is the case with the letters, SMS texts, and e-mails), or if there is a high probability that they do not know each other personally, then informants are assumed to be aware of the fact that they are engaging in personal correspondence (such as anonymous Web Chat and personal homepages). In the case of anonymous personal correspondence that takes place online, linguists have no other choice than to make sociolinguistic analyses and interpretations based on textual material of which they do not know the authors and (potential) readers.

The corpora collected for the empirical part of this research are subject to all of the implications outlined above. However, they do not claim to be "representative" of the five

types of personal written communication investigated in this study. As has been pointed out in the introduction, the main aim of this study is twofold in that it combines empirical research with testing a self-designed theoretical framework. It is in this sense that the samples do not attempt to represent the five different text types investigated in this study, but give insight into the way authors of personal written communication contextualise their writing. In order to meet academic requirements with regard to corpus design, the data collection was carried out according to predefined guidelines and in view of the ethical considerations outlined above.

3.2. Non-electronically produced text corpus

The Oxford Text Archive (cf. online source 6) offers, among other data, free access to the Corpus of Early English Correspondence Sampler (CEECS, 1999 version, available on CD-ROM). The CEECS was created from the larger Corpus of Early English Correspondence (CEEC), which has already been introduced in section 2.1. According to Nevalainen and Raumolin-Brunberg (1996: 39), one of the main requirements of the material collected for the Corpus of Early English Correspondence was that "the language used should represent private writing and relate closely to the spoken idiom" and thus qualifies the data from the CEECS for my type of research.

3.2.1. Early English correspondence corpus

The CEECS contains 23 letter collections (approximately half a million words) included in the CEEC that are no longer copyrighted and cover the years 1418 – 1680. The CEECS consists of 1147 letters by 194 informants and is divided into two parts[28]: $CEECS_1$, covering the 15th and 16th centuries, and $CEECS_2$ that consists of mainly 17th-century material. For reasons of scope I concentrated on samples from the $CEECS_2$ that covers the years 1580 – 1680 and consists of 13 letter collections (204'030 words).

Initially, I planned to focus on a particular time span (twenty to forty years) within the $CEECS_2$ for reasons of comparability. But wanting to limit the number of letters to three texts per person (in order to avoid bias) meant that the female authors were outnumbered

[28] Cf. *CEECS Manual*, online source 7.

by far within a particular time span of twenty to forty years.[29] Yet increasing the number of letters per person would have brought about the problematic issue of bias. It is for this reason that I settled for a sample from the CEECS$_2$ that covers a hundred years, making sure it contains no more than three texts per author from the same collection. In order to counterbalance the male vs. female discrepancy in letter-quantity, six letter collections consisting of letters composed exclusively by male authors were excluded. On the other hand, two letter collections comprising letters by female authors only were considered. All in all, the sample letter corpus for this study consists of 91 letters from five mixed and two female letter collections form the CEECS$_2$. Table 3.1 below gives an overview of the letter corpus (N_1).

Table 3.1: Overview of the sample taken from CEECS$_2$ (N_1 = 91 Early English letters).

LETTERS COMPOSED BY MALE AUTHORS		
Name of letter collection	**Time span of collection**	**N**
Original_3	1580 – 1665	13
Cornwallis	1613 – 1644	19
Harley	1625 – 1666	6
Charles	1634 – 1678	5
Basire	1651 – 1666	5
Total letters male		**48**
LETTERS COMPOSED BY FEMALE AUTHORS		
Name of letter collection	**Time span of collection**	**N**
Original_3	1580 – 1665	4
Royal_2	1612 – 1614	2
Cornwallis	1613 – 1644	22
Harley	1625 – 1666	5
Charles	1634 – 1678	1
Basire	1651 – 1666	3
Tixall	1656 – 1680	6
Total letters female		**43**
TOTAL LETTERS **TOTAL WORD COUNT**		**91** **31'077**

The letters in the CEECS$_2$ collections are assembled in chronological order, and aside from limiting the number of letters per author, the sample was randomly taken without regard to possible dialogue structures between the letters. As can be seen below, both the e-mail and

[29] Nevalainen and Raumolin-Brunberg (2003: 40) divided their data into twenty-year periods, which were taken to represent contemporaneous language, but found out during their investigation "that in many cases it was better to double this time in order to acquire reliable results on overall developments and correlations with external factors."

SMS corpora were collected according to the same pattern which ensures comparability. As has already been pointed out in the section on research ethics, the data from the CEECS was not made anonymous prior to investigation because of the long time span between the production of the letters and the publication this work.

3.3. Electronically produced text corpora

Contrary to the letters from the 17th century, all collected material from 21st century authors was made anonymous prior to corpus compilation. If informants decided to make the data anonymous before contributing it to one of my corpora, they were required to indicate whether the author(s) of the data were male or female, as well as indicating the age of the author(s). With respect to the authors of the personal homepages, it was possible to determine the gender of the persons as presented online, but not their age. Regarding Web Chat, neither gender nor age of the informants could be determined beyond doubt. Therefore, no gender-related comments will be made in connection to Web Chat. The online survey, on the other hand, while executed anonymously, contained two questions on the demographics of the informants in order to determine their gender and age.

3.3.1. SMS text corpus

Before introducing the SMS text corpus in more detail, it is important to highlight two important implications. First, when I collected the SMS messages during the years 2003 and 2004, limitations of mobile phone software meant that message length was restricted to 160 alphanumeric characters per text message. This is no longer the case with many of today's models, where the message length for SMS texts may be longer (or unlimited). Many of the modern mobile phones also offer Internet access to write e-mails. Further research into written communication via mobile phone will be needed to show whether or not users take advantage of these new features. However, I addressed this issue in one of the open questions of the questionnaire.

Hence, the SMS text corpus (N_2) used for this study consists of messages that were limited to 160 alphanumeric characters by the technological constraints at the time of their transmission. It was collected between February 2003 and May 2004 with the help of fellow students of the English Department at the University of Zurich, as well as friends and acquaintances of my age group (early twenties to early thirties). N_2 (18'426 words) consists of 340 SMS texts written by male authors aged 21-46 years, and 660 messages by female authors aged 20-41 years, with a similar distribution across native as well as non-

native English speakers.[30] Again, there are not more than three messages per author in the SMS text corpus.

3.3.2. E-mail corpus

Similar to the SMS text corpus, the personal e-mail corpus was collected with the help of fellow students and friends. Not only did they send me their own data, but also contributed to the corpus with e-mails they had received from other English speaking acquaintances (having informed them beforehand of the undertaking). The e-mail corpus (N_3) consists of 140 personal English e-mail messages (25'733 words) of which the oldest e-mail dates back to August 2002, and the most recent message is dated April 2006. Not more than three e-mail messages per person were considered in order not to bias the results of this investigation. N_3 is comprised of 77 e-mail messages authored by males aged 21-50 years, and 63 e-mail messages from female authors aged 23-80 years, including both native as well as non-native English speakers.

3.3.3. Web Chat corpus

The Web Chat data (N_4) was collected according to the following three requirements to ensure an uninfluenced and spontaneous use of language by the Chatters: (1) data must stem from Chat rooms without discussion topics, (2) data must stem from unmoderated Chat room discussions, (3) the data collector must remain quiet so as to not to influence the discussions in any way. N_4 consists of 30 Web Chat sessions (28'404 words), collected in six different Chat rooms (Rooms I-VI, 5 Chat sessions per room). The Chat sessions were recorded for 15 minutes and for each session the number of participants was noted, in order to be able to make deductions between number of Chatters and text volume. The Chat logs[31] of the individual sessions were collected between February and March 2006 (details on the individual Chat room sessions can be found in appendix 15.3.). Since nicknames of Chatters are in most cases of an anonymous nature, there was no possibility to determine the gender of Chatters. Therefore, when discussing examples from the Web Chat corpus, the participants will either be referred to as "Chatter(s)", or both genders will be given,

[30] Of the 1000 SMS texts, 400 were authored by English native speakers and 600 were composed by English non-native speakers.

[31] A service offered by most Chat rooms where the chronological dialogue (from the moment of login) is stored as a text file. However, once a Chatter logs out there is no possibility to retrieve that data at a later point.

such as "him/her" or "he/she", regardless whether or not the nickname contains (seemingly) gender-indicative terms such as *girl* or *boy*. The reason for this is that gender should not be assumed based on the nickname, as there are males opting for female nicknames and vice versa.

3.3.4. Personal homepage corpus

Döring (cf. 2002c) suggests three search methods to sample small or large numbers of personal homepages systematically: via Web directories, or with the help of Web rings or link lists. For reasons of heterogeneity, I decided to adopt the sampling via the Web directory *Yahoo* (cf. online source 8). First, this is a widely known portal and second, the homepages are not listed in terms of a similar milieu[32] or organisational features[33]. The corpus of personal homepages (N_5) was compiled over a three-month period (November 2006 – January 2007), accessing each site five times in order to record updates and changes. The first of these visits forms the basis for the textual analysis. N_5 consists of 60 personal homepages (16'030 words), of which 30 were written by male authors and 30 by female authors.

However, one has to bear in mind that a persona, as portrayed in the Internet, need not necessarily be congruent with the person who "paints the virtual picture". This may concern all information published on a personal homepage, or just extracts thereof, such as not being accurate about age, profession and so forth. However, for the current research, gender is only important from a corpus design point of view. Since all authors were contacted personally for approval previous to analysis, all of them confirmed their gender, as portrayed online, in their e-mails. This distinguishes the personal homepage corpus from the Web Chat corpus, where no information on gender could be determined. Still, there is no way of being 100% sure about gender without having met the informants of the personal homepages in person. Based on the e-mail exchanges it is assumed, however, that out of the 60 informants, 30 are in fact male and another 30 female.

The personal homepage sample was collected randomly based on the letters of the alphabet. Considering ethical aspects, an alphabetical reference system was adopted. Table

[32] As in Web rings, where some people who run personal homepages join together with respect to common interests. Web rings can be accessed over a Web ring catalogue (e.g. *WebRing* http://dir.webring.com/rw [28.10.2008]) and are an ideal search method to find homepages of members of specific groups such as Hindus, feminists, teachers, and many more (cf. Döring 2002c).

[33] As in Link lists, where references to personal homepages of members of different organisations are often compiled in thematic manner, for example, the homepages of companies (e.g. www.employees.org [28.10.2008]) or of university members (cf. Döring 2002c).

3.2 below illustrates the composition of the personal homepage corpus. The capitalised letters represent the initials of the last names (A-Z), the lowercase letters indicate the gender (m/f) of the informants, and the subscript numbers refer to the number of homepages per initial.

Table 3.2: Overview of personal homepage corpus (N_5 = 60 personal homepages).

INITIALS	PH COMPOSED BY MALES	PH COMPOSED BY FEMALES	TOTAL PER INITIAL
A	2 (Am_1, Am_2)	2 (Af_1, Af_2)	4
B	2 (Bm_1, Bm_2)	3 (Bf_1, Bf_2, Bf_3)	5
C	1 (Cm_1)	1 (Cf_1)	2
D	1 (Dm_1)	2 (Df_1, Df_2)	3
E	1 (Em_1)	1 (Ef_1)	2
F	2 (Fm_1, Fm_2)	2 (Ff_1, Ff_2)	4
G	2 (Gm_1, Gm_2)	2 (Gf_1, Gf_2)	4
H	1 (Hm_1)	1 (Hf_1)	2
I	1 (Im_1)	1 (If_1)	2
J	1 (Jm_1)	1 (Jf_1)	2
K	1 (Km_1)	1 (Kf_1)	2
L	1 (Lm_1)	1 (Lf_1)	2
M	1 (Mm_1)	2 (Mf_1, Mf_2)	3
N	1 (Nm_1)	1 (Nf_1)	2
O	1 (Om_1)	1 (Of_1)	2
P	1 (Pm_1)	x	1
Q	1 (Qm_1)	1 (Qf_1)	2
R	1 (Rm_1)	1 (Rf_1)	2
S	2 (Sm_1, Sm_2)	1 (Sf_1)	3
T	1 (Tm_1)	1 (Tf_1)	2
U	1 (Um_1)	1 (Uf_1)	2
V	1 (Vm_1)	x	1
W	1 (Wm_1)	1 (Wf_1)	2
X	1 (Xm_1)	x	1
Y	1 (Ym_1)	1 (Yf_1)	2
Z	x	1 (Zf_1)	1
TOTAL PERSONAL HOMEPAGES			**60**

Key: initial = initial of surname as given on the individual pages (A-Z); PH = personal homepages; m/f = male/female.
Please note: the 60 personal homepages were visited 5 times during the time period 14.11.2006 – 28.01.2007; the overview given here represents the first visit on 14.11.2006.

3.4. Online survey

The online survey was carried out between 13. August and 3. September 2007, and consisted of 31 thematically assembled questions, reflecting the main research areas of this study (the complete questionnaire can be found in appendix 15.2.). An e-mail, containing explanatory remarks concerning my dissertation and a link to the online survey, was sent to 92 potential informants who were asked to fill in the questionnaire as well as forward the e-mail to anybody else who may want to participate in the survey. The anonymous survey (N_6) generated 109 questionnaires that were filled in by 78 female informants and 31 male informants (native as well as non-native English speakers), providing me with valuable information on the communicative strategies used by 21st century authors. It should be kept in mind that questionnaire data is biased by the circumstance that informants, to a certain extent, answer the questions in hypothetical fashion, i.e. they answer what they *think* they would do if they were faced with a certain scenario as proposed in the questionnaire. Still, the results of the online survey are very interesting and will be discussed and compared with findings that the investigation into the text corpora yielded.

3.5. Overview of corpora

Table 3.3 below provides a summarised overview of all data collected, including average and total word counts for each of the text corpora (except for the online survey).

Table 3.3: Overview of corpora ($N_1 – N_6$).

Corpora	N	Average word counts per text entity	Total word counts
EEC (N_1)	91	* 341.5	31’077
SMS texts (N_2)	1000	18.4	18’426
E-mails (N_3)	140	183.8	25’733
Chat (N_4)	30	946.8	28’404
Personal homepages (N_5)	60	267.2	16’030
Online survey (N_6)	109	x	x
TOTAL			**119’670**

Key: * = average word count per text entity rounded to 1 decimal.

It is evident that the number of texts and total word counts for the five text corpora vary considerably. Frequencies and distributions of features investigated in this study will thus be given in either feature-to-word ratios or in percentages so as to ensure comparability of the findings across the five text corpora.

PART II: THEORETICAL SCOPE AND METHODOLOGY

4. Theoretical background

> The study of communication raises two major questions: first, what is communicated, and second, how is communication achieved?
>
> (Sperber & Wilson 1995: 1)

The 21st century has given rise to new types of interpersonal communication. The field of written, as opposed to spoken, personal correspondence in particular, has experienced changes in the way people communicate with each other. Both the correspondence via the Internet as well as the short message service of mobile phones have been recent additions in the field of interpersonal communication. New approaches are required to analyse these new types of personal telecommunication. Before introducing the analytical framework that attempts to be such a new approach, the theoretical background of this study, on which the analytical framework is based, will be introduced. The discussion of the theoretical background begins with the breakdown of personal communication into different components that influence message composition. This will be followed by comparing traditional and more recent approaches to the notion of "orality" (subsuming features typical of speech) and "literacy" (subsuming features typical of writing) in the light of the two production modes, "spoken language" and "written language".

Modern electronic communication has challenged the field of discourse analysis because it seems to be a hybrid of these two modes—not necessarily in terms of production, because most modern communication is still based on the written word, but definitely in terms of perception to the extent where it is said to have so many features of orality that it is reminiscent of speech. Furthermore, certain features of CMC, such as the hyperlink, cannot be defined in terms of orality or literacy and thus should be classified as idiosyncratic instead. However, because the immediacy of electronic communication has brought about changes in the communicative context, authors are hypothesised to contextualise their writing differently compared to older means of correspondence (such as the handwritten letter). In particular with regard to aspects tied to reference and relevance.

This chapter aims at giving an overview of the most important theoretical approaches that have been proposed as ways of conceptualising the contextuality of communication.

4.1. "Message" deconstructed

The desire to communicate a particular message always entails a choice of medium, and this inevitably has an influence on the message as a whole. Döring (2003: 128ff.) has developed a useful model labelled *Medienökologisches Rahmenmodell* ('media-ecological frame scheme') that illustrates in a simplified manner the various factors that come into play before, during, and after communication occurs:

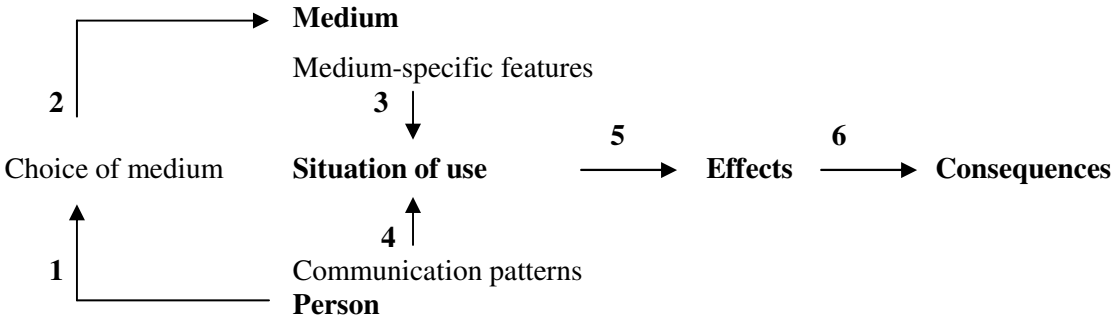


Figure 4.1: *Media-ecological frame scheme* (Döring 2003: 190, my translation).

In short, different means of communication allow users the possibility to choose which one of them seems most adequate for a particular task (step 1 in Döring's model). This is strongly connected to what sort of media users have access to, the connectivity of the addressee(s) (for example, if the addressee does not own a computer, then the choice of media decreases since sending an e-mail will not result in successful communication), as well as traditional norms and rules regarding media use (these are, of course, culture-dependent, but unless it is a case of emergency, phone calls at night time during the week are usually perceived as a breach of etiquette as opposed to sending an e-mail). Once the user has decided on a specific medium (step 2), which always entails the choice against all other available media[34], questions of media-specific aspects (step 3) and competence of the user regarding a particular medium arise (step 4). This influences the situation of use

[34] It should be noted that the 17th century handwritten letter is an exception as there were no other possibilities to personally communicate in written form at that time other than the letter. This means that the choice of medium was, first, limited and, second, did not entail the choice against all other available media.

which in turn creates short-term effects (step 5) and long-term consequences (step 6) regarding the use of media in particular circumstances.

However, Döring's pragmatic approach to communicative behaviour neglects the text, which happens to be the starting point of this investigation and Döring's (2003: 190) model has thus been adapted for the purpose of this study. Figure 4.2 below illustrates the connections between communicative context, message (and its contextuality), and the contextual effects that a message performs. The numbered steps from Döring's original scheme have been omitted deliberately, as this adaptation is not meant to be understood as a flowchart:

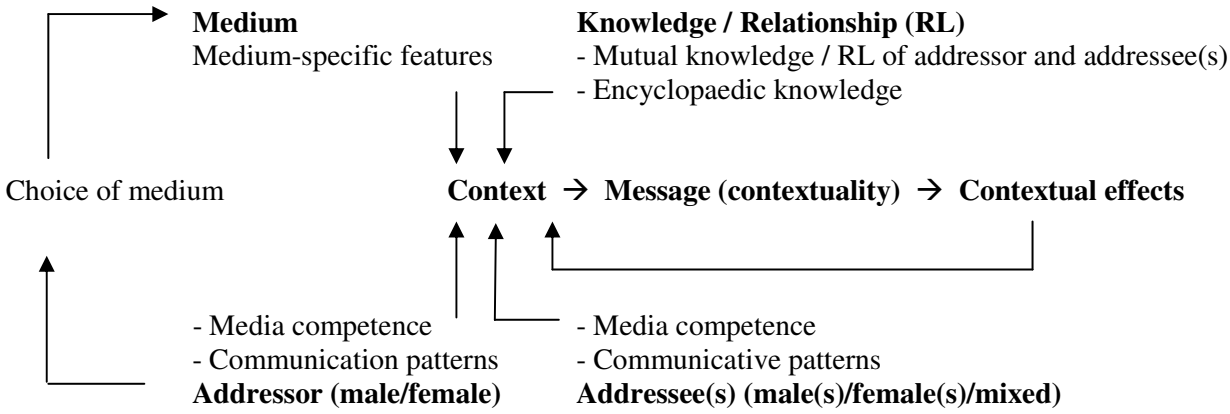


Figure 4.2: Personal written communication: context, message (contextuality), and contextual effects (adapted from Döring 2003: 190).

Although a piece of personal correspondence is written by the addressor, Figure 4.2 reflects how the addressor's act of communicating in writing is inextricably interrelated with addressee(s), medium, and context, all of which in turn influence its contextuality, i.e. how a message is contextualised. The three factors constitute the basic frame for any type of successful interpersonal communication: without a communication partner (real, fictional, or implied) there will be no dialogue, without medium there will be no written communication, and if no mutual communicative context can be established, then new information will not be successfully processable for the addressee(s) in terms of relevance and therefore not result in contextual effects.

Other influences on the context of any communicative interaction are the mutual as well as the encyclopaedic knowledge of addressor(s) and addressee(s). The better acquainted the correspondents are, the more shared (pre-)communicative context they have which can then be presumed when communicating with each other. For example, a sender

does not need to be as explicit about background information when communicating with a close friend, as opposed to an interaction with someone less acquainted. Similarly, the more aligned the correspondents' encyclopaedic knowledge is, the more shared (pre-)communicative context they have, such as common cultural values. Since the influence of both mutual and encyclopaedic knowledge on the communicative context is extremely difficult to measure, it will not be subject to this investigation. However, next to being in the position of knowing how to communicate in the English language (in other words, being familiar with the English lexicon and grammar—and being able to put this knowledge to use), it seems safe to assume that the participants of personal correspondence also share, or at least pretend to share in the case of anonymous communication, a certain amount of mutual as well as encyclopaedic knowledge.

This study focuses on communication in context, or rather, how personal written communication is contextualised in the 21st century in contrast to a time when the handwritten letter was the only means of private correspondence. While Figure 4.2 illustrates that context is of a fluid nature, the contextuality of a message is not. This is the reason why this study attempts to empirically investigate the "contextuality" of personal correspondence, rather than its "context". The only exception being the investigation into "contextual effects" because, although they will be classified based on written evidence (and thus qualify as an inside-of-text parameter, see also 1.3.2.), their effects are actually by definition aimed at going beyond the text: they are performed on the reader(s). However, it will be explained in more detail below how this situation was solved in view of the fact that the (intended) readership of the text corpora was in the majority of cases unknown to the researcher and could therefore not be contacted. Bearing the fluid nature of context in mind, it is believed that based on the empirical investigation into the different text types, educated conclusions can be drawn as to how changes in the communicative context have influenced the contextuality of personal written communication over time.

4.2. Speech vs. writing and the orality-literacy continuum

The tools for spoken conversation have not changed a great deal over the thousands of years of correspondence between humans (we are still using our vocal chords to produce sound). The media for written communication have, however, changed considerably: they have evolved from papyrus rolls and ink (3000 B.C.) to modern electronic texts of the present day (cf. Nickisch 2003: 63). Both SMS discourse and CMC are said to be reminiscent of spoken language, although the production mode in most cases is that of writing, which

may, in the case of digital communication, be supported by (audio-)visual tools. According to Baron (2000: 2), "most of us have been taught to maintain distinct styles for speaking and for writing" but it is observable that "increasingly, people are blurring these distinctions in the direction of the informal patterns of spoken language." But how can this phenomenon be captured in analytical terms?

Seminal work in reflecting on the notions of "orality" and "literacy" has been done by Ong (for example, 2002 [1982]). By exploring the profound changes in our thought processes, personalities and social structures in connection with speech, writing and print, he concludes that after a so-called "primary orality" (referring to cultures totally untouched by writing and print), the age of telecommunication and information technology has produced a "secondary orality". Consider the following quote by Ong (2002: 133-134):

> The electronic transformation of verbal expression has both deepened the commitment of the word to space initiated by writing and intensified by print and has brought consciousness to a new age of secondary orality . . . . This new orality has striking resemblances to the old in its participatory mystique, its fostering of a communal sense, its concentration on the present moment, and even its use of formulas . . . . [W]here primary orality promotes spontaneity because the analytic reflectiveness implemented by writing is unavailable, secondary orality promotes spontaneity because through analytic reflection we have decided that spontaneity is a good thing.

While this might sound a little mystical rather than analytical, it still strikes an important chord in that it captures two important aspects of "secondary orality": the connection between "spontaneity" and "immediacy", and the modern communication media fostering those two aspects.

Concerning the theoretical classification of language into speech vs. writing, Crystal (2001: 26ff.) gives an overview of the differences of the two modes. It becomes apparent that, apart from the production mode, the main differences are believed to be inherent in the structure of the texts: speech is usually less structured than written documents. Koch and Oesterreicher (1994: 588ff.)[35], on the other hand, have proposed ways of conceptualising the dimension of "speech vs. writing" and "orality vs. literacy" and state that although the production mode of utterances is either phonic or graphic, the

[35] It is felt that Koch and Oesterreicher's approach in capturing the differences between the production mode of utterances and how they are perceived by the hearer or reader in terms of communicative immediacy or distance is a very insightful contribution to the field of pragmatics. The source given here (Koch & Oesterreicher 1994) is, however, in German. It should be noted that Koch (1999) introduces the framework in English, albeit in a different context (applied to spontaneous dialogues in early Romance texts).

way that addressees conceive a specific utterance varies along a continuum of "communicative immediacy" (informality) and "communicative distance" (formality). Koch and Oesterreicher (1994) place a personal (i.e. intimate) face-to-face conversation at the one end of the continuum (communicative immediacy), and a text of law at the other (communicative distance).[36] In terms of perception it is thus possible that a written document may be perceived as more "oral" (i.e. containing more features otherwise typical of speech) than a spoken utterance. This means that, for example, a little note scribbled down on a piece of paper can contain more features of orality than does an oral presentation in an academic setting.[37] Haase et al. (1997: 61ff.) discuss the two poles of the continuum from a textual-pragmatic, syntactic, and lexical point of view. The main characteristics of communicative immediacy and distance as understood by Haase et al. are summarised in Table 4.1.

Table 4.1: Aspects of communicative immediacy and distance in language (following Haase et al. 1997: 61ff., my translation).

	Communicative immediacy	**Communicative distance**
Textual-pragmatic level	- structuring signals - turn-taking signals - hesitation phenomena - correction signals	- textual coherence - structured semantics - explicit connection between particular text segments
Syntactic level	A lack of syntactic adequacy is compensated by the use of non-verbal and paralinguistic signals between communication partners.	Syntactic adequacy.
Lexical-semantic level	The choice of words is as important as other factors that affect a conversation, such as extended knowledge about the context (of a particular communicative situation) or the present communication partners.	Missing paralinguistic contexts are compensated by a differentiation of the language used.

[36] According to Koch and Oesterreicher (cf. 1994: 588) the two poles of the continuum are characterised by specific parameters. For example, "communicative immediacy" is related to dialogue, familiarity between interlocutors, privacy, spontaneity, involvement etc., "communicative distance", on the other hand, corresponds to monologue, unfamiliarity between interlocutors, publicity, high level of reflection, detachment and so forth.

[37] Raible (1994) provides a short but insightful overview of the general aspects of writing and its use in connection to orality and literacy. Other approaches to the notion of orality and literacy can be found in Tannen (1984), Finnegan (1988, taking the technology of communication into account), Olson and Torrance (1991, assembling different approaches to oral and literate aspects of culture, cognition, and discourse) as well as Furniss (2004, with a particular focus on orality and "the power of the spoken word").

It is thus possible to analyse texts with regard to their features being distinctive of either communicative immediacy or distance, as illustrated in Table 4.1, and then classify them along an orality-literacy continuum as proposed by Koch and Oesterreicher (1994).

Another approach to the speech-writing dimension has been suggested by Baron (2000: 21) who observes that "the linguistic and historical/cognitive agendas both presuppose a dichotomous relationship between speech and writing" but that "anthropologically oriented studies of writing have found significant mismatches between forms of speech and writing." Baron (2000: 21) illustrates the traditional view of the two modes and their features by means of the following lists:

Writing is:	**Speech is:**
objective	interpersonal
a monologue	a dialogue
durable	ephemeral
scannable	only linearly accessible
planned	spontaneous
highly structured	loosely structured
syntactically complex	syntactically simple
concerned with the past	concerned with the present
formal	informal
expository	narrative
argument-oriented	event-oriented
decontextualised	contextualised
abstract	concrete

Opposed to this traditional and categorical separation of speech and writing mode, Baron (2000: 22) introduces the idea of "the cross-over view" claiming that "merely because a linguistic message looks as if it is designed to be spoken or written hardly ensures that this will be the medium through which everyone experiences it," referring, for example, to books that are read out loud, and speeches which are recorded. Still, the two lists provide a useful starting point for a discussion on the notions of "speech" and "writing", and while this study will not cover all aspects included in the two lists above, some will be addressed in more detail as we proceed.

In terms of literacy and writing, Bolter (2001) analyses the notion of "writing space" and its interrelations with computers, hypertext, and print and the developments in each of these fields. According to Bolter (2001: 21), "whenever a dominant technology is challenged, there may be a major refashioning of the culture's writing space." Hence, the traditional writing space of, for example, the handwritten letter and printed works is currently being challenged by the advent of electronic technologies, such as CMC and other

types of electronic correspondence. Or, in Bolter's (2001: 21) words, "electronic and digital technologies are helping to refashion the writing space again." How traditional notions of literacy contrast with (text-based) CMC will be of interest next.

4.2.1. Orality, literacy, and CMC (hypertext)

The discussion on the dimension of speech and writing in relation to the orality-literacy continuum is further enriched by the advent of CMC. In the majority of cases, CMC is produced in the mode of writing, but because of the frequent occurrence of features of orality is also said to be reminiscent of speech. Furthermore, CMC contains idiosyncratic features that cannot be assigned to either speech or writing mode, and can also not be captured by the notion of orality or literacy. Jacobsen (2002) thus suggests classifying CMC as a third discourse mode, featuring characteristics of "cyberdiscursivity". Table 4.2 below illustrates the principal elements of orality, literacy, and cyberdiscursivity according to Jacobson (2002: 8).

Table 4.2: Principal elements of orality, literacy, and cyberdiscursivity (Jacobsen 2002: 8).

Orality	**Literacy**	**Cyberdiscursivity**
embodied	disembodied	virtual
concrete	abstract	dynamic
aggregative	hierarchical	emergent
communal	individual	idiosyncratic

In his "theory of cyberdiscursivity", Jacobson (2002: ix) draws on Ong's approach where oral rhetoric is seen as embodied, concrete, aggregative, and communal, the literate rhetoric, on the other hand, is claimed to be disembodied, abstract, hierarchical, and individual in focus. Jacobson (2002: ix) postulates a theory of cyberdiscursivity which holds that "the more instantaneous, widespread, and individual discursive practices inherent in computer-mediated communication change the production, use, and conceptualisation of texts" which in turn renders "the familiar standards of print textuality only tangentially applicable to cybertexts." Jacobson (2002: 7) further claims that CMC contains features that go beyond features of orality and literacy (and thus support his theory of cyberdiscursivity), such as the hyperlink which adds another dimension to the notion of text, transforming it into hypertext.

The notion of "hypertext" is in fact not restricted to hyperlinked text on the Internet. Any text containing cross-references to other text segments in non-linear fashion can be seen as hypertext. Schlobinski (2005: 9, my emphasis), for example, suggests the terminological differentiation between hyper*text* and hyper*media* of which the former can also be used to refer to print works, but the latter cannot. As regards the terminology for the phenomenon of online hypertext, Snowden (2004: 2, my emphasis) argues that "hypertext would more accurately be described as '*hybrid*text' because it grafts new technology and the capabilities of electronic communications onto the traditional text-based, literate construction of information." In any event, it should be noted that for the purposes of this study, the term *hypertext* is understood to refer to text on the Internet that is (a) hyperlinked to other text segments in a non-linear fashion, and that (b) may or may not contain reference systems other than text (such as graphical representations of information).

4.3. Context and contextuality

> The significance of any speaker's communicative action is doubly contextual in being both context-shaped and context-renewing.
>
> (Heritage 1984: 242)

The quote above sums up the most essential property of language in use, namely that human communication stands in a reciprocal relationship with the notion of context. Essentially, communicative actions that are not context-shaped will be difficult to process in terms of relevance. How to make sense of something said or written "out of context"? Communicative actions that cannot be processed do not contribute to a context-renewal in Heritage's sense. Although Levinson (1983: 22) contends that "a . . . difficulty facing the definition of scope of pragmatics is that it calls for some explicit characterisation of the notion of context," this particular area of research has received comparatively little attention in relation to its significance in the past. This, however, seems to be changing as more and more recent works published in the area of pragmatics are dealing with the role of context in connection with language use and relevance.

In particular, several volumes of the *Pragmatics & Beyond New Series (BNS)*[38] take a detailed look at contextual issues from different angles. For example, Fetzer's (2004) contribution to *Pragmatics & BNS,* titled *Recontextualising Context,* is concerned with mutual intelligibility in conversation in connection with contextual information and meaning:

> Since one of the goals of natural-language communication is mutual intelligibility, co-participants tend to repair . . . communicative problems [such as deliberate or non-deliberate misunderstandings, mishearhings, non-acceptance of a prior contribution which is seen as untrue, insincere, or inappropriate] by making explicit their speaker-intended meaning. In the case of a mishearing, they paraphrase their contribution, and in the case of a non-acceptance, they spell out the controversial presupposed contextual information in order to make the argument more acceptable to their co-participants while in the meantime strengthening its validity.
>
> (Fetzer 2004: 1)

Fetzer (2004) also looks at different notions of context and differentiates between three types: "linguistic context", "social context", and "sociocultural context". With respect to linguistic context, Fetzer (2004: 4) points out that "in traditional linguistic accounts of context, context is conceived of as comprising the immediate features of a speech situation in which an expression is uttered, such as time, location, speaker, hearer, and preceding discourse." However, I agree with Fetzer (cf. 2004: 4) that context is a much wider and a far more transcendental concept than what these accounts imply. Fetzer (2004: 5) argues that a current contribution "is not only anchored to a prior contribution, but also constrains the form and function of upcoming talk." Or, in other words, "the language produced (formulated) and interpreted (decoded) by the co-participants is assigned a dual function" in that "it invokes linguistic context by constructing itself" as well as providing "the context for subsequent talk and recovery of intended meaning" (Fetzer 2004: 5-6, following Akman & Alpaslan 1999).

The constituents of social context are, according to Fetzer (cf. 2004: 7), for instance, the co-participants of a given communicative exchange, the immediate concrete, physical surroundings including time and location, as well as the macro contextual institutional and non-institutional domains (i.e. speaker, hearer, and audience are

[38] *Pragmatics & Beyond New Series* (John Benjamins Publishing Company) is a continuation of *Pragmatics & Beyond* and its *Companion Series*. *Pragmatics & BNS* "provides a forum for scholars in any area of pragmatics" and "aims at representing the field in its diversity covering different topics and different linguistic and socio-cultural contexts, including various theoretical and methodological perspectives" (cf. online source 9).

represented by multiple social roles). Also, social context is jointly constructed and negotiated by the co-participants of a given communicative exchange (cf. Fetzer 2004: 9). Furthermore, social context "subcategorises into different types of sociocultural context which are defined by a particular perspective on social context in general" (Fetzer 2004: 9). Hence, social context can be conceived of as an unmarked type of context, or as a "default context", whereas sociocultural context is more of a marked type of context in which particular variables, such as time or location, are interpreted in a particular mode (cf. Fetzer 2004: 9). In other words, "culture provides us with a filter mechanism which allows us to interpret social context in accordance with particular sociocultural-context constraints and requirements" (Fetzer 2004: 10).[39] This in turn also means that the notion of (social) context may vary considerably from one culture to another which makes a "universally effective" definition of context virtually impossible.

There are, however, two main dimensions in which the notion of context can be understood: as a static setting of affairs or a dynamic one. For example, the *New Oxford Dictionary* provides the following definition for "context": "The circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood and assessed," a description that paints a rather static picture of what context is supposed to constitute. In comparison, Sperber and Wilson (1995: 15), for example, understand context to be of dynamic character: "The set of premises used in interpreting an utterance . . . constitutes what is generally known as the context. A context is a psychological product, a subset of the hearer's assumption about the world." Both the "set of premises" and the "psychological product" are subject to change during a conversation as each new utterance (re-)shapes, or to use Heritage's (1984) terminology, re-news the context that has been established until that point. And context as understood by Sperber and Wilson (1995) is also constitutive of "relevance", i.e. the relevance of an utterance can only be interpreted against a particular context (Sperber and Wilson's Relevance Theory will be discussed in more detail in section 4.4. below).

While I agree with Sperber and Wilson's (1995) claim that context is dynamic in character, it is at the same time quite a discouraging observation for an empirical linguist: how to investigate a feature that is subject to constant change induced by different sources

[39] See hereto also another contribution to the *Pragmatics & BNS* by Fetzer (2007, ed.), titled *Context and Appropriateness: Micro Meets Macro*. The volume addresses the notion of context and appropriateness from a variety of theoretical and applied perspectives, and "the papers range from the research paradigm of philosophy of language speech act theory, sociopragmatics, cognitive pragmatics, and critical discourse analysis" (Fetzer 2007: 3, Introductory comments).

(physical setting, cultural background, participants' demographics and their contributions, and so forth)? This is the reason why this study investigates the *contextuality*, rather than the *context*, of personal written correspondence. The notion of contextuality is understood as the static "inside-of-text evidence" that is shaped by the communicative context and can be traced in written correspondence, as opposed to the notion of context, which is understood as a dynamic mix of a "both inside- and outside-of-text setting" (working definitions for both concepts can be found in 1.3.).

Context is to a certain extent a psychological construct, and its establishment between discourse partners is a combination of several factors (as illustrated in Figure 4.2). While it is difficult, if not impossible, to measure the dimension of mutual knowledge of addressor/addressee(s) from textual data alone, there are other factors, such as media-specific features, where it is possible to evaluate direct influences on contextuality based on the textual evidence provided. Furthermore, it should be noted that the treatment of context in this study is a purely textual and structural one and, apart from the investigation into contextual effects, the far-reaching effects of content and meaning will be largely ignored. It is assumed that "the utterer expects the interpreter to be able to process the information contained in the combination of words and context in order to infer the meanings intended" (Fitzmaurice 2002: 222). In this sense, it is not only the content of the messages as such that is at the heart of interest, but also the textual, visual, and, in the case of some types of modern electronic communication, audible features and strategies applied by the authors in order to achieve successful transmission/reception of their messages on the pragmatic (rather than technological) level.

This is, of course, related to what sort of communication channels a certain medium offers, also referred to as the "bandwidth" or "richness" of a medium. It should be noted that the term *communication channels* is used in this study as a cover term for all the different options that are available to the author in composing a message (both verbal and non-verbal). This includes, for example, the written word as well as visual features (such as pictures), or the compensation of missing paralinguistic cues by means of symbols (such as emoticons) and so forth. Thus, the more available communication channels a medium offers, the richer it is. Or, in other words, the sum of all communication channels constitutes the medium. Herring et al. (2003: 4) propose the following continuum of media richness in consideration of different digital, electronic, audio and print (handwritten) media that can be used to communicate interpersonally.

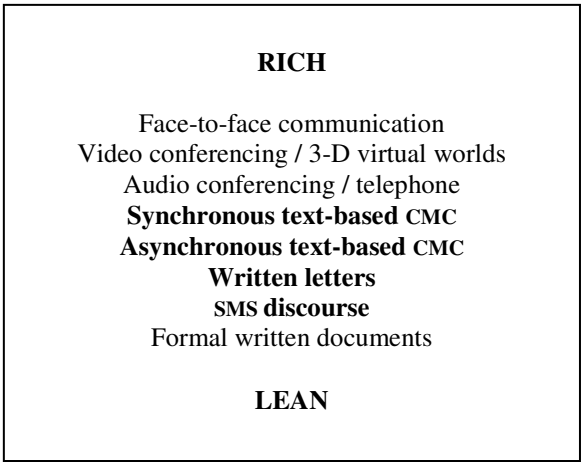


Figure 4.3: A continuum of media richness (adapted from Herring et al. 2003: 4, who in turn adapted it from Daft & Lengel 1984, my emphasis).

The continuum, which Herring et al. (2004) have adapted from Daft and Lengel (1984), has been supplemented by the text type SMS discourse that is, based on its available communication channels, leaner than letter writing (see also Table 5.1). Media-richness can be brought into relation with the communicative immediacy-distance continuum proposed by Koch and Oesterreicher (1994), as they also place a face-to-face conversation at the one end of the continuum (communicative immediacy) and the formal document at the other (communicative distance). However, the availability of communication channels in rich media for written communication does not mean that users also make use of all of them. This stands in contrast to face-to-face settings, where the physical presence of interlocutors makes it more difficult to ignore certain paralinguistic communication channels, such as tone of voice and facial expressions. It will thus be particularly interesting to see which communication channels are made use of in written interactions where the reader is absent or, in the case of Web Chat and the personal homepage, unknown. How authors of personal written correspondence contextualise their writing in the light of these circumstances will be the subject of chapter 6.

4.3.1. Indexicality and deixis

Personal written correspondence, of the type investigated in this study, is by definition directed at one or more addressee(s).[40] From a contextual point of view, this inevitably raises the issue of reference, in particular indexical and deictic reference. Although the terms *indexicality* and *deixis* denote the same linguistic phenomenon, they have a slightly different scope: the notion of "indexicality" is used to label the broader phenomena of contextual dependency, whereas "deixis" denotes the narrower linguistically relevant aspects of indexicality (cf. Levinson 2006: 97). According to Levinson (cf. 2006: 97) one of the core features of deixis is that it introduces subjective, attentional, intentional, and, particularly interesting for this investigation, context-dependent properties into human language.

Furthermore, deixis and the linguistic forms that subserve it, can be said to be "[the] single most obvious way in which the relationship between languages and context is reflected in the structures of languages themselves" (Levinson 1983: 54). Deixis in Levinson's terms encompasses a wide range of referential as well as non-referential functions of speech, from pronouns to regional accents (cf. Hanks 1992: 46). Hence, deictic reference covers a vast field of referential as well as non-referential expressions. However, since an analysis of written communication does not require the accommodation of accent or prosody, I will follow Hanks (1992: 46) in that deixis "designates a special variety of reference, sometimes called 'demonstrative reference', which is limited both formally and functionally" and does not include extra-verbal forms of indexicality such as non-referential functions. For the purposes of this study I will concentrate on verbal deixis that either refers to or indexes personae, space, time, or other texts (whole entities or segments therefrom). Table 4.3 below illustrates the different types of deixis relevant to this study.

[40] Most personal correspondence is directed to one or more addressee(s), but with some exceptions, such as the personal diary. Although the diary itself is sometimes used as an implied addressee ("dear diary"), it may also not be addressed at all (instead of starting the entry with "dear diary", one may choose to start the entry with the date). In this case a particular diary entry may be seen as a monologic type of personal correspondence without being directed to an (implied) addressee.

Table 4.3: Different types of deixis (following Levinson 2006: 111ff.).

Types of deixis	**Definitions**	**Illustrations**
Personal	The traditional person paradigm can be captured by the two semantic features of speaker inclusion (S) and addressee inclusion (A): first person (+S), second person (+A, -S), and third person (-S, -A).	(+S, -A): Pronominal self-address of speaker (S) (e.g. *I am here*), or other types of self-address, quite frequent in written correspondence (e.g. farewell section in a letter: *Yours truly, Paul*). (+A, -S): Pronominal address of addressee(s) (A) (e.g. *It's so nice you are here*) or other types of reader-address (e.g. greeting section in a letter: *Dear Andrea*) (-A, -S): Pronominal address of third party (e.g. *He is nice*) or other types of third party address (e.g. *The students are late again*).
Spatial	Deictics often contain information in an allocentric frame of reference hooked to geographical features, or abstract cardinal directions.	Geographical spatial deixis: Particular references to geographical locations, e.g. *Peter is in Paris at the moment*. Abstract cardinal directions: Spatial references that are unhooked from geographical deictic centres, e.g. *It all went downhill from there* (ideological spatiality rather than geographical).
Temporal	The most pervasive aspect of temporal deixis is tense. Tenses are traditionally categorised as ABSOLUTE (deictic) versus RELATIVE (anterior or posterior to a textually specified time). Other units of time measurement may either be fixed by direct reference to the calendar or not.	Absolute vs. relative tense: If the simple English past (e.g. *He went*) is classified as absolute, then the past perfect (e.g. *He had gone*) is relative. Fixed by direct reference to the calendar: E.g. *I will see you next Thursday at 2pm*. Not fixed by direct reference to the calendar: E.g. *Whenever he calls, I get nervous*.
[1]**Textual**	*Intra*-textual deictic reference denotes the reference to portions within the origo-text itself (or to the whole text entity). Discourse marker (DM) relates a current contribution to the prior utterance or text. *Inter*-textual deictic reference denotes the reference to portions of text (or whole text entities) outside the origo-text.	Intra-textual (endophoric) deictic reference: E.g. *As mentioned in the chapter above*. Connecting a textual unit to prior text: E.g., *He fancied the cake. So he ate it*. Inter-textual (exophoric) deictic reference: E.g.: *Didn't you read my last e-mail?*

Key: [1] = Levinson (cf. 2006: 118ff.) subsumes "textual deixis" under "discourse deixis" and restricts it to *intra*-textual reference. Textual deixis as understood for the purposes of this study includes both *intra*-textual reference as well as *inter*-textual reference. The definition for "textual deixis" presented in Table 4.3 thus deviates in this respect from Levinson (cf. 2006: 119).

Before further remarks on the contents of Table 4.3 will be made, I would like to draw attention to a type of deixis that is missing in the above overview: the notion of "social deixis". As a subtype of personal deixis, social deixis "involves the marking of social relationships in linguistic expressions" (Levinson 2006: 119) as in *hi honey* (intimate, i.e. close, social relationship) vs. *Dear Professor Miller* (professional, i.e. distant, social relationship). Since the social relationships between authors of personal written correspondence can be assumed to be close, this special type of deixis will not be treated in detail in this study. However, the empirical investigation into the use of personal deictic expressions includes forms of address, which in turn also reveals certain tendencies relevant to the social level of personal reference.

Table 4.3 illustrates the types of deixis as investigated in this study, and it emerges that language, and particularly language in use, is permeated with deixis. Essentially, everything we say or write involves reference tied to the communicative context; otherwise it would not be possible to decode a message as intended by the producer. I agree with Hanks (1992: 70) that "verbal deixis is a central aspect of the social matrix of orientation and perception through which speakers produce context" and thus deixis is an immensely interesting field for an investigation into the contextuality of personal written correspondence. With regard to the four types of deixis investigated in this study, key aspects were determined for each type of deixis.

The main focus of the investigation into personal deixis will be forms of address. As illustrated in Table 4.3, Levinson's "traditional person paradigm" (2006: 112) of personal deixis is based on speaker (S) and addressee(s) (A), and reference depends on the in- or exclusion of either speaker or addressee(s), or both (in which case it would be a reference to any other third party). Thus, the basic person paradigm includes the references to either speaker (+S, -A), or addressee(s) (-S, +A), or any other third party (-S, -A). In addition, Levinson (2006: 113) points out that "although the traditional notions first, second, and third persons hold up remarkably well, there are many kinds of homophony, or different patterns of syncretism, across person paradigms" and that much of this complexity is due to the distinctive notions of plurality, the "we" notions being especially troubling in many languages. In the case of the English language, the plurality of the grammatical first person allows references to speaker and addressee(s) excluding another third party (+S, +A, - third party), as well as references to speaker and any other third party excluding the addressee(s) (+S, -A, + third party). In addition, these first person plural references may be expressed by pronouns (for example, *we*) or noun phrases (for example,

our group). In any event, to accommodate the notion of plurality in the pronominal system of English, as well as the circumstance that it is the author (as opposed to speaker) that is the main focus of this study, the terminology as proposed by Levinson has been slightly adapted. All modifications will be illustrated in detail in chapter 5 (*Methodology*).

The key aspect of the analysis into spatial and temporal deixis will be the notion of "deictic origo", first formulated by Bühler (1934), who observed that the speaker and the place and time of his/her utterance, along with the role of recipient or addressee, recurs at the core of deictic distinctions in grammar after grammar (cf. Levinson 2006: 111). With respect to personal deixis outlined above, this origo is located with the author figure for the purposes of this study, for the reason that all analysed material was written by such author figures. The same is true for temporal origo, which is also tied to the author figure for similar reasons. Because even if the author, at the time of message production, takes into account when the message is assumed to be read by the recipient, the timeframe is still dependent on the author's perceptions of temporal cornerstones such as date or time, as well as his/her assumptions about the timeframe of the recipient. Thus, a comment like *It will probably be evening by the time you read this*, written by an author and his/her timeframe of early morning, the temporal origo is tied to the author's perceptions of time, regardless whether or not the author considers the reader's (presumed) timeframe. This is true for time deixis expressed by means of tense, references to the calendar as well as any other use of temporal expressions that are unfixed to the calendar.

While the spatial deictic origo is similar to personal and temporal deictic origo with respect to the dominance of the author's perception of space at the moment of message composition, it differs in one crucial aspect: the deictic spatial origo can be tied to the physical location of the author figure, or it can take the text as deictic spatial origo. Of course, the notion of text origo is tied to the author's perception of the spatial aspects of a particular text. However, there is a difference to personal and temporal origo (which remain with the author) in that the text provides another realm and offers a different matrix for spatial references—which is not the case for both personal and temporal origo.[41] In any case, the advent of the Internet and the notion of "mobile communication" via SMS

[41] This still holds even if temporal deictic expressions are sometimes used to refer to other text segments within the same text (for example, *as has been pointed out earlier*), because although these expressions are borrowed from temporal deixis, they nevertheless point to a spatial destination somewhere else in the text. Hence, the reader is faced with spatial orientation rather than temporal information, because looking something up that was previously written means going back to another location in the text, not going back in time.

discourse pose interesting scenarios for the conceptualisations of space. These concern, in particular, the notion of deictic origo as assigned to the text in contrast to the physical location of the author at the time the message was composed. Furthermore, text located in virtual space deals with yet another dimension, and the investigation into how authors contextualise their writings in terms of space in the light of these circumstances generated interesting results.

The discussion on textual deixis will mainly focus on the notion of textual networking via the use of *intra*-textual (i.e. "endophoric") as well as *inter*-textual (i.e. "exophoric") references to connect text segments or whole textual entities, respectively. As illustrated in Table 4.3, any given utterance can be deictically connected to other text by means of backward indexing (to prior text segments within the same text or other text entities) and forward indexing (to upcoming text segments within the same text or other text entities). It should be noted that for the purposes of this study the following terminological distinctions will be made with respect to textual deixis. First, the terms *text-internal* and *endophoric* will be used interchangeably to denote textual references to text (segments) within a given text, and the same is true for the terms *text-external* and *exophoric* to denote textual references to text (segments) outside a given text. Second, endophoric reference within a given text entity can be "anaphoric" (backward indexing) or "cataphoric" (forward indexing). As these terms are commonly used to refer to deictic connections between text segments within the same text entity, i.e. endophoric references, (see, for example, Levinson 2006), references to text segments or entities outside a given text entity will be referred to as text-external or exophoric "backward reference" and "forward reference", respectively.[42]

In addition, another important area of discourse deixis will be considered: the discourse marker (DM). In his contribution to the *Handbook of Pragmatics*, Levinson (2006: 119) refers to Blakemore in the same volume for the discussion of DMs, and it becomes evident quite quickly that DMs are difficult to capture. An exhaustive discussion

[42] Paraboni and Van Deemter (1999: 44) introduce the useful notion of "document deixis" in their paper, accommodating the fact that deictic references within a document (or to other documents) may be directed at pictures or graphical representations (rather than text segments or text entities).

of the ongoing debate on the meaning and function of DMs[43] would exceed the scope of this study. However, this linguistic phenomenon (if not to say chameleon) will be briefly introduced and the main function of the DM[44], as understood for the purposes of this study, will be determined.

According to Blakemore (2006: 221, emphasis original) "the term DISCOURSE MARKER is generally used to refer to a syntactically heterogeneous class of expressions which are distinguished by their function in discourse and the kind of meaning they encode" but at the same time, there is not yet a "generally accepted list of discourse markers in English" (Jucker 1993: 436). Furthermore, there seems to be some disagreement regarding terminology to denote the same phenomenon, amongst others: "pragmatic marker", "discourse particle", "discourse connective", "discourse operator", or "cue marker" (Blakemore 2006: 221).[45] And while Lenk (1998: 49) agrees that "no common definition of discourse markers exists," she raises the issue that it is in fact questionable whether a common definition is at all desirable, because "numerous different items are employed in all kinds of different discourse functions on various discourse planes beyond the propositional content." In a similar vein, Schiffrin (1987: 328) argues that the "try to find common characteristics of these items [DMs] to delimit what linguistic conditions allow an expression to be used as a marker," at the same time would not only require the discovery of the shared characteristics of an extremely diversified set of expressions, but that it would also "require analysis across a wide body of typologically diverse language to discover what other linguistic resources [such as paralinguistic features and non-verbal gestures] are drawn up upon for use as markers." On a more general note,

---

[43] Aijmer's (2006: 1) states that research on "pragmatic markers" (alternate terminology to denote the same linguistic phenomenon, but see also footnote 45) has exploded in the last 20 years with various important books published in this area (she mentions, among others, Schiffrin 1987 and Jucker & Ziv 1998). However, there are still major theoretical and descriptive challenges in this field of research, particularly in view of the fact that "there is a lack of a generally accepted terminology and useful taxonomies in this area" (Aijmer 2006: 2). See also Fraser (1996 and 1999) in connection with the challenge to capture the notion of pragmatic/discourse markers.

[44] It should be noted that the function of DMs is not tied to their form. Discourse units such as *oh*, can function as a DM (as in *Oh well, I'll do it another time*) but may also take on other discourse functions such as onomatopoeic interjection to express surprise (as in *Oh! This is so nice of you*!). See hereto 6.1.2., 8.1.2., and 11.2.5.

[45] See Pons (1997) for a broad and detailed set of alternative terminology for what is referred to as "discourse marker" (following Blakemore 2006) in this study. For example, Jucker (1993: 435ff.) points out that "sometimes the term 'discourse particle' is used interchangeably with the term 'discourse marker'," and that "sometimes the two terms are used to refer to different phenomena" (see hereto also Lenk (1998: 37ff.), who argues that discourse markers are in fact a subgroup of particles). Obviously, such terminological inconsistency across papers and studies by different linguists further complicates the issue.

Jucker (1999: 437) points out that any description of discourse markers must attend to their "polyfunctionality", that is to say the range of different uses and contexts in which they can occur.

The meaning of DMs is subject to constant discussion as it is, for example, difficult to determine the lexical meaning of the DM *so*. Also, the truth-conditional aspects of DMs have not yet generally been agreed upon, nor whether DMs contribute to truth conditions at all. And Blakemore (2006: 3, emphasis original) raises the question that "if they do not contribute to truth conditions, what *do* they contribute to?" Although there seems to be dissension regarding the definition and function of DMs, they are undoubtedly an important discourse structuring device. Thus, for the purposes of this study I decided to adopt Blakemore's (2006: 221-222, emphasis original) definition based on their pragmatic function:

> It seems that we can say that the term discourse is intended to underline the fact that these expressions must be described at the level of DISCOURSE rather than the sentence, while the term MARKER is intended to reflect the fact that their meanings must be analysed in terms of what they indicate or mark rather than what they describe . . . . The property generally considered to distinguish DMs from other discourse indicators is their function of marking relationships between units of discourse.

In this sense, the main pragmatic function of DMs is to connect units of discourse. Yet they do not only function as connectors *between* textual units but also *within* a textual unit (cf. Blakemore 2006: 232). Now, it is indeed problematic that there is also "considerable disagreement about what exactly a textual unit is" (Blakemore 2006: 232). However, for the purposes of this study, the term "textual unit" will denote a string of words that form a coherent chunk of language as produced by a human being in written fashion.

Coming back to the pragmatic function of DMs as connectors between and within textual units, I believe that based on their frequency, assumptions can be made with regard to the degree of intra-textuality of a written text. In other words, any given textual unit "B" that is fronted or introduced[46] by a DM is thereby connected to the prior textual unit "A", which in reverse relates textual unit "A" to the subsequent textual unit "B", both of which

[46] Levinson (1983: 87-88, my emphasis) labels this phenomenon as "utterance-**initial** usages of [for example] *but*, *therefore*, *in conclusion*, *to the contrary*, *still*, *however*, . . . , and so on" and argues that "it is generally conceded that such words have at least a component of meaning that resists truth-conditional treatment" and "what they seem to indicate, often in very complex ways, just how the utterance that contains them is a response to, or a continuation of, some portion of the prior discourse."

can be said to contribute to discourse structure in the connective sense. Or, to sum it up in Lenk's (1998: 52) words, DMs are "used with a pragmatic meaning on a metalingual level of discourse in order to signal for the hearer [or reader] how the speaker [or writer] intends the present contribution to be related to preceding and/or following parts of the discourse."

After discussing different aspects of person, time, space, and text deixis, it becomes apparent that deixis and language in use are inextricably connected.[47] Yet, according to Levinson (2006: 97), "deixis, despite its theoretical importance, is one of the most empirically understudied core areas of pragmatics," and this study seeks to shed more light onto this area of research. While the issue of deixis will be approached from different angles, it is important to notice that the property of indexicality is not exhausted by the study of indexical expressions, "for just about any referring expression can be used deictically" (Levinson 2006:101). Thus, it should be kept in mind that the treatment of deixis in this study focuses on (i.e. is limited to) the use of verbalised deictic expressions in written language as outlined above.

4.4. Relevance Theory and contextual effects

After having introduced the theoretical background on contextual features that can be tied to the different media and are traceable in the types of messages they foster, I would now like to address an area of contextual pragmatics that also takes both inside- and outside-of-text parameters into account.

> To communicate is to claim an individual's attention; hence to communicate is to imply that the information communicated is relevant. This fundamental idea, that communicated information comes with a guarantee of relevance, we call the *principle of relevance* . . . . [T]he principle of relevance is enough on its own to account for the interaction of linguistic meaning and contextual factors in utterance interpretation.
>
> (Sperber & Wilson 1995: vii, emphasis original)

The above quote by Sperber and Wilson highlights a constitutive element of human communication: that communicated information comes with a guarantee of relevance. Relevance Theory (RT) was first formulated in 1986 by Sperber and Wilson, and they further commented on the theory themselves (for example, Wilson 1994, Sperber & Wilson 1995, 1997, 2004, and Wilson & Sperber 1993, 2006) as well as inspiring other

[47] One could argue that a discussion of language isolated from context is less concerned with indexicality and deixis, such as a purely syntactic treatment of sentence structure. However, from a pragmatic point of view it seems there is no area of language in use where questions of indexicality and deixis do not arise.

works by proponents (for example, Blakemore 1987, 2001, 2002, Carston 1998, 2002, Noh 2001, Ifantidou 2001, Unger 2006) but also critical evaluations of certain aspects of RT (for example, Giora 1997, who argues that RT cannot account for speakers' intuition regarding coherence). An elaborate discussion of RT would indeed exceed the scope of this study, but a brief outline of the main tenets of this theory, with a particular focus on the concept of "contextual effects", is felt to be beneficial in order to better understand the importance of relevance in communication.

According to Sperber and Wilson (cf. 2004: 249), RT—which may be seen as an attempt to work out in detail one of Grice's central claims that an essential feature of most human communication, both verbal and non-verbal, is the expression and recognition of intentions—, a communicator provides evidence of his/her intention to convey a certain meaning, which can be inferred by the audience on the basis of the evidence provided. Or, in other words, RT builds on Grice's fundamental assumptions that "(a) the primary domain of pragmatic theory is *overt* intentional communication and (b) utterance interpretation is a non-demonstrative inference process, where hearers [or readers] *infer* the intended interpretation using contextual assumptions and general principles of communication" (Ifantidou 2001: 59, emphasis original).[48] This means that if the aim is to "successfully communicate", i.e. a speaker/writer works intentionally towards making the hearer/reader understand what he or she intends to communicate, then the inference of a given utterance by the hearer/reader should not cause major problems provided the hearer/reader is competent in the language used and, should the utterance be highly context-dependent, has access to enough contextual information to infer the intended interpretation.

However, as has been discussed previously in this study, much (if not all) interpersonal communication is context-dependent to a certain degree and the relevance of an utterance and its inference are closely tied to the communicative context, or, to use Sperber and Wilson's (cf. 1995: 119) terminology, an utterance is relevant if and only if it produces a *contextual effect* and the more contextual effects, the greater the relevance. The intuitive idea behind the notion of a contextual effect is the following: to modify and

[48] However, it should be noted that there are many differences between RT and Grice's approach. For example, Sperber and Wilson (1995: 161) point out themselves that "the principle of relevance is much more explicit than Grice's co-operative principle and maxims," and that "Grice assumes that communication involves a greater degree of co-operation than we do." This means that for Sperber and Wilson (1995: 161), the only purpose that a communicator and his/her audience necessarily have in common is to achieve uptake, "that is, to have the communicator's informative intention recognised by the audience," whereas Grice (1975: 45, quoted in Sperber & Wilson 1995: 161) assumes that "communication must have 'a common purpose or set of purposes, or at least a mutually accepted direction' over and above the aim of achieving uptake."

improve a context is to have some effect on that context—but not just any modification will do. Because "the addition of new information which merely duplicates old information does not count as an improvement; nor does the addition of new information which is entirely unrelated to old information" (Sperber & Wilson 1995: 109). The sort of effect that is of interest, "is a result of interaction between new and old information" (Sperber & Wilson 1995: 109). In the 1986 edition of *Relevance: Communication and Cognition* (henceforward referred to as *Relevance*), Sperber and Wilson describe three different types of contextual effect: (1) "contextual implication" (the conclusion is deducible from the input and the context together, but neither from input nor context alone), (2) "contextual strengthening of existing assumptions", and (3) "contextual contradiction/elimination of existing assumptions". In the postface to the second edition of *Relevance* (1995)[49], they consider a fourth type of contextual effect, namely (4) "weakening of existing assumptions". For the purposes of this study, a fifth type, (5) "non-contextual effect", was formulated, the properties of which will be explained shortly.

Sperber and Wilson (cf. 1995: 118) claim that the notion of a contextual effect is essential to a description of the comprehension process. As a discourse proceeds, the hearer/reader retrieves, or constructs, and then processes a number of assumptions which then form a gradually changing background against which new information is processed. In this sense, interpreting an utterance involves more than merely identifying what has been explicitly (or overtly) expressed: "it crucially involves working out the consequences of adding this assumption in a context determined, at least in part, by earlier acts of comprehension" (Sperber & Wilson 1995: 118). It emerges that the principle of relevance and the notion of contextual effects are very complex concepts that do not lend themselves easily for empirical research. In their contribution to the *Handbook of Pragmatics* (Wilson & Sperber 2006: 609), they claim that "relevance is not just an all-or-none matter but a matter of degree" and that "there are potentially relevant inputs all around us, but we cannot attend to them all." On another occasion, Sperber and Wilson (2004: 284) point out that "much pragmatic research has been carried out in a philosophical or linguistic tradition in which the goal of achieving theoretical generality, with a tendency to rely on intuitions, has created a certain reluctance to get down to the messy business of experimentation." I

[49] In the second edition of Relevance (1995), the original text from 1986 was preserved, but a postface was added that is aimed at "surveying recent developments in the theory and proposing several changes of terminology or substance" (Higashimori & Wilson 1996: 1).

took both these comments as encouragement to adapt RT's concept of contextual effects to the pragmatic purposes of this research.

For example, a strengthening of existing assumptions can be achieved in various ways, from soft-spoken to intense argumentation. Also, contextual effects can be defined in terms of production cost (with regard to speaker/writer) and processing cost (with regard to hearer/reader). Since the authors and recipients of the correspondence investigated in this study are unknown, it is virtually impossible to classify a particular contextual effect with regard to its strength. This means that for the purposes of this study, the notion of contextual effects was simplified to whether or not a certain type of contextual effect could be determined, regardless of its strength and nuances. Yet, as with other theories of comparable scope, RT's most general tenets, the concept of contextual effects being one of them, can be tested only indirectly, by evaluating some of their consequences (cf. Wilson & Sperber 2006: 625ff.). The textual evidence in the form of different types of correspondence, comprised of a certain number of textual units, is taken as such a "testable consequence"; a consequence in the sense of a textual product that expresses the results of an author's thought processes and intentions (that are directed at one or more readers) in written form.

Another adaptation to the theory concerns two types of contextual effects, the contextual weakening, and the fifth type of effect as formulated for the purposes of this study, the non-contextual effect. With regard to contextual weakening, Wilson (cf. Higashimori & Wilson 1996: 4) points out that a weakening of assumptions is never achieved directly, but only as a by-product of some basic contextual effect, and weakening of assumptions is therefore not treated as a distinct type of contextual effect (and thus contradicts the 1995 postface addition to *Relevance*, which introduces weakening as a forth type of effect). Whilst I agree that contradiction and elimination of existing assumptions might also produce a weakening effect on existing assumptions, I still believe that a weakening effect can be achieved independently. For example, authors frequently employ contextual weakening strategies in order not to insult their addressee(s) with something they wrote earlier without completely contradicting or eliminating the previous statement. For example, it was observed that the use of emoticons showed the tendency of contextually weakening existing assumptions, as illustrated in the Web Chat excerpt below:

(2) <Spasqo>: i'll threat you a drink at Axis!!!
<PunkyBrewster>: you better ;)
(Chat excerpt / Room II, session 4)

In example (2) above, the emoticon ;) (meaning 'winking smiley') need not be implicated, neither does it contradict or eliminate the preceding assumption *you better*, but it also does not strengthen it. What the "winking smiley" in fact does is weaken the statement *you better* in terms of illocutionary force.

In addition, for new information that is (or seems) totally unrelated to the context established, and thus does not perform a contextual effect in the strictest sense, a fifth, non-contextual effect, was formulated. This decision is based on the observation that (seemingly) unrelated information that is contributed to an established context still produces some sort of effect—albeit not necessarily on the communicative context, but on the reader(s), as for example, when a certain message cannot be processed and leads to requests for clarification. How the theoretical concept of contextual effects was applied to empirically investigate data for the purposes of this study will be subject of chapter 5 (*Methodology*) and will also be discussed and illustrated in more detail in chapter 12 (*Contextual effects in personal written communication*).

4.5. Some additional remarks on grammar

Due to reasons of scope, some grammatical (syntactical) aspects that will be investigated in this study, such as "tense" and "mood", have not received a theoretical discussion in this chapter. However, relevant terminology will be introduced and explained as the study proceeds. Furthermore, the main research categories and parent codes of the code system, "tense" and "mood" each constituting such a parent code, will be discussed in more detail in the upcoming chapter (*Methodology*). Reference work for all issues related to English grammar is, if not otherwise stated, Biber et al.'s (1999) *Longman Grammar of Spoken and Written English* (LGSWE), which is undoubtedly one of the seminal works in the field. The LGSWE is particularly useful for the purposes of this study, because it "adopts a corpus-based approach, which means grammatical descriptions are based on the patterns of structure and use found in a large collection of spoken and written text" (Biber et al. 1999: 4) of contemporary American (AmE) and British English (BrE). The LGSWE corpus contains over 40 million words of language from the following four core registers: conversation (BrE), fiction (AmE & BrE), news (BrE), and academic prose (AmE & BrE) (cf. Biber et al. 1999: 24ff.). Since this study is also concerned with features of orality and

literacy in personal written communication, Biber et al.'s (1999) findings and discussions in the field of spoken conversation are of particular interest for this study. They will provide a point of comparison for results that the investigation into the written text corpora of this study generated, particularly in those areas where empirical previous work is rare or not (yet) available.

5. Methodology

In order to be able to analyse the contextuality of personal written communication (on a theoretical as well as empirical level), it first needs to be established which factors influence and, to a certain extent, constitute the contextuality of written text that is produced under particular media-related circumstances. This chapter first presents the main research categories that subsume those factors that are felt to take on a crucial role in the contextualisation of written text (5.1.), to then introduce the analytical framework that was designed based on those research categories (5.2.). The chapter concludes with a section (5.3.) that introduces the main research hypotheses which have been formulated against the background of previous work, theory, and methodological reasoning.

5.1. Main research categories

At the heart of a dialogue is the reciprocal communicative exchange between two or more participants. In spoken conversation, an interlocutor is needed in order to establish a dialogic structure (either in physical proximity or on the telephone). It is quite a different matter with written conversation, where the addressee is physically absent in most cases.[50] Yet even if the (intended) addressee is absent, a written account can still be investigated in consideration of both sender and receiver because the writer constructs the reader in the process of writing, "for the necessary task of interpreting [the] addressee's interpretation of [the addressor's] utterance in order to address that interpretation" (Fitzmaurice 2002: 177). This leads to a contextualisation of the author's writing that takes the (implied) audience into account. This in turn, is believed to be traceable in the written evidence of personal messaging.

Obviously, an empirical investigation into the contextuality of personal written communication first requires the definition of the features that are thought to be constitutive of contextuality (of written text) on a theoretical level. The next step is then to classify these features into different research categories, which form the corner stones of any empirical investigation. For the five types of correspondence investigated in this study, the following six categories are believed to represent the most important formative factors

[50] Since the advent of telecommunication, however, there can be exceptions to this state of affairs. Consider, for example, two people in the same room where it is inappropriate to have a private conversation (for example, in a library) who send each other SMS messages.

with respect to how authors contextualise their messages: (1) "relationship of participants" (who communicates with whom?), (2) "number of participants" (how many participants communicate with each other?), (3) "immediacy of dialogue" (how immediate is the communicative exchange?), (4) "available communication channels" (through what channels is the message communicated?), (5) "reference" (what are the deictic cornerstones of the communicated content?), and (6) "relevance" (in how far is the content of the message relevant to the context established?).

The six research categories are introduced in more detail in Table 5.1 below, and each category will receive more elaborate discussion as we proceed. With respect to Table 5.1, it should be noted that the double line separating EEC from the other types of correspondence signifies that the handwritten letter is produced and transmitted under significantly different technological circumstances than the other types of electronic communication presented in the Table.

Table 5.1: Main research categories (generalised): formative factors on contextuality in a selection of personal written communication.

Research categories	**SMS discourse**	**E-mail**	**Web Chat**	**Personal homepage**	**EEC (handwritten letter)**
1) Relationship of participants	- acquainted - non-anonymous	- acquainted - non-anonymous	- mostly unacquainted - mostly anonymous	- mostly unacquainted - sometimes anonymous	- acquainted - non-anonymous
2) Number of participants / range	in general: one-to-one	optional: - one-to-one - one-to-many	optional: - one-to-one (private Chat) - one-to-many (spin Chat) - many-to-many (spin Chat)	in general: one-to-many	in general: one-to-one
3) Immediacy of dialogue	asynchronous: from minutes to days, depending on time of receipt / urgency	asynchronous: from minutes to weeks, depending on time of receipt / urgency	(near-)synchronous: within seconds	asynchronous: - if webmaster can be contacted: within minutes - if webmaster cannot be contacted: monologue	asynchronous: days to weeks or longer between letters
4) Available communication channels	- written word - upper- and lowercase writing - emoticons - capitalisation for reasons of emphasis	- written word - different fonts/colour - emoticons - upper- and lowercase writing - attachments of any kind (files, pictures, video sequences, sound etc.) - hyperlinks - capitalised/bigger/italicised/ bold writing for reasons of emphasis	- written word - upper- and lowercase writing - emoticons, - hyperlinks - depending on Chat room facilities: - different fonts/colour - sound (icons), capitalised/ bigger/italicised/bold writing for reasons of emphasis	- written word - different fonts/colour - emoticons - upper- and lowercase writing - sound - pictures, (moving) graphics - video sequences - hyperlinks - capitalised/bigger/italicised/ bold and moving writing for reasons of emphasis	- written word - upper- and lowercase handwriting - drawings, sketches - underlined/bigger/italicised writing for reasons of emphasis
5) Reference	- textual deixis: inter-/intra-textuality - personal deixis - spatial deixis - temporal deixis	- textual deixis: inter-/intra-textuality - personal deixis - spatial deixis - temporal deixis	- textual deixis: inter-/intra-textuality - personal deixis - spatial deixis - temporal deixis	- textual deixis: inter-/intra-textuality - personal deixis - spatial deixis - temporal deixis	- textual deixis: inter-/intra-textuality - personal deixis - spatial deixis - temporal deixis
6) Relevance	- contextual implication, strengthening, contradiction/ elimination, weakening of existing assumptions (textual evidence)	- contextual implication, strengthening, contradiction/ elimination, weakening of existing assumptions (textual evidence)	- contextual implication, strengthening, contradiction/ elimination, weakening of existing assumptions (textual evidence)	- contextual implication, strengthening, contradiction/ elimination, weakening of existing assumptions (textual evidence)	- contextual implication, strengthening, contradiction/ elimination, weakening of existing assumptions (textual evidence)

Key: private Chat = term used among Chatters to refer to chatting performed in private, usually involving two Chatters, in a "separate room"; spin Chat = term used among Chatters to refer to chatting performed by all Chatters in the "main room".

The first research category illustrated in Table 5.1 concerns the social relationship of the participants, and it becomes apparent that the nature of the relationship among participants varies considerably across the five types of correspondence. Still, the communicative exchanges that are produced in those different settings are all classifiable as personal, and are for this reason also comparable. Regarding the range of the different types of written communication, shown in research category 2, it emerges that Web Chat has the most diverse range, offering one-to-one, one-to-many, and many-to-many discourses. This is connected to the circumstance that Chatters frequently retreat in pairs to so-called "private rooms", as opposed to the main room where everybody is chatting, to be undisturbed and converse on a one-to-one basis, also referred to as "whispering". Although these one-to-one conversations are without a doubt very interesting, they could not be considered for this research because it would have meant that the researcher herself had turned into an informant. This is the reason why this study looks into Web Chat as performed in the main room, also referred to as spin Chat, generating one-to-many or many-to-many types of discourse, depending on the discourse structure.

All types of personal written communication illustrated in Table 5.1 are of dialogical character in that they are typically addressed to one or more (implied) reader(s). In the case of SMS discourse, e-mail, and EEC, the reader(s) will be known, whereas Web Chat and the contents of personal homepages are in most cases directed at a readership that is unknown to the author and thus to a certain extent implied. In any case, it is the possibility for the reader(s) to be able to respond that qualifies these five types of correspondence as dialogues. The only exception being personal homepages if they do not provide contact details, in which case they would be less of a dialogue and more of a monologue. However, the personal homepages considered for this research all contained contact details at the time of data collection and were thus classified as (potential) dialogues.

With respect to the immediacy of the dialogues, research category 3, it becomes apparent that all types (except Web Chat) are of an asynchronous nature. The period of time that elapses in asynchronous communication between messages may range from minutes to weeks or even longer. Web Chat on the other hand, is classified as (near-) synchronous because all Chatters are present at the same time. Hence, time lapses in between contributions are minimal, which is reminiscent of spoken discourse. Web Chat is usually categorised as synchronous communication because all participants are online at the same time (for example, Crystal 2001: 11). However, the decision to classify it as

*(near-)*synchronous is based on the time lapse between the typing and sending (and maybe re-reading) of a message, before it appears on the screen as uploaded by the server. This terminological differentiation seems adequate in view of the fact that those time lapses do distinguish Web Chat in this respect from a spoken face-to-face interaction. Spoken conversation can in fact be classified as synchronous because all interlocutors are present at the same time, and there will be no time delay between production of a message (i.e. verbally uttering it) and its transmission. However, in all types of correspondence shown in Table 5.1, the immediacy of the dialogue is to some extent media-related, having an influence on how authors contextualise their writings.

Another media-related feature that influences the contextuality of written correspondence concerns the available communication channels[51] that a certain medium provides (research category 4). One communication channel that unites all types of correspondence produced by humans is the written word. But there are various channels beyond the written word to communicate meaning. For example, paralinguistic cues such as tone of voice and facial expressions are immensely important in face-to-face communication. The same string of words has a different communicative meaning if shouted compared to softly spoken. Evidently, written correspondence lacks many of the paralinguistic cues that are available in spoken discourse, but authors do find ways of compensating for those missing paralinguistic cues. Of course, the types of communication channels that authors may employ are also dependent on their availability. For example, while the personal homepage is rich in its available communication channels, SMS discourse is not (as illustrated in Table 5.1). It will be interesting to see to what extent authors of personal homepages actually make use of the available communication channels, and in what ways authors of SMS discourse compensate for the lack of them.

The main research category 5 is concerned with reference and in particular with the degree of inter- and intra-textuality of the different types of communication. An additional aim is to investigate whether or not it is possible to identify a text entity, such as a particular SMS text or e-mail message, as a member of a larger discourse structure based on the textual references it contains. As has been outlined in the theoretical discussion above, deixis, whether textual, personal, spatial, or temporal, not only permeates language at every level, but is also context-dependent and thus constitutive of contextuality. It is for

[51] It should be noted that the terminology *communication channels* as understood for the purposes of this study includes all possible means of expressing verbal or non-verbal language in personal written communication (see hereto also section 4.3.).

this reason that the majority of codes (7 out of 12) in the analytical framework, which will be introduced shortly, are concerned with different types of deictic reference.

Research category 6 is the last category introduced in Table 5.1, and it deals with the notion of relevance in relation to context. Context is of dynamic nature and does not lend itself easily to an empirical investigation. In particular, the notion of contextual effects as brought forward by Sperber and Wilson (1995) is quite complex in its entirety, taking into account how the production and processing of linguistic meaning is connected to mental processes. In order to meet the requirements of an empirical investigation into data of which neither author nor readership is known, the notion of contextual effects, as pointed out previously, has been adapted for the purposes of this paper. However, to my knowledge, RT (or excerpts thereof) has so far not been empirically tested on written communication. Adaptations to the framework (see Figure 5.1 below and section 12.2.1.) were made in the light of these circumstances.

After having introduced the different research categories and factors believed to influence, and to a certain extent constitute, the contextuality of personal written correspondence, I would now like to turn to the analytical framework that was designed for the purposes of this study.

5.2. Analytical framework: code system

In order to be able to investigate the features subsumed under the six main research categories outlined above, a code system that would allow for the systematic investigation of those features was needed. It should be noted that the first two research categories listed in Table 5.1 (1: relationship of participants, and 2: numbers of participants / range) were not included in the code system due to the following reasons. The relationship of participants (research category 1) could not be systematically investigated because in all cases the authors and addressee(s) were unknown to the researcher. Since all material was viewed prior to investigation and classified as personal, it can be assumed that the writer/reader-relationship was of an informal nature at the time of message production for all types of correspondence. The second research category (number of participants / range) was neglected because most of the types of correspondence either have a one-to-one or one-to-many range, Web Chat being the only exclusion by providing the possibility to communicate many-to-many. However, the number of logged-in users was recorded for each session, and it will thus be possible to consider the text in proportion to the number of participants where applicable. The remaining research categories introduced in Table 5.1

(research categories 3 – 6) were carefully incorporated into the coding system, which consists of twelve parent codes and various subcodes that further subsume sub-subcodes (not shown in Table 5.2 for reasons of space).

The twelve parent codes (along with their subcodes) are aimed at investigating twelve main areas in the contextualisation of personal written communication. Through what kind of communication channels the communicative exchanges take place, and how certain content is emphasised, is subject of analysis in parent codes 01 ("communication channels") and 02 ("means of emphasis"). Issues in connection with reference and deixis are dealt with in parent codes 03 ("intra- / inter-textuality", i.e. the use of textual deictics), 04 ("personal deixis", i.e. the use of personal deictics), 05 ("spatial deixis", i.e. the use of spatial deictics), and 06 – 08 ("temporal deixis", i.e. use of the tenses and temporal deictics). More of a grammatical approach is parent code 09 ("mood and syntactic structures"), which investigates how authors of personal written draw on the notions of grammatical mood and syntactic structures to contextualise their writing. Parent code 10 ("contextual effects"), on the other hand, deals with the notion of relevance, and how the content of messages interrelates with the context. The last two parent codes, 11 ("organisation of text") and 12 ("word count"), are aimed at obtaining information on how messages are structured in connection with text length.

Table 5.2 provides a simplified overview of the code system, which will be discussed in more detail below. As has been pointed out above, only the twelve parent codes and corresponding subcodes are shown, the sub-subcodes are not included in Table 5.2 for reasons of space. The complete code system, however, can be found in the appendix (15.1.).

Table 5.2: Code system (simplified).

PARENT CODES		**Subcodes**
01	COMMUNICATION CHANNELS	1.1. written word 1.2. onomatopoeia 1.3. emoticons 1.4. signs 1.5. hyperlink 1.6. photograph 1.7. picture (drawing, physical or digital) 1.8. logo 1.9. graphics 1.10. attachment, enclosure (other than picture, photograph) 1.11. video sequence 1.12. sound (spoken or sung text segment) 1.13. other, borderline case
02	MEANS OF EMPHASIS	2.1. capitalisation 2.2. bigger font 2.3. bold font 2.4. different font 2.5. italicising 2.6. underline 2.7. moving (text) segment 2.8. sound (alert) 2.9. use of colour 2.10. repetition of letter 2.11. repetition of word 2.12. repetition of punctuation 2.13. repetition of other 2.14. different combinations of m.o.e. 2.15. no use of m.o.e.
03	INTRA- / INTER-TEXTUALITY	3.1. acknowledgment of receipt 3.2. ref. to previous writing specified 3.3. ref. to current writing specified 3.4. ref. to upcoming writing specified 3.5. ref. to contact face-to-face specified 3.6. ref. to contact telephone specified 3.7. ref. to writing unspecified 3.8. ref. to oral contact unspecified 3.9. ref. to contact unspecified 3.10. quotations 3.11. discourse markers 3.12. contact postal address 3.13. contact e-mail address 3.14. contact information other 3.15. other, borderline case
04	PERSONAL DEIXIS	4.1. greeting section 4.2. farewell section 4.3. author (other than pron.) 4.4. readership (other than pron.) 4.5. impersonal *you* / *one* 4.6. any other 3. party (other than pron.) 4.7. ellipsis 4.8. selected terms of non-pronominal personal reference 4.9. pronominal ref.
05	SPATIAL DEIXIS	5.1. spatial ref. to (specific) physical places 5.2. spatial ref. to (specific) destinations in virtual spaces 5.3. selected spatial expressions

06-08	TEMPORAL DEIXIS	06 TENSE 6.1. present tense 6.2. past tense 6.3. future tense 6.4. infinitive constructions 6.5. constructions with "let" 6.6. constructions with modal verb (no main verb) 6.7. constructions with periphrastic *do* / *did* 6.8. passives without tensed verb 6.9. other, borderline case 07 CALENDRIAL REFERENCES 7.1. (official) holidays 7.2. decade, century, millennium 7.3. year 7.4. season 7.5. month 7.6. week 7.7. day 7.8. time of day 08 SELECTED TEMPORAL EXPRESSIONS
09	MOOD + SYNTACTIC STRUCTURES	9.1. indicative mood (declaration, interrogation, imperative, exclamation, insert) 9.2. hypothetical mood (declaration, interrogation, exclamation) 9.3. (subjunctive mood) 9.4. other, borderline case
10	CONTEXTUAL EFFECTS	10.1. contextual implication 10.2. contextual strengthening 10.3. contextual contradiction/elimination 10.4. contextual weakening 10.5. non-contextual effect 10.6. other, borderline case
11	ORGANISATION OF TEXT	11.1 greeting section 11.2 farewell section 11.3 number of paragraphs (turns)
12	WORD COUNT	12.1. 0001 – 0010 words (continues in 10-word-steps) (…) 12.6. 0051 – 0100 word (continues in 50-word-steps) (…) 12.9. 0201 – 0300 words (continues in 100-word-steps) (…)

Key: ref. = references; m.o.e. = means of emphasis; pers. = person; pron. = pronoun.

Parent code 01, COMMUNICATION CHANNELS, aims at identifying the different communication channels that are, first, available in the different media and second, investigates their actual employment by the authors of the respective text types. The one communication channel that unites all types of correspondence is, of course, the written word. Yet, similar to a spoken conversation, the contents of a written message can also be communicated beyond the word. The subcodes listed under parent code 01 (subcodes 1.2. to 1.12.) in Table 5.2 can thus be seen as different ways of communicating beyond the written word.

The same is true for the second parent code, MEANS OF EMPHASIS, which is another type of communication channel to communicate beyond the written word. However, in practice, there are so many different realisations of the channel "emphasis", that "means of emphasis" was given its own parent code rather than subsuming it under parent code 01. Thus, the subcodes subsumed under parent code 02 are aimed at analysing what types of means of emphasis, in consideration of media-related constraints, are employed by the authors of the different types of correspondence. The results gained from the investigation into the uses of available communication channels by authors of different text types will shed light onto how strong media-related constraints take effect on the contextuality of personal written communication.

Different types of deixis are the main foci of parent codes 03 – 08. Parent code 03, INTRA- AND INTER-TEXTUALITY, concentrates on how authors integrate textual references into their messages. The scope of those textual references includes references to text segments within the same text entity (*intra*-textuality), as well as to all kinds of texts (both written and spoken) outside the text entity (*inter*-textuality). A further differentiation is made whether or not those textual references are specified, i.e. tied to a particular medium. The main focus is placed on references that concern text (or text segments) produced by author- and/or readership. Another subcode (3.11.) examines the use of selected discourse markers, such as *so* or *well*, so as to find out more about the internal structuring of the five text types.

PERSONAL DEIXIS, parent code 04, examines the different types of personal references as employed by authors for authorial self-address, addressing readership, or the forms of address used to refer to other third parties. Having introduced Levinson's (2006: 111ff.) system of personal deixis in the theoretical discussion in the previous chapter (see also Table 4.3.), I would now like to discuss the modifications that were made to that system for the purposes of this study. It should be noted that for all types of personal reference, the author figure constitutes the deictic origo, accommodating the fact that this investigation analyses texts from the author's point of view.

Levinson's (2006) traditional person paradigm is based on speaker (S) and hearer (A), and terminology has been modified to accommodate the circumstance that written correspondence is concerned with author (A) and reader (R). In addition, references to a third party will be referred to as (3P). Further modifications include the accommodation that person deixis is typically expressed by means of pronouns (P), or other person deictic expressions (O), in most cases nominal (such as *Helen*) or noun phrases (as in *our group*).

The inclusion or exclusion of one or more of those parameters will be indicated by (+) and (-), respectively. This set of symbols will allow expressing all forms of address relevant for this study.[52] For example, a code such as (+A, -R, -3P / +P, -O) denotes a form of authorial self-address that includes the author but excludes readership and any other third party (+A, -R, -3P), and is realised by means of a personal pronoun (+P), which excludes any other reference by means of name or noun phrase (-O). In the case of this particular example, only the first personal pronouns singular would qualify as potential reference (*I*, *me*, *mine*, *myself*). This both facilitates reference to different types of forms of address and avoids lengthy wording. In any event, the investigation into the use of pronominal and other forms of address also takes into account that some messages may not contain forms of address to author and/or readership. Furthermore, the uses of the impersonal pronouns *you* and *one* are also attended to (subcode 4.5.). In addition, a selection of non-pronominal personal references (subcode 4.8.), such as *people* or *everybody*, has been determined in order to be able to compare their frequency across the five text types.

How the notion of space is dealt with in the different types of correspondence is the subject of parent code 05, SPATIAL DEIXIS. Of interest are spatial references in connection with the physical location of the author (at writing time), opposed to spatial references that are oriented towards the text as deictic origo. Special attention is paid to what type of physical (geographical) locations authors refer (subcode 5.1.) and, in the modern types of correspondence, how destinations located in virtual space are identified (subcode 5.2.). Another subcode (5.3.) examines the frequency of selected spatial expressions, such as *here*, across the five text types in consideration of the deictic origo author vs. text.

Parent codes 06-08, TEMPORAL DEIXIS, investigates how authors contextualise their writings by embedding them in a timeframe. There are three main foci: the use of the tenses (parent code 06), calendrial references (parent code 07), and the employment of selected temporal expressions (parent code 08). Parent code 06 looks into different verb phrase constructions in consideration of tense, aspect, voice, and modality. How calendrial references, such as date and time, are employed by authors of personal written communication is subject of parent code 07. Parent code 08 analyses the distribution of selected temporal expressions (unfixed by the calendar), such as *now* and *always*, in order to observe whether or not the immediacy of the correspondence can be linked to certain temporal expressions.

[52] Table 15.2 in the appendix illustrates all possible code variations for author (A), reader (R), and any other third party (3P).

The parent code 09, MOOD, moves away from deixis and into the realm of grammar. It is important to point out that the grammatical field of "mood" has been adapted for the purposes of this study. Often, the notion of "mood" is used solely to refer to the use of modal auxiliary verbs (e.g. *may*, *shall*, *can*) and how they modify the main verbs of a sentence or phrase (cf. Biber et al. 1999: 483ff.). Also, the term *mood* is sometimes used interchangeably with *modality*, a terminology based on the function of modal verbs. However, for the purposes of this study, the scope of reference of "mood" has been extended in that it goes beyond the use of modals. As can be seen in Table 5.2 above, the notion of "mood" has been adopted in this study to refer two different types of propositional meaning as expressed by main verbs and auxiliaries: indicative and hypothetical mood.

The indicative mood expresses, with main verbs and/or in combination of modal auxiliaries, facts and factual conditions, as in *I will drive this car*, or, *I am the owner of this car*. The hypothetical mood, on the other hand, expresses, with the aid of modal auxiliaries, that a proposition is thought to be counterfactual, but otherwise possible (cf. Chung & Timberlake 1985: 243), as in *I would like to drive this car*. There exists, however, a third type, the subjunctive mood. It typically signals potentiality, uncertainty, prediction, obligation, and desire by using the base form of the verb (present subjunctive) as in *I suggest he drive this car* or the past simple form of the verb (past subjunctive) as in *If I were you I would drive this car*, which is also sometimes referred to as "were-subjunctive" (Quirk et al. 1985: 155). Yet it was observed that the subjunctive is practically non-existent in the modern types, and that only EEC features verb phrases in subjunctive mood on a regular basis. It has thus been decided that this particular grammatical mood will receive limited attention (and is therefore also listed in parentheses in the code system above). Hence, the investigation into grammatical mood will be focused on the indicative and hypothetical mood, analysing different syntactic structures (the properties of which will be explained in detail in chapter 11) in relation to their distribution and function.

Of a more experimental approach is parent code 10, CONTEXTUAL EFFECTS, which tries to capture what types of contextual effects are prevalent in the processing of the textual units of the five types of personal written communication. There are two problematic constraints to this undertaking. First, not knowing the authors of the data means no possibility to check whether or not an author in fact intended to generate a particular contextual effect. Second, not knowing the addressee(s) means no possibility to

check if such an (intended) effect was actually performed on the reader or not. Because of these constraints, it was necessary to limit the analysis into the contextual effects to the written evidence (i.e. linguistic context) and base the assignment of a particular type of contextual effect on the textual evidence provided and the comprehension process of the researcher(s), acting as a kind of stand-in for the reader originally intended by the author. This is one of the more crucial adaptations to the concept of contextual effects as formulated by Sperber and Wilson (1995). I will, however, refrain from discussing these adaptations any further at this point as they will be dealt with in detail in chapter 12. For the moment it suffices to know that the coding procedure of contextual effects is connected to the how a textual unit (or symbol) relates to the context established in terms of its content and how the readership is most likely to process this piece of information in terms of relevance. The coding procedure of contextual effects is illustrated in Figure 5.1 below.

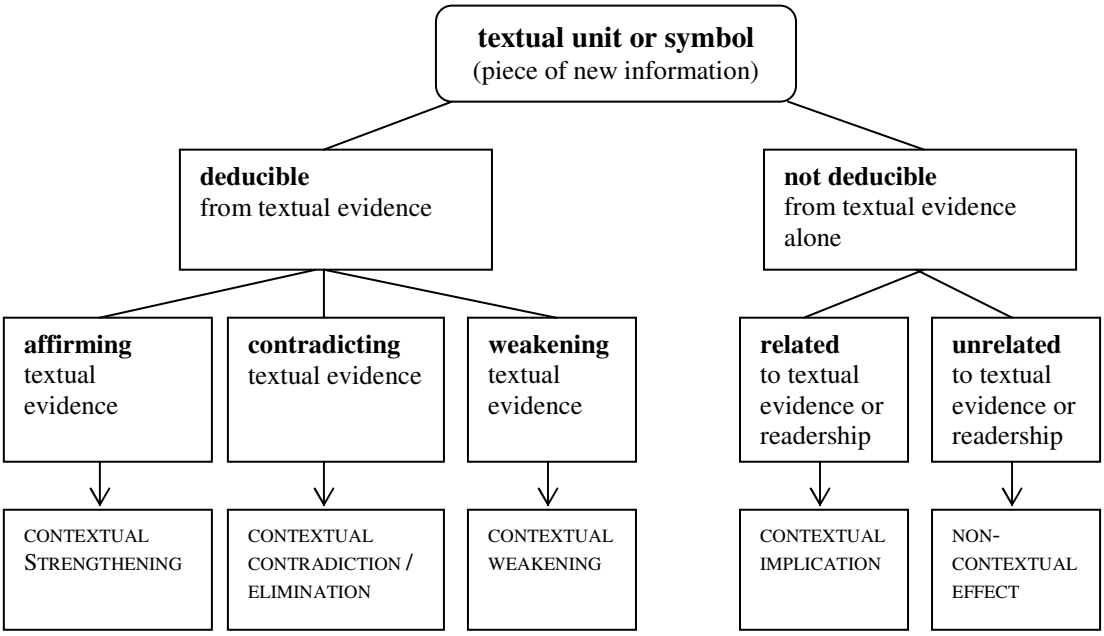


Figure 5.1: Coding procedure "contextual effects".

Apart from the five types of contextual effects illustrated in Figure 5.1, a sixth category "other, borderline case" was formulated for text segments that are not assignable to any of the five types of contextual effects. It was found that the "other, borderline" category was necessary for the Web Chat data. Since data recording started upon my login, the Chat was *in medias res* when data collection began. Thus, on average, it was not possible to classify the first 3 to 7 turns (depending on the number of Chatters and their contributions) into

different types of contextual effects because the preceding textual evidence was missing. Those turns were then coded as "other, borderline case".

Of course, it has to be kept in mind that the text entities of the other four types of personal written correspondence (SMS text, e-mail, personal homepage, and letter) can be, and often are, part of a larger discourse structure. However, the difference being that those text entities are "complete messages" within a larger discourse pattern, which means they can in fact be seen as entities, as opposed to the Chat sessions that are 15-minute-extracts from an ongoing discussion.[53] In any event, the evaluation of textual units and their classification into different contextual effects is tied to the discourse structure of a given communicative interaction. With regard to the five text types investigated in this study, the particularities in connection with the evaluation and classification of textual units will be explained in the discussion of the findings (chapter 12).

Last but not least, Parent codes 11 and 12 are aimed at investigating the structural features of personal written correspondence. ORGANISATION OF TEXT (11) looks into the textual arrangement of greeting section, farewell section, and main body. Of main interest is the question how authors visually structure their writings in the light of media-related constraints. Parent code 12, WORD COUNT, concludes the code system and provides important information regarding text volume, as otherwise none of the results generated by the other codes could be put into quantitative proportion. It is for this reason also that the same word count criteria were applied to each of the text corpora. The two main criteria being that no machine-generated language was considered, such as automatic machine-generated turns in Web Chat, and the same was true for contributions worded by transcribers or annotators of existing corpora, such as the EEC sampler.

5.2.1. Test-coding

Previous to its application, the entire code system was test-coded by three independent coders to check its reliability. The samples for test-coding were selected randomly from each of the five text corpora. It took two test-runs, carried out in June and July 2006,

[53] The idiosyncrasy of Chat correspondence in connection to the notion of "text entity" is the fact that, in the strictest sense, a Chat discussion is completed when everyone ceases contributing. With many Chat rooms being frequented on a twenty-four-seven basis this would result, at least in theory, in "never ending" discussions. However, in practice, Chat room discussions show frequent topic changes where the discussion of one topic by a group of Chatters could be classified as a smaller discourse entity within a larger discussion. Still, it can be and often is quite difficult to determine when exactly a discussion on a certain topic starts and ends. This is the reason why I prefer to speak of Chat "extracts" rather than "entities" when referring to Chat communication.

before the three coders reached an average agreement above 70% for each of the codes across the five text types. The code system was then classified as applicable and used to code all data. Only minor adaptations were made for the investigation into Web Chat to accommodate the fact that several authors, as opposed to one, were "visibly" involved in the text production.[54] Those adaptations all concern parent code 04, PERSONAL DEIXIS, and will be discussed in chapter 7 (*Personal reference*).

5.3. Main research hypotheses

Parts I and II have been aimed at outlining the essential properties of personal written communication and have proposed ways of investigating the contextuality of written text. Part III will now present the findings that the empirical investigation into the five text corpora, based on the analytical framework introduced in 5.2., generated. In order to capture some of the basic tenets discussed in Parts I and II, five main research hypotheses (H1 – H5) have been formulated. H1 – H5 will be critically discussed in connection with the empirical findings in the chapter summaries of Part III.

H1: Computer-mediated communication

Computer-mediated communication (CMC), being immediate in nature and multifaceted in its available communication channels, is hypothesised to be, first, rich in contextual features typical for spoken discourse and orality and, second, to show idiosyncratic features beyond orality and literacy (i.e. features of cyberdiscursivity).

H2: Short message service discourse

Short message service (SMS) discourse, being immediate in nature but limited in its available communication channels, is hypothesised to be rich in contextual features typical for spoken discourse and orality.

H3: Early English correspondence

Early English correspondence (EEC), being asynchronous in nature and limited in its available communication channels, is hypothesised to be rich in contextual features typical for written discourse and literacy.

[54] "Visibly involved" refers to the fact that the different contributions by the Chatters are assignable to the individual Chatters. In comparison, if an author from a one-to-one or one-to-many type of correspondence sought help in formulating the content of his/her message, then this would not be "visible" unless the author quotes the source.

H4: Electronic communication vs. Early English correspondence

Modern types of electronic personal written communication, CMC and SMS discourse, are hypothesised to show considerable differences with regard to contextuality (as expressed in writing) in comparison to an older, i.e. more traditional, form of personal written communication, the Early English handwritten letter.

H5: Acquainted vs. unknown readership

It is hypothesised that the difference of acquainted vs. unknown readership shows in distinctive ways in the contextualisation of personal written communication regardless of the production mode (electronic vs. manual).

PART III: DISCUSSION

6. Written communication: media-related features and constraints

> Each culture and each period has had its own complex economy of writing, a dynamic relationship among materials, techniques, genres, and cultural attitudes and uses.
>
> (Bolter 2001: 21)

The contextuality of personal written communication not only depends on the participants that write and read a message, but is also tied to the medium (or "materials") with which it is composed and transmitted. This chapter is aimed at investigating in how far media-related features and constraints influence the contextuality of personal written correspondence. Of particular interest will be how authors make use of the different communication channels that are available in SMS discourse, e-mail, Web Chat, the personal homepage, and the letter. The chapter opens with a discussion of structural aspects of the five types of correspondence that are subject to different media-related constraints in this respect (6.1.). Moving away from structure, the attention then shifts to the investigation of other available communication channels and their employment (6.2.), to then focus in on one particular communication channel, namely means of emphasis (6.3.). The chapter closes with a summary that highlights the most important findings (6.4.).

6.1. Organisation of text in personal written communication

The decision to include a section on the structural elements of written communication is based on the fact that text structuring devices, as employed by authors, can be classified as means of contextualisation. Also, the structure of a particular discourse may contribute considerably to its classification into a certain discourse type (see hereto also 1.3.1.). This in turn indicates that "discourses are not simply arbitrary collections of utterances" (Kehler 2006: 241), but, on the contrary, compositions of utterances that are, in most cases, carefully assembled by the author in order to meet particular communicative goals. The field of discourse structuring devices subsumes a wide array of linguistic strategies. For reasons of scope, only a few selected strategies will be examined in this study. It should be

kept in mind, however, that they are by no means the only strategies available to structure discourse.[55]

How text is organised in written communication is not only a matter of taste or choice, but is also, to a certain extent, media-related. For example, while the personal homepage gives its users a lot of freedom with regard to the organisation of text, the short message service in mobile phones is limited in this respect. In the following, two particular types of textual structuring will be explored: the incorporation of paragraphs and headings in personal written correspondence, and the arrangement of greetings and farewells (if they are included at all in a message). It should be noted that the treatment of greeting and farewell sections in this chapter is of purely structural character. Their content, however, will be subject of the next chapter (*Personal reference*).

6.1.1. The paragraph as a text structuring device

One way to visually structure text is to incorporate paragraphs. Most interesting about the paragraph is the fact that it does not occur randomly in a running text, but denotes the end of one particular piece of text and the beginning of another. The OED online provides the following definition for the entry "paragraph": "A distinct passage or section of a text, usually composed of several sentences, dealing with a particular point, a short episode in a narrative, etc." If understood literally, this definition would imply that the more paragraphs a text entity contains, the more different topics are dealt with. This is, of course, not always the case, but the above definition by the OED online provides an interesting starting point into the discussion of paragraphs.

From a purely quantitative point of view, the five text corpora show differences with regard to how often paragraphs occur as a text structuring device. It should be noted, however, that it was at the time of data collection not possible to insert paragraphs, that is press an "enter key", in SMS discourse due to limitations of software. Thus, research into SMS texts composed on newer mobile phone models would be required to be able to investigate this particular feature. However, probably also connected to the restricted text length of 160 characters at the time of data collection, many SMS messages in the corpus address only one main issue or topic of discussion, as illustrated in example (3):

[55] See Smith (2003) to learn more about discourse and internal structure of text. Smith (2003: 1) proposes "a local level of discourse, the Discourse Mode, which has linguistic properties and discourse meaning" and posits five modes: "Narrative, Report, Descriptive, Information, and Argument." Smith's (cf. 2003: 258) enquiry into local text structure concentrates on linguistic forms and categories as well as looking at organising principles of discourse (such as hierarchical structure, functional units, and discourse relations).

(3) WHERE ARST THOU?!?
(SMS text / author: female, 24)

The SMS message shown in example (3) is an enquiry into the addressee's whereabouts and, as will be seen in the course of this study, the whereabouts of author and readership seems to be of great importance in personal written communication. Notably, the content of the SMS text in example (3) consists of one topic. In such cases, the definition of "paragraph" by the OED online would hold (meaning that authors get around the issue of paragraph insertion by dealing with only one topic at a time), but since there are also quite a few SMS messages that address more than one topic, it seems the explanation does not seem to be so straightforward.

To come back to the frequencies of paragraphs, it emerged that Web Chat features the highest paragraph-to-word ratio (1:5.2) of all the text types that are not subject to media-related constraints in this respect. In comparison, the personal homepages feature on average one paragraph in every 10.7 words, the ratio in e-mail is 1:31.8, and the lowest frequency was determined for EEC (1:62.9). The paragraph-to-word ratio of 1:5.2 in Web Chat means at the same time that the average turn consists of 5.2 words. Hitting the enter button in Web Chat not only inserts a paragraph, but generates a new turn as well. Although the length of a turn would in principle not be limited by the medium, users have in general a strong tendency to compose short turns. For example Dittmann's (2001: 65) empirical analysis and comparison of German and French Chat showed that the average numbers of words per turn are 4.8 and 5.6, respectively. Thus, short turns that in most cases make up a complete textual unit seem to be characteristic of this medium (the connection between turn and textual unit is re-addressed in more detail in 8.1.2).

One of the main reasons for this is most probably the fact that communicative interactions in Web Chat are subject to speed in two ways. First, the more Chatters participate in a discussion, the more contributions there will be. This also means that if a Chatter wants to engage in a discussion, then he/she has to be quick in contributing in order to contribute something relevant. Second, regardless of how much time a Chatter needs to compose the next turn, the discussion continues. As new turns are added at the bottom of the screen, older ones disappear at the top of the screen. Of course, one can always scroll up or down to search for a particular turn, or even consult the Chat log. This entails, however, missing even more of the ongoing discussion. The only way to successfully participate in a Chat discussion thus really seems to be with short turns that

are quickly contributed. According to the definition of "paragraph" provided by the OED online, the high frequency of paragraphs in Web Chat would thus indicate topic changes with every turn. This is, however, not the case. While topic changes are undoubtedly frequent in Web Chat, it is often difficult to pin down where they occur. Also, a topic that seems to be fully discussed might be picked up upon later in the Chat session, because of an input by another Chatter that has joined the Chat at a later point. It is in this sense that the paragraph-topic-equation proposed by the OED online does not apply for Web Chat.

Interestingly, the text type with the second highest paragraph-to-word ratio is the personal homepage (1:10.7). On the one hand, this is unusual since authors are under no pressure to publish their material. Yet on the other hand, the personal homepage is the medium with the most freedom regarding text arrangement and it seems logical that authors make use of this feature. Many of the 1503 paragraphs counted in the personal homepage data are connected to hyperlinks that stand free from the running text as illustrated in example (4):

(4) My name is <Full Name> and this page was created to let the world know I exist and to make contact with living relatives, and maybe even find some of the dead ones. If you haven't figured it out yet, I am into genealogy, and want to collect information on as many of my family as I can. Also, those of my wife's families. If you have information on any names I have listed, please contact me. (…)

My Family Names

<Family Name 1>
<Family Name 2>
<Family Name 3>
<Family Name 4>

(Personal homepage excerpt / author: Fm_2)

The personal homepage excerpt shown above begins with a paragraph that introduces the author, along with one of his main initiatives to run the homepage. He then lists four hyperlinks that are connected to the endeavour outlined in the introductory remarks, and all of them stand free (see also example (29) below). This is presumably done to attract attention to the hyperlinks as well as facilitate reading. In any event, the large number of hyperlinks found in the personal homepage corpus (a total of 1153 hyperlinks, see also Table 6.6) thus has a considerable influence on this high paragraph-to-word ratio.

Another interesting observation is that all personal homepages arrange their text in linear fashion, which means top-down from left to right hand corner of the screen. In some

homepages, the navigation bars are located on the left side of the screen (most feature them on the top), which breaks up the screen into two windows, resulting in parallel text fields. Other than that the arrangement of text is linear in the personal homepages. However, the (potential) reading process is an entirely different matter. If the reader decides to follow up one or more hyperlinks, which may be presented in a linear fashion, the reading process of the homepage as a whole ceases to be linear. With respect to text arrangement, the homepages in the corpus show a strong tendency to include paragraphs in order to visually break up text. This usually serves the incorporation of hyperlinks, or the systematic placement of information, such as family background and personal interests, in different text formats.

Of the modern text types, the e-mail corpus features the smallest number of paragraphs per words (1:31.8). If put into the relation that thirty words make up about two to three lines of an A4 page, this is still quite a high paragraph-to-word ratio. This means that authors of e-mail seem to prefer to break up their messages into smaller chunks, so as to not overwhelm their recipients with a screen full of block text. This would attend to the widely held view that the entire e-mail message should be visible within one screen. The majority of messages in the e-mail corpus follow this "unwritten law" as most contain 50 to 300 words, which is perfectly visible in one screen view. But there are also a few texts that are rather long (3 e-mails of more than 800 words). The longest e-mail message in the corpus counts no less than 1065 words (as a comparison, the previous page contains 425 words), which is unusually long. Interestingly, the majority of the letters from the EEC corpus turned out to contain between 100 and 400 words, and are thus comparable to e-mail in this respect. While the word count is comparable, the number of paragraphs per text entity is lower in the letters when contrasted with e-mail. Overall, however, it can be asserted that the paragraph is a typical text structuring device in personal written correspondence. It is in fact so frequent in the Web Chat corpus, it almost has a contrary effect: rather than being of a structuring nature, it breaks up discussions and makes it more difficult to assign turns to each other (cf. hereto also example (51b.) in chapter 7).

6.1.2. The heading as a text structuring device

While paragraphs are frequent in all of the text types, headings in the function of titles are not. Although it could be argued that greeting sections, separated by paragraph, also classify as headings, they do not qualify as a typical heading because they do not front the following text passage in terms of content, but front the whole message in the sense of an

"opener"—regardless whether or not the message covers different thematic aspects. Headings are in the majority of cases separated by paragraph from the text passage for which they serve as some sort of title. To accommodate the fact that SMS discourse is subject to media-related constraints in this respect, punctuation aimed at compensating for the lack of this feature, such as the colon, was also considered.

Instances where a certain expression could be identified to front a thematically correspondent unit of text in SMS messages are few. A total of eight occurrences of "p.s." ('post scriptum') were observed, and another 15 instances of other expressions that front another unit of text, as illustrated in example (5):

(5) (…) I'm at hmv, tons of dvds on sale for 8pounds: monthy python's the meaning of life,true romance,the jackal, goodfellas… (…)
(SMS text excerpt / author: female, 28 yrs)

The expression *tons of dvds on sale for 8pounds* in example (4) serves as a kind of fronted heading for the listing of a selection of those DVDs. Text structures of this type are rare in SMS discourse and are most probably related to software limitations and restrictions in text length.

SMS discourse is, however, not the only corpus where headings are infrequent. It was observed that headings, apart from the personal homepage corpus, are extremely rare in all of the text types. Both the Web Chat and EEC corpora do not contain one instance of a heading, and the e-mail corpus features a mere total of four headings. Only the personal homepage frequently features this text structuring device. A total of 296 headings were found (headings that denoted a "welcome section" and thus fronted the whole text entity, rather than a thematic unit, were, similar to the greeting section, not considered). This means that quantitatively, each homepage contains an average of 5 headings, which stands in stark contrast to the other text types in terms of frequency. All headings found in the homepage corpus front thematically correspondent text, often in combination with a selection of hyperlinks. Example (6) shows the headings found in one of the personal homepages, each of them separated from the information it fronts (summarised in square brackets) by paragraph.

(6) **Vital Stats**
[demographics, no hyperlinks]
Other bits
[likes and dislikes, no hyperlinks]
Why not have a look at:
[4 hyperlinks to different compilations of photographs]
Wish I:
[personal wishes for the future, no hyperlinks]
But at least I have already:
[a list of things the author has already done, no hyperlinks]
(Personal homepage excerpt (headings), emphasis original / author: Bf_3)

The emphasis (bold font) shown in example (6) is both original and very typical. Most of the 296 headings in the homepage corpus are emphasised by means of bold font, italics, colour, and/or bigger font. It can be asserted that the text structuring device "heading" is typical for the personal homepage, as opposed to the other text types, where headings are infrequent or altogether absent.

6.1.3. Greetings and farewells

Greetings and farewells are of particular interest for a study into the structure of personal written correspondence, because, if at all, it is *interpersonal* communication (as opposed to other genres of writing) that opens with a greeting and/or closes with a farewell section. More conventionalised forms of written correspondence, such as the business letter, for which templates have been suggested in order to standardise its form (cf., for example, Stegbauer 2003: 87), traditionally features greeting and farewell sections that are separated from the main body of the text. Personal written communication, on the other hand, is different in this respect because it is less standardised. Nevertheless, EEC, albeit personal in its character and content, is hypothesised to show more features of a conventional letter due to the time and age it originates from. In this particular case, this would mean that EEC is assumed to feature greetings and farewells that are separated from the main body of the text, as opposed to the modern text types, which are expected to be less conventionalised in how greetings and farewells are incorporated, if at all, into the messages. Table 6.1 shows the distribution of greeting and/or farewell sections across the five text corpora. It should be noted that the text type Web Chat is not included in Table 6.1. The reason for this is that the discourse structure of Web Chat is considerably different compared to the other text types. The frequency of greetings and farewells identified in the Web Chat data will be discussed separately below.

Table 6.1: Distribution of greeting and/or farewell sections in personal written communication (corpora-based results).

	SMS discourse (1000)	**E-mail** (140)	**Personal HP** (60)	**EEC (letter)** (91)
+GS, +FS	* 261 ** 26.1	106 75.7	7 11.7	88 96.7
+GS , -FS	85 8.5	9 6.4	24 40.0	0
-GS, +FS	332 33.2	20 14.3	4 6.7	2 2.2
-GS, -FS	322 32.2	5 3.6	25 41.6	1 1.1

Key: +GS/-GS = greeting section (GS) included (+), not included (-); +FS/-FS = farewell section (FS) included (+), not included (-); HP = homepage; * = total number of text entities with indicated characteristics per text type; ** = percentages (rounded to 1 decimal) are given in relation to the total number of text entities per text type (in brackets).

Table 6.1 highlights several trends regarding the inclusion or exclusion of greeting and farewell sections. First, SMS discourse has a tendency to omit greetings. Second, the majority of e-mail messages tend to contain both greetings and farewells. Third, personal homepages are likely not to feature farewells, and fourth, letters from the EEC corpus contain in the vast majority of the cases a combination of greetings and farewells.

Thus, the assumption that the oldest text type among the corpora features the most conventional structure with regard to the incorporation of greeting and farewell sections was confirmed. This is further supported by the online survey where more than 90% of the informants stated that they always include both greetings and farewells in their letters. Interestingly, the modern text type e-mail comes a close second and is reminiscent of the handwritten letter in this respect. It is a different picture with SMS discourse, which shows a strong tendency for the omission of the greeting section, but includes the farewell section in more than a third of the messages. This is confirmed by Schmidt and Androutsopoulos (2004: 63) with regard to the greeting section (no quantitative or qualitative analysis into the farewell section was carried out), as they found only 26 greeting sections it a total of 703 SMS messages. With respect to the SMS corpus of this study, it was also observed that another third of SMS texts do not feature either greeting or farewell. The omission of these sections can thus be seen as a rather typical feature of this type of correspondence.

This is probably connected to the circumstance that communicative exchanges held via the short message service tend to involve several SMS texts that are sent back and forth quickly. Users are thus likely to reply to an incoming SMS message as soon as they see it, which may prompt another response and so forth (see also Androutsopoulos et al. 2001: 17, Schlobinksi et al. 2001: 30, Schmid & Androutsopoulos 2004: 53). This may account for the frequent skipping of greetings, because it would seem odd to include them every

time in a longer exchange. However, it is obviously a different matter with farewells, which could also be omitted, if one knew that the dialogue were to continue. And this is where an idiosyncrasy of SMS discourse emerges: while users are likely to reply to incoming SMS texts without a greeting section, they have a tendency to indicate the end of their part of the communicative exchange by including a farewell section. This is supported by the findings from the online survey, where the peak answer to the question how often users felt they include farewell sections in their SMS discourse was in fact "always".

The personal homepages also show an idiosyncratic employment of the greeting and farewell section: out of 60 homepages, 24 (40%) feature greetings but miss a farewell section. Another 25 (41.6%) do not contain either a greeting or a farewell section. Also, the majority of greeting sections features the word *welcome*, and without a farewell this invites the visitor to stay and surf around, rather than reading the homepage and leave again. This is further confirmed by the fact that if farewell sections are included in the personal homepage corpus, they encourage the reader to come back again, as illustrated in examples (7) and (8):

(7) Hope to see you again soon!
(Personal homepage excerpt / author: Of_1)

(8) Thanks for visiting Hope you enjoy yourself! ***Don't forget Come Back Soon!***
(Personal homepage excerpt, emphasis original / author: Ff_2)

However, the low numbers of farewells in personal homepages could also be related to the homepage being, with respect to the reading process, a non-linear text type. This means that the author does not know at what point in the homepage the reader decides to leave by following a hyperlink, or because he/she loses interest. In the light of these circumstances, the author might make the decision not to include a farewell section in the first place.

As indicated above, Web Chat is an entirely different matter and with regard to frequencies of greetings and farewells not comparable to the remaining text types. This is connected to two main reasons: first, the Chat sessions are 15-minute-extracts from ongoing discussions, during which other Chatters logged in and out again, and second, several authors (as opposed to one) were involved in the discussions, which means that there were also more possibilities for Chatters to greet or say goodbye to each other. A total of 487 greetings were found in the 30 Web Chats, which translates to one greeting in every 58.3 words. It is a different picture with the farewell section as only 93 farewells were counted in the whole corpus, resulting in a ratio of one farewell in every 305.2 words.

This bias could be connected to the fact that Chatters in general say hello to everyone who enters the room but tend to omit the farewells, or, Chatters generally stay longer than 15 minutes in Chat rooms. Whatever the reason, it was observed that the number of greetings per Chat session are proportionally connected to the number of Chatters that were chatting at the time, but the farewell sections are not (too few). Yet, as example (9) illustrates, the possibility to say goodbye does seem to be of importance to certain Chatters:

(9) <lady_grey>: bye <zong>
<lady_grey>: hate it when I miss sayin bye lol
(Chat excerpt / Room IV, session 4)

The statement *hate it when I miss sayin bye lol* by the Chatter named <lady_grey> would support the above assumption that Chatters do not omit farewell sections conventionally, but have a tendency to stay longer than fifteen minutes in Chat rooms. However, Runkehl et al. (cf. 1998: 93) observed a similar tendency in a sample of German Chat, where greeting sections are also strikingly more frequent than farewells. Their conclusion regarding this bias is that Chatters are intent on finding Chat partners and thus greet everyone that qualifies upon entering the room, but if they are ignored, then they tend to leave the room again without a farewell. According to Wirth (cf. 2005: 78), the greeting enjoys more prestige in Chat rooms than the farewell. This means that the more greetings a Chatter receives upon login, the more welcome he/she is to the community. In any event, this feature is characteristic for the discourse structure of Web Chat.

Next to the quantity of greetings and farewells, another aspect was investigated, namely their structure. These sections can be incorporated in various ways into a message (if they are incorporated at all). They may stand free from the main body (separated by a paragraph) of the text, or start separated from the main body of the text and continue into the main body, or be altogether part of the main body. Table 6.2 below illustrates the results gained from the investigation into the structure of greetings and farewells in the five text corpora. Percentages are given for the frequency of one particular type of greeting or farewell in relation to the total number of greetings or farewells per text type.

Table 6.2: Structure of greeting and farewell sections in personal written communication (corpora-based results).

Greeting section	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
separated by PG from MB	x	86.9	90.6	77.4	73.6
starts separated by PG and continues into MB	x	0.9	0.6	3.2	0
not separated by PG from MB	* 100	12.2	8.8	19.4	26.4
Farewell section	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
separated by PG from MB	x	89.7	71.9	66.6	36.3
starts in MB and is separated by PG at some point	x	4	14.6	16.7	60.4
not separated by PG from MB	100	6.3	13.5	16.7	3.3

Key: PG = paragraph; MB = main body (of message or turn); HP = homepage; x = insertion of paragraphs not possible due to media-related constraints; * = percentages (rounded to 1 decimal) are given in relation to the total number of greetings and farewells found in each of the text types.

Table 6.2 demonstrates one main trend in all the media that allow for the insertion of paragraphs: to separate the greeting section by paragraph. Since greetings (and farewells) are one of the characteristics of personal written communication, this means that if authors decide to include a greeting, they are very likely to visually separate (and thereby identifying the message as interpersonal) from the main body of the text. As SMS discourse was subject to software limitations at the time of data collection, no quantitative investigation into the use of paragraphs could be carried out. However, the vast majority of 93.3% of the greeting sections in SMS discourse are separated from the main body of the message by either punctuation or symbols, such as emoticons. This shows that authors of SMS discourse clearly have an inclination to visually isolate the greeting section from the main body of the text. Most of the greeting sections in SMS discourse are separated by means of the exclamation mark or the comma, as illustrated in examples (10) and (11) below.

(10) Hi <Nickname>! (…)
(SMS text excerpt / author: female, 23 yrs)

(11) Hey babe, (…)
(SMS text excerpt / author: male, 31 yrs)

One interesting observation concerning the visual separation of the greeting or farewell sections in SMS discourse is the fact that the messages, where the greeting or farewell is not separated from the main body of the text, turned out to contain, in the majority of the cases,

no further information that contrasts from the greeting or farewell. Thus if the paragraph is defined to comprise "a particular point" (OED online), then SMS discourse, by substituting the lack of paragraphs with punctuation, accords punctuation a similar effect. It was observed that punctuation is most likely to be omitted when the whole SMS message contains only one main piece of information. This implies that in those cases where greetings and farewells are not separated by punctuation from the main body of the messages, they are in fact the "particular point" of the whole message: to say "hi" or "bye". Examples (12) and (13) illustrate instances of such a "hi" and "bye" message, respectively.

(12) good morning and have a nice day :-)
(SMS text / author: male, 33 yrs)

(13) until soon <nickname>
(SMS text / author: female, 23 yrs)

Notably, this type of message was not found in any of the other discourse types and is thus idiosyncratic to SMS discourse.

Regarding the structure of the greeting section in personal written correspondence in general, it can be asserted that if authors decide to open their message with a greeting, then they are very likely to do so in a conventional fashion by separating it from the main body of the text. In fact, the most traditional of the five text types, the handwritten letter, shows the biggest deviation from this general tendency as about a fourth (26.4%) of the greeting sections are not separated from the main body by paragraph. However, similar to SMS discourse, those greetings are then isolated from the main body of the text by punctuation. These results are confirmed by the online survey, where the question how often the informants thought they separate greeting sections (if they included them) in their personal correspondence by either paragraph or punctuation, generated the peak answer "always" for each of the five media.

With regard to the farewell section, the majority of farewells in e-mail, Web Chat, and the personal homepage are separated by paragraph. Similar to the greeting section, SMS discourse compensates for the lack of this feature by isolating the majority of farewells (91.3%) by means of punctuation (mainly exclamation marks and commas) from the main body of text. Only the letter shows a different trend in that the majority of farewells (60.4%) start their pre-closing formulas in the main body and then separate the remaining farewell by one or more paragraphs at a later point:

(14) (…). So in haste, going to bed, we humblie take our leaves and rest
Your Majestie's most humbe and obedient sone and servant
Charles
Paris the 22 of Feb. 1622/23
(EEC excerpt (farewell section), ca. 1623 / letter collection Charles, author: male)

In example (14), the author, Prince Charles, first writes about a dance that the Queen had attended, to then, all of a sudden, switch to the pre-closing formula *So in haste, going to bed, we humbly take our leaves and rest*, to then insert two paragraphs to visually isolate the remaining farewell. As almost two thirds of the letters are closed in this style, it is strongly assumed that this is connected to writing conventions operative at the time.[56] According to the 21st authors of personal written communication who answered the online questionnaire, they always separate the farewell section by paragraph or punctuation in their personal messages.

Thus, the five types of correspondence vary considerably in the frequencies in which messages feature greetings and/or farewells, but if they are included, then they tend to be separated from the main body of the text by paragraph or punctuation. This is interesting in connection with contextual issues because it shows that authors of all the discourse types investigated, both modern and traditional, are interested in visually structuring their correspondence and by including greetings and farewells, they at the same time identify their writings as *inter*personal communication.

6.2. Available communication channels and their employment

Different types of media offer different types of communication channels. However, just because there are a certain number of communication channels on offer, this does not mean authors automatically make use of all of them. Having control over this matter distinguishes the authors of written communication from the participants of a spoken interaction in a face-to-face setting. For example, paralinguistic cues in a face-to-face conversation, such as gestures and mimic, will inevitably be noticed by the interlocutor. Furthermore, notions of origin, gender, and bodily appearance also come into play if communication is performed in physical proximity—whether this is desired or not.

[56] According to Bergs (2004: 209, quoted from Ferguson 1994: 21), most letters in Early Modern English "have a clear, identifying internal structure, differentiated from other message types in the repertoire of the community." It follows that "they are commonly realised with very fixed formulae and structures" and "thus leave very little room for personal choices" (Bergs 2004: 209). However, opposed to this view, Nevala (2004: 274) argues that "although one would think that people in early England obediently followed whatever form of instruction they could get in the art of letter-writing, in reality there was room for personal preference, or, as Irving (1955: 16) expresses it, 'sturdy individuality'." Obviously, opinions differ among scholars on this subject matter.

Thus, meeting in person and talking with each other involves many different communication channels that are inevitable rather than subject to choice. Writers, as opposed to speakers, are in the position to choose a particular medium to compose and transmit their message. They have thus more freedom with respect to the communication channels they want to take advantage of. As has been mentioned previously, the one communication channel that is shared across all five types of correspondence investigated in this study is the written word. The question is what other types of communication channels do authors of personal written communication make use of?

6.2.1. Emoticons

The term *emoticon* is a blend of the terms *emotion* and *icon* and emoticons serve to express emotions, or states of mind, beyond the written word by emulating facial expressions. In most cases, those "facial glyphs" are to be viewed sideways, as for example, the "smiley" which is also referred to as "basic" or "classic" smiley: :-) or (-: (one can see a "smiling face" if one tilts the head to the left or right, respectively). More recent uses include variations of emoticons that imitate typical shapes of heads, often combined with "facial accessories" such as facial hair, as well as other items such as hats and glasses, as for example: o-(:-{o{{{{ ('Santa Claus'). A more recent phenomenon is the use of emoticons that do no longer require head tilting, such as variations of the basic smiley: ^_^ (a somewhat more wry smile than that of the classic smiley, but still classified as a smile among users).

In any case, emoticons find their origin of use in the advent of CMC and have also been welcomed into the language of SMS discourse. They came into existence to compensate for the lack of emotion in CMC, which, in its beginnings, did not have many features (such as colour or different fonts) to express paralinguistic content. As is the case with the "Santa Claus emoticon" above, which does not exactly stand for a particular type of emotion, emoticons may also be employed to replace expressions and convey creativity. Although there are literally hundreds of different emoticons designed to express all kinds of emotions, states of mind, or typical head shapes,[57] the investigation into the different text corpora showed that most authors in fact make do with a selected few. Table 6.3 gives an overview of the most frequently used emoticons in each of the five text corpora. They

[57] The website http://www.chatropolis.com/emoticons.html [28.10.2008] gives an impressive overview of different emoticons that are listed alphabetically according to the emotion, state of mind, or head shape they are supposed to convey or portray (for example, under "S" one finds *Santa Claus*, *Sarcastic*, *Scared*, *Scuba diver*, and so forth).

include :-) ('smile'), ;-) ('smile + wink'), :-P ('tongue out'), :-o ('surprise'), :-* ('kiss'), :-(('sad') and :-/ ('wry smile'). It should be noted that only the "basic emoticon" for each of the types listed is shown in Table 6.3 and that variations of each type are subsumed under the basic type. This means that the category of the basic emoticon :-) ('smile') also subsumes its variations, such as, =), :o), :-D, :), and so forth. The same applies to all other types of emoticons listed in Table 6.3.

Table 6.3: Most popular emoticons in personal written communication (corpora-based results).

Types of emoticons (basic shape)	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
:-) 'smile'	* 270 ** 71.6	42 85.7	145 47.2	9 100	0
;-) 'smile + wink'	71 18.9	5 10.3	32 10.4	0	0
:-P 'tongue out'	0	0	54 17.6	0	0
:-o 'surprised'	3 0.8	0	31 10.1	0	0
:-* 'kiss'	22 5.8	1 2.0	0	0	0
:-('sad'	8 2.1	0	15 4.9	0	0
:-/ 'wry smile'	0	1 2.0	3 1.0	0	0
other	3 0.8	0	27 8.8	0	0

Key: HP = homepage; * = total number of emoticon with indicated characteristics per text type; ** = percentages (rounded to 1 decimal) are given in relation to the total number of emoticons found in each of the text types.

As emoticons are a recent phenomenon, the older form of correspondence, EEC, does not contain any types of emoticons. With respect to the modern text types, the personal homepages stand out because the only emoticons that were found are the classic smiley or variations thereof. Furthermore, the total number of emoticons found in the personal homepages amounted to a mere 9, which indicates that emoticons are not a typical feature of personal communication performed via this medium. Overall, the classic smiley, :-) ('smile'), is by far the most popular emoticon, followed by the emoticons ;-) ('smile + wink'), and :-P ('tongue out'). Although sticking the symbolic tongue out seems to be popular among authors of Web Chat only, as none of the other text corpora contains this type of emoticon. Also not very frequent are the emoticons :-o ('surprised'), :-* ('kiss'), :-(('sad'), and :-/ ('wry smile') with each having a frequency of below 6% in relation to the total number of emoticons.

If the total number of emoticons are put into relation to the total number of word counts for each of the corpora, it emerges that they are not a very frequent feature of most of the text types investigated. The exception is the SMS text corpus, where the ratio between total number of emoticons and total number of words amounts to 1:49. Taking the average SMS message length into account (18.4 words), this translates to an average of every second to third SMS text containing an emoticon, as illustrated in example (15):

(15) I am finally home… thanks for the perfect time table, dinner beers, etc. :-) <name> (SMS text / author: male, 27 yrs)

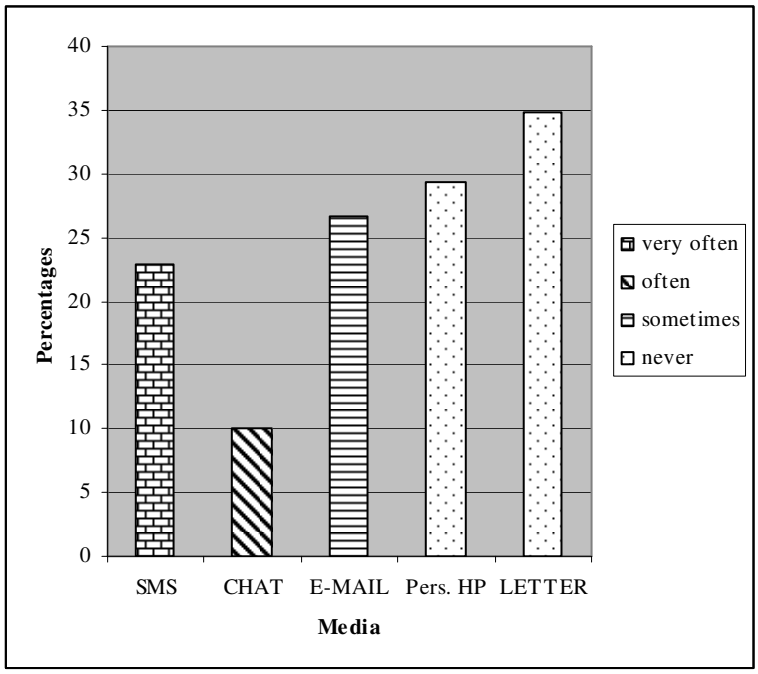
The author in the SMS message shown above lets the addressee know that he has reached home and is thanking the recipient of the message for dinner and so forth. He concludes the message by inserting an emoticon of the type "classic smiley", which supports the content of the message in affirmative and joyful manner. Example (15) is prototypical in that more than 70% of all emoticons found in the SMS corpus are variations of the classic smiley or a variation thereof.

It is quite a different picture with the emoticon-to-word ratio in the remaining text corpora: Web Chat features a ratio of 1:93 (an average of 10 emoticons per Chat), emoticons in e-mail result in a ratio of 1:525 (one emoticon in every third e-mail), the personal homepage has a ratio of a mere 1:1781 (translating into every sixth homepage containing an emoticon), and the EEC corpus, as mentioned above, does not contain any emoticons at all. Notably, Runkehl et al. (1998: 96) come to a similar conclusion concerning the use of emoticons in a sample of German Chat. The corpus they investigated featured a distribution of one emoticon in every 100 words with the "standard smiley", as they call it, being by far the most frequent type of emoticon used. This is further confirmed by Dittmann (2001: 73), whose investigation into German and French Chat yielded similar results. Furthermore, Schmidt and Androutsopoulos (cf. 2004: 53) point out that in their empirical investigation of 703 German SMS texts, emoticons occur "every now and then", thereby indicating that they are in fact not very numerous. Thus, although the various versions of different emoticons available could be used to compensate for the lack of paralinguistic cues, and would in this sense be reminiscent of spoken discourse, the corpora investigation showed that they are not overabundantly frequent. If emoticons are featured at all, then they are most likely to be variations of the classic smiley.

Interestingly, there is a more or less gradual decrease of emoticon-to-word ratio from SMS discourse to Web Chat, to e-mail, and to the personal homepage, suggesting that the desire of authors to express emotion via "facial glyphs" decreases as the potential delay between sending a message and its receipt (and potential response) by the addressee(s) increases. The mobile phone is often expected to be with the owner at all times and thus an SMS message sent by the addressor stands good chances of being read quite quickly upon receipt. The same is, of course, true of Web Chat where all participants are online at the same time. E-mail, on the other hand, might be read quickly as many people check their e-mail accounts several times a day, but it may also experience delays due to night time or problems connected to network and accessibility. The updates in personal homepages may be read soon after being published, but in most cases there will be a certain delay between the publication of new material on a homepage and the moment until it is being read. In the most extreme case, it will not be read at all if another, more recent update, has replaced it in the meantime. The use of emoticons, however, may also be related to the type of readership. Because authors of personal homepages write for complete strangers they may hesitate to include emoticons in their writing, as opposed to SMS discourse or e-mail. But Web Chat, which also takes place between strangers, features them more frequently than e-mail and deviates in this respect from the personal homepage. Thus, a connection between decrease of emoticon-to-word ratio and the increase of asynchrony of the different types of communication seems plausible.

Furthermore, the results gained from the online survey confirm the results from the corpora analysis. The 109 informants were asked how often they thought they use emoticons in the five different types of personal correspondence and they could choose between the given answers "always", "very often", "often", "sometimes", "rarely", "never", and "no answer" if they did not use the respective medium to communicate at the time the survey was taken. It should be noted that for the media Web Chat and personal homepage, the peak answers were in both cases "no answer", but these have been neglected in favour of the runner-up peak answers so as to be able to illustrate what the majority of informants answered who in fact used Web Chat and the personal homepage to

communicate at the time.[58] Graph 6.1 below illustrates the peak answers given by the respondents to the question how often they thought they use emoticons in their personal written correspondence.



Key: Pers. HP = Personal homepage

Graph 6.1: Use of emoticons in personal written communication (results from online survey, peak answers in percentages).

With regard to frequency of use of emoticons, a picture similar to the corpora-based results emerged. While the majority of informants said that they incorporate emoticons very often in their SMS discourse, the peak answer for Web Chat was that they use them often. Users of e-mail communication felt they sometimes use emoticons, as opposed to the majority of personal homepage and letter authors, who stated that they never make use of emoticons in their writing.

It appears that maybe emoticons are not as typical a feature of modern communication as has been suggested by other researchers (among others, Dittman 2001, Schlobinski et al. 2001, Dürscheid 2002a/b) and that next to not being frequent, their use is

[58] Just above 50% of the informants stated in the first question that they never use Web Chat and 53% said the same of the personal homepage. As informants were required to chose "no answer" for all remaining questions in relation to the medium or media they do not use, peak answers for Web Chat and personal homepage were inevitably "no answer" for all remaining questions (because the percentages for "no answer" exceeded 50% in both cases). It is for this reason that, if not otherwise indicated, the runner-up peak answers will be considered throughout the whole study when referring to Web Chat and the personal homepage.

also coupled to the immediacy of message transmission and potential response, as well as the type of readership that is addressed. On a different note, maybe the investigation into the use of emoticons would have yielded other results if older data had been analysed. In particular the language of CMC (as opposed to SMS discourse) can be expected to show more frequent uses of emoticons in its early days. In the beginning of 1990, it was all of a sudden possible to communicate in written fashion with the potential immediacy reminiscent of speech. It can thus be assumed that users took full advantage of the new features, in particular in their personal correspondence. Nowadays, users are more relaxed with the medium and see it as just another way to communicate, rather than a revolution. This might have a considerable influence that the use of emoticons has a tendency to be regressive. Furthermore, that the SMS texts in the corpus used for this study feature emoticons more frequently than the other types of communication, can probably be associated with the restriction to 160 characters per SMS message and the intention to economise on text length (for example, the classic smiley, :-), takes up less space than, for example, *this makes me smile*, or even the word *smile*). It would be immensely interesting to see whether or not newer SMS messages of unlimited size would show a decrease in the use of emoticons.

6.2.2. Onomatopoeia

While emoticons emulate facial expressions or looks, onomatopoeia is a rhetorical device that is aimed at imitating the sounds associated with the objects or actions it refers to. Well known examples are animal sounds such as, *meow* (cat) or *oink* (pig). Yet it also applies to terms that denote everyday objects such as, *buzzer* (imitating a humming, itself an onomatopoeic expression, sound) and *ping pong* (mirroring the sound the ball makes when going to and fro between players). Also, onomatopoeic expressions are frequently used in writing to emulate certain sounds of spoken language, or imitate paralinguistic signals, such as surprise and laughter, and are in this sense heavily oriented towards spoken language and orality. Since it was hypothesised that both CMC and SMS discourse are rich in contextual features typical for spoken language and orality (main research hypotheses H1 and H2), they should both feature onomatopoeia on a frequent basis. However, it was observed that onomatopoeic expressions are overall even less frequent than emoticons, which are not very frequent themselves.

The sounds that onomatopoeic expressions attempt to simulate can be categorised into different types. Based on the occurrences of onomatopoeia in the five discourse types,

the following seven categories of "sound imitations" were determined: (1) "imitating laughter" (such as *hahaha*), (2) "contentment, elation" (as in *mmh* in the sense of 'nice' or 'tasty'), (3) "discontentment, anger" (as in *grrrr*), (4) "surprise, astonishment" (for example, *aha*), (5) "incertitude, hesitation" (such as *mmh* in the sense of 'not sure')[59], (6) "sympathy" (as in *aaaaw*), and (7) "other" (subsuming other imitated sounds such as *booom*). The most popular sounds to be imitated in personal written communication are those of laughter and surprise or astonishment. Table 6.4 below illustrates the results that the investigation into the uses of onomatopoeic expressions generated in more detail.[60]

Table 6.4: Use of selected onomatopoeic expressions in personal written communication (corpora-based results).

Types of onomatopoeia	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
imitating laughter (e.g. *haha, hehe, hihi*)	* 15 ** 1:1228.4	8 1:3216.6	72 1:394.5	0	0
contentment, elation (e.g. *mmh* ('nice'))	8 1:2303.3	2 1:12'866.5	15 1:1893.6	1 1:16'030	0
discontentment, anger (e.g. *grrrr*, *eeeeew*)	4 1:4606.5	5 1:5146.6	32 1:887.6	0	0
surprise, astonishment (e.g. *oh*, *aha*)	2 1:9213	3 1:8577.6	81 1:350.6	1 1:16'030	0
incertitude, hesitation (e.g. *mmh* ('not sure'))	1 1:18'426	5 1:5146.6	37 1:767.7	1 1:16'030	0
sympathy (e.g. *aaah*, *aaaw*)	0	0	14 1:2028.9	0	1 1:31'077
other (e.g. *whew*, *boom*)	3 1:6142	1 1:25'733	27 1:1052	3 1:5343.3	0
TOTAL	**33** **1:558.4**	**24** **1:1072.2**	**278** **1:102.2**	**6** **1:2671.7**	**1** **1:31'077**

Key: HP = homepage; * = total number of instances of indicated type of onomatopoeia per text type; ** = indicated type of onomatopoeia-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

With respect to the frequency of onomatopoeic expressions, Table 6.4 shows that Web Chat is the only text type that features them on a regular basis, approximately one onomatopoeic expression per 100 words. Most frequent in the Chat data are sound-oriented

[59] The difference of meaning for the onomatopoeic expression *mmh*, classifiable as pertaining to either category (2) or (5), was determined depending on context. For example, whenever *mmh* occurred in a context of delight or appetite, it was classified as a member of category (2) "contentment, elation" (e.g. *mmh I love cake very much*). However, if *mmh* was used in a context of incertitude or a kind of hesitance (often coupled with three or more full stops), it was classified as a member of category (5) "incertitude, hesitation" (e.g. *mmh… I'm not sure about this*).

[60] It should be noted that the treatment of onomatopoeic expressions in this chapter is focused on the sounds they imitate. How discourse units like onomatopoeic expressions behave from a structural and functional point of view will be discussed in chapter 11 (in particular 11.2.5.).

expressions of surprise and astonishment. In example (16), a Chatter nicknamed <girly_gal> fronts his/her turn with *oh*, signalling surprise. Another Chatter in example (17) fronts the turn with *ahh*, which indicates astonishment:

(16) <girly_gal>: oh cool
(Chat excerpt / Room IV, session 1)

(17) <alex_in_mini_kilt>: ahh I see <sam>, now I understand, hahahahahah
(Chat excerpt / Room III, session 4)

Also, the majority of onomatopoeic expressions in Web Chat either make up a single turn, or they are fronted or turn final; rarely do they appear in turn medial position. The same is true for another type of onomatopoeia that aims at imitating the sound of laughter. Example (18) illustrates a typical placement of the imitated sound of laughter, namely in a separate turn.

(18) <24mcanada>: alrighty then, ttyl [talk to you later] <tristan>
<24mcanada>: haha
(Chat excerpt / Room V, session 2)

In a similar manner, SMS discourse and e-mail, where the imitation of laughter is the most frequent type of onomatopoeia, show patterns with regard to the placement of onomatopoeic expressions. They are predominantly placed so as not to break up textual units with a strong tendency for final positions:

(19) Could you give me your address again? haha. I am so bad at this stuff.
(E-mail excerpt / author: female, 23 yrs)

(20) <Name> thinks that y. [you] have a really beautiful british english! He he-take care, (…)
(SMS text excerpt / author: female, 20 yrs)

Both instances of onomatopoeia in examples (19) and (20) are carefully placed between two textual units and additionally separated by punctuation. And although onomatopoeia can be said to be reminiscent of spoken conversation, the strategic placement of onomatopoeic expressions as shown in examples (16) to (20) is not (cf. also examples (21) and (22) below) .

Another trend in Table 6.4 is that both the personal homepages as well as the letters feature hardly any onomatopoeic expressions at all. This is unusual with regard to the

personal homepage, but quite expected in EEC. However, for reasons of completeness, examples (21) and (22) illustrate the use of an onomatopoeic expression in the personal homepages (imitating the sound of being uncertain or hesitant, also referred to as "hesitators", see 11.2.5.), as well as the only occurrence of onomatopoeia in EEC (imitating the sound of sympathy):

(21) Right now I am in the process of undergoing a ... umm…upgrading course, if you like (…)
(Personal homepage excerpt / author: Xm_1)

(22) ah poore Ladies, such was their screechings, teares, and distractions, (…)
(EEC excerpt, 1628 / letter collection Original_3, author: male)

Analogue to the emoticons, a decrease of onomatopoeic expression-to-word ratio can be observed as the degree of (potential) asynchrony of the types of communication increases: Web Chat, being the most synchronous communication, features the most onomatopoeic expressions, followed by SMS discourse, e-mail correspondence, the personal homepage, and the letter. However, media-related constraints, such as tiresome typing and limitation of text length in SMS discourse, have to be taken into account. Onomatopoeic expressions have a tendency to be elongated by the repetition of certain letters, which might be a reason why they are not as frequent as emoticons in the SMS text corpus (cf. hereto also Androutsopoulos & Schmidt 2001: 20, and Schlobinski et al. 2001: 15). Overall, it can be asserted that except for Web Chat, onomatopoeia is not a typical feature of personal written communication.

6.2.3. Saying it with signs

Another available communication channel of the paralinguistic type, available to written correspondence, is the substitution of certain text segments with signs. This is comparable to gestures, as for example the "thumbs up" (symbolising something along the lines of 'alright', 'good', 'ok') in a face-to-face conversation. There exist different types of signs, and this investigation looks into the use of "iconic signs" (resembling the referent they represent), "symbolic signs" (having an arbitrary relationship with their referents), and "indexical signs" (having a causal relationship with their referents) as used in personal written communication. Of course, emoticons are also types of iconic signs since they are aimed at emulating facial expressions or head shapes. As they are a distinctive and quite popular means of compensating for the lack of paralinguistic cues in written

communication, it was decided to discuss them separately at the beginning of this section (see 6.2.1. and Table 6.3 for their distribution across the text corpora). This subsection, however, looks into the use of other types of signs and will pay no particular attention to emoticons.

Signs can be of quite complex shape and it can thus be hypothesised that the fewer the available communication channels, the smaller the number of different types of signs the correspondence contains will be. Also, signs are often not classifiable as being exclusively symbolic, iconic, or indexical, because they may also be a combination of different types of signs. For example, an arrow is an indexical sign (pointing towards something) but if, for example, it contains the word *new* in it (words, except for onomatopoeic expressions, have an arbitrary relationship with their referents and are thus symbolic), to indicate that what is being indexed is "new", then the arrow/word sign is of combinatory character (indexical and symbolic at the same time). It was observed that out of the three types of signs, symbolic signs are most frequent in their distribution across the five text corpora. Iconic and combinatory signs predominantly occur in the personal homepages, and the indexical signs are overall rather infrequent. Table 6.5 gives an overview of the distribution of the different types of signs (if not self-explanatory, their main meaning is given, within the respective contexts, in brackets) as found in the five discourse types.

Table 6.5: Types of signs used in personal written communication (corpora-based results).

	Symbolic signs	**Iconic signs**	**Indexical signs**	**Combinatory signs**
SMS discourse	# (‘number’) x (‘kiss’)	* (‘star’) (+ emoticons)	--> (‘forward arrow’)	0
E-mail	x (‘kiss’)	(+ emoticons)	--> (‘forward arrow’)	0
Web Chat	x (‘kiss’) @ (‘at’) (((( )))) + << >> (‘hug’) $ (‘money’) ♥ (‘love’) * (‘correction’) (^) (‘birthday cake’) (C) (‘copyright’) ^^ (‘happiness’)	0 (+ emoticons)	>> (‘fast forward’) << (‘fast backward’) <-- / <== (‘backward arrow’)	0
Personal homepage	National flags: (‘Singapore’) (‘Canada’) (‘U.S.A.’) (‘Hong Kong’) (‘Taiwan’) (‘Russia’) (‘copyright’) (‘red ribbon’) (‘U.S.A. ribbon’) (‘Ying & Yang’, used as a heading for “miscellaneous hyperlinks”)	(‘home’) (‘resources’) (‘pictures’) (‘essays’) (‘about me’) (‘write to me’) (‘sign guestbook’) (‘view guestbook’) (‘search’) (‘computer’) (‘gamble’) (‘videos’) (‘awards’) (+ emoticons)	0	(‘Ying & Yang’ in a Chinese style of house) (‘talk’) Not Online right now Search (‘street sign’) (‘ideas’) (‘links’) home E-MAIL (‘U.S. mail’) GUESTBOOK from HTML GEAR (‘get your own guestbook’) NEW
EEC (letter)	0	0	0	0

The assumption that more available communication channels generate a greater variation of different types of signs could only be partly confirmed. While the personal homepage, which is rich in its available communication channels, features a great variety of signs, it is

at the same time unusual that it lacked signs of indexical nature. As illustrated in Table 6.5, only one combinatory sign shows qualities of indexing: an arrow (indexical) that contains the word *new* (symbolic). This is astonishing because the Internet is often referred to as "virtual space", through which users have to navigate themselves. This would invite the use of indexical signs. Indexical signs are, however, almost non-existent in the homepage corpus collected for this study, but symbolic and iconic signs are all the more frequent. In comparison, Web Chat contains, emoticons excluded, relatively few different signs and the e-mail corpus even fewer (only two different signs, the "x" for 'kiss' and the "forward arrow", were found). SMS discourse, on the other hand, due to being limited in its communication channels, was expected to feature few different signs and this expectation was confirmed. Yet the Early English letter corpus, although richer in its communication channels than SMS discourse, does not contain one single instance of a sign.

Regarding the variety of signs, the e-mail corpus features a small selection of different signs. Although e-mail users have access to different formats that would allow the incorporation of all sorts of signs and symbols, they do not seem to take full advantage of these formats. The best-known e-mail formats are "plain text", "html" and "rich text format", which all offer individual possibilities for editing an e-mail message. Although Dürscheid (cf. 1999: 28) claims that the medium of e-mail enables the user to communicate on various textual levels at once, as it makes it possible to combine writing, graphics, pictures, films, and sound, the most common format among users of e-mail is the plain text format. Yet the fact is that the plain text format leaves little room for layout (font is predetermined and no signs or symbols can be inserted). It thus seems as if most people seem to be aware that "ability to send e-mail messages that incorporate unusual elements, is only as good as [the recipient's] e-mail software," and for this reason tend "to stick with tried-and-true format settings" (Flynn & Flynn 2003: 67). Obviously, the authors of the e-mail corpus attended to this credo, because most messages turned out to be composed in plain text format and thus lack features connected to rich text format.

With respect to the frequency of signs, as opposed to variety (excluding emoticons), the picture is slightly different. The e-mail corpus contains the lowest number of signs (11, or a sign-to word ratio of 1:2339), followed by SMS discourse (17 instances, or 1:1084). Thus, the two media with the smallest variety also contain the fewest signs. However, although the personal homepage corpus features a greater variety of signs than Web Chat, both corpora result in almost identical sign-to-word ratios with 1:286 (or 57 signs) for the personal homepage corpus, and 1:287 (or 99 instances) for the Web Chat

data. The most frequently used type of (symbolic) sign in both SMS discourse and e-mail is "x", which stands for 'kiss', being in most cases part of the farewell section. Examples (23) and (24) illustrate typical farewell sections of e-mail and SMS discourse that contain such symbolic kisses:

(23) Speak soon,
xx <name>
(E-mail excerpt (farewell section) / author: female, 30 yrs)

(24) With love & hugs, <Nickname> XXX
(SMS text excerpt (farewell section) / author: male, 31 yrs)

In comparison, Chatters have a tendency to use the (symbolic) sign "*" in the sense of 'correcting a previous error' as illustrated in example (25), where a Chatter nicknamed <Hannahbanana87^292> first misspells *always* as *alays* in one of his/her turns, to then correct the mistake (indicated by the asterisk) a couple of turns later, by contributing *always**. Another frequent feature is the employment of (symbolic) brackets to act out "virtual hugs", as shown in example (26).

(25) <Hannahbanana87^292>: im alays last!!
(…)
<Hannahbanana87^292>: always*
(Chat excerpt / Room III, session 4)

(26) <Dream>: (((((((((<milly1>))))))))) ['Dream hugs <milly1> nine times']
(Chat excerpt / Room IV, session 3)

With respect to the types of signs used in the personal homepages, a tendency for expressing citizenship via national flags could be determined. Also, signs related to writing are among the more frequent ones.

All in all, the findings show that the use of signs is connected to media-related features, but not exclusively so, as both e-mail and EEC deviate in this respect. Yet the fact that EEC does not feature any signs is not necessarily related to the letters originating in the Early English Period. The results from the online survey show that the majority (53%) of 21st century letter writers claimed that they never incorporate signs to substitute words in their epistolary correspondence. With regard to the other media, it emerged that using signs as a substitution for words is generally infrequently employed by the participants of the online survey. Peak answers for the use of signs in SMS discourse are "sometimes" (22%)

and "rarely" (22%), for e-mail it is "rarely" (31%), and for both Web Chat and personal homepage it turned out to be "never" (22% and 32%, respectively). To my knowledge, no empirical investigation into the use of signs (other than emoticons) in personal written communication has been carried out so far, and there would thus be ample room for future investigations.

6.2.4. Hyperlinks

After looking into different types of communication channels that would in principle be available to all five types of communication investigated in this study, I would now like to turn to a feature that is quite new to the field of personal correspondence (which means that EEC is not affected by the phenomenon). With the advent of the Internet, there came the possibility to connect text segments internally or externally with other text segments located on the Internet by means of so-called "hyperlinks"[61]. A text located on the Internet that contains such (functional) hyperlinks is commonly referred to as hypertext (see hereto also 1.3.1. and 4.2.1.). In contrast to the traditional structure of text, which is arranged in linear-hierarchical fashion, the notion of hypertext stands for multidimensional and dynamic presentation of text where the reading process is no longer sequential but to a certain extent optional (cf. Faulstich 2006: 177) because the documents linked by hyperlinks are non-linear in character (cf. Dittmann 2001: 19).

However, this is only true for hyperlinks that are, first, functional, and second, transfer the user to the intended destination. As will be seen below, the incorporation of hyperlinks is not only limited to text located on the Internet, but may be part of other text types as well—albeit with limitations in usability. But before taking a closer look at the qualitative side of hyperlinks, I would first like to say a few words about their incorporation into text. Hyperlinks may be part of the running text, as illustrated in example (27), or they may be embodied by a picture (to which I refer here as "pictogram-hyperlink"), as shown in example (28), or they may also stand free, as can be seen in example (29):

(27) I have now moved to Grenoble in France where I am doing a Masters in International Business. (…)
(Personal homepage excerpt / author: Tm_1)

[61] For reason of space no technological discussion of the hyperlink phenomenon can be included here, but see Runkehl et al. (1998: 127) for background on the technological aspects of hyperlinks.

(28) 

(Personal homepage excerpt / author: Im_1)

(29) Check out this link with pictures I found to show you:
<URL of website>
(E-mail excerpt / author: male, 37 yrs)

Hyperlinks that are part of the running text or that stand free are typically of blue colour and underlined.[62] This identifies text segments as hyperlinks among users and clicking on such a blue and underlined text segment will transfer a user, if the hyperlink is functional, to another destination inside or outside the text entity. In the case of example (27), this is another website dedicated to the French city Grenoble, whereas the hyperlink in example (29) has a website of photographs as its destination. Pictogram-hyperlinks, on the other hand, may take on any form. Example (28) shows the picture of a house, in combination with the word *home* (hence, a combinatory sign), that contains a hyperlink in the form of an "image map" (cf. Runkehl et al. 1998: 127), transforming the picture into a pictogram-hyperlink. Moving the mouse arrow over this picture transforms the arrow into a pointing hand, indicating that this picture is in fact a pictogram-hyperlink. In this particular case, if the visitor decides to surf around the website, clicking on this pictogram-hyperlink will take the user back to the *home*page.

An investigation into hyperlinks used to communicate linguistic content limits the undertaking, as far as the text corpora are concerned, to the modern types of correspondence. However, hyperlinks, as understood in this study, can be incorporated in handwritten text, but they would have to be classified as dysfunctional within the medium because they require Internet access. In any event, for information-technological reasons EEC does not contain any hyperlinks and will thus be neglected for the moment. With regard to the modern media, the following media-related constraints apply: an SMS text or e-mail message may only contain hyperlinks to destinations outside of the text entity, whereas the personal homepage and Web Chat, being located on the Internet, may also contain hyperlinks to destinations inside the text entity. Also, mobile phone software

[62] Opposing to this view, Boardman (2005: 19) argues that "the graphology of the hyperlink is now liberated from the default underline version" and "that any area of the page, text or graphics, is potentially a hyperlink, and an accepted component of the reader's stance in relation to web pages is now the assumption that the reader has to actively look for hyperlinks." With respect to the homepage corpus investigated in this study, however, it was observed that the vast majority of hyperlinks that consist of textual units (as opposed to pictogram-hyperlinks) were in fact of blue colour and underlined—most probably for reasons of facilitating navigation. This in turn does not support Boardman's argument.

limitations at the time of data collection meant that sending a hyperlink in an SMS text was possible, but clicking on it was not. Thus, sending a hyperlink via the short message service was useless unless the receiver was known to have Internet access via computer. This is probably also the reason why the SMS corpus contains only one hyperlink. Table 6.6 gives an overview of the hyperlinks found in the corpora of the modern text types.

Table 6.6: Hyperlinks in electronic personal written communication (corpora-based results).

Types of hyperlinks	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**
hyperlinks to a destination INSIDE text entity	x	x	0	* 634 ** 55.0
hyperlinks to a destination OUTSIDE text entity	1 100	10 100	7 100	466 40.4
hyperlinks to a NON-IDENTIFIABLE destination	0	0	0	53 4.6
TOTAL	**1** **100**	**10** **100**	**7** **100**	**1153** **100**

Key: HP = homepage; x = respective type of hyperlink not possible in the respective text type due to media-related constraints; * = instances of the different types of hyperlinks per text type; ** = percentages (rounded to 1 decimal) are given in relation to the total number of hyperlinks found in each of the text types.

Only the personal homepage corpus contains hyperlinks of the type that link a text segment with another text segment inside the same text entity. Web Chat does not feature any hyperlink of said type although it would be technically possible. As mentioned above, both SMS text and e-mail corpora do not feature this type of hyperlink due to media-related constraints. It is a slightly different picture concerning hyperlinks that take the user to destinations outside of the text entity. As has already been mentioned, the SMS text corpus features only one SMS message that contains a hyperlink, most probably connected to media-related constraints at the time of data collection. However, this situation has changed considerably because contemporary mobile phone models provide access to the Internet. Thus nowadays, receiving a hyperlink in an SMS text means that the user can actually click on the hyperlink and be directly transferred to the website, without having to change the medium. It seems safe to assume that SMS messages sent to and from newer models contain more hyperlinks than the ratio 1:1000.

What seems unusual is that both e-mail and Web Chat feature few hyperlinks to destinations outside their text entities. In 140 e-mail messages only 10 hyperlinks of this type were found, and on average, only every fourth Web Chat contains a hyperlink to a destination outside the Chat room. It seems as though personal homepages are the only discourse type that makes very frequent use of hyperlinks. This is, however, not surprising

in view of the fact that it is the hyperlinks that turn "text" into "hypertext". Yet almost a fifth of the hyperlinks found in personal homepages are linked to non-identifiable destinations, such as password protected sites or, in the majority of the cases, sites that could not be found. The original destination is then non-identifiable, as the hyperlink takes the user to another site with an error message that may or may not contain other hyperlinks to help the user find the original destination. More often than not the site containing the error message does not provide any search aides and can thus be seen as a dead-end and "the most common message to appear when you follow a dead hyperlink is 'Error 404—Page not found'," which is "a code generated by a web server when it is asked for a page that it does not have" (Boardman 2005: 77).[63] Hyperlinks that lead to dead-ends are fairly frequent in the fast-changing world of the Internet and their destination looks something like this:

Figure 6.1: Destination of a dead-end hyperlink.

Such "dead-end hyperlinks"[64] are seen by many as indicators of neglect on the webmaster's behalf who is held responsible for the content of the website, including the functionality of the hyperlinks which would require frequent verification.

This may be a reason why the majority of informants that participated in the online survey stated that they do not frequently include hyperlinks in their personal correspondence. Around 60% said they never include hyperlinks of any type in their SMS

[63] Interestingly, the "404 experience" is, according to Boardman (2005: 77), "such a common occurrence when Web browsing, that the term '404' is starting to be used in everyday life as a metaphor for a person who is clueless or generally unaware," as in, for example, *Don't bother asking him. He's 404, man* (example taken from Boardman 2005: 77).

[64] Technologically speaking, "the search engine's Web spider will find new pages and index them, but if a page is deleted the spider will not know until it visits again—so the URL remains in the index and a dead hyperlink is the result" (Boardman 2005: 77).

texts and letters. Surprisingly, the majority of users who maintained a personal homepage at the time the survey was taken stated that they never include hyperlinks. According to the informants, the medium most likely to include hyperlinks is e-mail: 25% indicated that they often include hyperlinks and another 24% that do so very often. Thus, the results from the survey do not confirm the outcome of the corpora-based investigation.

Not many papers deal with the functions of hyperlinks in personal homepages or websites in general. Askehave and Ellerup Nielsen (2005: 6), however, observed that hyperlinks in personal homepages can be distinguished into "generic hyperlinks" and "specific hyperlinks". While the former provide access to the main topic of the website, the latter function as appetizers or previews of what is to come. According to Askehave and Ellerup Nielsen (2005: 6), "generic hyperlinks are always of descriptive nature" and mainly serve to navigate the site (as in *home*). Specific hyperlinks, on the other hand, introduce new topics and function as gateways. They take the user to another destination, where more information on the topic, as introduced by the hyperlink (such as *read more*), can be found. Askehave and Ellerup Nielsen (2005: 7) come to the conclusion that hyperlinks are significant additions to what has so far been perceived as the traditional text typologies (or "writing genres"). More research will be needed in order to be able to better understand the phenomenon of hyperlinked text types.

6.2.5. How the medium influences the production of text

The communication channels that have been looked at so far, emoticons, onomatopoeia, use of signs, and the employment of hyperlinks, have all been available, with certain restrictions, to the five types of media investigated in this study. There are, of course, other communication channels that are only available to particular media. Some of these more exclusive communication channels will now be examined in consideration of their actual occurrences in the respective discourse types.

One way of communicating beyond the words of the actual message, is to attach documents or files of any kind or, in the case of the letter, enclose them. The SMS corpus does not feature any attachments due to limitations of software at the time of data collection. However, nowadays it is quite common to attach photographs (or other kinds of attachments) to an SMS message and thereby turning it into a so-called MMS (Multimedia Message Service) message. Again, trends into that direction would have to be investigated with a more recently collected corpus. Speaking of photographs and pictures (drawings, physical or digital), the saying "a picture says more than a thousand words" comes to

mind. And while this saying may be true for some types of correspondence, it is an entirely different matter for others. Between the 60 personal homepages 22 pictures and 75 photographs were found and the majority of photographs show the (supposed) homepage owners or their families and friends. In total, 51 of the 60 personal homepages contain pictures and photographs, only 9 are based on text only. These results are contradictory to Miller and Arnold's (2003: 81) observation that 30 out of the 70 homepages (35 pages composed by males and 35 pages authored by females) they investigated did not feature any images at all. It seems as though more research into this medium is needed in order to be able to make claims with respect to self-representation on personal homepages by means of pictures and photographs.

However, the homepages investigated in this study also stand in stark contrast with the 30 Web Chats, which do not feature a single picture or photograph, as well as the 140 e-mails where only 4 attachments were found, all of which are photographs. Neither SMS discourse, due to above-mentioned media-related constraints, nor EEC have pictures or photographs attached or enclosed (physical drawings only in the case of EEC, as photography had not yet been invented in the 17th century). It seems as though pictorial expression is a communication channel predominantly used in personal homepages, although the attachment or enclosure of pictures (physical or digital drawings) would be available to all types of correspondence (except SMS discourse).

The only corpus featuring attachments other than pictures or photographs is EEC. Of a total of 91 letters, 14 mention an enclosure of the following kinds: "letter(s)", "a type of written document other than letter" (e.g. a testimonial), "key(s)", "book(s)", "food", or some sort of "gift(s)". All in all, seven enclosures are of the type letter(s) or type of document other than letter, which would translate into a word document or PFD file in modern correspondence. However, the other half of enclosures is of the physical type, which cannot be attached electronically. These enclosures are thus idiosyncratic of the letter as transmitted by the postal service. All in all, it can be asserted that attachments and enclosures are connected to media-related constraints but not exclusively so, as they are infrequent in the e-mail corpus and non-existent in the Web Chat data.

Other communication channels that are exclusively available to the modern types of correspondence are the employment of sound and incorporation of video material into a message. Yet none of the modern text types contains video material and/or sound in the form of spoken or sung text segments. It is a similar picture with graphically enhanced writing and implementation of logos, which appear only occasionally in the personal

homepage corpus, but no other text types make use of those features, although both e-mail and Web Chat would offer a technological environment allowing the incorporation of graphics and logos. Thus so far, it can be stated that authors of personal written correspondence by no means fully exploit the available communication channels.

6.3. Means of emphasis in personal written communication

Although communication can be expected to come with a guarantee of relevance, it can also be assumed that not all of the communicated content is of similar importance. In a spoken conversation, the most important parts of a message are often emphasised by, for example, intonation and stress. In written communication, various means of emphasis can be employed to indicate that certain parts of a message are meant to receive particular attention. Some types of means of emphasis are available to all types of written correspondence, whereas others are subject to media-related constraints. The most frequent means of emphasis, as found in the five text types, will be discussed next and compared to the findings from the online survey.

6.3.1. "What's in a word?" – Emphasis of particular words and expressions

One of the most straightforward ways of emphasising certain content in a piece of writing is to highlight particular words or expressions. It is important to point out that Web Chat data was collected from Chat logs providing the text in black and white. Thus no quantitative analysis into the use of colour in Web Chat could be carried out. It was, however, noted down for each of the rooms what sort of tools were made available to the Chatters to enrich their contributions, and they will be discussed as we proceed. Also, it should be noted that hyperlinks that are part of the running text or stand free (as opposed to images), are automatically highlighted by blue colour and underlining and are for this reason not further considered. Table 6.7 presents the means of emphasis determined in the different corpora. They include the following seven types: (1) "capitals", (2) "italics", (3) "underline", (4) "use of colour", (5) "repetition of punctuation", (6) "repetition of letter(s)", and (7) "repetition of words". Table 6.7 gives an overview of the distribution of these seven types of means of emphasis across the five corpora.

Table 6.7: Selected means of emphasis in personal written communication (corpora-based results).

Types of MOE	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
capitals (e.g. *this is NOT nice*)	* 203 ** 1:90.8	83 1:310.0	187 1:151.9	74 1:216.6	0
italics (e.g. this is *not* nice)	x	1 1:25'733	0	69 1:232.3	0
underline (e.g. *this is not nice*)	x	0	0	7 1:2290.0	0
use of colour (e.g. *this is not nice*)	x	0	***	134 1:119.6	0
repetition of punctuation (e.g. *this is not nice!!!!*)	280 1:65.8	328 1:78.5	497 1:57.2	76 1:210.9	0
repetition of letter(s) (e.g. *this is sooooo nice*)	31 1:594.4	13 1:1979.5	464 1:61.2	1 1:16'030	0
repetition of word (e.g. *very very nice*)	23 1:801.1	11 1:2339.4	126 1:225.4	4 1:4007.5	1 1:31'077

Key: MOE = means of emphasis; HP = homepage; * = total number of instances of indicated type of MOE per text type; ** = indicated type of MOE-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types; x = respective type of MOE not possible in the respective text type due to media-related constraints; *** = no quantitative uses of colour could be determined for Web Chat due to constraints related to data collection.

Two trends become evident in Table 6.7: first, the repetition of punctuation is by far the most popular means of emphasis and second, a negligent number of the letters in the EEC corpus features any type of emphasis (in fact, only one letter contains a word repetition). After emoticons, use of signs, and incorporation of hyperlinks, it means that this is the fourth type of communication channel that is either not available to authors of EEC or was not employed by choice. Another aspect that stands out is that SMS discourse is constrained in its available means of emphasis by media-related features (indicated by "x" in Table 6.7 above).

Interestingly, even newer mobile phone software does not (yet) include the features "italics", "underline", or "use of colour" (for written text that is, emoticons are frequently coloured, mainly in yellow). Also, none of these features are popular in the other text corpora except for the personal homepage, where the use of colour is in fact the most frequently applied means of emphasis. As mentioned above, constraints related to data collection did not allow a quantitative investigation into the use of colour in Web Chat. However, the possibility to use colour as a means of emphasis (or otherwise, such as personalising turns) depends on the Chat room facilities and is thus connected to media-related constraints. Of the six Chat rooms (Rooms I-VI) that were visited to collect data, only two, Rooms I and III, offered their Chatters the possibility to make use of colour. Rooms II, IV, V, and VI, on the other hand, did not feature colour to be used as a means of emphasis. A special (software-generated) feature with regard to colour was observed in

Room IV, where the nicknames of Chatters, if used by other Chatters in one of their turns, appeared in pink colour on the computer screen of the Chatter with the corresponding nickname. This was probably aimed at helping Chatters identify turns directed at each other.

To return to the phenomenon "repetition of punctuation" as a means of emphasis, it should be noted that the repetition of the full stop was also counted as a means of emphasis. Repeated full stops create visible gaps following (and in some cases preceding) the word or word string, which in turn makes that particular word or word string stand out from the remaining text, as illustrated in example (30) below (see also example (21) above). This type of means of emphasis is most frequent in the four modern types of correspondence, followed by the exclamation mark, illustrated in example (31), and the question mark, as shown in example (32):

(30) I didn't move too far……. about 7 hours up northern (region), (…)
(E-mail excerpt / author: female, 21 yrs)

(31) hope you feel better!! (…)
(SMS text excerpt / author: female, 26 yrs)

(32) <desirable>: my wat????
(Chat excerpt / Room III, session 3)

SMS discourse features the repetition of punctuation as its favourite means of emphasis. This may be connected to two main reasons: first, capitals are often interpreted as shouting and thus come with a stigma (cf. Schlobinski et al. 2001: 7); second, the use of repeated punctuation might require less effort than switching the "all caps mode" on and off. However, since SMS discourse does not have many means of emphasis at its disposal, the use of capitals is still quite frequent, and the same is true for the other types of correspondence:

(33) I forgot to tell you all WHEN I'm leaving. Doh!
(E-mail excerpt / author: female, 24 yrs)

(34) 99 Dumb Facts about ME!
(Personal homepage excerpt / author: Bf_3)

(35) <catty^700>: oh right gud job I HAVE JUST LEFT COLLEGE
(Chat excerpt / Room III, session 4)

(36) BIG kiss <name>
(SMS text excerpt (farewell section) / author: female, 27 yrs)

Example (33) illustrates a typical case of emphasising the most important word. The author of this e-mail had informed her friends in a previous e-mail that she was due to leave the country for a longer period of time, but in fact forgot to include the date of her departure. It is a similar case in example (34), stemming from a personal homepage, where the capitalisation of the pronoun indicates that the 99 dumb facts are about the author, rather than about anyone else. The use of capitals in example (35) indicates new, and seemingly important, information. The first part of the turn makes a reference to what another user revealed about how he/she got him-/herself a good job. Chatter <catty^700> responds by referring first to the other Chatter's job situation, to then reveal something about his/her own, the new information being emphasised with capitals (*I HAVE JUST LEFT COLLEGE*). A slightly different scenario can be observed in example (36), where the word *big* is emphasised by capitals because it is important. It reconfirms at the same time its semantic meaning by being visually bigger than the remaining message content as well.

In comparison, both italics and underline are not very frequently used means of emphasis in the five text types. The same is true of the use of colour, which is only frequent in the personal homepage corpus. A total of 134 instances of use of colour were determined for the 60 personal homepages, mostly applied to titles and headings, which makes this quite a frequent feature of this text type. Table 6.7 concludes with two more types of means of emphasis as used by the authors of the modern text corpora, both of which are most frequent in the Web Chat data: repetition of particular letters and words. The repetition of letters is in fact so frequent that it can be classified as characteristic of Web Chat. Example (37) illustrates a use of this feature distinctive to Web Chat, namely the elongation of another Chatter's nickname by repeating a particular letter, usually word-final, and thus drawing attention to the turn:

(37) <divainross>: hey all
<cute_sami>: <diva<u>aaaaaaaaaaaaaaaa</u>>
(Chat excerpt / Room III, session 1)

Although the Chatter with the nickname <cute_sami> first shortens another Chatter's nickname from <divainross> to <diva>, he/she then decides to draw attention to it by repeating the "a" of *diva* no less than 16 times. Another phenomenon that was observed in relation to the repetition of letters is illustrated in example (38).

(38) <guccigirl>: <star> looooool
(Chat excerpt / Room IV, session 3)

The Chatter <guccigirl> refers to something the Chatter nicknamed <star> wrote beforehand, finding it so funny that he/she has to *loooool*, meaning 'to laugh out loud' or 'laughing out loud'. The origin of the acronym *lol* is generally assigned to Chat culture (cf. Schlobinski et al. 1998: 104ff.), and its use has become so frequent that it starts being treated like a single word (rather than an acronym consisting of the initial letters of three different words). This has been exemplified by <guccigirl's> attempt to emphasise *lol* by repeating the "o", as if it was the vowel of the word. But "o" stands for *out* in the acronym *lol*, and it would in fact make more sense if either "l" (the verb *laugh* or the adverb *loud*) would be repeated to emphasise that one is laughing, or doing so in a loud fashion. Furthermore, it was observed that words emphasised by repeating particular letters are often placed at turn final position, as opposed to turn initial or medial.

The repetition of particular words is not frequent in SMS discourse, e-mail, or in the personal homepage corpus; and in the letter corpus it is extremely rare. There is only one instance of a word that is repeated for emphasis in EEC, illustrated in example (39), and this is at the same time the only instance of any type of means of emphasis in the whole EEC corpus.

(39) Dearst dearst Sister,
(EEC excerpt (greeting section), 1656 / letter collection Tixall, author: female)

Here, the author writes to her sister, and in her greeting section decides to repeat the affectionate term of address *dearst* in order to emphasise this part of the greeting. The only text type where the repetition of particular words is somewhat frequent is Web Chat. A total of 126 word repetitions mean that this features occurs in an average ratio of 1:225 words. Predominantly, these repetitions were classified as affirmative response forms (see 11.2.5.), often fronting or constituting a turn, as illustrated in examples (40) and (41) below.

(40) <gother>: was only a friendly kiss
(…)
<cute_sami_devil>: yea yea <gother> it's ok!
(Chat excerpt / Room III, session 1)

(41) <Ladybug>: give him a big happy b-day from me
(…)
<Dish>: I will do that
<Dish>: yup yup
<Dish>: I will
(Chat excerpt / Room I, session 3)

In example (40), a Chatter nicknamed <gother> is fooling around with another Chatter, and it turns out that yet another Chatter with the nickname <cute_dami_devil> feels, or pretends to be, offended. At some point, <gother> defends him-/herself by saying *was only a friendly kiss*, and after a while <cute_sami_devil> solves the situation by fronting the turn *it's ok* with the affirmative response *yea yea <gother>*. In the excerpt shown in example (41), two Chatters named <Ladybug> and <Dish> converse about the birthday of another person, whom <Dish> identifies as his/her son. When <Ladybug> tells <Dish> to forward his/her happy birthday wishes <Dish> consents and then re-confirms this by means of *yup yup*. This feature is in fact highly reminiscent of spoken discourse and only logical that it is most frequent in the text type classified as the most synchronous of the five media.

Overall, it was found that the use of emphasis for particular words or expressions is most frequent in Web Chat and SMS discourse, followed by e-mail and the personal homepage. EEC, on the other hand, features hardly any means of emphasis. Other trends include a vast variation of signs in the personal homepages, as opposed to all other media. However, this is to a certain extent connected to media-related constraints. Also, the personal homepage shows several means of emphasis of particular words that can be classified as exclusive to that medium, because none of the other text types features them. Again, this is partly due to media-related constraints. They include the use of bigger, bold, and/or different fonts. There is a clear tendency for the combination of bigger (than the rest of the text in the homepage) font and bold font, sometimes also coupled with the use of colour, above all for titles and headings.

As can be gathered from the code system (Table 5.2, feature 2.4.) it was at first attempted to investigate into the use of different fonts for means of emphasis. However, this undertaking was neglected for the following reasons. First, it was impossible to determine whether the fonts found in the e-mails were intended by their authors, or whether this was connected to the software of the receiver. Although it can be assumed that it is in fact a mixture of the two. Second, due to media-related constraints, it was not possible to determine fonts as used in Web Chats. Even though different fonts might have

been used in the Chat rooms, the Chat logs were in all cases made available in the font type "Arial". And third, the fonts used in SMS messages were subject to a twofold media-related constraint: when composing a message, the font cannot be selected in the first place, and since most of the SMS texts were given to me transcribed on paper or transmitted electronically (via e-mail), they underwent another media-related constraint. Thus, the investigation into the uses of different types of fonts would have generated biased results and was therefore neglected.

6.3.2. Movement and sound

Movement and sound as means of emphasis in text are tied to the level of technology of a certain medium. It emerged that the personal homepage and Web Chat are the only media that feature movement and/or sound. Due to constraints in relation to data collection, no quantitative uses of features beyond the text level could be determined for Web Chat. However, two of the rooms, Room II and IV, offered the feature of "audio emoticons" to their Chatters, where the decision to use a particular emoticon also entailed the decision whether or not this emoticon should make a machine-generated sound based on the emotion it aimed at conveying. Another room, Room I, offered a feature called "sound alert" and Chatters could choose whether or not they wanted their turns to be accompanied by such a sound alert when appearing on screen. Furthermore, Room I also offered "moving emoticons", jumping up and down within a turn and thereby drawing attention to it. This shows that half of the rooms that provided the data for the Web Chat corpus offered sound as means of emphasis, and only one room featured movement as an available means of emphasis. Since no quantitative uses of either feature could be determined, more research is needed in this area of Web Chat. With regard to the personal homepage, no instance of a sound alert could be determined and only 35 moving text segments were found. This means neither can be classified as typical means of emphasis in personal homepages.

6.4. Chapter summary

The investigation into the uses of the available communication channels in the different types of personal written correspondence exposed trends that confirm the main research hypotheses (H1 – H5) in some areas, but also contradict them in others. It was observed that there seems to be a connection between the decrease of the frequency of emoticons, onomatopoeia, and the employment of different means of emphasis, as the degree of

asynchrony of the media increases. Since all these features qualify for the emulation of spoken conversation, the hypothesis that modern correspondence, owing to its immediacy, contains more features of orality, has been confirmed (H1, H2). However, the investigation into the use of signs (comparable to gestures in spoken conversation) that would also allow compensation for the lack of paralinguistic cues, showed that their employment is to a certain extent media-related (having more possibilities concerning bandwidth, the homepage also features a greater variety of signs compared to SMS discourse), but not exclusively so, as both e-mail discourse and EEC do not include a vast number of signs to substitute for words.

While the lack of signs in EEC confirms the hypothesis H3, which claims the letters to be rich in contextual features typical for written discourse and in reverse to contain few features typical for spoken conversation, the lack of signs in e-mail contradicts the opposed claim (H1), i.e. CMC being rich in the compensation of paralinguistic features. Also, the second part of H1, that CMC has characteristics idiosyncratic to that medium and thus contains features of cyberdiscursivity, is only partly supported by the personal homepages that incorporate hyperlinks and headings in each of the text entities. However, both e-mail and Web Chat do not. Another contradiction concerns research hypothesis H4, which hypothesises that modern correspondence shows considerable differences compared to EEC regarding its contextuality, because the employment of paragraphs as text structuring device does not show great differences between the text types (apart from SMS discourse, where the lack of this feature is subject to media-related constraints).

It was further observed that the letters differ from the modern text types with regard to the frequency of text entities that conventionally include both a greeting and farewell section, confirming two of the hypotheses (H3, H4), but with regard to how those sections are incorporated into the messages, all text types show similarities, which contradicts, again, H4. With respect to H5, predicting differences between messages aimed at an acquainted vs. unacquainted readership, no tendencies could be determined. Thus, so far it can be asserted that H1, H2, and H3 hold (with minor exceptions), whereas H4 and H5 could not be confirmed.

7. Personal reference

As has been pointed out in chapter 6, greetings and farewells are of particular interest because they are distinctive of interpersonal communication, as opposed to other writing genres (such as manuals or creative writing). Yet the quantitative investigation into greetings and farewells showed that by no means does the correspondence subsumed in the different text corpora feature greetings and farewells in all cases. If greetings and farewells are incorporated into a message, however, then the two sections are highly likely to contain a distinctive kind of personal reference, namely forms of address. Of course, forms of address are not restricted to the greeting and farewell sections. They may permeate personal written communication throughout. It will thus be of interest in this chapter how the authors of personal written communication address themselves, their readership, and/or any other third party, and whether or not tendencies can be observed that are idiosyncratic for any of the five discourse types. As the use of personal pronouns is typical of spoken conversation (see also Biber et al. 1999: 333), it can be hypothesised that the more synchronous text types feature them more frequently. Furthermore, while this chapter looks into the use of different forms of address by the authors of the different text types, it also accommodates the circumstance that explicit personal references may also be disregarded in personal written communication.

7.1. Forms of address

One interesting aspect comes into play concerning personal references as employed by the author: two of the five text corpora were produced by authors that did not (personally) know their readership at the time of message composition. Both personal homepages as well as Web Chat are communicative settings where the relationship between participants is typically of an unacquainted nature (see also Table 5.1). The SMS text corpus, e-mail messages, and EEC data, on the other hand, were composed with an acquainted and thus specific readership in mind. This means that authors of messages aimed at an unknown readership can be assumed to construct their readership to a certain extent. Although all five types of correspondence are of personal character, never mind whether they are aimed at an acquainted or unknown readership, it can still be hypothesised that these differences in readership reflect in the use of personal references, both quantitatively as well as qualitatively.

It should be noted that the discourse type Web Chat will be included in the Tables illustrating the frequencies of different types of forms of address, even though varying numbers of authors were involved in the production of the Chat extracts. However, contrary to the investigation into the frequencies of greetings and farewells, where the employment of those discourse units was felt to be biased due to the varying numbers of authors, the frequency of forms of address is comparable. The reason for this is that all private Chats (8 in total), taking place between the researcher and other Chatters in private rooms (it seemed rude to reject invitations for one-to-one Chats whilst collecting data in the main rooms), generated frequencies for forms of address that were comparable to the main rooms. This is connected to the discourse structure of Web Chat, which is highly reminiscent of spoken conversation in that participants take alternate turns. It was found that whether this be two or two hundred participants, those turns contain comparable frequencies of forms of address. However, if the difference in participant numbers is felt to potentially bias the results for Web Chat in comparison to the other text types, then this will be pointed out and discussed alternatively.

7.1.1. Authorial self-address

Personal correspondence of dialogical structure is understood as involving at least one addressor that writes to one or more addressee(s). All types of correspondence investigated in this study were classified as (potential) dialogues in that the addressees are all in a position to respond. What is rather new to the field of personal written communication is the possibility to engage with an unknown readership. Ever since the wide public gained access to the Internet in the early 1990's, people from all over the globe have been communicating with each other on a personal level, even though they have never met and will most probably never meet. While it is highly likely that all those people have experience in personally communicating with their friends and families, and are most likely to draw on this experience when communicating with strangers, it can still be hypothesised that there ought to be differences in how authors refer to themselves in anonymous settings compared to settings where everybody is familiar with each other.

As has been pointed out in the theoretical discussion (see 4.3.1.), authorial self-references may vary in scope as well as in form. With regard to scope, authorial self-references may refer exclusively to the author (+A), or also include the reader (+R), and/or any other third party (+3P). With respect to form, authorial self-references may be pronominal (+P), for example the exclusive authorial self-reference by means of the first

person singular pronoun *I*, or they may be some other kind of other non-pronominal (usually nominal) reference (+O), such as *Paul*. The former is also frequent in spoken discourse, the latter, however, is characteristic of written correspondence (in particular of the farewell section) and quite rare in spoken interaction.[65] Hence, there are six different types of authorial self-references: (1) exclusive authorial self-reference by means of a pronoun (+A, -R, -3P / +P, -O) such as the first person singular pronoun *I*, or (2) some other type of authorial non-pronominal self-reference (+A, -R, -3P / -P, +O) such as *Peter*, (3) authorial self-reference that includes the readership by means of a pronoun (+A, +R, -3P / +P, -O) as for example the first person plural pronoun *us*, or (4) by means of some other non-pronominal reference (+A, +R, -3P / -P, +O) such as *our group*, or (5) authorial self-reference that includes another third party by using a pronoun (+A, -R, +3P / +P, -O) as for example the first person plural pronoun *we* (where the pronoun *we*, based on contextual information, excludes the reader but includes another third party), or (6) some other non-pronominal reference (+A, -R, +3P / -P, +O) like *our team* (where the possessive adjective *our*, based on contextual information, excludes the reader but includes another third party).[66] It was found that all five types of correspondence favour the exclusive authorial self-reference, by means of a personal pronoun, over the other types of references.

[65] Communicative settings where people do not know each other well, or have just met, are insofar an exception that participants may include their own name in the farewell, if they realise that other participants do not remember it. However, this study is concerned with personal communication, and it can be assumed that interlocutors in both spoken and written communicative settings know each other's names and thus generally refrain from referring to themselves by their own names (apart from the farewell section in written exchanges).

[66] It should be noted that there exist two more possibilities for authorial self-reference in personal written communication: (+A, +R, +3P / +P, -O) and (+A, +R, +3P / -P, +O), where the former could be any first person plural pronoun that, based on contextual information, refers to author, reader, and any other third party and the latter any non-pronominal reference that includes all three parties. These types of authorial self-address were, however, not observed in any of the text corpora and were therefore neglected.

Table 7.1: Authorial self-address in personal written communication (corpora-based results).

Types of authorial self-address	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
1. (+A, -R, -3P / +P, -O)	* 971 ** 1:19	1356 1:19	950 1:29.9	650 1:24.7	1989 1:15.6
2. (+A, -R, -3P / -P, +O)	462 1:39.9	127 1:202.6	280 1:101.4	230 1:69.7	184 1:168.9
3. (+A, +R, -3P / +P, -O)	83 1:222	45 1:571.9	41 1:692.8	2 1:8015	15 1:2071.8
4. (+A, +R, -3P / -P, +O)	4 1:4606.5	6 1:4288.8	4 1:7101	0	20 1:1553.9
5. (+A, -R, +3P / +P, -O)	39 1:472.5	203 1:126.8	23 1:1235	15 1:1068.7	121 1:256.9
6. (+A, -R, +3P / -P, +O)	6 1:3071	29 1:887.3	5 1:5680.8	15 1:1068.7	42 1:739.9
TOTAL	**1565** **1:11.8**	**1766** **1:14.6**	**1303** **1:21.8**	**912** **1:17.6**	**2371** **1:13.1**

Key: A = author; R = reader(s); 3P = third party; P = pronoun; O = other; HP = homepage; * = total number of indicated type of authorial self-address per text type; ** = indicated type of authorial self-address-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

The frequencies for the first type of authorial self-address (+A, -R, -3P) illustrated in Table 7.1, concern the uses of first person singular pronouns (+P, -O), i.e. *I*, *me*, *mine*, *myself*. It was observed that the media with an acquainted readership have higher frequencies of this type of authorial self-address compared to Web Chat and the personal homepages. However, these results run the risk of being biased by one crucial factor: ellipsis. It is above all the modern text types that feature frequent ellipses of first person pronouns. This is not only reminiscent of spoken conversation, but also ensures quick typing in, for example, Web Chat (cf. Dittmann 2001: 67ff.), or economises on text length in SMS discourse (cf. Dürscheid 2002a: 109). These instances of first person pronoun omission are not considered in Table 7.1 because they are not by definition authorial self-references, for the simple reason that the textual evidence is missing. However, ellipses of first person pronouns were recorded and amounted to 745 in SMS discourse, 258 in the e-mail corpus, 609 in the Web Chat data, 59 in the personal homepages, and 67 in EEC. For unknown reasons the personal homepage features fewer instances of first person pronoun ellipsis than EEC. Both Web Chat and SMS discourse show frequent omission of first person pronouns, most probably due to the above-mentioned reasons. It can be stated, however, that even in consideration of first person pronoun ellipses, Web Chat and the personal homepages still show a lower frequency of authorial self-address than the other text types.

The second type of authorial self-address illustrated in Table 7.1 concerns references to the author figure other than with pronouns (-P, +O). It was distinguished

between the following types of non-pronominal authorial self-address (+A, -R, -3P / -P, +O): "nominal references" (name, surname, full name, nicknames derived from name, initial(s), or any kind of title), "anonymous nicknames" (i.e. the use of nicknames in Web Chat), "pet names" (such as *darling*, *honey*), "kinship terminology" (for example, *brother*, *sister*), as well as the category "other" for references that do not fit any of the above-mentioned categories. As can be gathered from Table 7.1, SMS discourse shows the highest frequency of this type of authorial self-address compared to the other text types. The initial, probably owing to its brevity, is the most favoured variety of nominal authorial self-address in SMS discourse. It is in most cases part of the farewell section, as illustrated in example (42):

(42) (…) thinking of you & hope to cu soon my darling, love <n.>
(SMS text excerpt / author: female, 24 yrs)

Nominal authorial self-address is also quite frequent in the personal homepage corpus, but opposed to the SMS texts, it is not the initial but the name or full name that was found to be most popular among the homepage authors. Predominant uses of this type of authorial self-address were observed in the greeting section or the main body of the homepages, as illustrated by examples (43) and (44), respectively.

(43) Hey there, and welcome to
<Full name's> Home Page!!!
(Personal homepage excerpt (greeting section), emphasis original / author: Mf_1)

(44) I'm <full name> and I'm in Seattle, Washington, USA.
(Personal homepage excerpt / author: Df_1)

As has been pointed out before, personae as portrayed on the Internet need not necessarily be real. However, this study looks into idiosyncrasies of how personal written communication is contextualised on the level of language. It is therefore less concerned with identity issues. Furthermore, all homepage owners were contacted previous to data collection and analysis, and it turned out that details given in the e-mail exchanges were congruent with those presented on the homepages. It is for this reason that the nominal references in personal homepages were coded with the same procedure as SMS discourse, e-mail, and EEC. For Web Chat, however, the additional category "anonymous nickname" was formulated, in order to accommodate the fact that the identity of the Chatters was unknown. To come back to the types of nominal authorial self-address as found in the

personal homepage corpus, it was found that most references are of the nature illustrated in examples (43) and (44), whereas pet names and kinship terminology are non-existent.

Nominal authorial self-address in Web Chat is most frequent in the form of anonymous nicknames, predominantly in so-called "action turns", in which Chatters write about (verbalised) actions they perform from a third person point of view (cf. Dittmann 2001: 76). Example (45) illustrates such an action turn transacted by a Chatter nicknamed <friendseveryone>:

(45) <friendseveryone> flies up and slams down onto <shakies>
(…)
<shakies>: owww <friends>!
(Chat excerpt / Room IV, session 2)

In example (45), <friendseveryone> "attacks" another Chatter called <shakies> by "slamming down" onto him/her. The Chatter nicknamed <shakies> reacts a couple of turns later by exclaiming *owww <friends>!* With respect to <friendseveryone's> turn, he/she refers to him-/herself by means of his/her own nickname, and this is not only indicated by the syntactic structure of the turn (third person inflection of the verb phrases), but also because no colon appears between the nickname and the content of the turn. While "normal" turns are fronted with the nicknames of the Chatters, in most Chat rooms separated by colons, the nicknames used in action turns are not fronting the turn but are in fact part of the turn, typically in subject position. This kind of nominal authorial self-address is not only the most frequent in the Web Chat corpus, it was also found to be idiosyncratic to the text type Web Chat because it is not present in any of the other corpora.[67]

Although e-mail and EEC do not feature nominal authorial self-address as frequently as the other three text types, both corpora feature them, similar to SMS discourse, on a regular basis in their farewell sections. Authors predominantly refer to themselves by means of their names, full names, or nicknames (derived from the "real name"), although authorial self-reference by means of nicknames are more frequent in e-mail than EEC. Quite rare in both e-mail as well as EEC, and non-existent in all other corpora, is the authorial

[67] On a side note, both Wirth (2005: 77) and Dittmann (2001: 47) mention the circumstance that nicknames, above all the more extravagant ones, often serve as vehicle to facilitate contact in the sense that the meaning and/or significance of a "special nickname" is a first topic to talk about. See Bechar (1995: OD) for a typology of nicknames as observed in a sample of Internet Relay Chat (Bechar defined 14 categories and comes to the conclusion that nicknames of the category "nicknames related to the self" (such as <*shydude*> or <*baddady*>) were the most favoured among the Chatters of her data.

self-address with the term *friend* (pertaining to the category "other"). On a side note, the use of the term *friend* to address the readership is also infrequent. Nevala (2004: 288), looking at forms of address in 17th and 18th-century letters, observed that the term *friend* is commonly used in the material of the 17th century, but slowly vanishes by the beginning of the 1680's. Although the term *friend* is still very much in use in the 21st century in other communicative settings, to refrain from addressing each other by means of this term in personal written communication seems to have its origins in epistolary correspondence.

However, besides referring to themselves by means of their names and full names, authors of EEC show another trend in that they quite frequently refer to themselves by means of kinship terminology, as illustrated in example (46):

(46) (…) And thus I rest,
Your lovinge brother, ready to doe you service,
Ed. Bacon
(EEC excerpt (farewell section), 1628 / letter collection Cornwall, author: male)

Other than kinship terminology, which is rare in all other text types, authors of EEC also frequently opt for the term *servant* for self-address (chiefly used by male letter writers), a terminological idiosyncrasy tied to the time from which the letters originate.

The third type of authorial self-address listed in Table 7.1 concerns the use of pronouns (+P, -O) that refer to both author and readership (+A, +R, -3P), such as *we*, *us*, and *ours*. This form of authorial self-address is infrequent in its distribution across the five text types, as are in fact all other forms of authorial self-address listed in Table 7.1. There are, however, some trends worth paying attention to. Although generally rare in its distribution, authorial self-address by means of pronouns that include both author- and readership is still employed considerably more often by authors of SMS discourse, e-mail, and Web Chat compared to the personal homepage and EEC. From the three pronouns mentioned above, it is *we* that occurs with the highest frequency. Example (47) shows a typical employment of the pronoun *we* (+A, +R, -3P) in SMS discourse:

(47) do you remember? we've got a date on Friday eve, (…)
(SMS text excerpt / author: female, 26yrs)

Most of the instances of *we* observed in the SMS text corpus and e-mail data are used in similar contexts as shown in example (44), namely to arrange a date/meeting between author- and readership that is yet to happen, or to discuss a date/meeting that took place in

the (near) past. Although less frequent, the pronoun *us* (+A, +R, -3P) is predominantly used in the same way in both text corpora.

Another trend in Table 7.1 concerns the circumstance that non-pronominal authorial self-address that includes both author and readership (+A, +R, -3P / -P, +O) is extremely rare to literally non-existent (the personal homepage corpus lacks this type of form of authorial self-address completely). The same is true for non-pronominal authorial self-address that includes both authorship and any other third party (+A, -R, +3P / -P, +O). However, another type of authorial self-address concerns the reference to both author and another third party by means of pronouns (+A, -R, +3P / +P, -O), and this is in fact quite frequent in both e-mail as well as EEC. Similar to the pronominal authorial self-address including readership above, the pronouns *we* and *us*, in another context referring to authorship including any other third party, are most frequently employed. Of course, in order to be able to determine whether a first person plural pronoun refers to author and readership, or to authorship and any other third party, the context in which the pronoun is used must be known. In the case of this investigation, this context is provided by the textual environment in which the pronouns occur. Examples (48) and (49) illustrate the employment of the pronouns *we* and *us* (+A, -R, +3P) in e-mail correspondence and EEC, respectively.

(48) <Name> and I have made a promise to each other that we will play more often and together as much as we can. I wished she lived closer.
(E-mail excerpt / author: female, 40 yrs)

(49) Deare Sister, in my other letters I did desire you to send us ouer your monie; which, if you please, should bee very wellcome vnto us. (…)
(EEC excerpt, 1636 / letter collection Cornwall, author: female)

Even from the small textual evidence provided in examples (48) and (49) it becomes clear that the uses of the pronouns *we* and *us* exclude the intended readership and refer to another third party instead. This is particularly frequent in longer texts, indicating that this specific use of these pronouns is tied to a certain text volume. Interestingly, this form of authorial self-address (+A, -R, +3P / +P, -O) is most infrequent in the two media that generate correspondence for an unknown readership, the personal homepage and Web Chat. This implies that this particular form of authorial self-address is not only tied to text volume, but also to the relationship between participants and the shared contextual background the correspondence is performed against.

Overall, it can be asserted that authorial self-address occurs on a more frequent basis in personal written communication that is addressed to an acquainted, as opposed to unknown, readership. Further, if the frequencies of authorial self-address in the Web Chat corpus and personal homepage data are compared, then it becomes apparent that they are more frequent in the personal homepages. This most probably has to do with the circumstance that the authors present their personae on those pages and, by referring to themselves, manifest these personae. This stands in stark contrast with Web Chat, where users are a lot less likely to inform fellow Chatters who they are. On the contrary, they are in most Chat rooms repeatedly reminded not to reveal their identities and share personal details, as illustrated in example (50):

(50) <The operator of Room III> reminds everyone NOT to give out email addresses, phone numbers or any other contact details especially in public chat rooms. This is for your own safety. Thank you.
(Chat excerpt, emphasis original / Room III, session 2)

In all of the Chat sessions recorded, there is not one instance of a Chatter ignoring the above warning issued by the operator of Room III. It seems as though Chat room users are aware of the policy to remain anonymous in Chat rooms. This is probably also the reason why the frequency of authorial self-address is lowest in the Web Chat data. In how far the parameter anonymity can be brought into connection with the uses of forms of address referring to the readership, will be of interest next.

7.1.2. Authorial address of readership

The advent of the Internet has introduced a completely new perspective into the field of personal written correspondence: the readership need no longer be narrow or specific. It has become possible to personally communicate with an unlimited number of known as well as unknown addressees at any given time. While the authors of EEC compose writings intended for a distinct readership, participants in anonymous online settings, such as Chat rooms, and authors of personal homepages, who have in most cases no idea who will read their publications, write for an audience they can only imply because they do not know them. While authors of e-mails know their recipients, this medium has also brought changes in connection with readership, because an e-mail can be sent to one or 500 recipients without much additional effort. The messages that compile the e-mail corpus for this study, however, are directed at one recipient or, in a few cases, a small group of

recipients. SMS discourse, on the other hand, similar to the handwritten letter, is generally aimed at one specific recipient. In any event, it can be assumed that the changes in the communicative context regarding readership have an influence on how authors address the readers, implied or specific, of their personal written communication.

Similar to authorial self-address, the readership can be exclusively addressed (-A, +R, -3P) or together with another third party (-A, +R, +3P). As the combined address of author/readership has already been discussed in connection with the author, this variety will not receive further attention. However, concerning the two types of authorial address of readership mentioned above, they can both be realised by means of pronouns (+P) or other non-pronominal references (+O). For example, exclusive readership address can be achieved by the use of the second person singular pronoun *you*, or by a non-pronominal reference such as *Anna*, or *mate*. Readership address that includes another third party may be realised through the use of second person plural pronoun *you* (as in 'you and someone else'), or by means of a noun phrase such as *your group*. Analogue to the authorial self-address, exclusive readership address by means of personal pronouns turned out to be the most frequently used form of address. Table 7.2 summarises the results that the investigation into different types of readership address generated.

Table 7.2: Authorial address of readership in personal written communication (corpora-based results).

Types of authorial address of readership	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
(-A, +R, -3P / +P, -O)	* 1137 ** 1:16.2	719 1:35.8	873 1:32.5	136 1:117.9	1057 1:29.4
(-A, +R, -3P / -P, +O)	408 1:45.2	139 1:185.1	1521 1:18.7	35 1:458	353 1:88
(-A, +R, + 3P / +P, -O)	4 1:4606.5	5 1:5146.6	0	0	0
(-A, +R, + 3P / -P, +O)	12 1:1535.5	14 1:1838.1	6 1:4734	0	32 1:971.2
TOTAL	**1561 1:11.8**	**877 1:29.3**	**2400 1:11.8**	**171 1:93.7**	**1442 1:21.6**

Key: A = author; R = reader(s); 3P = third party; P = pronoun; O = other; HP = homepage; * = total number of indicated type of authorial address to readership per text type; ** = indicated type of authorial readership address-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

There is one trend in Table 7.2 that stands out: the high frequency of authorial nominal address of readership (-A, +R, -3P / -P, +O) in Web Chat. While the pronominal address of readership (-A, +R, -3P / +P, -O) is nothing unusual in a conversation-like discourse structure like Web Chat, the reiterative use of nominal address of readership is noteworthy. It was found that this feature is tied to the habit of Chatters to repeatedly include the

nickname of the addressee in their turns throughout a discussion (see hereto also Storrer 2000: 443ff.). Consider example (51a.) below.

(51a.) <*machine-generated turn*>: <*HuggyBear*> *has joined the chat.*
(…)
<StarDust>: Excuse me <huggybear>?
(…)
<StarDust>: <Huggybear>????
<HuggyBear>: yes <star> ?
(…)
<StarDust>: asl [‘age, sex, location’] <huggybear> please?
(…)
<HuggyBear>: 34m <star>
(…)
<StarDust>: In a relationship or looking <huggy>???
(…)
<HuggyBear>: always looking <star>
(Chat excerpt / Room IV, session 3)

Example (51a.) shows a Chat excerpt with a typical exchange between two Chatters, including in their every turn the nickname of the intended addressee. Upon login of the Chatter nicknamed <HuggyBear>, another Chatter, <StarDust>, asks for <HuggyBear’s> *asl*, meaning his/her ‘age, sex, location’. <HuggyBear> responds to this question by providing <StarDust> with his/her alleged age and sex (*34m* meaning ‘34 years, male’). The conversation then continues with <StarDust> asking whether <HuggBear> is single or not (*In a relationship or looking <huggy>???*) to which <HuggyBear> responds *always looking <star>*, which implies that he/she is single or in a relationship, but nonetheless looking for companionship. Particularly interesting about this exchange is that both Chatters include the other Chatter’s nickname in each and every turn. In example (51a.) this may seem superfluous because it is clear at whom the turns are aimed. The reason for this is that all other turns in between have been removed. However, in a setting where up to 200 Chatters are chatting at the same time, not only in binary structures but also in a “cross-over” fashion, it becomes more of a necessity to identify which turn is directed to whom. This becomes obvious in example (51b.), where all of the turns that (co-)occurred between the exchanges of <HuggyBear> and <StarDust> are shown. The backward arrows indicate their turns as illustrated in (51a.).

(51b.)
01 <*machine-generated turn*>: <*HuggyBear*> *has joined the chat.* ←
02 <Dream>: omg what is her job <jo>.....i can't remember....lmao
03 <iz79>: oh you sound like someone i know...sorry
04 <*machine-generated turn*>: <*nantida*> *has left the chat.*
05 <kcham32>: so what is my job then?
06 <abby>: goodnight <milly> tc * hugs
07 <Dream>: ((((((((((((<huggybear>))))))))))))wb
08 <HuggyBear>: <dreamy>!!! xxx
09 <*machine-generated turn*>: *scrivo24 has joined the chat.*
10 <jo>: oh oh...<huggybear> is the next bloke in the room! sic em <dream>…
11 <gucci> ...<greatmum>!!
12 <HuggyBear>: ((((<jo>))))
13 <*machine-generated turn*>: *milly1 has left the chat.*
14 <guccigirl>: wb <huggy>
15 <Dream>: ((((((((((<milly1>))))))))))))))))))
16 <jo>: ((((((((((<huggy>))))))))))
17 <HuggyBear>: OMG....wot is going on???????????
18 <StarDust>: Excuse me <huggybear>??? ←
19 <Dream>: bugger too slow
20 <*machine-generated turn*>: *quad_guy_uk has joined the chat.*
21 <StarDust>: <Huggybear>???? ←
22 <HuggyBear>: yes <star> ? ←
23 <kool_kat>: hey <huggy> coming on tomorrow
24 <StarDust>: asl <huggybear> please? ←
25 <liz79>: if you don't like missionary <kcham> how about tied to the bed by
26 your ankles
27 <quad_guy_uk>: ((<gucci>))
28))))))))))))))))))))))))
29 <abby>: hi quad
30 <quad_guy_uk>: ((((((((((((((((((<abby>))))))))))))))))))
31 <HuggyBear>: 34m <star> ←
32 <jo>: lmao <liz79>! go girl!!!!
33 <StarDust>: In a relationship or looking <huggy>??? ←
34 <quad_guy_uk>: ((((((((((((((((((((((((((((<greatmum>))))))))))))))))))))))))))
35 <Lady_Grey>: ((((((((((<quad guy>)))))))))))))))) :/
36 <HuggyBear>: hi kool...prob
37 <guccigirl>: ((((((((((((((((((((((((((((((<quaddles>))))))))))))))))))))))))))))))
38 <greatmum>: ((((((((((((((((((<quad>))))))))))))))))))
39 <HuggyBear>: always looking <star> ←
(Chat excerpt / Room IV, session 3, number of logged in Chatters: 22)

If one considers that the above Chat excerpt unfolded in under a minute, it becomes apparent that identifying the intended addressee, or addressees (as is the case in line 10), of a turn is crucial in order to be able to keep track of who is writing what to whom. It also illustrates the usefulness of the "coloured nickname feature" as observed in Room IV (if a Chatter uses a particular nickname in one of his/her turns, it appears in pink on the screen of the Chatter with the respective nickname, cf. 6.3.1.). The contents of example (51b.) will not be discussed any further, but it serves as a prime example to illustrate and explain why the frequency of nominal readership address in the form of (anonymous) nicknames is

so high in Web Chat. This is, similar to the action turns identified in connection to authorial self-address above, idiosyncratic to the text type Web Chat as none of the other text types feature this kind of nominal readership address. However, it needs to be kept in mind that the frequency of this particular authorial readership address is tied to the number of Chatters present in a given Chat discussion.

Another trend in Table 7.2 includes a high frequency of pronominal readership address (-A, +R, -3P / +P, -O) in SMS discourse. Since all personal pronouns are short, this is probably connected to the intention to economise in text length. Another reason for this could be that authors of SMS texts have a tendency to omit the greeting section (cf. Table 6.1 in the previous chapter), which would in fact be predestined to address the readership with a pet name or his/her name (or a variation thereof). E-mail and EEC, on the other hand, although showing lower frequencies for nominal readership address than SMS discourse, incorporate the majority of nominal readership address into their greeting sections. While authors of e-mail show a tendency to opt for nicknames (derived from names) and pet names, authors of EEC almost exclusively address their readership in the greeting section with their names, full names, or some kind of kinship terminology. Interestingly, Tieken-Boon van Ostade (2006), who looks into the use of abbreviations in 18th-century letters, observes that abbreviations are also sometimes used in addressing the recipients of the letters. Tieken-Boon van Ostade (2006: 244) comes to the conclusion that the main function of abbreviations in letters "is to speed up the process of writing" and that there is a "higher occurrence of them in letters in which appearance is subordinate to contents," such as draft letters and letters to close friends. For unknown reasons abbreviations emerged to be in general extremely uncommon in the corpus of 17th-century letters investigated for this study. Whatever the reasons may be, it implies that for these authors the appearance of the letters is not subordinate to contents.

Coming back to Table 7.2, it emerges that frequencies of readership address that includes another third party, both pronominal (-A, +R, + 3P / +P, -O) as well as nominal (-A, +R, + 3P / -P, +O), are very low for all five discourse types. However, as will be discussed shortly, the references to any other third party separate from readership are considerably more popular. But as far as readership address is concerned, it can be asserted that authors of personal written communication, modern or traditional, predominantly address the recipients of their correspondence on a one-to-one basis.

With regard to differences in connection with the communicative context, it becomes apparent that messages intended for an acquainted readership (SMS discourse,

e-mail, EEC) show comparable distributions across the different types of authorial readership address, whereas both Web Chat and the personal homepage show peaks that deviate—albeit in different directions. While Web Chat features the highest frequency of nominal readership address (with a total of 1521 instances, or 1 instance in every 18.7 words), the personal homepage features the lowest (with a total of 35 instances, or 1 instance in every 458 words). The majority of non-pronominal readership address in Web Chat is connected to the above-explained high frequency use of (anonymous) nicknames, the personal homepage corpus, on the other hand, features other types of non-pronominal readership address such as *visitor*, *explorer*, and *surfer*. This, of course, is tied to the circumstance that authors of personal homepages in most cases do not know their readers. It is assumed that the pronominal readership address is for similar reasons kept to a minimum in the homepages.

Thus, because Chatters engage in communicative exchanges with many strangers that are online at the same time, they have to make sure their messages reach the right destination and thus tend to include the intended recipient's nickname in their turns (an observation supported by Runkehl et. al (1998: 86) and Dittmann (2001: 53), results based on empirical investigations into German and French Chat). Furthermore, of all turns directed to the readership (i.e. the other Chatters) in Chat rooms, it was observed that five times as many references are directed to one specific Chatter by means of his/her nickname, as opposed to the whole room or another group of Chatters. The personal homepage authors, on the other hand, have no idea who is going to read their material and authorial address of readership is thus less frequent. Both of which can be said to distinguish these two media from media intended to communicate with an acquainted readership.

7.1.3. Authorial address of third party

The last type of specific personal reference of interest to the current study is how authors include references to any other third parties perceived to be separate from both author as well as readership (-A, -R, +3P). These may, analogue to authorial self-address and authorial address of readership, be realised by means of third personal pronouns (singular or plural (+P), such as *he* or *them*) or through other, non-pronominal (+O) references in singular, such as *Carl*, or plural, as in *Mr. and Mrs. Brown*. The use of third party references presupposes a shared (pre-)communicative context or, if the third party is expected to be unknown to the reader, requires some sort of introduction within the

message in order to identify this other third party. It can thus be hypothesised that frequencies of authorial address of third parties will be lower for the anonymous communicative settings Web Chat and personal homepage, as well as the text type SMS discourse for reasons of brevity. The empirical investigation into the five text corpora, however, only partly confirms these assumptions. The results are summarised in Table 7.3 below.

Table 7.3: Authorial address of third party in personal written communication (corpora-based results).

Types of authorial address of third party	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
(-A, -R, + 3P / +P, -O)	* 58 ** 1:317.7	323 1:79.7	212 1:134	66 1:242.9	1020 1:30.5
(-A, -R, + 3P /-P, +O)	193 1:95.5	335 1:76.8	145 1:195.9	301 1:53.3	831 1:37.4
TOTAL	**251** **1:73.4**	**658** **1:39.1**	**357** **1:79.6**	**367** **1:43.7**	**1851** **1:16.8**

Key: A = author; R = reader(s); 3P = third party; P = pronoun; O = other; HP = homepage; * = total number of indicated type of authorial address of third party per text type; ** = indicated type of authorial third party address-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

As can be gathered from Table 7.3, SMS discourse and Web Chat reveal low total frequencies regarding authorial address of any other third party. Hence, the results from the analysis into Web Chat and SMS discourse confirm the above assumptions. In connection with SMS discourse, this is thought to be tied to the circumstance that authors of SMS texts faced restrictions in text length at the time of data collection. It is feasible that authors thus concentrated on the addressee, rather than introduce other third parties into the discourse. Web Chat, on the other hand, was expected to feature few references to other third parties because the shared communicative context is in most cases limited to the Chat Room.

The personal homepage, on the other hand, contains more third party references than assumed. Interestingly, the vast majority of these references are of the nominal type (names and full names) and accompanied by further information on these people. It seems as though personal homepage authors aim at briefing their unacquainted readership, so that they can in fact make sense of such third party addresses. Example (52) illustrates a typical incorporation of third party address as found in the personal homepages.

(52) Other cool people
- The official Lynda Barry Page, set up by yours truly. Ms. Barry is a brilliant cartoonist, artist and writer.
- Joel & Ethan Cohen, the brothers responsible for such great films as *Raising Arizona* and *Fargo*.
- <Full Name>, my best friend since the fifth grade.

(Personal homepage excerpt / author: Mf_1)

Not only does the author of the above homepage excerpt provide background information on the third party references in the accompanying running text, each of them is further equipped with a hyperlink to other websites that contain yet more information on these people. This type of third party address is so predominant in the personal homepage that it can be classified as characteristic. In comparison, Chatters are in most cases not able to provide background information about other Chatters because they do not know them personally. Furthermore, providing background information on "real people" (that are not present in the Chat Room) is advised against in most Chat rooms.

It is an entirely different matter with e-mail and EEC, where third party references are not only more frequent than in all other text types, authors also often do not provide any background information on these people. The reason for this is that they presuppose this kind of knowledge in their readership. Example (53) exemplifies a prototypical third party address in EEC:

(53) My most honoured Lady,
There was one question my Lady Barrington asked me when she was in towne that I would desire to be provided of answer from you for her against she come. She desired to know, if Sir William Curteen asked what portion you would demand, what she should say. (…)
(EEC excerpt, 1629 / letter collection Cornwall, author: female)

If the intended addressee of the above example (53) does not share the author's contextual background on *Lady Barrington* and *Sir William Curteen*, then this passage would not make much sense to the reader. However, it is strongly assumed that the intended reader of the above letter was in fact familiar with these two people and that for this reason no background information is given. On the contrary, it would seem bizarre if the author had provided personal details for people most likely acquainted with the reader.

Overall, the assumption that authorial address of third parties is less frequent in anonymous settings was confirmed. However, in the case of the personal homepage, such references are likely to be accompanied by background information on these third parties, in order to familiarise the readership with all the people mentioned on the homepage. This

stands in stark contrast with Web Chat, where users are advised not to reveal their personal details, or the details of anybody they know. Although there is the occasional reference to another Chatter who is absent, Chatters are in general not able to provide any background information about each other for the simple reason that they do not have any.

7.1.4. Missing forms of address

In principle, every message written for an (implied) readership can be classified as addressing that readership on a general level. However, as we have seen above, there are distinct ways of how authors can explicitly address themselves or their readership. Messages may, however, also lack specific forms of address. Yet it was observed that, apart from SMS discourse, text entities that do not contain authorial self-address and/or readership address are extremely rare. For example, none of the Web Chat sessions lacks authorial self-address or references to the readership. The same is true for EEC, where all text entities contain both authorial self-address as well as references to the readership. And while the e-mail corpus contains three text entities that do not feature any type of authorial self-address, the personal homepage corpus includes three homepages that do not feature any kind of readership address. These numbers are small compared to the number of messages that lack author and/or readership address in SMS discourse.

In the SMS text corpus, almost a fifth of the messages (196 texts or 19.6%) lack any form of authorial self-address, and another 13.7% (137 texts) do not feature any kind of authorial readership address. All in all, a total of 61 SMS texts (6.1%) lack both authorial self-address and readership address. Two typical SMS messages of this kind are shown in examples (54) and (55) below:

(54) ;-) hugs!
(SMS text / author: male, 28 yrs)

(55) Ok, no worries… by the way the bed is not here yet :-(
(SMS text / author: male, 26 yrs)

As pointed out above, the circumstance that both SMS texts in examples (54) and (55) are intended for specific recipients, qualifies them as readership-oriented communicative acts. However, neither of the two SMS texts contains specific forms of address with regard to either author or readership. The tendency to omit forms of address was found to be connected to SMS discourse being also the text type most likely to omit greetings and/or farewells (sections predestined to contain authorial readership address and self-address,

respectively). Since none of the other text types shows trends in this direction, this can be classified as idiosyncratic to SMS discourse.

7.1.5. Impersonal *you* and *one*

Messages with missing author or readership address have shown to be infrequent, except in SMS discourse. Another way of not specifically addressing author or readership can be achieved by either using the second person impersonal pronoun *you* (as in *You would not expect something like this to happen*), or the third person impersonal pronoun *one* (as in *One is surprised by the outcome*). The scope of reference of impersonal *you* or *one* is large, they both include author and readership to a certain extent. However, by being impersonal, direct address (or confrontation) is avoided. Generally, the use of impersonal pronouns is more common in formal registers (such as academic writing) or fiction, where they are often employed to outline subject matters or describe events without too much author or reader involvement (impersonal *one* is often perceived as more formal than impersonal *you*). Since personal written communication is aimed at being *personal*, frequencies of impersonal pronouns can expected to be low.

This expectation was confirmed by the investigation into the five text corpora. While the impersonal pronoun *you* occurs rarely in all the text types, the impersonal pronoun *one* is non-existent. No specific tendencies could be determined for the use of the impersonal pronoun *you*, except that it is slightly more frequent in the e-mail corpus and EEC (appearing approximately once in every 600 and 900 words, respectively). This, however, has less to do with the e-mails and letters being more formal than the other text types, but more with the narrative style of some of the messages (this issue will be re-addressed in chapter 12), as texts that contain narrations of events are more likely to feature the impersonal *you* (reminiscent of fiction writing). However, overall it can be asserted that personal written communication, being of an informal nature, contains in general few impersonal pronouns.

7.2. Chapter summary

The investigation into personal reference exposes several trends in connection with how authors of different text types address their readership and other third parties in their personal written communication. While the frequencies of the employment of available communication channels (chapter 6) have shown to correlate with the degree of synchrony of the dialogues (the more synchronous the dialogue, the more communication channels

authors tend to employ), the investigation into forms of address shows the main differences between correspondence with an acquainted readership (SMS discourse, e-mail, EEC) and communication with an unknown readership (personal homepage, Web Chat).

Overall, it can be stated that authorial self-address and references to any other third parties occur more frequently in personal written communication that is addressed to an acquainted, as opposed to unknown, readership. This confirms H5, which claims that the type of readership ought to foster differences in how personal written communication is contextualised. It is a slightly different picture with authorial readership address, where the text types aimed at acquainted readers (SMS discourse, e-mail, EEC) show more or less comparable distributions, whereas Web Chat (highest frequency together with SMS discourse) and the personal homepage (lowest frequency) deviate into different directions. Furthermore, no distinctive patterns in connection with the uses of the impersonal pronouns *you* and *one* could be determined. They were found to be generally infrequent in personal written communication. Collectively, these results contradict the main research hypothesis H4, which claims that modern types of communication ought to show considerable differences in their contextualisation compared to EEC.

However, certain findings (partly) confirm research hypotheses H1 and H2. For example, the high frequency of readership address in Web Chat by means of (anonymous) nicknames to identify the intended recipient of the turns, and the tendency of personal homepage owners to accompany third party references with a hyperlink can be classified as idiosyncratic to CMC (confirming the second part of H1, which claims CMC to show features of cyberdiscursivity). Furthermore, the inclination of SMS texts authors to use pronouns to address their readership is reminiscent of spoken discourse (confirming H2). How the authors of EEC incorporate forms of address into their letters could not be classified as typical of literacy. On the contrary, pronominal readership address is high in frequency, which is reminiscent of a spoken conversation (contradicting H3).

As pointed out above, the text type letter does not show any trends with regard to how authors contextualise themselves on the level of personal reference that isolate it from the modern types of correspondence (contradicting H4). If at all, then the only difference that was observed concerns the type of (acquainted) readership. Many of the letters in the EEC corpus contain references in the form of kinship terminology, which is an indicator that the intended readers are family members. In comparison, the uses of kinship terminology in the modern text types are negligible. This would imply that they are in fact not intended for family members, but to other people close to the authors, such as friends

and partners. However, family, friends, and partners are comparable addressees with regard to shared (pre-)communicative knowledge. This factor is therefore felt to be too weak so as to confirm the hypothesis that the contextualisation of EEC is considerably different compared to the modern types of correspondence.

8. Textual reference: texts as networks

While personal reference, as analysed in chapter 7, is tied to the people that compose and receive correspondence, textual reference is concerned with the correspondence itself. Textual deixis is relevant on two levels: text-*internal* and text-*external*. While text-internal, i.e. endophoric, references provide information on textual structures within a text, text-external, i.e. exophoric, references place text in relation to other text(s) and can thus be seen as indicators of larger textual networks that go beyond a particular text. Both types of textual references are important means of contextualisation because they allow authors to internally organise their messages, as well as identifying them as discourse units within larger discourse structures.

The primary function of deictic terms in textual deixis is to navigate the reader through the text, and "the most common terms used in this kind of discourse deixis are taken from the semantic spheres of time and place" (Claridge 2001: 55, based on Fillmore 1997: 103-104). Most interesting for this discussion is Clardige's (2001: 56) observation that temporal expressions used in discourse deixis to organise text internally, such as *earlier* and *now*, are potentially more oral in nature than spatial deictic terms, such as *above* and *below*, which reflect the physical aspects of text. It is important to point out that the treatment of discourse deixis in this chapter touches upon issues dealt with in upcoming chapters 9 (*The spatiality of text*) and 10 (*Text is time-bound*). It should be noted, however, that the main focus of this chapter is placed on the contextualisation on the "textual level" by means of deictic expressions. Chapters 9 and 10, on the other hand, will be concerned with the notions of space and time on the textual level and its relations to the "dimension of space" beyond the text (physical and virtual) and "real time" (in reference to the calendar).

In any event, the type of inter-textual networking that explicitly relates to other (past or future) messages, and thereby identifies a particular message as part of a larger discourse structure, is more typical of written correspondence than spoken conversation, which tends to be less organised and more concerned with the present (cf. Baron 2000: 21). It can thus be assumed that the more synchronous and conversation-like discourse types, such as Web Chat, contain more temporal expressions for endophoric discourse deictic purposes. Furthermore, it can also be hypothesised that less extra-textual networking occurs in messages composed on the more synchronous media. In reverse, the

asynchronous text types EEC and personal homepage are expected to show higher frequencies of spatial deictic terms that organise discourse on a text-internal level, and that they feature more indicators for extra-textual networking. In the following sections, different features of intra-textuality (8.1.) and extra-textuality (8.2.) of personal written communication will be discussed. It will be interesting to see whether or not the investigation into the five text corpora will confirm the above-made assumptions.

8.1. Intra-textuality of personal written communication

When authors give directions in written text, it implies two things: first, the text is organised according to certain criteria and second, the text is written for an intended readership that can in fact make use of such directions. On another level, text-internal discourse deixis is also an indicator in how far a particular text can be perceived as a coherent and, to a certain extent, independent unit. With respect to personal written communication, this means that an investigation into endophoric discourse deixis sheds light onto how authors contextualise their messages from a text-internal structure point of view. Textual deixis of this type does not require knowledge of past or the anticipation of (potential) future messages, but concentrates on the current message itself as a realm of reference. It can thus be expected that text-internal references are less frequent in shorter messages compared to longer ones, because authors can assume that the audience of shorter texts require fewer structural aids to orient themselves.

8.1.1. Text-internal (endophoric) reference

As pointed out above, text-internal discourse deixis is likely to draw on the vocabulary from temporal and spatial deixis. While (spontaneous) spoken discourse is said to have a preference for temporal deictics, written discourse is more likely to feature spatial deictic expressions in connection with text-internal reference. With respect to written discourse, there are two main directions to which endophoric references may point to in a text that is arranged in linear fashion: anaphoric (backward) reference and cataphoric (forward) reference to particular text segments. In addition, an author may also refer to the text (or message) as a whole.

Any given message in personal written communication can thus be internally contextualised in terms of (1) anaphoric (such as *see above*) and (2) cataphoric references (such as *see below*) to text segments within the same discourse unit, as well as (3) references to the message as a whole (for example, *I've finally come round to write this*

e-mail), which at the same time identifies a particular message as an independent discourse unit. Table 8.1 gives an overview of the frequencies with which the authors of the five text corpora included those three types of endophoric reference in their correspondence.

Table 8.1: Endophoric reference in personal written communication (corpora-based results).

Types of endophoric reference	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
anaphoric ref. to particular text segments	0	* 21 ** 1:1225.4	312 1:91	30 1:534.3	33 1:941.7
cataphoric ref. to particular text segments	0	11 1:2339.4	89 1:319.1	62 1:258.5	33 1.941.7
ref. to the message as a whole	12 1:1535.5	38 1:677.2	6 1:4734	162 1:99	52 1:597.6
TOTAL	**12** **1:1535.5**	**70** **1:367.6**	**407** **1:69.8**	**254** **1:63.1**	**118** **1:263.4**

Key: ref. = references; HP = homepage; * = total number of occurrences of indicated type of text-internal reference per text type; ** = indicated type of endophoric reference-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

It is apparent in Table 8.1 that SMS discourse completely lacks both anaphoric and cataphoric text-internal references. The only type of endophoric reference found in the SMS text corpus is when authors refer to the message as a whole, as shown in example (56) below:

(56) *this*is*an*I*beam*you*sunshine*sms* (…)
(SMS text excerpt / author: female, 22 yrs)

By referring to the message as a whole, the author of example (56) objectifies her SMS text in that it has become part of the message, rather than just transmitting it. The low frequencies of endophoric reference in SMS discourse is definitely tied to the length of the text units that are, at an average of 18.4 words per text entity, extremely short. No preferences with regard to particular deictic expressions could be determined other than the use of *this (SMS)*, which occurs three times in the corpus.

With respect to the remaining text types, e-mail and EEC show comparable frequencies and the same is true for Web Chat and the personal homepages. Text-internal references are, however, more frequent in the messages aimed at an unknown readership. Those differences do not only seem to be connected to text length, because, although the average word count of e-mail (183.8 words) is considerably lower than for the Web Chat sessions (946.8 words), the average word count for EEC (341.5 words) is higher than for the

personal homepage (267.2 words). It can thus be assumed that the qualitative difference in readership plays a certain role.

It was found that e-mail and EEC show a slight preference for endophoric reference that refers to the message as a whole. Examples (57) and (58) illustrate prototypical uses of this type of reference for both discourse types:

(57) (…) i am sorry this is so short but i will write again soon.
(E-mail excerpt / author: female, age unknown)

(58) (…); and I hope you are nowe well at Heariford, wheare it may be, this letter will put you in minde of me, and let you knowe, all your frinds heare are well; (…)
(EEC excerpt, 1625 / letter collection Harley, author: female)

In example (57), the author uses the demonstrative pronoun *this* to refer to the whole e-mail, for which she apologises because it is so short. The author of example (58), on the other hand, makes use of the noun phrase *this letter* to refer to the message as a whole. The letter is thought to let her husband know that she is thinking about him, and that all his friends are well. Similar to SMS discourse, constructions with *this* or *these* are more frequent than others. In the e-mail corpus, 10 out of 38 references to the whole message consisted of the demonstrative pronoun *this*, or the noun phrases *this e-mail* or *this note.* The letter corpus shows even higher frequencies: 17 references consist of the demonstrative pronouns *this* or *these*, and another 20 references to the message as a whole are noun phrases of the type *this letter* (or: *bearer*, *messenger*), or *these lines*.

With regard to anaphoric and cataphoric text-internal reference, it was observed that neither e-mail nor EEC feature them in abundance. On the contrary, these types of endophoric references are few and far in between. In the rare cases of endophoric reference in EEC and e-mail, there is a tendency to use spatial deictics to refer to text segments in the same text entity:

(59) Way back in December 2005, it took me about 8 hours to do my first perfect pyramid using the procedure described above.
(E-mail excerpt / author: male, 61 yrs)

In the e-mail excerpt shown in example (59), the anaphoric spatial deictic term *above* refers to a previous text passage in the same e-mail with a description of the procedure "how to do a pyramid". Next to the low frequencies of anaphoric and cataphoric text-internal references for both e-mail and EEC, no distinctive preference for one over the other

could be determined. Interestingly, Claridge (2001: 61), who analysed how authors of Early Modern English tracts use discourse deixis to direct their readership, comes to different conclusions as "the most striking difference between backward and forward reference terms is the greater frequency of backward reference terms" and "that the authors are not particularly concerned with pointing their readers to upcoming parts of the discourse" (Claridge 2001: 62). However, this may be connected to the circumstance that the *Lampeter Corpus* is compiled of published, as opposed to personal, texts in areas such as politics, religion, or economy. It seems as though that great differences exist with regard to endophoric reference between published and personal texts from the Early English Period.

The investigation into the use of text-internal references in the media fostering personal written communication with an unknown readership, namely Web Chat and the personal homepage, yielded different results. As can be gathered from Table 8.1, text-internal reference is considerably more frequent in Web Chat than in SMS discourse, e-mail, and EEC. Furthermore, the distribution of the different types of text-internal reference is not as equal as with the discourse types directed at an acquainted readership. While the majority of text-internal references in Web Chat are of the anaphoric type, authors of personal homepages show a preference to refer to the text entities as a whole. Examples (60) and (61) illustrate typical endophoric references as found in the two text types.

(60) This page is no longer under construction. Of course the opinions herein are my own, (…)
(Personal homepage excerpt / author: Xm_1)

(61) <zxkzxk>: do u like climb the mountains?
(…)
<zxkzxk>: any good equipments u wanna tell me
<MYNAMEIS>: good equipment? what do u mean?
(…)
<zxkzxk>: i mean gears
(…)
<zxkzxk>: such as pack or shoes
<MYNAMEIS>: gears what do u mean? sorry
(…)
(Chat excerpt / Room IV, session 1)

The author of the excerpt shown in example (60) refers to his homepage twice as a text entity, first with the noun phrase *this page* and then again with the spatial expression

herein. References to the homepage as a whole by means of the preposition *here* are very frequent (the spatial deictic term *here* will be discussed in more detail in chapter 9, in particular 9.1.1.). However, it is a different picture with Web Chat, which does not contain many spatial deictics to refer backward or forward within the same Chat session. Example (61) illustrates prototypical anaphoric text-internal references by a Chatter nicknamed <MYNAMEIS>. As a matter of fact, most anaphoric text-internal references in Chat concern text segments that are misunderstood, or not understood at all. Chatters often express such misunderstandings by repeating the concerned text segment along with a question mark. In example (61), the Chatter nicknamed <MYNAMEIS> further illustrates his/her puzzlement by adding *what do you mean?* to the requests for clarification regarding the meaning of *good equipment* and *gears*. This is reminiscent of a spoken conversation and is connected to the (near-)synchronous discourse structure of Web Chat (we will return to the issue of misunderstandings in Web Chat in connection with contextual effects in chapter 12).

Overall, it emerged that endophoric reference is to a certain extent tied to text length as well as the type of readership the messages are aimed at. SMS discourse features hardly any kind of text-internal reference, which is most probably connected to the brevity of the text entities. Also, the messages aimed at an acquainted readership (SMS discourse, e-mail, and EEC) contain considerably fewer text-internal references than Web Chat and the personal homepages. Furthermore, the (near-)synchrony of Web Chat further fosters the use of anaphoric references that are neither spatial nor temporal in nature, as there is nothing inherently deictic about the repetition of a text segment. However, coupling it with a question mark indicates that something is not quite clear with a previous text segment and that clarification is necessary. The assumption that the (near-)synchronous text type Web Chat contains more temporal discourse deictics, owing to its reminiscence of spoken discourse, could not be confirmed. Yet the anaphoric text-internal reference by means of repetition is not only reminiscent of spoken discourse, but at the same time distinguishes Web Chat from all other types of correspondence where nothing analogous could be observed. Also, the investigation into the more asynchronous personal homepages confirmed expectations in that they show a clear tendency for the use of spatial terminology to refer to other text segments within the same discourse entity. This is, of course also tied to the circumstance that the text type personal homepage is located in virtual *space* (see hereto chapter 9). However, of interest next is a discourse unit that connects text segments in specific ways with each other, namely the discourse marker.

8.1.2. Discourse markers

As has been pointed out in the theoretical discussion of DMs (see 4.3.1.), the main function of DMs (such as *and*, *however*, *furthermore*), on a pragmatic level, is to establish relations between textual units. Or, in Fraser's (1999: 938) words, DMs "function like a two-place relation, one argument lying in the segment they introduce, the other lying in the prior discourse." DMs are of interest for this study because the contextuality of personal written communication is also connected to questions of discourse cohesion, towards which DMs contribute considerably. Discourse that is directed at a readership needs to take into account that there will be expectations on the reader's behalf regarding the content (and coherence) as well as the structure (and cohesion) of a given message. It is in this sense that the frequency of DMs is believed to contribute to the endophoric contextualisation of personal written communication.

One peculiarity of DMs is that their form does not equal function. For example, the discourse unit *so* can function as a DM (see also example (64) below), or as a modifier in an adjective phrase (as in *so pretty*). The same is true for the discourse unit *oh*, which may function as a DM when it fronts a textual unit, but at the same time qualifies as an onomatopoeic expression emulating the sound of surprise or astonishment (see also 6.2.2.). Furthermore, a DM need not consist of one word only, but can be made up of a combination of words (such as *oh well*), or longer string of words (as in *last but not least*). The multi-functionality of a word like *so* and the non-conformity of DMs with regard to their form requires careful coding practices in order to gain meaningful results (a circumstance that has been attended to in this study, cf. appendix 15.1.).

It was observed that the DMs *and*, *but*, and *so* are most frequent in all of the text types investigated. Although this study is less concerned with the meaning of DMs, or how DMs contribute to the meaning of the textual units they connect (but see the discussion of the DMs *and*, *but*, *so* and, in particular, *well* below), and more with their function, I would still like to briefly draw attention to how the DMs *and*, *but*, and *so* differ in their impact on the meaning of textual units that are connected by these DMs. Consider examples (62) – (64) below:

(62) The same to you. Sleep well and sweet dreams!
(SMS text / author: male, 31 yrs)

(63) November 11, 1006: Things are going well here but I'm really busy at University.
(…)
(Personal homepage excerpt / author: Tm_1)

(64) (…) I'm at work right now, so I can't write much now. (…)
(E-mail excerpt / author: female, 24 yrs)

While the DM *and* in example (62) signals that the textual unit *sweet dreams* is an addition to the prior textual unit *Sleep well*, the DM *but* in example (63) indicates a contrast between the preceding textual unit *Things are going well here* and *I'm really busy at University*. Here, the DM *but* signals that being busy is seen as something separate from things that are going well, in other words, the *but* in *but I'm really busy at University* implies something like 'being really busy at university contrasts with (other) things that are going well'. Had the author of example (63) opted for the DM *and* instead, then the overall meaning would have differed along the lines of 'being busy at university is an addition to the (other) things that are going well'. It is yet another situation with the DM *so* in example (64), which connects the textual units *I'm at work right now* and *I can't write much now*. In this case, the DM *so* signals that the preceding textual unit can be seen as a causal explanation for the connected one, and the overall meaning of example (64) is something like 'the author cannot write much because she is at work'.[68]

Table 8.2 provides an overview of the most frequent DMs found in each of the five text types. For reasons of scope, it will not be possible to discuss how all of them contribute to the meaning of the textual units they connect. However, in their function as discourse connectives they all contribute comparable pragmatic meaning in personal written communication, namely that of contextualising it with regard to its intra-textuality. It should also be noted that only those types of DMs that occurred a minimum of ten times, in at least one of the text types, are listed in Table 8.2. These include (in alphabetical order): *and* (including the symbols "&" and "+"), *after all*, *alright*, *anyway/anyhow*, *besides*, *but*, *now*, *oh/ah* (alone or in combination with other units, such as *well*, forming the combinatory unit *oh well*), *ok* (alone or in combination with *then*, forming the combinatory unit *ok then*), *or*, *so*, *then*, *therefore*, *though*, *thus* and *well*. Please note that the double line separating the "total" from the listed DMs in Table 8.2 indicates that this total concerns all DMs determined for each of the text types. DMs that are not listed in Table 8.2, but included in the total, will be discussed shortly.

[68] The discussion of examples (62) – (64) is based on Blakemore (2006: 224) and her considerations on Grice's (1989) claim that from a speech act theoretical point of view, each DM can be assumed to correspond to a speech act that is in turn individuated by its content. These considerations are, however, by no means exhaustive as "Grice's characterisation of the meanings of these expressions [DMs] fails to account for all their uses" (Blakemore 2006: 224, based on Wilson & Sperber 1993). For reasons of scope this study will not take "all uses" of DMs into account, but focus on their function as discourse connectives as outlined in the theoretical discussion in chapters 4 and 5.

Table 8.2: Discourse markers in personal written communication (corpora-based results).

[1] **Types of DM** (listed alphabetically)	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
and (incl. symbols "&","+")	* 170 ** 1:108.4	689 1:37.3	222 1:128	302 1:53.1	924 1:33.6
after all	0	13 1:1979.5	0	2 1:8015	0
alright	0	0	14 1:2028.9	1 2:16'030	0
anyway / anyhow	13 1:1417.4	21 1:1225.4	8 1:3550.5	2 1:8015	0
besides	0	0	0	0	10 1:3107.7
but	99 1:186.1	176 1:146.2	66 1:430.4	38 1:421.8	212 1:146.6
now	0	7 1:3676.1	4 1:7101	4 1:4007.5	11 1:2825.2
oh/ah (+ combinations)	20 1:921.3	11 1:2339.4	58 1:489.7	2 1:8015	1 1:31'077
ok (then)	29 1:635.4	4 1:6433.3	18 1:1578	3 1:5343.3	0
or	46 1:400.6	58 1:443.7	49 1:579.7	37 1:433.2	67 1:463.8
so	27 1:682.4	134 1:192	47 1:604.3	16 1:1001.9	53 1:586.4
then	15 1:228.4	0	28 1:1014.4	2 1:8015	28 1:1109.9
therefore	2 1:9213	0	0	0	11 1:2825.2
though	1 1:18'426	16 1:1608.3	6 1:4734	0	32 1:971.2
thus	0	0	0	1 1:16'030	30 1:1035.9
well	11 1:1675	25 1:1029:3	41 1:692.8	8 1:2003.8	0
[2] **TOTAL (all DMs)**	**461** **1:40**	**1226** **1:21**	**582** **48.8**	**474** **1:33.8**	**1425** **1:21.8**

Key: [1] = only those types of discourse marker (DM) listed that occurred a minimum of 10 times in at least one of the text types; [2] = total given for all DMs that occurred in each of the text types; incl. = including; HP = homepage; * = total number of indicated DM per text type, ** = indicated type of DM-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

It was found that *and*, *but*, and *so* are not only the most frequent DMs, they are, together with *or*, also the only DMs where more than 10 occurrences per text type for all the five text types could be determined. Also, some DMs that are quite frequent in the modern text types, such as *ok (then)*, *oh/ah* (occurring on their own or in combination with *shoot*, *great*, or *sorry)*, and *well*, are rare or non-existent in EEC. The same is also true in reverse for the DM *besides* that occurs 10 times in the EEC corpus, but is not present in any of the modern text types.

Although non-existent in EEC and not overabundantly frequent in most modern text types, I would nevertheless like to include at this point a discussion on the DM *well*. Partly

because of the extensive literature that can be found on *well* (cf. Jucker 1999: 437), which makes it one of the most researched DMs, and partly because the different uses of *well* illustrate the difficulty to try and capture the meaning and function of DMs beyond the pragmatic level where they operate as discourse connective (as previously touched upon in 4.3.1.). Since it was observed that the DM *well* is more frequent in Web Chat (with a ratio of 1:692.8) than in any other text type investigated, I will discuss the different uses of *well* in connection with this particular type of discourse.

Jucker (1999: 438), who analyses the DM *well* using Sperber and Wilson's (1995) Relevance Theory as a descriptive framework, proposes the following four main uses of the DM *well*:

1. It can be used as a marker of insufficiency, indicating some problems on the content level of the current or the preceding utterance.[69]
2. It can be used as a face-threat mitigator, indicating some problems on the interpersonal level.
3. It can be used as frame marking device indicating a topic change or introducing direct reported speech, and
4. it can be used as a delay device.

With respect to the first point above, Jucker (1993: 440) argues that "this use is very well attested in the literature on *well*" and, according to Lakoff (1973: 458/463, quoted in Jucker 1999: 440), the discourse marker *well* is often used "in cases in which respondents know that they are not giving directly the information which the questioner has requested," and thus sense "'some sort of insufficiency' in their replies." Example (65) below gives a relevant example of the use of *well* as a marker of insufficiency in such a question/answer pair:

(65) <Some_Guy>: wanna be an actress <NeGaRiSh~AsS~*>?
(…)
<NeGaRiSh~AsS~*>: <u>well</u> thatsss my hope
(Chat excerpt / Room V, session 4)

[69] In the discussion on *well* as a marker of insufficiency, Jucker (1999: 442) also mentions the use of *well* as a "qualifier" (based on Svartvik 1980 and Carlson 1984), where "*well* is said to qualify the content of the previous move or the move which it introduces." For example, if speaker A says "That man speaks extremely good English" and speaker B replies "Well, he is American", then *well* indicates that B is not surprised by the command of 'that man's' English because he is American (and thus most likely a native speaker of English)—and B thereby qualifies the content of A's utterance as being "unsurprising", indicating that there is nothing remarkable about an American that speaks extremely good English (example taken from Jucker 1999: 442).

The above example is taken from a Web Chat session in which participants are conversing about the possibility of making a film together, discussing different responsibilities (such as the part of director, film star, caterer and so forth) and story lines. At some point the Chatter nicknamed <Some_Guy> asks fellow Chatter <NeGaRiSh~AsS~*> if he/she wants to be an actor/actress, to which <NeGaRiSh~AsS~*> replies *well thatsss my hope*. The use of *well* as shown in example (65) marks two different types of insufficiency in the particular conversational setting of virtual Web Chat. First, it hints at the circumstance that while the whole discussion about making a film together would not be impossible in "real life", the conversation is nevertheless based on a role play and is thus fictitious. It is in this sense that <NeGaRiSh~AsS~*'s> use of *well* may indicate this insufficiency, i.e. the lack of truth conditions. <NeGaRiSh~AsS~*'s> reply would then translate into something like 'I guess in the framework of this role play becoming an actress would be my hope'. Second, the use of *well* in example (65) also indicates that <NeGaRiSh~AsS~*> is aware that his/her reply is not exactly the information that <Some_Guy's> requested with his/her yes/no-question *wanna be an actress <NeGaRiSh~AsS~*>?*. Under ideal circumstances, a yes/no-question would generate an answer along the lines of *yes* or *no*—rather than the elusive statement *well thatsss my hope*.

The second use of the DM *well* outlined above, i.e. the use of *well* as a face-threat mitigator, is found in exchanges where "either the face of the speaker or the face of the hearer is threatened" (Jucker 1999: 444).[70] The DM *well* then serves the main function of mitigating some sort of confrontation between speaker (writer) and hearer (reader), such as disagreement on a certain topic or the rejection of a request (cf. Jucker 1999: 444). Example (66) illustrates a communicative situation that is (potentially) confrontational. A Chatter nicknamed <PunkyBrewster's> reprimands another Chatter called <UnitedStates32730>, who claims to be underage, that underage chatting should be subject to parental control. <PunkyBrewster's> use of *well* in example (66) mitigates the potential face-threat that such a reprimand might otherwise have had on <UnitedStates32730>:

[70] The notion of "face" (a person's public self-image) goes back to Goffman's (1963) Face Theory, which claims that everybody has "face wants", i.e. is concerned with their public self-image. Brown and Levinson (1987) draw on Goffman's Face Theory in their Model of Politeness (offering different politeness strategies), where essentially, all (speech) acts can be classified as either "face-threatening" (i.e. infringing on another person's concern with public self-image) or "face-saving" (i.e. respecting another person's concern with public self-image).

(66) <UnitedStates32730>: just 2 let u know im <dani> and im 10 years old!
<PunkyBrewster>: <u>well</u> then you better have a parent around while you are chatting.
(Chat excerpt / Room II, session 4)

By introducing his/her turn with the DM *well*, <PunkyBrewster> mitigates a confrontation, namely the circumstance that underage Chatters should not be chatting unsupervised in this Chat room, because the use of *well* in this example softens the illocutionary force of the imperative *then you better have a parent around while you are chatting*—hence, <PunkyBrewster's> wish for parental control is less confrontational.

As listed under the third point above, the DM *well* can also be used as a frame marking device indicating a change of topic or introducing direct reported speech (cf. Jucker 1999: 438). In the following example (67), a Chatter nicknamed <babez7> first discusses family matters with fellow Chatters to then change the topic quite abruptly to his/her intentions to *best go have an early night* (meaning 'to leave the Chat room'), the topic change being introduced by *well*:

(67) <babez7 >: i wouldnt bable [be able] to cope on my own i know my b/friend [boyfriend] wouldnt come he loves it here sooo many friends n family
(…)
<babez7 >: i know i dont really have a big family my fella family is massive
(…)
<babez7 >: <u>well</u> best go have an early night got bad cold but still have to go to work in mornin
(Chat excerpt / Room III, session 4)

If <babez7> had not fronted his/her last turn with *well* in the above example (67), the shift in topic focus would have been a lot brusquer. However, by introducing the topic change with the DM *well*, <babez7> not only connects the change of topic to his/her previous contributions, but thereby also provides the frame that the fellow Chatters need in order to follow his/her train of thought.

Next to the uses of *well* as a marker of insufficiency, face-threat mitigator, or frame marking device, there is also the possibility of using the DM *well* as a delay device. However, as pointed out by Jucker (cf. 1999: 448), the use of *well* as a delay device may in some cases overlap with the category of *well* used as a face-threat mitigator, because the delaying tactic may also indicate problems on the interpersonal level. Example (68) is illustrative of this use of *well*:

(68) <AdamBomb>: chickens are great, ever hear of a dish called KFC [Kentucky Fried Chicken]?
(…)
<DanTheAutomator>: ever heard of the chicken flu ???
(…)
<AdamBomb>: fuckin puke, thanks alot <DanTheAutomator>
(…)
<AdamBomb> <--- sctratches KFC from speed dial
(…)
<DanTheAutomator>: well ….imagine all the hormones that chook eats before we get it …….errr yuk
(Chat excerpt / Room V, session 1)

While <AdamBomb> starts out as being a fan of eating chicken, <DanTheAutomator> seems to have quite strong aversions to the consumption of this particular type of poultry. Towards the end of the exchange on this topic, <DanTheAutomator> contributes a turn fronted with *well* in combination with four full stops. It can be assumed that in a spoken conversation, the use of the DM *well* as delay device serves the need of buying time to think about what to say next (cf. Jucker 1999: 448)—and to keep the floor while doing so. However, this is different in written exchanges; even if they are as near-synchronous as Web Chat, there is more time to think about what to write—without losing the floor. In more asynchronous modes there is the possibility to delay the sending of a message at one's own discretion if there is a need to make revisions of some sort. In near-synchronous Web Chat, while there is time to think about what to write in advance and maybe even change one or the other comment prior to sending, it is customary to send off contributions quickly, otherwise one risks losing the thread (rather than the floor).

Still, instead of using the DM *well* as a delay device, that is write it out, it would also be possible to just *think* the delay, i.e. refrain from writing it out, and start (somewhat delayed) with the actual contribution straight away. This suggests that while the use of *well* in example (68) looks like a proper delay device, in particular because of the four full stops which create a gap between *well* and the remaining turn, the function might be more that of a face-threat mitigator. After all, <DanTheAutomator's> viewpoint is confrontational to what <AdamBomb> initially claims, namely that *chickens are great*. By fronting the turn with *well*, <DanTheAutomator> softens the impact of the successive imperative assessment *imagine all the hormones that chook eats before we get it*.

On a more general note, it was observed that out of all the four uses, the DM *well* is most frequently used in Web Chat as a frame to signal topic changes and as a face-threat mitigator to avoid offending fellow Chatters and avert potential confrontation. The former

is connected to the varying number of participants that address different topics in a given Web Chat session (resulting in frequent topic changes). The latter is believed to be connected to Chat room regulations that advise their participants to be friendly to each other (otherwise participants risk to be kicked out of the Chat room)—in other words, participants are encouraged to minimise potential face-threats to each other.

After having discussed different uses of the DM *well* in the text type Web Chat, I would now like to come back to Table 8.2, or rather, pay attention to the DMs that are not listed, but are subsumed in the total Table 8.2. They include (in alphabetical order, with their total frequencies across all text types in brackets): *by the way* (11), *despite* (5), *finally* (2), *generally* (1), *hence* (2), *however* (18), *I mean* (17), *if you like* (1), *in any case* (8), *in other words* (1), *last but not least* (1), *like* (6), *meanwhile* (1), *obviously* (4), *of course* (21), *other (than that)* (7), *otherwise* (16), *personally* (1), *so far* (8), *so to speak* (1), and *yet* (10). These 22 DMs, along with the 16 types listed in Table 8.2, amount to a total of 38 variations that occur in the five text corpora. However, as has been pointed out above, of these 38 variations only four (*and*, *but*, *or*, and *so*) were classified as frequent.

It was found that e-mail features DMs most frequently of all the five text types, with approximately one DM in every 21 words. Web Chat and SMS discourse both show lower frequencies in the use of DMs than e-mail and EEC, and this is believed to be connected to the short discourse units of these text types. For example, the average word count for SMS texts in this corpus is 18.4 words and a mere 5.2 words for a turn in Web Chat. This means that the need to connect textual units is lower because the entire message might consist of just one textual unit. Examples (69) and (70) below further illustrate this point:

(69) You're really damn cool!
(SMS text / author: female, 26 yrs)

(70) <Geux_Tigers>: thats right
(Chat excerpt / Room V, session 4)

Interestingly, Chatters show tendencies to split a particular message into textual units that then occur as two (or more) separate turns—rather than connect two (or more) short textual units by means of DMs in one turn:

(71) <Mr._Kind>: in brazil the women are better
<Mr._Kind>: i'll go there
(Chat excerpt / Room V, session 2)

In example (71), the two turns by <Mr._Kind> occurred immediately after one another in the original discussion, but they could have been easily combined to form one turn by means of, for example, the DM *so*: *in brazil the women are better so I'll go there*. The breaking up of messages into different turns is extremely frequent in Web Chat. This is assumed to be the main reason why the DM-to-word ratio is lowest for this text type.

Furthermore, the DM-to-word ratio in personal homepages is slightly lower than in e-mail and EEC, but slightly higher than SMS discourse and Web Chat. It is at the same time the text type with the greatest variation of DMs. Overall, it can be asserted that DMs are frequently used devices in personal written communication to connect textual units in order to support the internal structure of a given message. However, it emerged that messages that are made up of shorter text units also feature lower DM frequencies. In the case of SMS discourse, this supports findings that the investigation into text-internal references produced. In any event, opposed to the different types of endophoric reference discussed above, the uses of DMs do not correlate with the type of readership (acquainted vs. unknown), but have more to do with the discourse structure of the text types.

8.2. Inter-textuality of personal written communication

The reciprocal relationship of addressor and addressee(s) in personal written communication entails that the author of a particular message directed at a distinctive (or implied) readership, will turn into a reader him-/herself if his/her readership decides to respond, taking over the role of the author. In other words, personal written communication can be seen as a succession of information exchange, and the individual messages that make up this sequential information exchange are the links in this chain. This also means that a particular message most likely stands in relation to messages that were written in the past and precede it. At the same time, a given message may also anticipate messages that will be written in the future and are to follow. Whether a message refers to previous ones, or anticipates potential future messages is, of course, not always explicitly expressed. For example, authors may also (and often do) make implicit references to something that was written (or said) in the past, presupposing that the reader already knows what is being referred to (we will return to this issue in the discussion of contextual effects in chapter 12).

This type of textual reference is difficult to investigate if the messages that preceded and/or followed a given message are not available for research. Constraints in data collection for this study effectuated that the text corpora are comprised of "isolated"

messages, and it was thus not possible to put them into direct relation to other messages. This is the reason why the investigation into references to preceding messages (text-external, or exophoric, backward reference, 8.2.1.) and references to future messages (text-external, or exophoric, forward reference, 8.2.2.) can only take into account explicitly formulated references. And while the investigation into explicit textual references sheds light onto the inter- and intra-textuality of personal written communication, it should be kept in mind that they do not make up the whole picture because implicit textual references could not be considered.

8.2.1. Text-external (exophoric) backward reference

> The function of backward reference is refreshing the reader's or listener's mind through a reminder that something has been mentioned or discussed in the preceding discourse, in case he or she has forgotten or—in the case of a written text—might want to look it up. It is conceivable that an additional function might be the emphasis of the point referred to again, the quasi-repetition implying to the audience that this is an important point to be kept in mind.
>
> (Claridge 2001: 58)

In addition to the points made by Claridge in the above quote, that backward reference may function as a reminder and/or emphasis of something that was said or written in the past, it can also be seen as an attempt to contextualise a particular message (or parts of it) by means of backward reference in the sense that the author explicitly indicates to the reader that the message (or parts of it) is (are) to be understood in relation to something that was said or written in the past.

Exophoric backward reference may vary in what it refers to, meaning what type of texts or text segments are indexed, and how it refers to those texts and text segments. For example, an author of an SMS message may refer to another message (or its parts) produced on the same medium (another SMS text), or he/she can also refer to another message (or its parts) produced on another medium (such as an e-mail message). In addition to exophoric backward reference to written text, an author of written correspondence can, of course, also refer to something that was said in the past over the telephone or in a face-to-face setting. Text-external backward reference of either category may also be unspecified in the sense that it is not clear with which medium the indexed written or spoken text (segment) was produced. Hence, we have two main categories and six different types of exophoric backward reference: indexing backwards to *written* text (produced on the (1) "same

medium", or (2) "other medium", or (3) "unspecified medium") and backward reference to *spoken* text (reference to a (4) "face-to-face meeting", or (5) "telephone call", or (6) "unspecified medium"). Before discussing these categories in more detail and illustrating them with examples, I would first like to introduce another, seventh type of reference, as illustrated in examples (72) and (73) below:

(72) Sorry man - I don't think we got this message!!? I don't know why....Anyway - thanks for the message - I had a good birthday and had a fun weekend at home with the family etc. (…)
(E-mail excerpt / author: male, 36 yrs)

(73) Hi <nickname> *thanks* for your b-day wishes!I'm back :-) (…)
(SMS text excerpt / author: female, 27 yrs)

In example (72), the author makes a backward reference to a message he should have received, but believes that he did not, and a second reference to the message that presumably informed him about the message that he thinks he did not receive. The author of example (73), on the other hand, thanks the addressee for his/her birthday wishes that she has obviously received. However, in all cases, the text-external backward references do not identify the discourse mode (writing or speech), or the medium, with which the previous messages were produced and transmitted. In example (72), the author could make a reference (in both cases) to a handwritten letter (or card), e-mail or SMS text, or a spoken message left on an answering machine. The same is true for example (73), where it is not clear how the birthday wishes reached their destination. For this kind of indefinite reference a seventh category was formulated, namely (7) "unidentified exophoric backward reference".

This category is not included in Table 8.3 because it is unclear to what sort of text type these references index to, and, as illustrated in example (72), whether such an indexed past message exists at all. Table 8.3 thus illustrates the frequencies and distributions for the six identifiable types of exophoric backward reference. Percentages are given in relation to the total number of identifiable exophoric backward references found in each of the text corpora.

Table 8.3: Exophoric backward reference in personal written communication (corpora-based results).

Exophoric backward ref. to WRITTEN text (segments)	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
SAME medium	* 22.7	49.5	81.0	84.0	85.8
OTHER medium	30.9	10.1	16.7	16.0	x
unspecified	1.0	12.2	2.3	0	0
TOTAL (written)	**54.6**	**71.8**	**100**	**100**	**85.8**
Exophoric backward ref. to SPOKEN text (segments)	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	
face-to-face meeting	25.8	17.0	0	0	14.2
telephone	11.3	9.0	0	0	x
unspecified	8.3	2.2	0	0	0
TOTAL (spoken)	**45.4**	**28.2**	**0**	**0**	**14.2**

Key: ref. = references; HP = homepage; * = percentages (rounded to 1 decimal) given in relation to the total number of identifiable exophoric backward references (per text type) to both written and spoken text (segments); x = respective type of exophoric backward reference not possible due to temporal origin of data.

Table 8.3 reveals several trends across the five discourse types. First, apart from SMS discourse, the majority of text-external backward references concern texts that were written on the same medium. Second, again except for SMS discourse, this type of textual reference is not very frequently made to spoken text (segments). In reverse, SMS discourse is the only text type that contains more text-external backward references to texts composed on another medium (predominantly e-mail), and where backward references to spoken text (segments) are almost as frequent as references to written text (segments). Furthermore, the two text types aimed at an unknown readership do not feature one single reference to a spoken text (segment) from the past.

Before illustrating those two main categories and the six types with examples, I would first like to shed light onto how frequent text-external backward references are in each of the text types. Web Chat and the personal homepage feature the lowest exophoric backward reference-to-word ratios (1:676.3 and 1:641.2, respectively), followed by EEC (1:293.2), e-mail (1:289.1), and SMS discourse (1:190). These ratios represent the six types of exophoric backward reference as listed in Table 8.3. If the seventh type of unidentifiable reference is also considered, then the reference-to-word ratios are higher, but the distribution across the five text corpora remains similar: Web Chat and the personal homepage corpus still feature the lowest exophoric backward reference-to-word ratios (1:591.8 and 1:616.5, respectively), followed by EEC (1:239), e-mail (1:169.3) and SMS discourse (1:142.8). This shows that text-external backward reference is in general not extremely frequent in personal written communication, but still occurs in every second to third message.

With regard to the main trends illustrated in Table 8.3, it becomes apparent that SMS discourse seems to be different in its exophoric backward reference compared to the other four discourse types. Not only do the 1000 messages in the corpus contain more backward references to text (segments) produced on media other than the short message service, but there are also frequent references to spoken text (segments), both of which are considerably less frequent (or non-existent) in the other text corpora. Examples (74) and (75) are relevant examples of the two trends as observed in the SMS text corpus:

(74) (…) could you please tell me again which film you're showing? I've already deleted the email, sorry. *thanks* kisses
(SMS text excerpt / author: female, 27 yrs)

(75) Sorry, I couldn't talk freely, that is why I sounded distant… (…).
(SMS text excerpt / author: male, 34 yrs)

In example (74), the author asks her addressee if he/she could tell her again what film he/she is planning to show, she justifies her question by explaining that she deleted the e-mail message which most likely contained this information. Example (74) is prototypical, in that e-mail is by far the medium that authors of the SMS corpus referred to the most. In Example (75), the SMS text is aimed at apologising for something that was said to the receiver, or rather the way in which this conversation took place (*I sounded distant*). It was found that the frequencies, with which these two types of exophoric backward reference occur in SMS discourse, distinguish this medium from the other types of correspondence.

With respect to the other media, it can be asserted that the majority of text-external backward references are made to text (segments) produced on the same medium. Examples (76), (77), and (79) illustrate such typical exophoric backward references. Please note that example (78) is a special case.

(76) Hey <forename>!
I'm going to make this a quick one because the last Email I sent went to the wrong address AGAIN !
(E-mail excerpt, emphasis original / author: male, 22 yrs)

(77) **Archives**
- August 2006
- October 2006

(Personal homepage excerpt, emphasis original / author: Cm_1)

(78) <tracyt>: he said i was a slag b4
(Chat excerpt / Room III, session 3)

(79) Worthy Prince and my dearest brother:
I received your most welcom Letter and kynd token by Mr. Hopkins, highly esteeming them as delightfull memorialls of your brotherly love. (…)
(EEC excerpt, ca. 1600 / letter collection Royal2, author: female)

Example (76) is prototypical for e-mail correspondence because most of the text-external backward references concern one of the following two scenarios. First, a certain e-mail message did not reach its intended destination, either because it was sent to the wrong address (as in the example above), or because it was sent to the correct address, but it seemingly got lost. Or second, the author thanks the reader for an e-mail that he/she sent in the past, and which was in fact received. Examples (77) and (78) are characteristic for the personal homepage and Web Chat, respectively, because more than 80% of the references are of this type. In the personal homepages, most backward references concern older versions of certain parts of the website which are stored in so-called "archives". In most cases those archives are titled with calendrical references as illustrated in example (77).

Web Chat, on the other hand, caused some problems with regard to exophoric backward references. The Web Chat sessions are extracts from discussions that were in full flow at the moment when data recording of the 15-minute-sessions commenced. As has been pointed out in connection with contextual effects in the theoretical discussion in chapters 4 and 5, it was thus not possible to consider the first few turns due to lack of (preceding) textual evidence (which is needed in order to determine what kind of effect a certain textual unit could have been aimed at performing). The same procedure was also applied in connection with exophoric backward references. Thus, references occurring at the beginning of the Chat sessions were classified as "unidentified exophoric backward references" (and subsumed in the seventh category illustrated above) if it was not possible to identify the textual unit to which they indexed. And although the excerpt shown in example (78) occurs late into the 15-minute-session, no textual evidence to which the Chatter named <tracyt> could have been referring to was found. It is also believed that <tracyt> was at the time not referring to something that was *said* with her comment *he said i was a slag b4*, but that he/she indexed backwards to another textual unit in Web Chat (hence a *written* comment). Nevertheless, due to lack of textual evidence, this type of reference was also classified as pertaining to category 7, "unidentified exophoric backward reference" (as opposed to a backward reference in the same medium). However, with

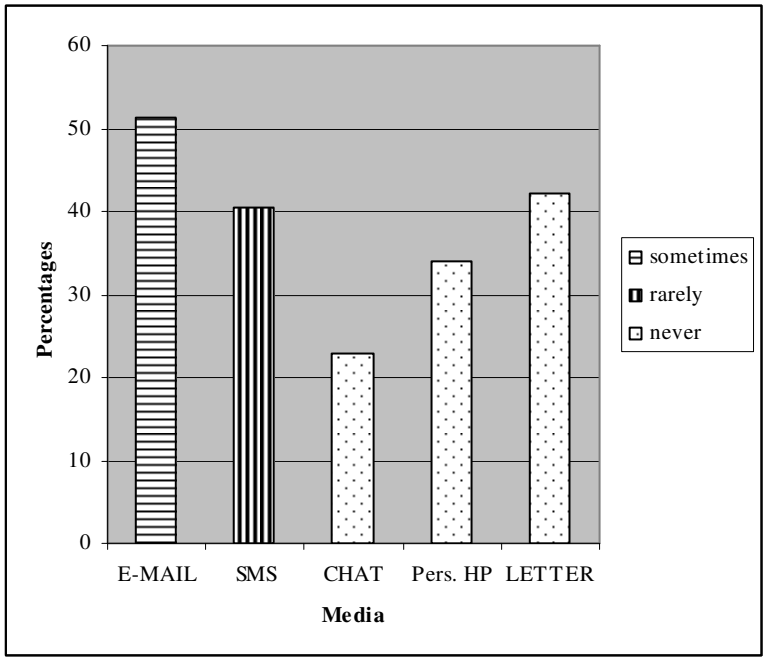
frequencies of about one text-external backward reference in ca. every 600 words for both Web Chat and the personal homepages (including unidentified exophoric backward references), this type of reference is not very popular in either text type.

With respect to the handwritten letter, illustrated in example (79), the vast majority of text-external backward references concern the acknowledgment of the receipt of previous letters. Due to the time and age from which the data originates, no references to electronic media were possible at the time (indicated by "x" in Table 8.3). However, even though face-to-face meetings were very much possible in the 17th century, references to something that was said at such occasions are kept to a minimum in the letters of this corpus. Example (80) illustrates such a rare backward reference to spoken text:

(80) Sister,
This eveninge the messenger I sent to London is retourned; what I have learnd of his sendinge ys, that the busynes wherin my cosin dealt was the treaty of a match betweene Mr Philip Woodhouse and the daughter of the L. Lovelace, <u>wherof I sayde somewhat to you when we last mett</u>. (…)
(EEC excerpt, 1633 / letter collection Cornwall, author: male)

In example (80), the author begins the letter by telling his sister about his cousin's business, to then at some point include the text-external backward reference *wherof I sayde somewhat to you when we last mett*. This, in my opinion, serves the cause of contextualising a rather abrupt delving into a topic that might not be at the front of his sister's mind. By referring to what he said in their last meeting, the author is giving his reader more (implicit) background information on the topic. However, as has been pointed out above, this kind of reference is low in frequency in EEC. It seems safe to assume that this is connected to the time period that lies between saying something and the perusal of a letter that contains such a backward reference. In the case of EEC, the time lag caused by transmission time is considerably longer than for any of the modern media that foster correspondence with an acquainted readership. In other words, the more synchronous the personal written communication with an acquainted readership is, the more likely it contains text-external backward references to spoken text (segments). Personal written communication aimed at an unknown readership, however, tends not to contain text-external backward reference to spoken text (segments), regardless of its degree of synchrony, because the interlocutors have never met in person or have had a telephone conversation with each other.

Overall, it was found that text-external backward reference is not extremely frequent in the five text types. These findings are further supported by the results from the online survey. Informants stated that if they include text-external backward references in their messages, then they tend to do so in text types that are, first, aimed at an acquainted readership and second, prefer to do so in the more synchronous media. Graph 8.1 below illustrates the likelihood with which the informants thought they make use of exophoric backward references in their personal written communication.



Key: Pers. HP = Personal homepage

Graph 8.1: Exophoric backward reference in personal written communication (results from online survey, peak answers in percentages).

While the majority of informants felt they never include exophoric backward references in their correspondence with strangers (Web Chat and personal homepages), they at the same time were more likely to include this kind of backward reference in the more synchronous correspondence e-mail and SMS discourse (in contrast to the letter, which is more asynchronous as regards transmission speed).

8.2.2. Text-external (exophoric) forward reference

As opposed to exophoric *backward* reference, which refers to text segments that were in fact produced (whether in writing or in speech) at some point in the past, exophoric *forward* reference refers to text (segments) that have yet to be produced. Interestingly,

text-external forward reference is more frequent in all text types except for EEC. Again, a differentiation was made between identifiable exophoric forward references (listed in Table 8.4 below) and those classified as unidentifiable such as examples (81) and (82):

(81) Take care, keep in touch.
<Name> (the Sebastian Kiwi)
(E-mail excerpt (farewell section) / author: female, age unknown)

(82) Ok <Nickname>, I'll let you know asap. (…)
(SMS text excerpt / author: female, 34 yrs)

In both examples (81) and (82), it is not clear what sort of text type the "future message" will be, or by means of which medium it will be transmitted. One can *keep in touch* via spoken and written text types, which can be transmitted via all sorts of media, and the same is true for *I'll let you know asap*. However, similar to the analysis into text-external backward reference above, the reference-to-word ratios including the unidentifiable text-external forward references (given in brackets below) are slightly higher for each of the text types, but the distribution across the five text types remains similar. Thus, SMS discourse is likely to feature an exophoric forward reference every 33.5 (or 38) words, e-mail includes such a reference in every 141.4 (or 202.6) words, the personal homepage corpus has a ratio of one reference in every 219.6 (or 254.4) words, Web Chat features this type of reference in every 424 (473.4) words, and authors of EEC are likely to include a forward reference to other text (segments) in every 398.5 (or 526.7) words.

Next to being more frequent (except in EEC) than backward references, text-external forward references also include more references to spoken text (segments). In fact, for all the media aimed at an acquainted readership, the majority of the exophoric forward references concern something that the author intends to tell the readership in an upcoming face-to-face meeting. Since exophoric forward reference is planned (rather than produced) at the moment of writing, there are even a few instances observed in text types directed at an anonymous readership. Table 8.4 summarises the distribution of the six different types of identifiable exophoric forward reference across the five text types: indexing forward to *written* text (produced on the (1) "same medium", or (2) "other medium", or (3) "unspecified medium") and forward reference to *spoken* text (reference to a (4) "face-to-face meeting", or (5) "telephone call", or (6) "unspecified medium"). Percentages are given in relation to the total number of occurrences of all types of exophoric forward reference.

Table 8.4: Exophoric forward reference in personal written communication (corpora-based results).

Exophoric forward ref. to WRITTEN text (segments)	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
SAME medium	* 1.5	21.3	21.6	27.0	35.6
OTHER medium	3.3	3.2	71.7	61.9	x
unspecified	1.2	4.7	3.3	6.3	0
TOTAL (written)	**6**	**29.2**	**96.6**	**95.2**	**35.6**
Exophoric forward ref. to SPOKEN text (segments)	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
face-to-face meeting	68.2	55.9	1.7	1.6	64.4
telephone	20.6	9.4	1.7	3.2	x
unspecified	5.2	5.5	0	0	0
TOTAL (spoken)	**94**	**70.8**	**3.4**	**4.8**	**64.4**

Key: ref. = references; HP = homepage; * = percentages (rounded to 1 decimal) given in relation to the total number of identifiable exophoric forward references (per text type) to both written and spoken text (segments); x = respective type of exophoric forward reference not possible due to temporal origin of data.

It was found that the majority of messages directed at an acquainted readership contain forward references to spoken text in upcoming face-to-face meetings, whereas messages aimed at an unknown readership are a lot more likely to feature references to future writings other than the medium with which the message was produced.

SMS discourse contains by far the most text-external forward references (one in every other message), the vast majority of which are aimed at spoken text (segments) to occur in upcoming face-to-face meetings, as illustrated in example (83):

(83) Hey babe. I cant leave today until 230. Do you want to come here for a coffee, hang out and talk about it all? (…)
(SMS text excerpt / author: female, 30 yrs)

The author of the SMS text excerpt shown above writes about a potential face-to-face meeting in the near future, offering to discuss a certain issue with the recipient of this message. Not as frequent as this type of exophoric forward reference, but still amounting to around 20% of all forward references in SMS discourse, is the mention of future telephone calls:

(84) Call me on my natel, I have a meal break before going to Geneva. :-)
(SMS text / author: male, 33 yrs)

In example (84) the author desires the recipient of the SMS text to call him on his *natel* (Swiss German term for 'mobile phone') before he is due to travel to Geneva. Most references of this kind are in fact made in connection with the mobile phone (as opposed to

the landline phone), which is not surprising in the light of the circumstance that SMS discourse is generally performed via the mobile phone (and only rarely via computer).

With respect to e-mail, a little more than half of the exophoric forward references (55.9%) concern spoken discourse in future face-to-face meetings. E-mail differs in its exophoric forward references from SMS discourse in two ways: first, a little more than 20% concern future messages composed on the same medium (as opposed to a mere 1.5% in SMS discourse), and second, references to future telephone calls are only half as frequent. Regarding forward references to upcoming e-mail exchanges, they in all cases concern "promises" made by the author to be in touch soon again via e-mail, illustrated in example (85), or authorial "requests" that the readership write an e-mail next, shown in example (86):

(85) (…) Don't ask me how but somehow I got talked into buying an HSV ['Holden Special Vehicles', i.e. a car] Clubsport. (…)
Anyway I will send you a photo via e-mail as soon as I do one and you can see for yourself. (…)
(E-mail excerpt / author: male, 50 yrs)

(86) (…). I am waiting for your e-mail!
Cheers,
<name>
(E-mail excerpt / author: female, 18 yrs)

The investigation into text-external forward references in EEC produced results comparable to e-mail, in that the highest frequencies were determined for forward references to other letters (35.6%), as well as references that concern spoken text (segments) in an upcoming face-to-face meeting (64.4%). Since no electronic media were available to the people in the 17th century, no forward references to messages produced on such media were possible at the time (indicated by "x" in Table 8.4). However, as pointed out above, exophoric forward references are overall less frequent than backward ones in EEC. There does not seem to be an obvious reason for this phenomenon, but maybe the authors of EEC felt constrained with regard to future letters because they did not know when they would be able to write a letter next, and, probably connected to several situational factors at that time and age, also have it delivered.

The Web Chat and personal homepage corpora, once more, differ from the messages directed at an acquainted readership, in that the vast majorities (over 90% for both text types) of exophoric forward reference are made to other pieces of writing as opposed to spoken text (segments). Of those forward references to other written text

(segments), references to texts produced on other media or services are the most frequent. In Web Chat this concerns in the majority of cases enquiries to change to "pm", an abbreviation for 'private messaging', or in other words, leave the main room and chat one-to-one, as illustrated in example (87). The personal homepages, on the other hand, contain in most cases text-external forward references referring to potential e-mail contact, as shown in example (88).

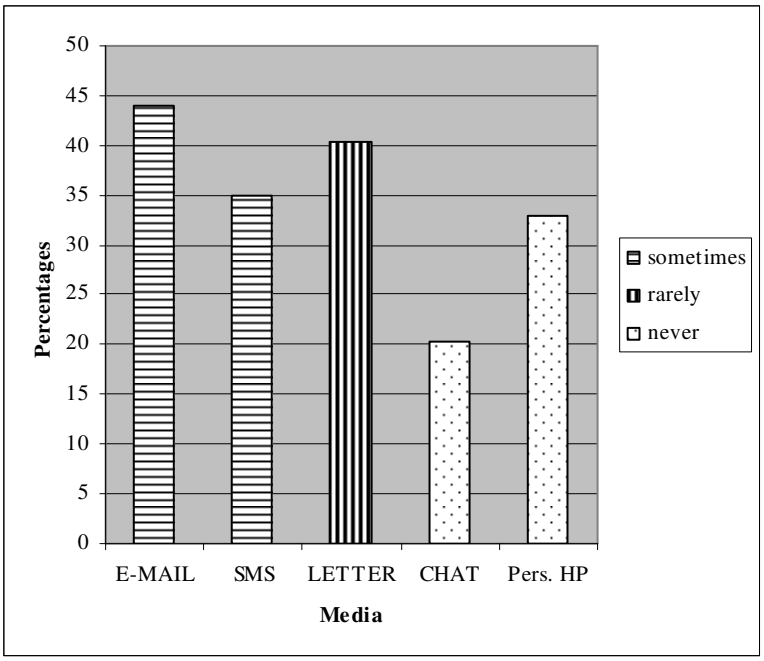
(87) <skyler>: can i pm you for a min <liza>
(Chat excerpt / Room IV, session 1)

(88) Email me <e-mail address> and tell me to get a life.
(Personal homepage excerpt / author: Mf_1)

With respect to the personal homepage corpus, all of the references to (potential) future messages, composed on other media, concern e-mail contact to be initiated by the readership, for the simple reason that the authors are in no position to initiate it themselves since they do not know their readers.

Overall, it can be asserted that exophoric forward references are, in all of the text types except in EEC, more frequently used to contextualise personal written communication than backward ones. This is interesting because exophoric forward reference is directed at text (segments) that have yet to be produced at the time when the forward reference is made, as opposed to exophoric backward reference, which concerns references to texts that were in fact produced at some point in the past. However, in the modern types of correspondence, in particular e-mail, there are quite a few references to messages that were produced, but for some reason did not reach their intended destination (cf. example (76) above). Another idiosyncrasy concerns the text-external forward references as found in the personal homepage corpus, which all refer to (potential) future e-mail contact to be initiated by the readership. Also, the media fostering personal written communication directed at an acquainted readership show a tendency for text-external forward reference to spoken interactions in upcoming face-to-face settings as well as telephone conversations, which is extremely rare in the two media engaging with an unknown readership.

These findings are in fact confirmed by the results from the online survey. Graph 8.2 below summarises the peak answers given by the 109 informants to the question how often they thought they include exophoric forward references in their personal written communication.



Key: Pers. HP = personal homepage

Graph 8.2: Exophoric forward reference in personal written communication (results from online survey, peak answers in percentages).

Graph 8.2 illustrates that 44% and 35% of the informants stated that they sometimes include this type of textual reference in their e-mail and SMS correspondence, respectively. However, with respect to the letter, it emerged that 40% believed they rarely include exophoric forward reference in their epistolary correspondence, and the majority of Web Chat and personal homepage authors claimed they never make use of this type of textual reference in their writings. If compared to Graph 8.1, which illustrates the frequencies of exophoric backward reference, it emerges that the informants also thought they include exophoric forward reference on a more frequent basis than backward ones, which in fact supports the findings from the corpora-based investigation. This indicates that personal written communication seems to be oriented towards the continuation of the communicative exchanges, rather than contextualise them based on past messages.

8.3. Chapter summary

The investigation into the intra- and inter-textuality of the five discourse types showed that authors of different kinds of personal written communication also use different strategies to contextualise their messages with regard to discourse deixis. While textual deixis is tied to text length and discourse structure of the messages, it also emerged that the type of readership, to which the messages are directed at, influences how authors contextualise

their messages on the textual level. SMS discourse, e-mail, and EEC feature fewer endophoric references than Web Chat and the personal homepages. Chatters' idiosyncratic use of anaphoric text-internal reference (the repetition of certain text segments with a question mark to indicate a need for clarification) is reminiscent of spoken discourse and thus partly confirms H1.

It is a slightly different picture with regard to exophoric references. The main trends include a preference for forward over backward text-external references, except in EEC, which differs from all other text types in this respect, confirming H4. However, with regard to the different types of exophoric forward references, it was observed that media directed at an acquainted readership refer to upcoming spoken conversations via telephone or, above all, potential face-to-face meetings more frequently. Since participants in anonymous communicative settings are highly unlikely to ever engage in telephone conversations or meet in person, such references are extremely infrequent (supporting H5, assuming differences connected to the type of readership). Another interesting observation concerns the fact that SMS texts feature text-external forward reference to upcoming telephone conversations more frequently than any of the other text types, which indicates that SMS discourse is seen as an "extension" of the telephone call by some of the users.

All in all, there is a clear trend towards contextualising personal written communication by means of endophoric structuring devices, as opposed to exophoric ones. This in turn indicates that messages are seen as isolated textual networks rather than being part of larger discourse structures. Furthermore, it can be stated that the investigation into different types of textual references highlighted tendencies tied to discourse structure and type of readership (confirming H5). This means, however, that none of the trends fully support the main research hypotheses H1 – H4 that all assume differences between the modern text types and the more traditional one, EEC (except EEC being the only medium favouring exophoric backward over forward references, thereby partly confirming H4). In this chapter, only H5 could be verified, because differences connected to acquainted vs. unknown readership are apparent in the contextualisation of personal written communication.

9. The spatiality of text

> In order for spatial deixis to communicate successfully, interlocutors need to have visual access to a common *context*, and share (or be able to interpret) one another's visual *perspective*. More fundamentally yet, interlocutors need to be aware of their physical *orientation* with respect to one another (whether they are face-to-face, next to one another, back-to-front, back-to-back, etc.).
>
> (Herring et al. 2003: 9, emphasis original)

The above quote concerns spatial deixis in face-to-face settings, Herring et al. (2003), however, deal with spatial reference in 3-D virtual, i.e. non-physical, worlds (such as video conferencing) in their paper. According to Herring et al. (2003: ii), "graphical multi-user environments such as 3-D virtual words aim to recreate the communicative affordances of face-to-face discourse" and if collaborative multi-media interfaces were transparent, "then deixis should work basically the same way in the mediated world as in the unmediated world" (Herring et al. 2003: 9). Yet Herring et al. (2003: 42) observed that multi-media interfaces are rarely transparent and state that the goal should be "to understand multi-media systems in their own terms, rather than as alternative forms of face-to-face communication." While this study is more concerned with text-based communication, as opposed to 3-D virtual worlds, it is nevertheless a good starting point to raise awareness that spatial deixis in communicative settings located on the Internet are subject to the dimension of virtual space.

As has been outlined in the theoretical discussion in chapter 5 (see also Figure 4.3), different types of communication face different challenges in connection with media-related constraints (in terms of media richness vs. media leanness) and the types of participants that engage with each other (in terms of acquainted vs. unacquainted correspondents). This chapter aims at capturing how authors of different types of personal written communication may compensate for the lack of shared physical space with their readership, and in what ways the perception of spatiality in text can be tied to the different types of personal correspondence. As was pointed out in the last chapter, text-internal reference to other text passages within the same text entity intersects with the notion of spatiality in text (as in *see above*). However, while chapter 8 concentrated on references on the textual level, and how these contribute to the internal structure of a text-based message, this chapter will be more concerned with the notion of space in connection to deictic origo

(9.1.) and how personal written communication is contextualised in terms of physical vs. virtual space (9.2.).

9.1. The notion of space and deictic origo: author vs. text

The investigation into the intra-textuality of the five types of correspondence has shown that endophoric references occur on a regular basis, above all in longer texts. Also, text-internal references are repeatedly realised by means of demonstratives and spatial deictic expressions (such as *this will be discussed below*). This indicates that textuality is predominantly expressed by means of spatial terms, as opposed to temporal ones, which are more common in spoken discourse (for example, *I will tell you later* as opposed to **I will tell you below*). However, the peculiarity of written communication is that the deictic origo of physical space is located with the author (location of author is deictic origo), while notions of textual and virtual space are associated with the text (location of text or text itself is deictic origo). While spatial references to either physical or textual origo can co-occur in text-based personal written communication, it was observed that certain text types have tendencies for favouring one over the other.

9.1.1. Where is *here*?

Not all spatial deictics lend themselves for a comparison of modern types of communication with Early English correspondence. Some vocabulary from 17th-century English is perceived as archaic by authors of the 21st century and is thus no longer used (for example, *whither* in the sense of 'whereto'), whereas other spatial terms have entered the vocabulary more recently, and were thus unknown to 17th-century authors (as is the case with *online*)[71]. However, the locative adverb *here* is, amongst others (see 9.1.2. below), popular in both EEC (on average occurring every 586.4 words) and is still very much in use today (between the four types of modern communication, it occurs ca. every 152 words). This qualifies the spatial deictic term *here* for the comparison of its use in the five types of correspondence. Most interesting about the spatial expression *here* is the circumstance that the term can be employed to denote the physical location of the author,

[71] In the OED online, the earliest record of the adjective *online* (with the meaning of 'situated on the route of a railway line' or 'in use on a railway line') is 1926. The use of the term *online* (as adjective or adverb) in connection with the computer (with an extended meaning of 'an operation or process carried out while connected to a computer and under its direct control') is recorded from 1950 onwards. Notably, the name of the reference work itself makes use of the term (OED *online*).

as well as referring to a specific passage in text as produced by such an author figure. Compare the following examples (89) – (91):

(89) Ok, easy, I'm already here, will start drinking :-)
(SMS text / author: female, 28 yrs)

(90) My schedule is a bit crazy these days and made even more so by an exam i have on friday. So, here's what would be good for me:
[author suggests 3 different dates and time slots to arrange a meeting]
(E-mail excerpt / author: female, 30 yrs)

(91) To go to <Full Name's> Groups, Representations and Cohomology Preprint Archive, click here
(Personal homepage excerpt / author: Bm_1)

The author of the SMS message shown in example (89) uses *here* to refer to her current geographical location. From what can be gathered of the message's content, it can be assumed (based on the response form *ok, easy*)[72] that the author reacts to an SMS text that she received from the person she was meant to meet in a certain place at a certain time. However, at the time this message was composed, this person had obviously not yet arrived. The author then refers to her physical location by means of *I'm already here*. As she does not include any geographical information as to where *here* is situated, she thereby implies that the recipient knows where *here* is. It is a different case with example (90), where the deictic expression *here* is used in indexical fashion to refer to an upcoming text passage. Another way of using the locative adverb *here* in personal written communication is illustrated in example (91), where *here*, being incorporated into a hyperlink, is multi-spatial in that it refers to (a) a specific location in the text, (b) a spot on the computer screen, which may be clicked on, and (c) to the word *here* itself (cf. Loehr 2002: 9). However, the spatial orientation of the deictic *here* as used in examples (90) and (91) is tied to the text (and its dimensions), which contrasts with example (89), where the physical location of the author determines where *here* is.

The uses of the deictic expression *here* in consideration of the deictic origo brought interesting results to light. They are summarised in Table 9.1, and percentages are given in relation to the total number of occurrences of the spatial deictic *here* in each of the five text types. On a general note, all instances of *here* as found in the five corpora were unambiguously classifiable in terms of deictic origo, i.e. author location vs. text

[72] "Response forms" (such as *ok*, *easy*) are discourse units that are used to briefly respond to a previous remark by a different speaker or author. They will be discussed more elaborately in chapter 11 (in particular in 11.2.5.)

(dimension). However, the exact scope of reference of *here* in connection with author as deictic origo could not always be determined because the necessary contextual information was missing, or, in other words, implied by the author.

Table 9.1: Deictic origo of *here* in personal written communication (corpora-based results).

Deictic origo of *here*	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
AUTHOR location	* 100	83.7	16.8	15.5	83.0
TEXT (dimension)	0	16.3	83.2	84.5	17.0

Key: HP = homepage; * = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of the spatial deictic *here* in each of the text types.

As illustrated in Table 9.1 above, spatial references by means of *here* are most frequently used to identify the geographical location of the author in SMS discourse, e-mail, and EEC, rather than spatially referring to a text passage within the message. From a mobility point of view, SMS discourse distinguishes itself from the other types of correspondence in that it can be performed regardless of the location, as long as there is a compatible telephone network available. Often, one may even move location whilst writing and transmitting the message (for example, whilst travelling on public transport or when walking around). CMC and EEC on the other hand, are often more static in this respect as people tend to refrain from writing letters in public spaces, and many do not carry a computer with them at all times, as opposed to the mobile phone, which is a constant companion for the majority of mobile phone users. This idiosyncrasy of SMS discourse, coupled with the tendency of authors to be brief in their messaging, means that while it is clear (based on linguistic context) that they are referring to some sort of geographical location, it is often unclear what exactly they are referring to by means of *here* (for example, a particular restaurant or cinema, a specific spot in town, and so forth).

The investigation into Web Chat and the personal homepage data showed that deictic references with *here* are in most cases used in connection with the text and its dimensions. This difference is believed to have less to do with the type of readership, which is addressed (acquainted vs. unknown), but more with how authors perceive the different media in terms of spatiality. It seems that SMS discourse and e-mail, although modern text types and, in the case of e-mail, even allowing for multi-medial correspondence, are perceived by their authors first and foremost as texts and not as "communicative spaces". They are in this respect reminiscent of traditional forms of correspondence, such as EEC. In contrast, authors of Web Chat and personal homepages

seem to be more aware of the spatiality of the texts they compose. This awareness is, of course, connected to the location of the texts on the Internet. For a start, both media work with metaphorical extensions of extremely familiar spatial expressions to denote their communicative settings: people converse in the Chat *room* with each other and/or publish material on their personal *home*page. In comparison to SMS discourse, e-mail, and EEC, Web Chat and personal homepages thus also deal with the additional dimension of virtual space. This also shows in idiosyncratic uses of *here* as illustrated in examples (92) – (93) below:

(92) <Liza>: <edge> can we talk out here please? thanks
(Chat excerpt / Room IV, session 1)

(93) WELCOME!!!! Great to have you here :-)
(Personal homepage excerpt / author: Of_1)

Example (92) is typical for Web Chat in that the Chat room is referred to as a spatial communicative setting. It is a similar case in example (93) where the use of *here* contextualises the homepage as a location (rather than a text). This kind of spatial reference by means of *here* is idiosyncratic to Web Chat and the personal homepage, as they were not observed in any of the other text types. However, this is, of course, connected to the circumstance that the communicative settings of Web Chat and the personal homepage are situated on the Internet, which inevitably has an influence on how authors contextualise their messages in terms of spatial perception. However, it is interesting how the perceptions of the "author in physical space" are translated into the conceptualisation of "text in virtual space" concerning word choice. However, before those two different notions of space are juxtaposed (cf. 9.2. below), I will first take a closer look at other spatial expressions and their employment in connection with deictic origo author vs. text.

9.1.2. Where is "elsewhere"?

The investigation into the use of the spatial expression *here* showed that authors of SMS discourse, e-mail, and EEC use this term predominantly to refer to the physical location of the author at writing time, whereas participants in Web Chat and authors of personal homepages mainly use it for spatial references on the textual level. In order to examine whether this distribution is tied to the spatial term *here*, or if other spatial terms are employed in a similar fashion by authors of personal written communication, the

distribution of other selected spatial expressions, which qualify for both physical and textual spatial reference in all of the five text types, was investigated. These expressions include the preposition *in*, the spatial adverb *there*, as well as the demonstratives *this*/*that* and *those*/*these*. The results are summarised in Table 9.2 below and percentages are given in relation to the total number of occurrences of these spatial expressions in each of the text types.

Table 9.2: Deictic origo of selected spatial expressions in personal written communication (corpora-based results).

Types of spatial expressions (listed alphabetically)	**SMS discourse**		**E-mail**		**Web Chat**		**Personal HP**		**EEC (letter)**	
	OA	**OT**	**OA**	**OT**	**OA**	**OT**	**OA**	**OT**	**OA**	**OT**
in	* 100	0	99.6	0.4	52.6	47.4	76.9	23.1	95.5	4.5
that	18.3	81.7	16.3	83.7	1.3	98.7	2.2	97.8	20.6	79.4
there	100	0	91.9	8.1	17.5	82.5	40	60	100	0
these	0	0	60.0	40.0	0	0	0	100	14.7	85.3
this	0	100	27.1	72.9	5.0	95.0	9.5	90.5	24.7	75.3
those	0	0	77.8	22.2	0	100	71.4	28.6	90	10

Key: OA = deictic origo author; OT = deictic origo text; HP = homepage; * = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of the indicated spatial deictic in each of the text types.

Table 9.2 illustrates that the distribution of other spatial deictics by no means always follow the pattern of *here*. Of all the spatial expressions listed, only the locative adverb *there* behaves akin to *here* in that it is in the majority of cases tied to the deictic origo of author in SMS discourse, e-mail, and EEC, as opposed to Web Chat and the personal homepages, where it is predominantly used to index another text passage or refers to the text dimension as located on the Internet. Since adverbs usually modify verb phrases, this may indicate that the uses of the locative adverbs *here* and *there* in SMS discourse, e-mail and EEC are tied to actions performed by the author (hence the author location is deictic origo), whereas their use in Web Chat and the personal homepage is more oriented towards action performed within the media (as in *click here* in personal homepages). However, the spatial deictic *in*, in comparison, is predominantly employed to denote locations in connection with the author's location as deictic origo regardless of the text type. Another pattern was observed with indexical demonstrative *this*, which is more frequently oriented towards the text (and its dimensions) as deictic origo, regardless of the type of correspondence. Last but not least, the remaining demonstratives *that*, *these* and *those* were observed to be, first, not as frequent as the other spatial expressions, and second, not distributed according to specific patterns.

As pointed out above, *there* is the only spatial expression with an analogue distribution pattern to *here*. This may also be related to the peculiarity of Web Chat and personal homepage authors to use *there* in linguistic environments where *here* would be just as suitable (if not better suited). Consider the following Web Chat excerpt:

(94) <jess>: <bryan> are you there
(…)
<bryan>: yes <jess> what is it
<jess>: wanna talk pvt
(Chat excerpt / Room IV, session 4)

In example (94) above, <jess> asks if <bryan> is *there*, however, since they share the same communicative setting, *there* could be replaced with *here* and the enquiry would still be understandable, if not more suited to the context: *<bryan> are you here [?]*. This stands in stark contrast with the employment of *there* in connection with locations tied to the author origo, as illustrated in example (95) below:

(95) We visited a humungous Ikea store in New Jersey on Sunday. I never want to go back there again… (…)
(E-mail excerpt / author: female, age unknown)

In this context, it would not be possible to replace the locative adverb *there* with *here* as **I never want to go back here again* does not really make sense.

With respect to the spatial expression *in*, it was observed that it is more frequently used to denote locations oriented towards the author as deictic origo regardless of the communicative setting. Unusual about this trend is that the authors in communicative settings located in virtual space also tended to use the preposition *in* tied to locations in relation with author origo, rather than the origo of the text and its dimensions. This tendency was found to be related to the desire to communicate where one is located in the real (as opposed to virtual) world, as shown in the following example (96), an excerpt taken from Web Chat:

(96) <gowiththeflow>: where r ya in the world <desireable>?
(…)
<desireable>: im in New Zealand baybee im a strict straight maori girl
(…)
<gowiththeflow> thats out of order. im in the uk
(Chat excerpt / Room III, session 3)

This excerpt is characteristic of Web Chat because it illustrates that the whereabouts of the participants are popular discussion topics. In example (96), a Chatter nicknamed <gowiththeflow> wants to know <desireable's> location, using the spatial expression *in* (*where r ya in the world <desireable>?*), to which <desireable> replies that he/she is located, again by means of the preposition *in*, in New Zealand. This prompts <gowiththeflow> to respond *thats out of order. im in the uk*, presumably hinting at the fact that the climate was milder in New Zealand at the time this Chat was recorded. Speaking of climate and the weather, which are generally seen as "safe topics" to open a conversation, it was interesting to observe that Chatters have a tendency to start many communicative exchanges by discussing each other's whereabouts in the real world. This in turn explains why *in* is more frequently used in connection with the author origo than the text. I will, however, come back to the importance of the Chatters' whereabouts in the discussion of physical vs. virtual space below (9.2.).

Authors of personal homepages use the preposition *in* in a similar fashion to the participants in Web Chat, as it predominantly serves to communicate the location of the homepage author in the real world:

(97) My name is <Full Name>. I was born in Hong Kong *in* 1972, but my family moved to San Francisco *in* 1984. I have lived in the Bay Area since.
(Personal homepage excerpt, italics my emphasis / author: Yf_1)

Next to illustrating the use of *in* functioning as a spatial expression, example (97) brings another factor to light that can potentially bias the results of an empirical investigation into functional aspects of discourse units. The italicised uses of *in* are temporal, whereas the underlined instances are spatial. This means that a discourse unit such as *in* can be used multi-functionally to index locations (with varying deictic origo), as well as providing temporal information. The only way to determine which of the possible functional slots such a discourse unit in fact occupies in a given utterance is by means of context. This has, of course, been carefully accounted for in the coding of the data.

Regarding the uses of *in* and its distribution across the remaining text types, it was observed that it is in the vast majority of cases employed to refer to the physical locations in connection with the author at the time of message composition. It is not very frequent in SMS discourse, but the authors of both e-mail and EEC make regular use of the spatial deictic *in*:

(98) Hello. It's <name>. How are you? I have a good time in Whistler now. (…)
(E-mail excerpt / author: female, 18 yrs)

(99) Mr. Durell is at Windsor, and will not be in town till next week. (…)
(EEC excerpt, 1665 / letter collection Basire, author: male)

Both examples (98) and (99) show typical uses of the spatial expression *in*, the indexed locations being connected to the author's position at the moment of writing time.

Compared to the spatial deictic *in*, the demonstrative *this* showed an inverted distribution pattern. It was found to be predominantly employed to index text passages (or the whole message, cf. examples (56) – (58) in chapter 8), and only rarely is the demonstrative *this* used to refer to geographical locations in the vicinity of the author. The same is true of the corresponding demonstrative *that*, which is commonly employed to index locations that are somewhat further away (calculable or perceived (emotional) distance) than locations referred to with *this*. Compare examples (100) and (101) on the uses of *this* and *that*:

(100) Deere Sister,
The sayme daye I receued your last by my footman *I fell extreame ill of a scouringe, which hath continued upon mee eauer sence*, but, I thanck God, it beegins somethinge to lessen it selfe of it selfe, which was the cause you haue not herd from mee till now; *haueinge eauer sence kept my lodgeinge by reason of a soreness in my throte, and horseness withall, that did macke mee holely vnfitt for company*; yett, for all this, I durst not adventuere to put my selfe in to the docktor's hands, but hath lett Nateure worcke his wille with mee, and I hope for the best.
(…)
(EEC excerpt, 1625, italics my emphasis / letter collection Cornwall, author: male)

(101) (…) It is my dayly comfort to remember how fast we approach each other, and that every moment dispatches some part of the way: this makes me content with present absence, nay even *love it*. Forgive the sound of that last word, and consider the sence of it. (…)
(EEC excerpt, 1656, italics my emphasis / letter collection Tixall, author: female)

In both examples, the demonstratives *this* and *that* are used to refer to a previous text passage, but there is a difference. In example (100), the author makes an anaphoric text-internal reference to a state of affairs by means of *this*, and from the content of the text passage that *this* refers to (in italics), one can gather that the state of affairs is a current one with which the author still identifies. It is a different situation in example (101), where the author makes an anaphoric text-internal reference to the verb phrase *love it* by means of the demonstrative noun phrase *that last word*. In this example, the content indicates that the author distances herself from the literal meaning of the verb *love*, because it is not the

absence that she loves, but the effect that every moment that passes diminishes the time span until she and the recipient of the letter next see each other. As pointed out above, uses of the demonstratives *this* and *that*, other than on the textual level, are rare in all the text types.

No dominant distribution pattern could be determined for the remaining demonstratives listed in Table 9.2, *these* and *those*. SMS discourse does not feature either of them and while the Web Chat corpus features a very low frequency of *those* (a mere 3 instances, all of them referring to another text passage), it also does not feature *these*. In the personal homepages, the demonstrative *these* is exclusively employed with the text as deictic origo, while *those* predominantly indexes locations in connection with the author as deictic origo. Whereas the e-mail corpus shows a tendency to index other text passages by means of *these* and *those*, EEC follows that trend only with the uses of *these*, the demonstrative *those*, on the other hand, predominantly indexes locations (or situations) in connection with the author as deictic origo.

All in all, it can be asserted that SMS discourse, e-mail, and EEC have greater tendencies to refer to locations in connection with the author as deictic origo than Web Chat and the personal homepages. This is believed to be related to communicative settings of Web Chat and the personal homepages, which are both located on the Internet. Hence, although both text types are in fact predominantly text-based, the communication is often perceived as taking place in space, virtual space that is. Or, as Loehr (2002: 29) points out in connection with Web pages, "the context of a hypertext document such as a World Wide Web page is the World Wide Web itself."

9.2. Physical vs. virtual spatial reference

Web Chat and the personal homepage are text-based communicative settings. However, because they are located on the Internet, the messages produced in those settings gain an additional spatial dimension that is also apparent in the language that is used to compose these messages. While the previous section was concerned with the deictic origo, this section is aimed at further revealing how authors of personal written communication refer to locations in virtual space (within the same text entity or outside), and in how far such virtual references are based on the conceptualisation of "conventional" spatial reference in the real world. I will first look at how authors of personal homepages and participants in Web Chat deal with the notion of physical vs. virtual spatial reference, before including the other three text types into the discussion.

9.2.1. *Nomen est omen*: the personal *home*page and the Chat *room*

As an entrance page, the homepage performs two overall functions: first, it introduces the visitor to the general content of the site in a preferably informative and enticing way, and second, it functions as the official gateway of the website because it enables the reader to access and navigate the site. Askehave and Ellerup Nielsen (2005: 2-3) compare this binary function of the homepage in virtual reality with a spatial analogy in real life:

> This duality inherent in homepages may best be described by conceptualising the homepage as a front door with a door sign. The door sign indicates the name of the residents (i.e. the "content" of the house) while the door itself is the gateway (the medium) which enables guests to enter the house and visit the residents inside.

In a similar vein, Arnold and Miller (2003: 81, italics original) purport that among the personal homepages investigated in their study, the majority of them used "buildings as metaphors for the structure being presented" and that the use of spatial metaphors for data is generally very common on the Internet—"it is cyber*space* after all."

The communicative setting of Web Chat is different to that of the personal homepage in that the setting is often perceived as a room, in the sense of a terminated space of some sort. This difference shows in the spatial terminology used by Chatters and personal homepage authors when they themselves refer to the two communicative settings. As illustrated in Figure 9.1 below, only the term *room* could be determined as shared terminology.

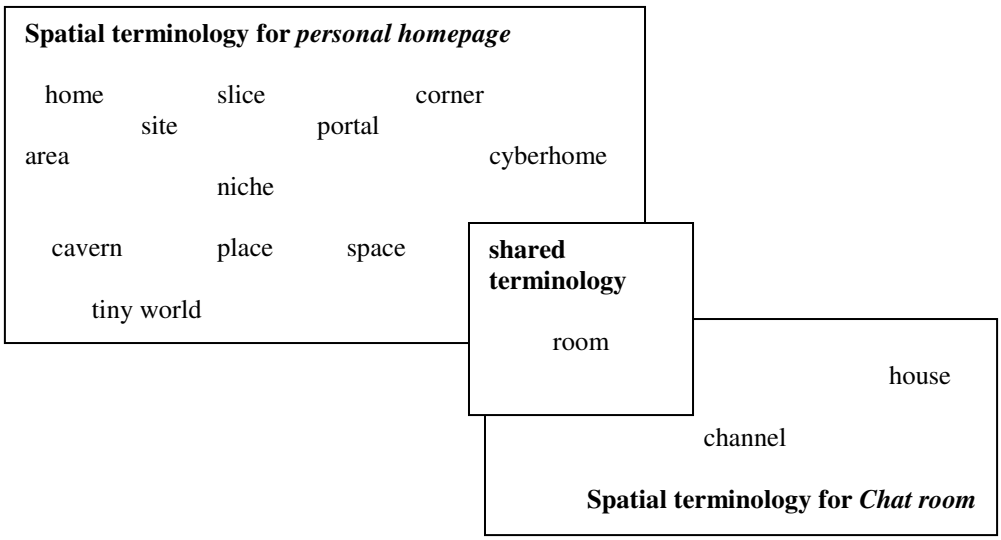


Fig. 9.1: Alternative spatial terminology for *personal homepage* and *Chat room*.

It was observed that the personal homepage is in most cases referred to with terminology reminiscent of a home or house. Notably, most terms indicate that the homepage is perceived as a defined space (for example, *niche*, *tiny world*, *corner*) rather than a hypertext that spreads in all directions without beginning or end. Only one homepage author refers to his homepage as a *room*, which in turn indicates that the homepage is more something of a house rather than a room. In Web Chat, on the other hand, the term *room* is by far the most prevalent expression among Chatters to refer to their communicative setting. The two other terms in Figure 9.1, *channel* and *house*, are extremely infrequent compared to *room*. However, apart from *cyberhome*, all terms used to denote these two virtual communicative settings draw on vocabulary from the "real world" and were adapted by means of metaphorical extensions to serve their cause in the "virtual world". This indicates that the conceptualisation of virtual space is met by means of analogies from physical space.

9.2.2. "Located" or "lost in space"?

In connection with the distribution of the spatial expression *in*, it emerged that most of its uses are oriented, regardless of the communicative setting, towards the physical location of the author at the time of message composition. One interesting observation concerns the circumstance that references to space and time are in proportion in EEC and the personal homepages, but the other three types of modern correspondence turned out to be more focused on spatial information as opposed to a communicated temporal timeframe. This is logical for two reasons: first, both SMS and e-mail messages are equipped with machine-generated temporal information of date and time in every message, and second, since Web Chat is (near-)synchronous with everybody present in front of their individual screens, the importance of a timeframe as a means of contextualisation decreases considerably (this will be discussed in more detail in the next chapter). The location of the author and/or communicative partner(s), however, seems to be of great interest in all types of correspondence. Consider the following example (102) from Web Chat:

(102) <god_of_gambler>: asl <gabby>
(…)
<gabby1>: 35 f d/dee
(…)
<god_of_gambler>: what dee ??????????????????????????????????????
<gabby1>: dundee
(…)
<god_of_gambler>: oh ok
(Chat excerpt / Room III, session 5)

Example (102) is similar to example (96) in that the main concern lies in finding out where another Chatter is located in the world. However, example (102) is somewhat more typical because the interest about location is coupled with an enquiry about age and gender, resulting in the popular Chat acronym *asl* (meaning 'age, sex, location', see also example (51.a/51.b, line 24) in chapter 7). Interesting to note is that example (102) also illustrates that the acronym *asl* is not just an "empty" opening formula without genuine interest in the answer (as is often the case with *How do you do?*), but a serious enquiry about another Chatter's basic personal particulars. At first, <god_of_gambler> does not understand the abbreviated location *d/dee* in <gabby1's> response to his/her *asl*-enquiry, and so he/she asks again a couple of turns later (*what dee*?), to which <gabby1> replies *dundee* (a city in Scotland). Although information about someone's age, sex, and location given out in anonymous Internet settings might (and often does) deviate from the truth (because there is no way to verify it), a lot of the correspondence revolves around the positioning of author and/or readership in the "real world".

Similar to Web Chat, authors of personal homepages also usually do not know their audience, but opposed to Web Chat, this audience is also not directly addressable, at least not until a visitor decides to contact the personal homepage author, who in turn would then be in a position to directly address his/her readership. It is for this reason that enquiries into the location of the readership are non-existent in the personal homepage corpus. Revelations about the author's coordinates, however, were observed to be extremely popular. All personal homepages include the author's country of residence, most also provide information on the author's location down to the city, and a total of eight homepages contain the full particulars of the authors' home addresses. One author even published a map to his apartment along with other relevant contact details:

(103) • Map to my Apartment – Come see me in person
• <postal address> – Mail me
• <e-mail address> – Email me
• <phone number> – Call me at IBM (where I work)
• <mobile phone number> – Call my cell phone
(Personal homepage excerpt / author: Fm_1)

Although the personal homepage excerpt shown in example (103) is unusually elaborate in its revelation of the author's contact details, it still shows that personal homepage authors are willing to share information on where they live with their readership. This stands in stark contrast to Web Chat, where many discussions revolve around countries and hometowns, but no postal addresses are revealed. This is insofar not unusual because Chatters are advised not to disclose their personal details (cf. hereto also example (50), a warning from a Chat operator not to give out contact details).

With regard to the types of communication that are aimed at an acquainted readership, the inclusion of the author's postal address is extremely rare in the messages of the respective corpora. A reason for this may be that the author's postal address in EEC can probably be found somewhere on the envelope (unfortunately, the corresponding envelopes to the letters of the EEC corpus were not available for research). In the case of e-mail and SMS discourse, it can be assumed that from a communicative point of view, the postal address takes a backseat to the e-mail address and the mobile phone number (to send SMS messages). Another factor comes into play concerning SMS discourse. Owing to the mobility of the mobile phone, it is not the home address that is of prime interest, but the current location of a given mobile phone owner. This has led to the adaptation of the conventional opening question *How are you?* to the spatially oriented *Where are you?* in mobile telephony as well as SMS messaging.

It emerges that the five types of correspondence deal with the notion of author localisation differently, and that authors in virtual communication settings, such as Web Chat and the personal homepage, seem to have a greater need to identify their location in order to contextualise the correspondence from a spatial point of view. In an attempt to answer the question in the title of this subsection, it appears that the interlocutors of all types of correspondence are by no means lost in space, as they have found different ways of locating each other. It will, in the following section, be interesting to investigate and compare the distribution of different types of spatial references to physical and virtual space.

9.2.3. Locations that matter

While the analysis of text-internal reference in chapter 8 focused on textual references (which may or may not be explicitly spatial), this chapter has so far dealt with the notion of deictic origo from a physical as well as textual point of view. It will now be of interest to look at spatial reference from yet another angle, namely to what sort of locations authors of personal written communication most frequently and specifically refer in physical and/or virtual space. Of course, the investigation into the text type EEC is constrained with regard to referencing to virtual destinations, as virtual space had yet not existed at the time the messages were produced. However, with regard to references to locations in the "real world", the analysis into EEC showed to be insightful and the results will be discussed shortly.

Table 9.3 below lists the different types of locations that the authors of the five types of correspondence referred to in physical and virtual space. The categories were not formulated prior to research, but determined after viewing the data. This means that all types of spatial references that occur in the text corpora are subsumed in Table 9.3. It should be noted that the categories are listed in accordance to their overall frequency in the five text corpora (except for the category "other, borderline case", which is listed last, regardless of frequency). The main parameters for locations in physical space are "geographical references" (from the *world* down to the notion of *region*), "public places" for leisure and entertainment (such as *hotel*, *restaurant*, *bar*, *pub*/*club*, *party*, *shop*), "public places" connected to work (terminology for workplaces such as *office*) or "transport" (for example, *public transport*, *train station*) and "private spaces" (referred to as *home*, *flat*, *house*, or *estate*). In the area of virtual space, the main dimensions include spatial references to the "Internet" or "Web" in general, the "current personal homepage" or "current Chat room", other "personal homepages" (or "Weblogs") and "Chat rooms", "public websites" (providing news or related to business), "discussion forums", and "any other service" available on the Internet. In addition, both physical and virtual references include a category "other, borderline case" for spatial references that do not fit into any of the above mentioned categories. Table 9.3 provides an overview of the frequency and distribution of all spatial references to physical or virtual space as observed in the five text types.

Table 9.3: Specific physical and virtual spatial references in personal written communication (corpora-based results).

Ref. to PHYSICAL space	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
region, town, city / canton, county	* 114 ** 1:161.6	232 1:110.9	52 1:546.2	142 1:112.9	178 1:174.6
state, country, continent	31 1:594.4	112 1:229.8	174 1:163.2	115 1:139.4	33 1:941.7
somebody's home (flat, house, estate)	73 1:252.4	96 1:268.1	13 1:2184.9.5	8 1:2003.8	89 1:349.2
hotel, restaurant, bar, pub, club, party	47 1:392.0	25 1:1029.3	3 1: 9468	1 1:16'030	2 1:15'538.5
public institution (university, museum etc.)	29 1:635.4	30 1:857.8	16 1:1775.3	20 1:801.5	24 1:1234.9
somebody's workplace	15 1:1228.4	30 1:857.8	13 1:2184.9	9 1:1781.1	2 1:15'538.5
the world, universe	1 1:18'426	10 1:2573.3	9 1:3156	10 1:1603	7 1:4439.6
shop, shopping	14 1:1316.1	3 1:8577.7	2 1:14'202	0	0
public transport, train/tram station, airport	31 1:594.4	12 1:2144.4	0	0	0
other, borderline case	56 1:329.0	82 1:313.8	80 1:355.1	66 1:242.9	81 1:383.7
TOTAL (physical)	**411 1:44.8**	**632 1:40.7**	**362 1:78.5**	**371 1:43.2**	**416 1:74.7**
Ref. to VIRTUAL space	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
current homepage or Chat room	x	x	313 1:90.7	246 1:65.2	x
public website (news, business, etc.)	1 1:18'426	8 1:3216.6	16 1:1775.3	58 1:276.4	x
Internet, Web	0	1 1:25'733	3 1:9468	41 1:391.0	x
other personal homepage or weblog	0	8 1:3216.6	2 1:14'202	26 1:616.5	x
discussion forum	0	2 1:12'866.5	0	15 1:1068.7	x
other Internet service	0	1 1:25'733	21 1:1352.6	0	x
other Chat room	0	0	13 1:2184.9	0	x
other, borderline case	0	2 1:12'866.5	36 1:789	63 1:254.4	x
TOTAL (virtual)	**1 1:18'426**	**22 1:1169.7**	**404 1:70.3**	**449 1:35.7**	**x**

Key: ref. = references; HP = homepage; * = total number of occurrences of indicated type of spatial reference per text type; ** = total number of occurrences of indicated type of spatial reference-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types; x = respective type of spatial reference not possible due to temporal origin of data or media-related constraints.

The frequencies highlighted in Table 9.3 show two main trends: while SMS discourse, e-mail, and EEC all most frequently feature physical spatial references to a region, town, city or canton/county, spatial references in Web Chat and the personal homepage, on the other hand, predominantly concern locations in virtual space, preferably within the same

communicative setting. Also fairly frequent are references to states, countries, or even continents. In this category, Web Chat scores highest, and this is definitely related to the inclination of Chatters to enquire into each other's whereabouts. However, because Chatters are advised by Chat room regulations not to give out specific personal details, they often draw the line at the country. This is probably also the reason why the frequency for physical spatial reference to cities and towns is lowest for Web Chat. Example (104) illustrates that this issue is indeed a delicate one:

(104) <24mcanada>: where in canada though?
(…)
<**Tristan**>: Ontario, thats all I'm going to say lol
(Chat excerpt / Room V, session 2)

The above example occurs early in session 2 of Room V, it can thus only be assumed that the Chatter nicknamed <**Tristan**> previously revealed that his/her location is somewhere in Canada. The Chatter <24mcanada> enquires further into the matter with the question *where in canada though?*. It seems as if <**Tristan**> does not feel too comfortable in narrowing down his/her location in the real world and is quick to point out that he/she will not reveal anymore than *Ontario* (a province located in the east-central part of Canada). Even though <**Tristan**> closes the turn with *lol*, aimed at making the "tone of writing" sound less harsh (or weakens the previous textual unit in terms of its contextual effect, see chapter 12), he/she does not engage with <24mcanada> anymore for the remaining Chat session. It seems as though being too nosey about someone's whereabouts is perceived as a breach of netiquette.

Specific spatial references to one's own or somebody else's home are rare in the Web Chat and personal homepage corpora and considerably more frequent in SMS discourse, e-mail, and EEC. This is definitely related to the kind of readership that the correspondence is addressed to. It seems safe to assume that references about one's own home are bound to be less frequent in messages addressed to strangers, as opposed to messages for friends or family. References to the remaining locations in the "real world" listed in Table 9.3 are altogether not very numerously referred to in any of the text corpora. However, it is evident that quite a lot of references are subsumed under the category "other, borderline case" and these references concern diverse places that do not neatly fit into any of the other categories.

For example, EEC features a lot of spatial references to journeys that are specific spatial references with regard to movement from one place to another, but the places themselves are in many cases not explicitly referred to (which does not qualify such references for the category "region, town, city / canton, county"). Also, it is not always clear (even in consideration of context), whether a certain place name is meant to refer to an estate (in which case it would have been classified as "somebody's home") or a town. Consequently, such ambiguous spatial references where then assigned to the category "other, borderline case". The same is true for spatial references in SMS discourse, in particular in connection with the locative adverb *here*, as discussed above.

With respect to Web Chat, such borderline cases include spatial references to the author's physical location in connection with his/her computer, as illustrated in example (105):

(105) <purpy_slurpy> is away from keyboard.
(Chat excerpt / Room VI, session 2)

Although example (101) contains a specific spatial reference (*away from keyboard*), it cannot be determined where the keyboard is in fact located (it could be in <purpy_slurpy's> home or office, in the library, in an Internet café or elsewhere). References like these were also assigned to the "other, borderline case" category.

The second trend highlighted in Table 9.3 concerns specific spatial references in virtual space. It was found that Web Chat and the personal homepages not only feature the most occurrences of this type of reference, but they are also predominantly endophoric spatial references rather than indexing other communicative settings that are located text-external in cyberspace. In comparison, SMS discourse and e-mail rarely contain spatial references to locations in virtual space and EEC, owing to the temporal origin of the data, does not feature them at all. As has been discussed above in connection with the distribution of spatial deictics, both Web Chat participants and the personal homepage authors show a tendency to use *in* or *here* to refer to the current communicative setting. Spatial terminology from the real world denoting *home* or *room*-like spaces are also frequently employed. Only rare spatial references to other virtual communicative settings were found, the only trend being references to other websites, of a public rather than a personal nature, in the personal homepages. Similar to the physical spatial references, there are also quite a few borderline cases in the references to virtual space. These

predominantly concern references where the destination of the reference is unclear (similar to the journeys in EEC). Compare hereto example (106) below:

(106) Jump on to other pages!
(Personal homepage excerpt / author: Qf_1)

In the personal homepage excerpt illustrated in example (106), it is unknown where those other pages are located, or what sort of information they might contain. Below the heading *Jump on to other pages!* the author lists various links, and if one clicks on them, they all lead to some kind of text-external destination (at the time of data collection). However, from the heading alone, the reader does not know to which pages the links will take him/her (other personal homepages, public websites, or else), and this is why spatial expressions as shown in example (106) were also categorised as borderline cases.

As illustrated in Table 9.3, spatial references to virtual space that qualify as borderline cases are most frequently observed in the personal homepage corpus, almost always in combination with hyperlinks. This implies that personal homepage authors generally do not feel the need to provide full details of spatial references in connection with hyperlinks because the reader is expected to do the follow-up him-/herself. Furthermore, the reading process of a text-based personal homepage is potentially non-linear. It seems as if the authors take this into consideration, because if spatial references only make sense by clicking a hyperlink, the reader is bound to deviate from the linear reading process—and, in case of hyperlinks that lead to text-external destinations, the authors also have to take into account that the reader may decide not to return to the homepage if the information on the hyperlinked website is of more interest. With regard to the other text types, borderline cases in connection with virtual spatial reference are rare or non-existent.

Furthermore, the investigation into specific physical and virtual spatial reference brought to light that the authors of the five types of personal written communication have different preferences: while SMS discourse, e-mail, and EEC contain mostly references to locations in physical space (predominantly smaller regions, towns, and cities), Web Chat and the personal homepages tend to favour references to their own communicative settings in virtual space. However, it should be noted that the second highest frequencies for Web Chat and the personal homepages concern references to states, countries, and continents located in physical space, indicating that the two communicative settings are by no means

solely oriented towards virtual space, but, on the contrary, include frequent references to the positions of their users as a means of spatial contextualisation.

All in all, most uses of spatial deixis were classified as unambiguous in the corpora-based investigation and this feature can be seen as typical of literacy, as indexing via gestures in a face-to-face conversation does not require such explicitness (compare, for example, pointing to a picture on a wall in face-to-face conversation and saying *I like this* as opposed to a written note without further information of what *this* is supposed to refer to). Of the few ambiguities found in connection with spatial deixis, they are most likely to occur in the more synchronous text types, which could be connected to Herring et al.'s (2003) concept of media richness (see also Figure 4.3 in chapter 4), where synchronous text-based CMC is closer to face-to-face communication than asynchronous media. However, Herring et al. (2006: 10) also point out that in collaborative multimedia settings (according to Herring et al. closest to face-to-face communication of all CMC), "spatial reference is complicated by the fact that users who move about must constantly adjust to shifts in perspective." There is thus a fundamental difference between text-based correspondence and multimedia settings, as shifts in perspective do not apply in the types of correspondence investigated in this study. This, however, also made it possible to classify the spatial deictics as oriented towards author or text.

9.3. Chapter summary

An investigation into space deixis requires a determination of the spatial deictic origo, as otherwise results would say more about the semantic properties of spatial deictics (which expressions are used?) rather than provide information about their pragmatic function (what are they used for?). From a contextual point of view, the notion of spatial deixis is immensely important as it not only localises author-/readership and discourse topics in terms of space, but also aids the organisation and thus contextualisation of the discourse itself. Text-based personal communication deals with two different types of spatial deictic origo: the physical location of the author at the time of message composition and the text with its (virtual) dimensions. Most interesting is the fact that both types of deictic origo can be traced in the text. However, this is, of course, a highly context-dependent process as nothing in the spatial expression *here* reveals anything about its deictic origo.

The investigation into the five different corpora (in consideration of contextual information provided in the messages) generated interesting results in connection with the use of deictic expressions and deictic origo. While spatial deixis in SMS discourse, e-mail,

and EEC tends to be oriented towards the author as deictic origo, Web Chat and personal homepages are more focused on the spatiality of the text and its dimensions. And although Chat rooms and the personal homepages investigated in this study are primarily text-based, they are nevertheless located in virtual space, which is bound to show in the texts produced in those communicative settings. It emerged that Web Chat and the personal homepages feature frequent spatial references to virtual space, however, these predominantly concern the current setting rather than referring to other locations in the Web. This feature not only distinguishes the two text types from the other types of personal written communication, but is also idiosyncratic of these types of CMC and in this sense confirms the second part of H1 (which claims CMC to show features of cyberdiscursivity).

Furthermore, the fact that the spatial origo in Web Chat and the personal homepages is predominantly oriented towards the text (as opposed to SMS discourse, e-mail, and EEC where the location of the author is more dominant with respect to spatial origo) is also to a certain extent related to the type of readership. Or, in other words, because authors of anonymous correspondence are aware that their readership is not able to identify their location (unless explicitly indicated), they tend to concentrate on the text as spatial origo instead. This supports H5, which predicts that differences in the contextualisation of personal written communication correlate with differences in readership (acquainted vs. unknown). With respect to H2 – H4, it was once more not possible to confirm the hypotheses that suggest differences between the modern types of communication and the more traditional Early English letter.

10. Text is time-bound

As has already been addressed in chapter 8, references on the textual level would invite the use of temporal expressions (for example, *as stated earlier*) as well as spatial deictics (for instance, *as stated above*). However, while written text is more likely to be organised by means of terminology borrowed from space deixis, spoken discourse has a tendency to be structured by means of temporal deictics (see hereto also Claridge 2001: 56). This is believed to be connected to differences in the production mode of spoken and written language. Written text is "materialised text" in the sense that the textual evidence is physically available. Hence, it is more likely to be indexed by means of spatial terminology. Spoken text on the other hand, is ephemeral (unless it is taped) and thus more of an abstract product, linearly and synchronically unfolding in temporal succession in a face-to-face meeting or throughout a telephone conversation. This, of course, invites the use of temporal expressions in connection with organising spoken interaction. Furthermore, in traditional perceptions of the two production modes, speech is also often said to be more concerned with the present, as opposed to writing that is more concerned with the past (see also the literacy-orality discussion in 4.2.). This in turn reflects in speech events as they are produced at present time, and writing as a product as completed at some point in past time. A challenge to these traditional notions of speech vs. writing are (near-) synchronous media, such as Web Chat or SMS discourse, that may be performed "on the go".

With respect to the use of discourse structuring devices, it was observed that personal written communication in general (including (near-)synchronous Web Chat) favours spatial deictics over temporal expressions. This, as discussed above, is believed to be related to the circumstance that writing is perceived as graspable, textual evidence. This chapter, however, is aimed at investigating how personal written communication is embedded in a temporal timeframe connected to the author's perception of time at the moment of message composition. These perceptions can be expressed in different ways and this chapter looks into the uses of tense (10.1.), being the most pervasive aspect of temporal deixis (cf. Levinson 2006: 114), time measurement fixed to the calendar (10.2.) and other selected temporal expressions (10.3.).

In consistence with previous discussions in this study, the analysis will focus on the author figure in connection with time deixis and the "temporal ground zero", the moment

at which an utterance is issued, also referred to as "coding time" (Levinson 2006: 114, based on Bühler 1934 and Fillmore 1997). In the case of written or recorded uses of language, the coding time can be distinguished from reading time with varying time lags in between. While these time lags are connected to the frequency with which the receiver of a message checks the respective medium (a user who checks e-mail once a day will produce a longer time lag than a user who checks it hourly), they are also to a certain extent media-related (the transmission of a letter takes considerably longer than the transmission of electronic messages). Largely ignoring reading time, this chapter will focus on the message at writing time (or coding time), that is the author's perceptions of time at writing time and how this is expressed in personal written communication.

10.1. Time deixis expressed through the verb phrase

> One might gloss the English present tense as specifying that the state or event holds or is occurring during a temporal span including the coding time, the past tense as specifying that the relevant span held before coding time, the pluperfect (as in *He had gone*) as specifying that the event happened at a time before an event described in the past tense, and so on. OBVIOUSLY, SUCH A SYSTEM FAILS TO CAPTURE MUCH ENGLISH USAGE.
>
> (Levinson 2006: 115, my emphasis)

What Levinson is driving at in the above quote is that an utterance with a verb phrase in present tense may not only be used to refer to events (or actions) that occur at the moment the utterance is issued, but that it can also denote an event (or action) that takes place in the (near) future, "as long as the event referred to is assumed to be prearranged, and there is an adverbial expression indicating a future time" (Green 2006: 411), as in *Team A plays against Team B tomorrow*. Furthermore, there is the notion of the "historic present tense, referring to past time, occasionally used in fiction (especially colloquial narratives) to produce a more vivid description, as if events were being enacted at the time of speech" (Biber et al. 1999: 454). Thus, it becomes evident that an investigation into the use of tenses needs to take into account that there can be (and often are) discrepancies between the tense used in a certain verb phrase and the point (or period) in time to which that verb phrase refers.

Before analysing the use of the three basic verb tenses of English (past, present, future) as used in personal written communication in more detail, I would first like to make a few introductory remarks. On a general note, the number of works that have been written

on the subject of tense in English is quite impressive (cf. Declerck 1991: 1)[73]. According to Declerck (1991: 1), "this shows both that the interest in the phenomenon of tense is great and that the problems in connection with it are numerous" and that "it is also striking how many different 'theories' of tense have been proposed." One of the main problems in connection with different theories of tense in English is the circumstance that "the linguistic literature reveals a total lack of consensus as to the question of how many tenses there are in English" (Declerck 1991: 8). While some suggest that we can speak of different tenses only if the verb form is morphologically differentiated, leading to the conclusion that there are only two tenses in English, others hold that tenses may be marked by auxiliaries (as is the case with future tense) as well as by inflectional morphemes (cf. Declerck 1991: 8-9). As will be explained in more detail below, this study follows the latter tradition in that future tense is understood to be a proper tense equal to present and past tense.

This study is focused on how tense is used by authors to place personal written communication into a temporal timeframe. It does not, however, aim at formulating yet another theory of tense. If not otherwise stated, this study will follow the classifications and terminology as proposed by Biber et al. (1999). Of main interest is the investigation into time deixis expressed through the verb phrase. It will concentrate on finite verb phrases (marked for tense or modality) and take into account the following four major structural distinctions (cf. Biber et al. 1999: 452ff.): "tense" (past, present, future[74]), "aspect" (perfect, progressive), "voice" (active, passive) and "modality" (marking verb phrases by means of (semi-)modals). It should be noted that the use of (semi-)modals will not be discussed in great detail in this chapter because a more extensive treatment of their function, in connection with grammatical mood, will follow in chapter 11.

[73] As regards the quantity of works published on tense in English, Declerck (1991: 1) refers to Schulze (1985), who "offers a bibliography which is twenty-five pages long in spite of the fact that it covers only a period of (roughly) ten years and does not aim at being complete." As this bibliography was compiled more than twenty years ago, one may rest assured that there have been many more publications in the field since.

[74] Biber et al. (cf. 1999: 452) include only present and past tense under the heading "tense", because future tense, as mentioned above, is not inflected but marked by modals (and is thus subsumed under "modality" rather than "tense"). However, for the purposes of this study it was decided to deviate from this classification and adapt it as outlined in section 10.1.

10.1.1. The present

Verb phrases in present tense constructions are by far the most frequent in all of the five text corpora, and they will thus receive a more extensive treatment than verb phrases in past and future tense constructions. As has been discussed above, a verb phrase in present tense need not always refer to a temporal span that is congruent with writing time. The present tense can also be used to describe an event that happened in the past or is planned to take place in the (near) future. This concerns in the majority of cases the uses of the simple present tense (cf. Biber et al. 1999: 454). However, according to Biber et al. (cf. 1999: 453), simple present tense that refers to present time has two major meanings: to describe a state at present time (for example, *I want a packet of crisps*), and to describe present habitual behaviour (as for instance, *She's a vegetarian but she eats chicken*). In addition, the simple present tense can also be used for general statements that do not only concern particular moments in time, but refer to actual "facts and figures" or "eternal truths" (cf. Leech 2004: 6) without specified beginning or end. This specific use of the simple present, commonly referred to as "neutral present" (cf. Bartsch 1995: 140), "is found in scientific, mathematical, and other statements made 'for all time'," (Leech 2004: 6) such as *Hydrogen is the lightest element* or *Two and three make five*.[75]

With respect to the empirical corpus analysis of this study, it was for reasons of scope not possible to make distinctions between simple present tense used (a) to describe a state at present time, (b) habitual behaviour, or (c) actual facts of permanent truth value. All of these uses were classified as pertaining to the category "simple present tense referring to present time" (see below). Although there are differences in scope of reference, all these uses of the simple present tense include, to a varying degree, the present time at coding time. In addition, the analysis of a smaller text samples from all corpora indicated that the frequencies of simple present tense referring to habitual behaviour and actual facts

[75] There exist differing viewpoints with respect to the referential scope of the present tense in English as presented in this study (proposing that the simple present tense can be used to refer to past, present, and future time). For example, Langacker (2001: 1) points out that "the English present is notorious for the descriptive problems it poses" and that "a characterisation in terms of present time seems hopelessly unworkable." The reason for this is, according to Langacker (2001: 1), partly because the simple present cannot be used to describe events that are occurring at the time of speaking (resorting to the present progressive instead), and partly because "many of the uses of the so-called present do not refer to present time at all, but to the future, to the past, or to transcendent situations where time seems irrelevant" (Langacker 2001: 1). Langacker (2001: 1, emphasis original) then makes the claim that "it appears, in fact, that the present tense can be used for **anything but** the present time." While this claim is certainly noteworthy, the empirical investigation into the five text corpora showed that the simple present is habitually used to index a time span referring to present time at writing (coding) time and less frequently serves to index time spans in the past and future time. It was thus decided to work with a threefold referential scope of the simple present tense (namely, past, present, and future time).

of permanent truth value were low. Hence, this generalisation with respect to simple present referring to present time was not rated as too strong a bias.

It should also be noted that for the purposes of this study, the simple present is defined, in grammatical terms, along the following guidelines: the simple present tense includes all uses of unmarked simple present (in active voice, for example, *I do it*) as well as "simple present + infinitive-constructions" (in active voice, either tensed verb + *to*-infinitive (for example, *I want to do it*) or tensed verb + bare infinitive (such as, *He helps me do it*)).[76] It excludes, however, all uses of marked present tense by means of aspect (perfect or progressive), passive voice, and modal verbs (the uses of marked present tense will be discussed afterwards). Based on these guidelines, the following three categories were defined: (1) "simple present tense referring to present time" (for example, *I am busy now*), (2) "simple present tense referring to past time" (as in *I received a letter from Peter yesterday and he writes about all sorts of things*), and (3) "simple present tense referring to future time" (for instance, *It is his turn next week*). Table 10.1 summarises the results that the investigation into the uses of the unmarked simple present in the different text corpora generated. Percentages are given in relation to the total number of occurrences of simple present verb phrases in each of the text types.

Table 10.1: Simple present tense referring to present, past, or future time in personal written communication (corpora-based results).

[1] **Simple present tense**	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
referring to PRESENT time	* 1309 ** 88.1	1354 94.6	2008 99.5	676 97.3	1184 92.8
referring to PAST time	5 0.3	19 1.3	2 0.1	11 1.6	32 2.5
referring to FUTURE time	172 11.6	59 4.1	9 0.4	8 1.1	60 4.7
TOTAL	**1486** **100**	**1432** **100**	**2019** **100**	**695** **100**	**1276** **100**

Key: [1] = includes all uses of unmarked simple present (active voice) and "simple present + infinitive"-constructions, excludes all uses of marked present tense (by means of aspect (perfect/progressive), passive voice, modality); HP = homepage; * = total number of occurrences of indicated type of time-reference of simple present per text type; ** = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of simple present found in each of the text types.

As highlighted in Table 10.1, the simple present is in the majority of environments used to refer to a time span that is congruent with writing time. Consider examples (107) – (109) below:

[76] For reasons of scope, no discussion on the functional and semantic differences between *to*-infinitive and bare infinitive can be included in this study. But see Duffley (1992) for an insightful and corpus-based examination on the contrast in function and meaning between *to*- and bare infinitives.

(107) Uff! It's over! The listening was awful but i think i did well in the speaking.We had really friendly examiners.Kisses, *now* i enjoy the sun!<N>
(SMS text, italics my emphasis / author: female, 20 yrs)

(108) Hi - yes:) This is still a *current* e-mail address.
(E-mail excerpt / author: female, 30 yrs)

(109) (…) I am not able to sett a sertain day when, by reason of my building and som busnes I have with the dilatory Chancellor of the Exchequer.
(EEC excerpt, 1616 / letter collection Cornwall, author: female)

In example (107), an SMS message, the verb phrases *(it)'s (over)* and *(i) enjoy* are in simple present tense and it can be gathered from the context that both uses correspond with a time span that occurs at the time of writing. In addition, the temporal adverb *now* further supports the assumption that *i enjoy the sun* is an action that took place at the moment the message was written. It is a similar situation in example (108), where the temporal adjective *current* supports the assumption that the verb phrase *(this) is* corresponds to a state of affairs that holds during the time span the message was composed. Example (109) is different in this respect because it does not contain any types of temporal expression tying the utterance to a temporal time span correspondent to writing time. However, based on the context provided, it can still be assumed that the underlined simple present tense constructions refer to present time at coding time.

All text types in the corpus show a clear preference for the use of the present tense referring to present time (as opposed to past or future time). In Web Chat, over 99% of the verb phrases in simple present tense refer to present time at writing time, which indicates that any other use is in fact an exception. Quite rare in all text types is the use of simple present referring to events or actions that took place in the past, as illustrated in examples (110) and (111) below:

(110) And this amazing piece of artwork *was drawn* especially for me by my special 9-year-old buddy, <link to pictures of the "9-year-old buddy"> of Bradford, England. As he points out, I live in downtown Seattle.
(Personal homepage excerpt, italics my emphasis / author: Df_1)

(111) (…) So to kill time we *took* a taxi to a beach bar at the end of the runway and *watched* the airplanes land, just over the beach... it's quite spectacular to see a B747 so close..(…)
(E-mail excerpt, italics my emphasis / author: male, 34 yrs)

As mentioned above, the historic present is mostly used to give an account of a past event a more contemporary feel. This is most probably also the incentive of the authors of both

examples (110) and (111), where the simple present (coupled with an infinitive in example (111)) is used to refer to events that took place in the past, but using present tense connects them to the present. The author of the personal homepage, from which the excerpt shown in example (110) is taken, first informs the reader about an artwork that *was drawn* for her by a 9-year-old boy, to then continue in present tense that this 9-year-old boy *points out* in his artwork that she lives in Seattle. While the simple present verb phrase *(he) points out* is not congruent with writing time (the pointing out took place before writing time), the other simple present verb phrase *(I) live* is presumably congruent with writing time (assuming that the author in fact lived in Seattle at the time this information was published on the homepage). Similarly, the author of example (111) first describes an event that happened in the past by means of two verb phrases in simple past tense (*took*, *watched*), to then switch to present tense for the evaluation of that event (*it's quite spectacular to see a B747 so close*). This not only gives the narration a more actual aura, but also implies that the impression to see a B747 so close still lingers. Still, it was observed that simple present referring to past time is not a very frequent feature in all five text corpora.

The same is true for simple present referring to future time, which is, apart from SMS discourse (where it makes up over 10% of all uses of the simple present), not very frequent. Interestingly, the simple present referring to future time is used to cover all sorts of time spans, and it is in all cases coupled with temporal expressions indicating future time:

(112) (…) You are welcome here any time! Let me know when you can come. I'm away *easter*.
(SMS text excerpt, italics my emphasis / author: female, 23 yrs)

(113) Just had 3 of my friends from London over last week, 2 German girls and a Brit. 2 friends from France come *in December* as does my girlfriend and another Kiwi friend from London.
(E-mail excerpt, italics my emphasis / author: male, age unknown)

(114) <ARMY_MP>: well i have to go to sleep talk to u all *tomorrow*
(Chat excerpt, italics my emphasis / Room I, session 5)

The simple present can be used to refer to a future event when this event is planned (and thus highly likely to take place) and is often coupled with the use of a temporal expression indicating future time. As can be gathered from the italicised terms or phrases in examples (112) – (114), this is the case in all three of them. The exact time when the message shown in example (112) was sent is unknown, however, it was definitely at some point before

Easter (or the utterance would not make sense), where the term *easter* functions as a temporal expression (tied to the "religious calendar" of Western culture who celebrates Easter on the first Sunday after the first full moon on or after 21st March). Thus, the simple present used in this particular lexical unit refers to something that has yet to take place at coding time.

According to the header of the e-mail excerpt shown in example (113), the e-mail was sent in September of the same year. Thus, the simple present verb phrase *come*, coupled with the temporal expression *in December*, refers to an event that was more than two months away at the time the message was written. The use of the simple present in this lexical unit suggests that the visit of the author's two friends, girlfriend, and another "Kiwi friend" was felt to be definite at writing time. In comparison, if the author had used "will" instead (*two friends from France will come in December as will my girlfriend and another Kiwi friend from London*), it would still have meant that the visit was planned at writing time, but from the sound of the utterance it would have moved more into the distant future.

The author of the Web Chat turn shown in example (110), on the other hand, refers to some action he/she intended to do the day after coding time (indicated by the temporal deictic *tomorrow*). By using simple present *talk (to u all)* to refer to an action that is planned to happen the next day, the Chatter <ARMY_MP> expresses that his/her return to the Chat room on the following day is very likely to happen. Many of the uses of simple present referring to future tense (above all in SMS discourse) were observed to be coupled with the temporal expressions *soon* or *tomorrow*.

All in all, it can be asserted that the simple present is predominantly used to refer to present time in all five text types, which in turn indicates that with regard to simple present, the authors prevalently contextualise their messages congruent to writing time. This assumption is further supported by the frequencies and distributions of different present tense constructions, as illustrated in Table 10.2. They are by far more frequent than past or future tense constructions. It should be noted that the different active voice constructions (in past, present, and future tense) are listed separately in the upcoming tables of this chapter. The frequencies for passive voice constructions, however, because of comparably low frequencies, will be presented in cumulated fashion. It should also be noted that the term "modals" as used in this study includes the central modals *can/could*, *may/might*, *should*, *would*, *must* and the semi-modals *need (to)*, *ought to*, *dare (to)*, *had better*, *have to, (have) got to*. The two central modals *will* and *shall* as well as the semi-modal *be going to* are excluded because they will be treated separately in connection

with the future tense (see 10.1.3. below). Furthermore, modals may be used to refer to present time in indicative or hypothetical mood (as in *I can read this* or *I could do this*, respectively) or they may be used in connection with past time in indicative or hypothetical mood (for example, *I could not read the road sign* or *I could have done this*, respectively).

Table 10.2 below presents the distribution of all present tense constructions that were found in the five text corpora. These include the following seven different types of present tense constructions: (1) "simple present" + "simple present + infinitive-constructions", (2) "simple present + periphrastic *do*", (3) "(semi-)modals referring to present time", (4) "present progressive", (5) "present participial phrases", (6) "present perfect", and (7) "present perfect progressive". As mentioned above, frequencies for (8) "present verb tense constructions in passive voice" will be presented in cumulated fashion.

Table 10.2: Present tense constructions in personal written communication (corpora-based results).

Types of present tense constructions	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
[1]simple present	* 1486 ** 1:12.4	1432 1:18.0	2019 1:14.1	695 1:23.1	1276 1:24.4
+ periphrastic *do*	3 1:6142	5 1:5146.6	2 1:14'202	4 1:4007.5	35 1:887.9
[2] (semi-)modals referring to present time	252 1:73.1	316 1:81.4	242 1:117.4	87 1:184.3	383 1:81.1
present progressive	129 1:142.8	184 1:139.9	111 1:255.9	33 1:485.8	11 1:2825.2
present participial phrases	121 1:152.3	148 1:173.9	170 1:167.1	103 1:155.6	176 1:176.6
present perfect	69 1:267.0	158 1:162.9	58 1:489.7	67 1:239.3	224 1:138.7
present perfect progressive	6 1:3071	25 1:1029.3	6 1:4734	5 1:3206	1 1:31'077
[3] present tense passives (all)	2 1:9213	23 1:1118.8	22 1:1291.1	26 1:616.5	113 1:275.0
TOTAL	**2068** **1:8.9**	**2291** **1:11.2**	**2630** **1:10.8**	**1020** **1:15.7**	**2219** **1:14.0**

Key: [1] = includes all uses of unmarked simple present (active voice), "simple present + infinitive"-constructions and constructions with periphrastic *do*, excludes all uses of marked present tense (by means of aspect (perfect/progressive), passive voice, modality); [2] = includes central modals *can*/*could*, *may*/*might*, *should*, *would*, *must* and semi-modals *need (to)*, *ought to*, *dare (to)*, *had better*, *have to, (have) got to*, excludes modals *will* and *shall* and semi-modal *be going to* for future tense constructions (see Table 10.5 below); [3] = includes all present tense constructions in passive voice; HP = homepage; * = total number of occurrences of indicated type of present tense construction per text type; ** = indicated type of present tense construction-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

The simple present does not only refer in most cases to a time span in present time (at writing time), but is at the same time also the most frequently used of all present tense constructions. As the different uses of the simple present have already been discussed above, I will now shift attention to the frequency and distribution of the simple present. It was observed that SMS discourse and Web Chat feature the unmarked simple present on the most frequent basis (one simple present verb phrase in every 12.4 and 14.1 words, respectively), the personal homepage and EEC feature it the least frequent (one simple present verb phrase in every 23.1 and 24.4 words, respectively), and e-mail is positioned somewhere in between (one simple present verb phrase in every 18.0 words). A similar distribution pattern emerges in the overall frequencies of present tense constructions shown in the total of Table 10.2 (although e-mail is somewhat less in the middle but closer to Web Chat and SMS discourse). It seems as though the frequencies of present tense constructions correlates with the degree of synchrony of the different media, where the more synchronous media feature present tense constructions more frequently than the more

asynchronous ones. However, we will come back to this issue after the discussion of all other tenses.

As can be seen in Table 10.2, the use of modals (on their own or in combination with other verbs) is also quite common. Again, SMS discourse features the highest frequency, the modals *can/could* and the semi-modal *(have) got to* (often realised as *gotta*) being by far the most frequent choices. With respect to the use of modals referring to present time, the discourse types SMS discourse, e-mail, and EEC show comparable frequencies, while Web Chat and the personal homepages feature considerably lower frequencies. In how far this can be brought into connection with the type of readership (acquainted vs. unknown), or is related to the text type, is unclear.

Another trend shown in Table 10.2 is that EEC and the personal homepages favour present participial phrases over the present progressive, and that overall, the present participial phrase is featured frequently in personal written communication. Consider examples (115) and (116) below:

(115) The cheldren, not knowing anything, ax me when you will com hom, (…)
(EEC excerpt, 1654 / letter collection Basire, author: female)

(116) (…) The hip surgery has been truly amazing, bordering on the miraculous! (…)
(E-mail excerpt / author: female, 80 yrs)

Both uses of the present participial phrase in the above examples express a connection with present time (at writing time), and the use of the present participle further suggests that the action described is of an ongoing nature (at writing time). This supports the above made assumption that authors of personal written communication tend to contextualise their messages based on events and states of mind that prevail at the time span during coding time. In addition, EEC shows the tendency to emphasise such events or states of mind by the use of "periphrastic *do*":

(117) (…) and in truth your being so far from me hath ben som sorow to me when I could not here from you: but oure cheldren do dayly pray for your prospering (…)
(EEC excerpt, 1654 / letter collection Basire, author: female)

Similar to example (117), the "periphrastic *do*" is often used in a contrastive manner (here indicated by *but*), to signalise that one is in fact doing a certain action or experiencing a particular state of mind, in most cases connected to present time (at coding time). This main focus on present time is further supported in that the present tense passives

predominantly concern simple present passives (for example, *is delivered*) and present progressive passives (for example, *is being delivered*), both suggesting connections to present time (at coding time). Another indicator are the low frequencies of both present perfect and present perfect progressive which both have a wider range into past time (albeit still connected to the present time) than all other tense constructions listed. These findings are confirmed by the investigation into past tense constructions, which are decidedly less prevalent than verb phrases in present tense constructions.

10.1.2. The past

Although past tense can be used to refer to present time (as in *I wanted to tell you* [and am hereby telling you] *that I won't be coming to your party next week*), it was observed that, opposed to the simple present (and simple future, see below), the uses of the simple past in the five text corpora do not show any trends towards marking another time. However, there are five occurrences of simple past marking present time in SMS discourse, and they all concern combinations with the verb phrase *wanted to let you know*:

(118) 5 different SMS text excerpts illustrating the use of simple past verb phrase *wanted to let you know* referring to present time:

(…) just wanted to let you know that I'll be there	(female, 24 yrs)
(…) just wanted to let you know I'm kinda fine	(female, 24 yrs)
(…) just wanted to let you know that I've arrived safely	(female, 25 yrs)
(…) wanted to let you know that our WG's finally set a date for our next party	(female, 27 yrs)
(…) just wanted to let you know we're not coming tonight	(female, 27 yrs)

All uses of *wanted to let you know* shown in example (118) could be replaced by the present verb phrase *I want you to know* without a drastic change in meaning, which indicates that the reference of *wanted to let you know* can be assumed to refer to present time (at writing time). According to Biber et al. (1999: 454), there are "functions of the past tense which relate more to present time, but with an added indication of stance." For example, with the verb *want*, among others, "past tense can indicate a present state of mind with a tentativeness that shows the speaker is being especially polite" (Biber et al. 1999:454). With respect to the excerpts shown in example (118), this is certainly the case as *wanted to let you know* is a polite way of beginning to inform someone about a certain matter. However, no other occurrences of simple past referring to another time span than past time were found in any the corpora. This subsection will thus be concerned with the

frequency and distribution of different past tense constructions and largely ignore potential distinctions in time reference of the simple past tense.

The main focus is placed on the simple past tense as "absolute" (i.e. textually specified time, for example, *Steven flew to New York yesterday*) and how other types of past aspects (such as the past perfect, *(before Steven flew to New York yesterday,) he had sorted out all pending business at his office*) are employed as "relative" to that textually specified time (cf. Levinson 2006: 115). Analogue to the present tense, seven different types of past tense constructions were investigated and these include: (1) "simple past" + "simple past + infinitive-constructions" (i.e. simple past + *to*- or bare infinitive), (2) "simple past + periphrastic *did*", (3) "(semi-) modals referring to past time", (4) "past progressive", (5) "past perfect", (6) "past participial phrases", and (7) "past perfect progressive". Again, it should be noted that all past tense constructions listed in Table 10.3 concern verb phrases in the active voice, the corresponding passive constructions have been subsumed under the category (8) "past tense passives (all)" for reasons of low frequency. With regard to the occurrences of modals, Table 10.3 refers to those modals that can be used to refer to past time. These include the central modals *could*, *might*, *should*, *would* and the semi-modals *dared (to)*, *had better*, and *had to*.

Table 10.3: Past tense constructions in personal written communication (corpora-based results).

Types of past tense constructions	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
[1] simple past	* 387 ** 1:47.6	650 1:39.6	268 1:106.0	210 1:76.3	696 1:44.7
+ periphrastic *did*	0	8 1:3216.6	1 1:28'404	1 1:16'030	31 1:1002.5
[2] (semi-)modals referring to past time	7 1:2632.3	8 1:3216.6	3 1:9468	6 1:2671.7	33 941.7
past progressive	12 1:1535.5	29 1:887.3	17 1:1670.8	4 1:4007.5	3 1:10'359
past perfect	4 1:4606.5	12 1:2144.4	1 1:28'404	2 1:8015	49 1:634.2
past participial phrases	20 1:921.3	8 1:3216.6	49 1:579.7	12 1:1335.8	3 1:10'359
past perfect progressive	0	1 1:25'733	0	0	0
[3] past tense passives (all)	0	22 1:1169.7	1 1:28'404	35 1:458	43 1:722.7
TOTAL	**430** **1:42.9**	**738** **1:34.9**	**340** **1:83.5**	**270** **1:59.4**	**858** **1:36.2**

Key: [1] = includes all uses of unmarked simple past (active voice), "simple past + infinitive"-constructions and constructions with periphrastic *did*, excludes all uses of marked past tense (by means of aspect (perfect/progressive), passive voice, modality);); [2] = includes central modals *could*, *might*, *should*, *would* and semi-modals *dared (to)*, *had better*, and *had to* that can be used to refer to past time, excludes all other modals and semi-modals; [3] = includes all past tense constructions in passive voice; HP = homepage; * = total number of occurrences of indicated type of past tense construction per text type; ** = indicated type of past tense construction-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

Table 10.3 highlights two main tendencies: first, past tense constructions are overall less frequent than present tense constructions, and second, e-mail and EEC feature more past tense constructions than all other text types. However, one feature that all text types have in common is that the simple past is by far the most popular type of verb phrase referring (in the majority of cases) to a fixed point in past time:

(119) <Hannahbanana87^292>: was it u who sed u won loads ov money *the ova day*?
(Chat excerpt, italics my emphasis / Room III, session 4)

(120) Did the earth move for you, too? There was an earthquake *at 2141* (…)
(SMS text excerpt, italics my emphasis / author: male, 33 yrs)

In both examples (119) and (120), the use of the simple past is coupled with temporal expressions (in italics) that indicate a particular point in past time. These are both textually specified uses of the simple past as absolute time, and it is interesting that other past tense constructions relative to the absolute simple past (such as, past perfect, or past perfect progressive) are extremely infrequent. It appears that when authors of personal written

communication decide to refer to past time by means of past tense constructions, they almost always settle for the simple past and in general steer clear of other types of verb phrases. Also, the use of (semi-)modals to refer to past time is rare, and the same is true for passive constructions. Furthermore, the past tense in combination with periphrastic (and often contrastive) *did* (as in *I forgot to clean the windows but I did do the dishes*) is also rare in all text types with the highest frequency in EEC (one instance in about every 1000 words). It can thus so far be asserted that all types of communication generally favour present tense constructions over past tense, however, the longer text types tend to contain less present tense constructions and more passives (present and past) than the shorter ones. It was also observed that authors of e-mail and EEC are more likely to make use of past tense constructions than any of the other three media. How future tense constructions fit into this picture will be of interest next.

10.1.3. The future

Similar to the present tense, which can be used to mark future time, the future tense can be used to refer to present time, usually by means of adverbial expressions indicating a present time, as in *John will be sleeping now* (cf. Levinson 2006: 115). However, as pointed out by Biber et al. (1999: 485), each modal (including *will*) "can have two different types of meaning, which can be labelled intrinsic and extrinsic (also referred to as 'deontic' and 'epistemic' meanings)." Essentially, "intrinsic modality refers to actions and events that humans (or other agents) directly control," whereas "extrinsic modality refers to the logical status of events or states, usually relating to assessments of likelihood" (Biber et al. 1999: 485). The future tense construction *John will be sleeping now* can be classified as epistemic in terms of speaker assessment, with the main function "relating to speaker stance rather than the marking of time distinctions" (Biber et al. 1999: 485). Hence, *John will be sleeping now* expresses a speaker stance along the lines of possibility or prediction—as opposed to the certainty that *John is sleeping now* would express. Nonetheless, the sentence *John will be sleeping now* still contains a verb phrase in future tense referring to present time at coding time.

As illustrated in Table 10.4 below, this is a sparingly employed use of the three future tenses (formed with modals *will*, *shall*, or semi-modal *be going to*) in the five text corpora. The frequencies shown in Table 10.4 include all uses of unmarked future tense constructions (active voice) with modals *will* and *shall*, and semi-modal *be going to* as well as "future tense + infinitive-constructions" (either future tense + *to*-infinitive (*I will pay*

you to do this) or future tense + bare infinitive (*I will help you do this*)), but excludes all uses of marked future tense (by means of aspect (perfect, progressive) and passive voice). Percentages are given in relation to the total number of occurrences of unmarked future tense found in each of the text types referring to either future or present time at coding time.

Table 10.4: Simple future tenses referring to future or present time in personal written communication (corpora-based results)

Unmarked future tenses	**SMS discourse**	**E-mail**	**Web Chat**	**Personal HP**	**EEC (letter)**
referring to FUTURE time	* 246 ** 98.4	155 95.7	81 90	33 100	232 99.1
referring to PRESENT time	4 1.6	7 4.3	1 10	0	2 0.9
[1] **TOTAL**	**250** **100**	**162** **100**	**82** **100**	**33** **100**	**234** **100**

Key: [1] = includes all uses of unmarked future tense constructions (active voice) with modals *will* and *shall*, and semi-modal *be going to* and "future tense + infinitive"-constructions, excludes all uses of marked future tense (by means of aspect (perfect/progressive) and passive voice); HP = homepage; * = total number of occurrences of indicated type of time-reference of simple future tenses per text type; ** = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of unmarked future tenses found in each of the text types.

Web Chat is the text type that most frequently features future tense that refers to present time. Similar to all other media, this effect is achieved in most cases with the future tense formed with the modal *will*, only rarely by means of the future tense formed with *be going to* and never with *shall*. Examples (121) – (122) illustrate, respectively, typical uses of future tenses formed with *will* and *be going to* referring to present time:

(121) (…) I will *nowe* bide you god night, for it is past a leauen [eleven] o cloke. (…)
(EEC excerpt, 1627, italics my emphasis / letter collection Harley, author: female)

(122) (…) I'm quite pissed and gonna sleep *now* (…)
(E-mail excerpt, italics my emphasis/ author: female, 20 yrs)

In both examples shown above, the verb phrases in future tense are combined with the temporal deictic *now* which "clashes" for two reasons: first, to couple *now* with a future tense is not in tune with a linear course of time and second, although *now* implies 'at this moment in time' both examples (121) and (122) show that the action referred to with *now* (*I will nowe bide you god night*, and *gonna sleep now*) are strictly speaking not happening at writing time (the letter writer continues to write and the e-mail correspondent is not sleeping yet). It may also be due to these types of paradoxes that authors of personal written communication in general steer clear of using future tenses to refer to present time.

With respect to using future tenses to refer to future time, it was observed that they are overall less frequently used than present or past tense constructions, however, the distribution across the five text types showed interesting tendencies. Table 10.5 illustrates the frequencies and distributions of future tense constructions with *will*, *shall*, and *be going to* as found in the five text types. These include (for all three types of future tense): (1) "simple future" and "simple future + infinitive" (simple future + *to*-infinitive or simple future + bare infinitive), (2) "future progressive", (3) "future perfect", and (4) "future perfect progressive". Again, the frequencies for different future tenses are listed for the active voice constructions only. Similar to present and past tense constructions, the passive is by far less frequently employed than the active voice and passive voice constructions are thus subsumed under category (5) "future tense passives (all)" in Table 10.5.

Table 10.5: Future tense constructions in personal written communication (corpora-based results).

Types of future tense constructions	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
[1]future tense WILL	* 215 ** 1:85.7	139 1:185.1	62 1:458.1	29 1:552.8	162 1:191.8
future progressive	8 1:2303.3	8 1:3216.6	2 1:14'202	1 1:16'030	1 1:31'077
future perfect	0	3 1:8577.7	0	0	0
future perfect progressive	0	0	0	0	0
[1] future tense BE GOING TO	29 1:635.4	20 1:1286.7	28 1:1014.4	3 1:5343.3	1 1:31'077
future progressive	1 1:18'426	2 1:12'866.5	1 1:28'404	0	0
future perfect	0	0	0	0	0
future perfect progressive	0	0	0	0	0
[1] future tense SHALL	6 1:3071	3 1:8577.7	0	1 1:16'030	71 1:437.7
future progressive	0	0	0	0	0
future perfect	0	0	0	0	1 1:31'077
future perfect progressive	0	0	0	0	0
[2] future tense passives (all)	0	2 1:12'866.5	11 2582.2	5 1:3206	10 1:3107.7
TOTAL	**259 1:71.1**	**177 1:145.4**	**104 1:273.1**	**39 1:411.0**	**246 1:126.3**

Key: [1] = includes all uses of unmarked future tense (active voice) and "future tense + infinitive"-constructions, excludes all uses of marked future tense (by means of aspect (perfect/progressive), passive voice); [2] = includes all future tense constructions in passive voice; HP = homepage; * = total number of occurrences of indicated type of future tense per text type; ** = indicated type of future tense-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

It emerges that future tense formed by means of the modal *will* is overall by far the most popular of all the future tense constructions. Of the remaining two future tenses, *be going to* is favoured by authors of modern correspondence, whereas the future with *shall* is distinctively more frequent than the future with *be going to* in EEC. It seems safe to assume that these preferences are most likely connected to the temporal origins of the modern vs. epistolary correspondence from the 17th century, meaning that the use of *shall* is not connected to the text type itself, but is more related to the time and age that the text type originates from. This suggests that letter writers of the 21st century are, similar to electronic communication, also less likely to opt for *shall* to refer to future time. As this chapter is more concerned with how authors contextualise their correspondence from a temporal point of view (past vs. present vs. future), and less with the particular distinctions in the vocabulary used, differences between *shall* and *be going to* future will not be considered any further.

In any event, the "*will*-future" is undoubtedly the most preferred choice to refer to some point in future time in all text types. Interestingly, many of these uses of *will* concern references to the establishment of contact in the (near) future, i.e. cataphoric text-external references. Consider examples (123) and (124):

(123) no worries 'cause of 16.8. we'll find another date, will write *e-mail* tomorrow (…)
(SMS text excerpt, italics my emphasis / author: female, 27 yrs)

(124) I will write (*a letter or email*) soon and give you all the details. (…)
(E-mail excerpt, italics my emphasis/ author: female, 23 yrs)

If the frequencies for cataphoric text-external references are recalled (as discussed in 8.2.2.), it emerges that they correlate with the frequencies of the uses of *will*. Future tense constructions with *be going to* (or *gonna*) are overall less frequent in all of the five text corpora, and although the reasons for this are not clear beyond doubt, it is assumed that it is connected to the difference in meaning. While both *will* and *be going to* can be used to express volition and/or prediction (cf. Biber et al. 1999: 485), the semi-modal *be going to* is often used for future events that are planned and are highly likely to take place. Future events referred to by means of *will*, on the other hand, are more of a predication rather than actual plans. Another aspect could be the brevity with which the *will*-future can be realised, above all the contracted forms of the *will*-future (such as *I'll* or *we'll*) are considerably shorter than *I'm going to* or even *I'm gonna*, which economises on text length and reduces typing speed. However, whatever the reasons may be, authors of personal written

communication show a clear tendency to opt for the *will*-future when referring to upcoming events.

10.1.4. Infinitive constructions

The last type of verb phrase to be discussed in this chapter is the infinitive construction that lacks a tensed verb (*to*- or bare infinitives in combination with tensed verb phrases have been subsumed in the respective verb phrase constructions referring to present, past, and future time). In *The English Infinitive*, Duffley (1992) examines, inter alia, the differences in meaning between *to*- and bare infinitives, the infinitive in combination with full and auxiliary verbs, as well as "the infinitive not incident to another verb" (Duffley 1992: 116ff.). Most useful to this study is his analysis of the *to*-infinitive as subject of a clause as it takes the temporal scope of reference of the *to*-infinitive into account. Duffley (1992: 127) hypothesises that "the support of the *to*-infinitive is always situated before the place in time (to be) occupied by its event." Duffley (1992: 126, example taken from Conrad 1982: 119) illustrates this circumstance with the example *To visit the poor is a Christian obligation*, where the notion of obligation implies that the place in time where the support is affected by the obligation "is prior to the realisation of the infinitive's action." In other words, the *to*-infinitive in this example refers to an action that has yet to be carried out. Hence, it is a temporal reference to future time.

As will be discussed below in connection with examples (125) and (126), where the *to*-infinitive does not take the subject position, the temporal reference of an infinitive without tensed verb phrase cannot always be determined beyond doubt. However, it appears that next to future time (see above), infinitives can also refer to present (see below), and past time. To illustrate the past time reference of the to-infinitive, Duffley (1992: 127, example taken from Conrad 1982: 137) gives the following example: *To read Gideon on Beet was a new literary experience* [the speaker refers to a man called Gideon who wrote a progress report on the work of another man called Beet]. Here, according to Duffley (cf. 1992: 126), the use of the *to*-infinitive implies that the speaker has actually realised the action denoted by the infinitive (otherwise the speaker could not evaluate the report as a new literary experience). In other words, the speaker has already read the progress report and thus refers to an action in the past by means of an infinitive construction.

Hence, although infinitive verb phrases that lack a tensed verb do not inherently refer to a particular time, their time reference can often be gathered from the context.

However, even in consideration of context it is sometimes not clear what type of tensed verb phrase (with a particular time reference) could be substituted for an infinitive construction in order to determine the time reference. Compare examples (125) and (126):

(125) Subject: just to say hi
(E-mail excerpt, subject line / author: female, 29 yrs)

(126) <ikle_sanj>: <URL of website> <==== to see my pic
(Chat excerpt / Room III, session 1)

In example (125), the use of the infinitive in the textual unit *just to say hi* is reminiscent of the past tense construction *just wanted to say hi* (see also example (118) above) referring to present time at writing time. However, *just to say hi* could also be paraphrased with the present tense construction *I want to say hi with this e-mail* which supports the above made assumption that the infinitive verb phrase in example (125) is referring to present time at coding time. Example (126), on the other hand, leaves more room for options with respect to possible tensed verb phrases that would fill the slot adequately. The infinitive verb phrase *to see my pic*, in combination with the backward arrow that points towards an URL of a particular website, could mean something along the lines of 'click on this link *if* you want to see my picture' and refer to present time in terms of a conditional probability. However, it could also be meant in the sense of 'click on this link *and* you will see my picture', referring to (near) future time in terms of an assertive statement. This is the reason why it was decided to discuss infinitive verb phrases separately, rather than trying to subsume them in present, past, or future tense constructions.

Overall, infinitive verb phrases are not extremely numerous in the five text corpora. EEC (with an infinitive verb phrase-to-word ratio of 1:217.3) featuring them the most frequently of all the text types, followed by e-mail (1:329.9), SMS discourse (1:409.5), the personal homepages (1:445.3), and Web Chat (1:887.6). The infinitive verb phrase-to-word ratios are not particularly meaningful by themselves. However, they are the last type of the verb phrases observed in the five text corpora, which means that overall frequencies of verb phrases can now be determined for each of the text types. It was found that SMS discourse features verb phrases on the most frequent basis (2802 occurrences, 1 verb phrase in every 6.6 words), followed by e-mail (3284 instances, 1:7.8), EEC (3460, 1:9.0), and Web Chat (3106 verb phrases, 1:9.1). The lowest frequency was observed in the personal homepages (1365 occurrences, 1:11.7) As has been pointed out above, the more synchronous media tend to generate texts with more verb phrases in present tense

compared to the more asynchronous ones. However, in the case of Web Chat (the most synchronous of all the media), this headstart with present tense constructions has been counterbalanced by lower frequencies in past as well as future tense constructions. SMS discourse, on the other hand, features high frequencies of verb phrases throughout all types of verb phrases, whereas the personal homepage (located towards the more asynchronous end of the immediacy continuum) features low total counts for all types of verb tense constructions.

It seems as though that the degree of asynchrony is not the only reason for the high frequency of verb phrases in SMS discourse. There is, indeed, another factor that has considerable influence regarding the number of verb phrases in a particular text type: text length in relation to the number of textual units. Compare examples (127), a 27-word excerpt from an SMS message, and example (128), a 26-word personal homepage excerpt, in connection with the number of verb phrases (underlined) and textual units (indicated by square brackets):

(127) (…) [i even remember the tune!] [i d forgotten this one,] [how beautiful!] [have an excellent nite] [&enjoy every minute],[once its all over it gets real bad] (…)
27 words, 6 textual units, 6 verb phrases
(SMS text excerpt / author: female, 24 yrs)

(128) [Our cozy little localpatron retail home on Valencia Street finally closed on May 19th.] [We had a great time,] [and hope to open again this autumn] (…)
26 words, 3 textual units, 3 verb phrases
(Personal homepage excerpt / author: Uf_1)

While the SMS text excerpt shown in example (127) consists of six textual units, containing a total of six verb phrases, the personal homepage text excerpt shown in example (128) is made up of three textual units and three verb phrases. It emerges that, regardless of its length, a textual unit is very likely to contain one or more verb phrases (although they may also lack them, of course, as in the exclamation *how beautiful!*) and consequently, more textual units per word count generate more verb phrases (see also the discussion on clausal and non-clausal units in chapter 11). And while turns in Web Chat are extremely short text entities as well, Web Chat is more likely to feature turns that lack verb phrases, partly owing to the tendency to split textual units into two or more turns rather than combining them into one turn (see also discussion of discourse markers in connection with textual units in 8.1.2.), and partly because Chatters in general have a tendency to omit as many

elements as possible in their contributions, which also reduces the number of verb phrases per words.

This trend is connected to typing speed (the quicker the turn appears on the screen, the more chances it has to be read, hence turns tend to be short), but is also made possible by the discourse structure of Web Chat, where every participant may contribute as many turns in a given discussion as he/she desires. In SMS discourse, on the other hand, users are in general hesitant to send, for example, five SMS messages in a row about the same topic and would be more likely to call or write an e-mail message instead. The same is true for topics that are seen as too delicate or complex to be discussed via SMS discourse (see also Höflich 2002: 45ff.). However, for issues that are seen as suitable for SMS discourse, users are aware that there is only a certain amount of space available to communicate these issues. It seems as though they adapt to this limitation by shortening the length of the textual units. Thus, overall it can be asserted that the discourse structure of the text types has an influence on the frequency of verb phrases, and that the authors communicating via the more synchronous media tend to favour present tense over past tense constructions. How another type of temporal reference, connected to the calendar, ties into these findings, will be of interest next.

10.2. The calendar as a means of temporal reference

There are differences, of course, in how temporal expressions are fixed in their reference to the calendar. For example, a specific date (such as *23.4.2007*) can only be used to refer to one particular day, whereas other calendrical units, such as *week* and *today*, "have a constant meaning, but systematically varying reference" because "they often have little descriptive content (and hence resist good paraphrase), but in their constantly changing reference they could hardly be more different" (Levinson 2006: 103-104). Thus, most temporal deictics (fixed to the calendar or not) are shifters and can be used to refer to a wide scope of potential referents. Another important factor that needs to be considered is that "the nature of calendrical units varies across cultures" (Levinson 2006: 114), and it should be noted that calendrical references investigated in this study are based on the notions of the calendar as perceived by Western culture. This is particularly important in connection with calendrical units like public (or bank) holidays that are connected to the calendar, but have a religious or political background (see also discussion of example (112) in connection with the religious holiday *Easter* above).

The investigation into the calendar as a means of temporal reference is based on the following main calendrical units that cover (more or less) fixed time spans: *year*, *month*, *week*, and *day*. All other temporal references in connection with the calendar are oriented towards these time units. All types of calendrical units considered in this investigation are listed in Table 10.6 below. Please note that no machine-generated time references were considered (such as *sent*/*received* information in e-mails and SMS texts), and that the references to a particular calendrical unit as listed in Table 10.6 include all variations that occurred in the text corpora (hence no distinctions were made between "full references", such as *Sunday* or *September*, or "abbreviated references" like *Sun* or *Sep*). The different types of calendrical units considered are listed under the following categories: (1) "year" (including *decade* and *century*), (2) "specific date" (*day*/*month*/*year* or any two of the three, as in *24.12.* or *June 2007*), (3) "season" (*spring*, *summer*, *autumn* (*fall*), *winter*), (4) "month" (as in *next month*, *in March*, *Apr.*), (5) "week" (including *weekend* and *fortnight*) and (6) "day" (such as *Sunday*, *in two days* as well as the spatial expressions *today*, *yesterday*, *tomorrow*). The last category listed in Table 10.6 is (7) "time of day" that includes specific indications of time (such as *2pm* or *21:14*) as well as particular sequences of the 24h-day (*morning*, *noon*/*midday*, *afternoon*, *evening*/*tonight*, *midnight*, *night*).

Table 10.6: Temporal references to calendrical units in personal written communication (corpora-based results).

Types of calendrical units	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
YEAR / DECADE / CENTURY (e.g., *in 2004*, *a year ago*)	* 8 ** 1:2303.3	83 1:310.0	18 1:1578	225 1:71.2	67 1:463.8
SPECIFIC DATE (*d/m/y or any 2 of the 3*)	19 1:969.8	28 1:919.0	0	91 1:176.2	66 1:470.9
public (or bank) holidays	35 1:526.5	28 1:919.0	0	1 1:16'030	3 1:10'359
SEASON (e.g., *next autumn*, *fall*)	4 1:4606.5	27 1:953.1	14 1:2028.9	10 1:1603	14 1:2219.8
MONTH (e.g., *in March*, *next month*)	10 1:1842.6	76 1:338.6	6 1:4734	138 1:116.2	77 1:403.6
[1] **WEEK / FORTNIGHT** (e.g., *in week 23*, *last week*)	58 1:317.7	80 1:321.7	4 1:7101	4 1:4007.5	22 1:1412.6
[2] **DAY** (e.g., *on Sunday*, *in 2 days*)	328 1:56.2	197 1:130.6	56 1:507.2	113 1:141.9	163 1:190.7
TIME OF DAY					
specific time of day (e.g., *2pm*)	118 1:156.2	26 1:989.7	0	21 1:763.3	3 1:10'359
particular sequence: morning	30 1:614.2	15 1:1715.5	5 1:5680.8	0	11 1:2825.2
noon / midday	2 1:9213	1 1:25'733	0	0	0
afternoon	13 1:1417.4	9 1:2859.2	0	0	0
evening / tonight	120 1:153.6	15 1:1715.5	10 1:2840.4	1 1:16'030	1 1:31'077
midnight	1 1:18'426	0	0	0	0
night	56 1:329.0	30 1:857.8	21 1:1352.6	4 1:4007.5	12 1:2589.8
TOTAL	**802** **1:23.0**	**615** **1:41.8**	**134** **1:212.0**	**608** **1:26.4**	**439** **1:70.8**

Key: [1] = includes the term *weekend(s)* as it refers to a time segment in relation to the calendrical unit *week*; [2] = includes the temporal expressions *today*, *yesterday* and *tomorrow* as they refer to time segments in relation to the calendrical unit *day*; HP = homepage; * = total number of occurrences of references to indicated type calendrical unit; ** = indicated type of calendrical unit-to-word ratio (rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

There are several trends in connection with the use of calendrical references in the five types of correspondence, the most obvious being that references based on the calendrical unit *day* are overall most frequent, only the category *year/decade/century* is more popular in the personal homepages. As pointed out above, several types of temporal expressions referring to time segments in relation to the calendrical unit *day* have been subsumed in this category. Consider the following examples (129) – (132) below:

(129) (…) I pray you remember that I recken [count] the days you are away; (…)
(EEC excerpt, 1625 / letter collection Harley, author: female)

(130) So, I've got a vacation from work, running from Wednesday through next Monday. (…)
(Personal homepage excerpt / author: author: Cm_1)

(131) <kool_kat>: see you guys tomorrow
(Chat excerpt / Room IV, session 3)

(132) How 'bout lunch in the sun today? Gimme a call if you feel like it
(SMS text / author: male, 27 yrs)

All examples above are illustrative of the fact that the time unit *day* is central to authors of personal written communication. It was observed, however, that references by means of the temporal expressions *today* and *tomorrow* are a lot less frequent in the more asynchronous dialogues EEC and the personal homepages, which in turn show preferences for the temporal expressions *day(s)* and weekdays (*Monday – Sunday*) as shown in examples (129) and (130). E-mail, Web Chat, and above all SMS discourse refer a lot more to points in time by means of the temporal expression *tomorrow*, almost always in connection with the arrangement of dates or the next Chat session (as shown in example (131)), and SMS discourse being the only medium where the temporal deictic *today* also occurs in relation to meetings as illustrated in example (132). This illustrates that the sender of the SMS text shown in example (132) assumes that the recipient will see this message before lunch time is over. It can thus so far be asserted that of the calendrical units listed in Table 10.6, all types of correspondence share the preference for the temporal reference by means of diurnal time spans. Another similarity concerns low frequencies for references to the four seasons. However, there is no obvious reason for this trend and it can only be assumed that authors of personal written communication steer clear of such references for reasons of personal taste.

Next to these similarities, there are also quite a few differences with regard to calendrical references and their distribution across the five text types. Notably, the incorporation of specific dates is more frequent in EEC and the personal homepages, whereas SMS discourse and e-mail are more likely to include references to public (or bank) holidays than all other text types. A reason why authors of personal homepages are hesitant to refer to public holidays could be their connotation with religion (for example, *Christmas* or *Easter*) or politics (such as the *Swiss national holiday* on August 1st). Yet the readership of personal homepages is potentially international, and maybe authors tend to avoid such

references because they may not make sense to everybody. Specific dates and references to the current year, on the other hand, are frequently incorporated in personal homepages. They are predominantly used to identify updates of a particular text segment, as shown in example (133) or, more commonly, the whole page, as illustrated in example (134):

(133) **What's new…**
- 8th August 2006: Draft of PhD proposal (in Writing section)
- 14th May 2006: A few more photos of <Name> (from when he was still very small)

(Personal homepage excerpt, emphasis original / author: author: Im_1)

(134) This page last updated on 07/26/2006
(Personal homepage excerpt / author: author: Df_1)

Example (133) and, above all, example (134) illustrate an important point in personal homepages: the need for an "absolute" temporal timeframe. Since authors have no idea when exactly their readership will read their published material, they identify the content of the page by means of the last update. Otherwise, the reader will not be able to aptly decode any of the other temporal references made in the page. Consequently, such "update references" generally refer to a date in the past (unless the content is read on the very same day it is published). All pages were visited once a week during a time period of six weeks, and it was observed that the majority of pages were only updated once or twice (if at all) during that time span. This could also be the reason why authors of personal homepages refer to the calendrical unit *month* more often than any of the other text types, as most update their pages on a monthly rather than a weekly or daily basis.

This is different in all other types of correspondence. Both e-mail and SMS discourse contain machine-generated temporal information and letters that are not dated can be classified in terms of delivery time. Web Chat, on the other hand, has participants assemble at the same time in front of their screens. Hence, it is clear for everybody that the content of the discussion is as current as it can get in written correspondence. Web Chat features overall the lowest frequency with regard to temporal references in connection with the calendar (one reference to a calendrical unit in every 204.3 words). This is not only related to the fact that all participants are present at the same time, and thus references to a specific time of (the current) day lose importance (see also 9.2.2.), but at the same time it indicates that this type of correspondence is very much focused "on the moment", rather than being oriented towards the past or the future (which is, again, reminiscent of spoken conversation).

Another discrepancy concerns the calendrical unit *week* (including the terms *weekend* and *fortnight*), which is considerably more popular in SMS discourse and e-mail than in all other text types. This is again related to the circumstance that these two media are often used to organise meetings. In particular SMS discourse, which features the highest frequency of specific times of day and references to the time unit *evening* (including *tonight*) for the same reason:

(135) I am on my way now, will be home 10:30ish
(SMS text excerpt / author: female, 23 yrs)

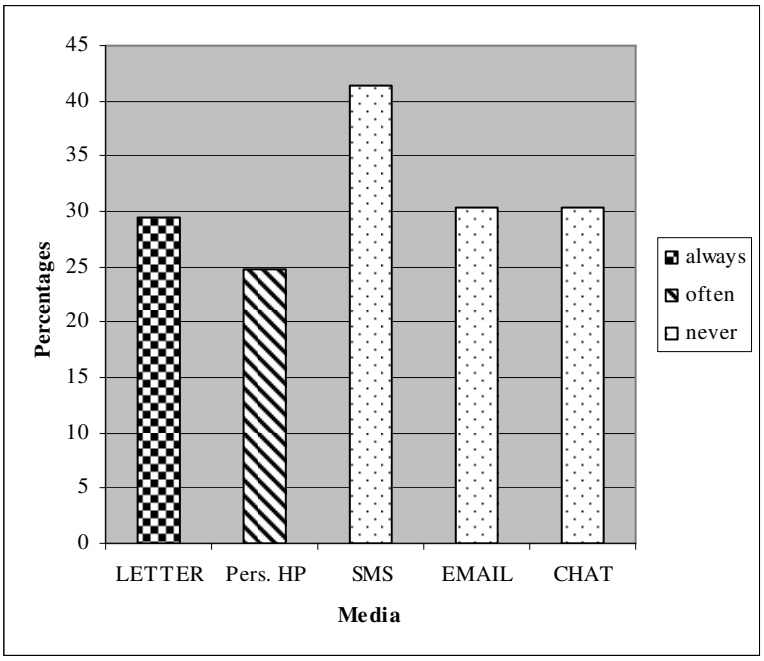
(136) (…) hope you're back early enough to pick me up around 20:15 & if no then I look forward to seeing you tonight! big kiss!
(SMS text excerpt / author: female, 26 yrs)

The type of temporal reference illustrated in examples (135) and (136) turned out to be idiosyncratic of SMS discourse, and while arranging dates via SMS discourse is one of the main reasons why users send SMS messages in the first place, it is also reminiscent of a telephone conversation (both examples (135) and (136) would seem displaced in the context of a face-to-face conversation).

Overall, it was observed that SMS discourse and the personal homepages feature the highest frequencies of temporal references fixed to the calendar—albeit for different reasons. While users of SMS discourse tend to organise and fix dates, authors of personal homepages aim at contextualising the content of their pages from a temporal deictic point of view. E-mail and EEC feature calendrical references less frequently. However, while EEC is most likely to include references to the calendrical units *day*, *year* and *specific dates*, e-mail tends to feature references to the calendrical units *day*, *week* and *month*. Authors of personal homepages, on the other hand, show a clear preference to include references to the calendrical units *day*, *specific date*, *month*, and *year*, which is definitely connected to the need for a temporal timeframe in order for the reader to be able to determine how current the content in fact is.

In order to determine how frequently authors of the 21st century think they incorporate calendrical references in their correspondence, I included a question on the temporality of personal written correspondence in connection with the calendar in the online survey. Informants were asked how often they thought they include a specific date (inserted by hand or by means of the keyboard, not mobile phone or computer-generated dates) in their messages composed on the different media. The results from the survey

revealed that the majority of informants felt that they always incorporate a date in their epistolary correspondence. With respect to the homepage, the question in connection with the insertion of specific dates generated the peak answer "often". Concerning the short message service, e-mail, and Web Chat, on the other hand, the majority of informants stated that they never incorporate dates into messages composed and transmitted via these media. Graph 10.1 below illustrates (in percentages) the peak answers of the 109 informants in connection with the insertion of specific dates.



Key: Pers. HP = Personal homepage

Graph 10.1: Insertion of specific dates in personal written communication (results from online survey, peak answers in percentages).

If the results from the survey are brought into relation with the findings that the empirical investigation generated, it becomes apparent that they propose similar trends. Although the empirical investigation showed that specific dates are more frequent in personal homepages than the letters, these two media feature overall the highest frequencies of specific dates and this is supported by the results from the online survey. Furthermore, it was found that occurrences of specific dates are generally infrequent in SMS discourse and e-mail, and non-existent in Web Chat. This, again, is confirmed by the results from the online survey as the peak answer is "never" for the insertion of dates in messages composed on these media. It can thus be asserted that the insertion of dates is connected to the text types and their media-related features.

As has been pointed out in the beginning of this section, references to calendrical units may vary in how they are tied to the calendar. However, they all share the property of denoting a more or less fixed time span. There exist, of course, other temporal expressions that denote certain points in time without being (explicitly) fixed to the calendar, such as *now* and *always*. Some of these expressions will be looked at in more detail in the next section.

10.3. Popular unfixed temporal expressions in personal written communication

While the investigation into the references to calendrical units generated results that can be tied to the communicative setting of the five different media, it will be interesting to see whether the same is true for temporal references that are not fixed to the calendar. For reasons of scope as well as comparability, only those temporal deictics that occurred at least five times within each of the corpora will be considered. These include (in alphabetical order): *already*, *at*, *first*, *in*, *last*, *late* (*later*, *latest*), *never*, *new*, *now*, *on*, *since*, *soon* (*sooner*, *soonest*), *still*, *then*, *time* and *when*. By far the most frequently used temporal expressions in all text types are *time* (318 occurrences) and *now* (315 occurrences). Examples (137) – (140) illustrate prototypical uses of *time*:

(137) easy, another time then, take care!
(SMS text / author: female, 23 yrs)

(138) Happy New Year! Hope you had a great time in Mexico!! (…)
(E-mail excerpt / author: female, 28 yrs)

(139) (…) have a good time in davos!!
(SMS text / author: female, 25 yrs)

(140) (…) We have a great time travelling through the States! Bet ya would like to join us … ;-) (…)
(SMS text / author: male, 32 yrs)

Similar to example (137), the uses of *time* mostly occurs in connection with the arrangement of dates. Another area of use concerns events in the past that were good (or are assumed to have been good (and only rarely bad)), as illustrated in example (138), or are expected to be fun (and in the rarest occasions bad), as shown in example (139), as well as events that fall into present time (at writing time), as illustrated in example (140). The temporal deictic *time* with its range of temporal references to a point (or time span) in

past, present, or future time occurs more frequently in SMS discourse and e-mail than in the other discourse types.

The use of the temporal expression *now*, on the other hand, is more limited in its time range as *now* is generally used to refer to present time (at coding time):

(141) (…) I'll say goodbye for now and catch you soon.
(E-mail excerpt / author: male, 50 yrs)

(142) pray Deare Mr. Hide forgive me for not writeing to you before now for the reasone is I have bin sick thre months & sinse I recoverd I have had nothing to interaine you (…)
(EEC excerpt, 1678 / letter collection Charles, author: female)

In example (141) the sender closes the e-mail with the words *I'll say goodbye for now* and as the message in fact ends immediately after this statement the *now* corresponds to present time. The same is true for example (142), where the author apologises for not having written before the point in time she refers to as *now*, and which happens to correspond with present time at writing time. Similar to the temporal expression *time*, the temporal deictic *now* is frequent in all text types, but most frequent in SMS discourse and e-mail. This is probably connected to the circumstance that authors using these media are aware of transmission speed and thus happily write about events or states of mind that hold at present time. Another term that was observed to be considerably more frequent in SMS discourse, compared to the other text types, is the temporal deictic *soon*, which in most cases accompanies plans to meet up with, or hear from, or write to the correspondent before long.

Expressions that turned out to be evenly distributed across the five text types include *at*, *in*, and *on* (with temporal as opposed to spatial meaning), which are regularly coupled with references to calendrical units, such as *in two weeks*, or *see you on Thursday*. Similarly, *then* and *when* also do not show distinctive distribution patterns for any of the text types, and the same is true for the remaining temporal expressions listed above. This implies that personal written communication, may it be modern or ancient, favours similar temporal expressions, although slight preferences towards temporal expressions referring to present time could be determined in texts produced on the new media. With regard to frequency, the temporal non-calendrical reference-to-word ratios determined for each of the corpora (including all temporal expressions found in each of the corpora, see appendix 15.1. for the complete list of all temporal expressions considered) correlate with the ratios

determined for calendrical references. It thus emerges that overall, authors of SMS discourse, e-mail, and the personal homepages contextualise their writing more frequently with temporal expressions than those of EEC and Web Chat.

As it is difficult to find out more on the use of a specific temporal expression (such as *now* or *never*) in a survey, the question into the uses of temporal expression was kept on a general level. It thus intersects with the previous discussion on calendrical references, as informants were asked how often they thought they use temporal expressions (besides specific dates) such as *now*, *yesterday*, *today*, *next week* (and so forth) to embed their messages in a timeframe. The results from the online survey insofar support the empirical findings, as the majority of informants answered that they use temporal expression most frequently in e-mail and SMS discourse, and less frequently in EEC and Web Chat. Yet the majority of informants also stated that they rarely include temporal expressions in their personal homepages, which contradicts the findings from the homepage corpus, where temporal expressions are overall very frequent. In any event, this deviation may also be related to the survey data being hypothetical to a certain extent, as it reflects on what authors *think* they do when composing their personal written correspondence. However, regardless of this inconsistency, the findings presented in this chapter suggest that the temporal contextualisation of personal written communication is media-related.

10.4. Chapter summary

The investigation into time deixis expressed through verb phrases showed that the more synchronous media favour present tense over past tense constructions, often coupled with temporal expressions (fixed or unfixed to the calendar) that point to present time at writing time. This is reminiscent of spoken discourse, which also tends to favour verb phrases in present tense constructions (simple, progressive, and perfect) over past tense constructions (see also Biber et al. 1999: 461ff.). However, although the personal homepage is a modern medium, but at the same time potentially asynchronous in nature, it features lower frequencies for verb phrases in present tense, and is thus comparable to EEC in this respect. Thus H1 (claiming all types of CMC to show features of orality) can only be partly confirmed. H2 (hypothesising features of orality in SMS discourse) and H3 (assuming that EEC shows features of literacy), however, are supported.

Furthermore, it was observed that overall, authors of SMS discourse, e-mail, and the personal homepages tend to contextualise their writings more frequently by means of temporal expressions (both calendrical and non-fixed to the calendar) than authors of EEC

and Web Chat. This means that neither the degree of asynchrony (the (near-)synchronous Web Chat features the lowest frequencies, but SMS discourse, which is close to the synchronous pole, features the highest), nor type of readership (EEC is directed towards an acquainted readership, whereas Web Chat participants do in general not know each other) influence the number of temporal expressions used. Thus, H4 (hypothesising that the modern text types show considerable differences in how they are contextualised compared to EEC), as well as H5 (claiming that the type of readership shows in distinctive ways in the contextualisation of the messages) could not be confirmed. However, it can be asserted that time deixis is connected to factors that are not primarily related to modern vs. handwritten correspondence, but are more concerned with text length (limited text length in SMS discourse is believed to foster shorter textual units, resulting in more verb phrases) and to what sort of effect the messages produced on the different media are aimed at. Hence, if SMS messages predominantly serve to exchange information about current events and states of mind as well as organising dates, then temporal expressions are bound to refer to the present time or near future time at writing time. The personal homepage, on the other hand, is in the majority of cases aimed at narrating the life of a particular person and is thus also more likely to be oriented towards past time.

11. Grammatical mood and syntactic structures as indicators for authorial intention

The investigation into grammatical mood and specific syntactic structures of any language is faced with several difficulties. Crucially, one needs to determine the most suitable research entity for this undertaking, which is not as straightforward as it might seem at first glance. The disagreement among grammarians on both the notion of grammatical mood and the classification system of specific syntactic structures into higher level functional categories further complicates the issue. Yet both grammatical mood and syntactic structures contribute considerably to the meaning of utterances and their illocutionary force. They are thus important carriers for authorial intention and even more important still, they are carriers of authorial intention that can be investigated on the textual level. This in turn sheds light onto how these features are used in the contextualisation of personal written communication. As has been pointed out in the theoretical discussion (see 4.5.), this study takes Biber et al.'s (1999) LGSWE as point of reference in issues connected to grammar. However, there are some areas where this study deviates, for either reasons of simplification or specification, from Biber et al. (1999). It was thus decided to include a (brief) discussion at the beginning of this chapter to clarify how terminology connected to grammatical mood and syntactic structures will be used in this study.

11.1. Terminology

11.1.1. Grammatical mood

The notions of "mood" and "modality" are often used interchangeably. However, while the two concepts are certainly related, they do not have the same scope. According to Biber et al. (1999: 483), "English verb phrases can be marked for either tense or modality, but not both" and the term "modality" then refers to the range of use and actual uses of central and semi-modals. The concept of "mood", on the other hand, has a larger scope in that it subsumes all types of verb phrases (tensed in active or passive voice, or untensed in active or passive voice, and marked by aspect and/or modality). In addition, the notion of mood also subsumes language use that does not contain a verbal phrase. Hence, for the purposes of this study, all language output is subject to grammatical mood, and there are three main types: indicative mood, hypothetical mood, and subjunctive mood. As has already been stated earlier in this study, the subjunctive mood (where the verb phrase indicates that an

act or state is contingent or hypothetical, expressing doubts, hopes, and wishes) is practically non-existent in the modern text types and only occurs on a regular basis in EEC. Hence, it is not possible to compare uses of the subjunctive across the five text corpora. It is for this reason that the subjunctive mood will not receive the same amount of attention as indicative and hypothetical mood.

11.1.2. Research entity

In the discussion on discourse markers in relation to text-internal structure (chapter 8), and the investigation into verb phrases in connection with temporal reference (chapter 10), the relevant research entity was the textual unit, denoting a string of words that forms a coherent chunk of language (see also 4.3.1.). From a syntactic point of view, the textual unit can be further specified according to specific structures and, in written language, this is also connected to punctuation. On a general level, syntactic analysis is concerned with sentence structure and how sentences are composed. Although the sentence is often cited as the main anchor of linguistic syntax (also referred to as "the study of sentence structure"), it is in fact "difficult to give a good linguistic definition of a sentence which applies equally well to writing and spontaneous speech" (Biber et al. 1999: 202).

In the light of these circumstances it seems useful to look for a research entity that lends itself more easily to syntactic research than the sentence, and is at the same time more specific than the notion of a textual unit. Biber et al. (1999: 120) propose the clause which they define as "a unit structured around a verb phrase," where the verb phrase is accompanied by one (usually the subject) or more elements. The core of the clause can thus be divided into two main parts: the subject and the verb phrase (cf. Biber et al. 1999: 122). Clauses are traditionally distinguished whether they are finite (contain a verb form which specifies their tense or modality) or not. Finite clauses are independent clauses and "correspond to what is generally defined as sentences in other grammars" (Biber et al. 1999: 202). Clauses that are embedded in such independent clauses are referred to as dependent clauses. However, this need not concern us any further as this distinction is not central to the research aim of this study.

Thus, so far it emerges that the notion of clause facilitates the investigation and definition of syntactic structures. However, if language in use should be analysed purely in terms of clauses, then it would quite quickly become apparent that there is a lot of language in spoken conversation, as well as in writing, that does not conform to the notion of a clause. This is where Biber et al.'s (1999: 224) notion of "non-clausal material" is

most useful. Non-clausal material is defined by two main characteristics: first, the unit cannot be analysed internally in terms of clause structure, and second, it is not analysable as part of any neighbouring clause. According to Biber et al. (1999: 1082, emphasis original), non-clausal material can be broadly divided into two categories: "(a) single words (like *Hi*) known as **inserts** (sometimes occurring with added modifiers, as in *Hi there*), and (b) **syntactic non-clausal units**, such as *My turn?*" that are capable of entering into syntactic relations with other units, as opposed to inserts, which cannot. Although the notion of inserts is introduced in the section on the constructional principles of *spoken* grammar (Biber et al. 1999: 1082ff.), it turned out to be an immensely useful concept if applied to personal written communication (above all with regard to the more synchronous media). The category inserts subsumes, for the purposes of this study, the following five types: (1) "interjections", (2) "expletives", (3) "greetings/farewells", (4) "hesitators", and (5) "response forms" (the properties of which will be explained as we proceed). Thus, the investigation into syntactic structures will focus on, and make a general distinction between, the notions of clause and non-clausal material, where the non-clausal material subsumes inserts and all other (syntactic) non-clausal units. Hence, the research entities for the investigation into the syntactic structures, and their functional aspects, will be the clausal unit or clause (independent and dependent clauses) and the non-clausal unit or non-clause (inserts and all other non-clausal material). However, as will be seen below, all units, both clausal and non-clausal, are classifiable in their performance of specific functions.

11.1.3. Syntactic structures and functional categories

The five syntactic structures of interest for this study (declarative clause, interrogative clause, imperative clause, exclamative clause, and non-clausal material) are listed under different entries in Biber et al. (cf. 1999: 202ff., 1082ff.) and they distinguish between different types of independent clauses, dependent clauses, non-clause material in writing, and inserts, non-clausal constructions found predominantly in spoken language. As has been discussed above, the classification used in this study draws on these concepts, albeit in an adapted as well as simplified manner. First, no distinction will be made between independent and dependent clauses, and second, selected types of inserts, originally formulated for the grammar of spoken language, will be applied to written communication. Words or string of words that do not qualify for the categories "clause" or "insert", are classified and referred to as "other non-clausal material".

Furthermore, although the ellipsis of a first person pronoun (such as *I*, *me*, or *we*, *us*) would normally render a clause syntactically anomalous, such clauses were not distinguished from other independent/dependent clauses as long as the remaining syntactic structure (i.e. the verb phrase and its complements) was "intact". This decision is based on the observation that the vast majority of first person pronoun ellipses (which occurs predominantly in the more synchronous text types, see also 7.1.1.) concerns formulaic sequences[77], such as *see you soon*. In this particular example, it would even be unusual, at least in personal written communication, to include the first person pronoun, as in *I see you soon*. Apart from first person pronoun ellipsis, no further adaptations were made in distinguishing clausal from non-clausal material. This means that all non-clausal material lacks elements of the verb phrase, or the whole verb phrase. It should be noted, however, that this investigation is more concerned with the functional aspects of different structural clause types, and less with the structure itself.

The syntactic structures of interest include (with their main functions in brackets): (1) "declarative clause" (statement), (2) "interrogative clause" (question), (2) "exclamative clause" (exclamation), (3) "inserts" (subsuming "interjections", "expletives", "response forms", "greetings/farewells", and "hesitators"), as well as all other (4) "non-clausal material". Table 11.1 gives an overview of these different syntactic structures with illustrating (self-chosen) examples.

[77] The notion of a "formulaic sequence" (strings of words that appear to be stored and retrieved whole from memory) is reminiscent of idioms (see Wray & Perkins 2000). This will, however, be discussed in more detail in connection with contextual effects in chapter 12. For the moment, it suffices to know that formulaic sequences, such as *see you later*, are idiomatic to a certain extent, and the omission of the first person pronoun *I* is thus not an ellipsis in the strictest sense, but a common usage of this particular sequence.

Table 11.1: Structural and functional classification of different types of clauses and non-clausal material (adapted from Biber et al. 1999: 202).

[1] **Structural types**	**Functional types**	**Examples**
DECLARATIVE CLAUSE	Statement	*This restaurant serves delicious food.* *I don't like this restaurant.*
INTERROGATIVE CLAUSE	Question	*Are you going to the cinema?* *Why weren't you there?*
IMPERATIVE CLAUSE	Command, request	*Close the window.* *Please do not open the door.* (softened)
EXCLAMATIVE CLAUSE	Exclamation	*What a wonderful restaurant!* *That's just not right!*
INSERTS + OTHER NON-CLAUSAL MATERIAL	Expressive (interjection, expletive) Responsive (response forms) Systematic (greeting/farewell) Pause filler (hesitators) Various (other non-clausal material)	*Wow!* (interjection), *Damn!* (expletive) *Yeah*, *ok*, *nope*, *no* (response forms) *Hi Tom* (greeting), *Bye bye!* (farewell) *Erm*, *uhm* (hesitators) *Sorry!*, *Why?* (other non-clausal material)

Key: [1] = It should be noted that Biber et al. (1999: 202) exclusively refer to independent clauses in their distinction of the different structural types as presented in Table 11.1. However, for the purposes of this study the structural types refer to both independent as well as dependent clauses. The definition of "structural types" thus deviates from the system proposed by Biber et al. (1999: 202). Also, to include inserts as a fifth structural type deviates from the classification as proposed by Biber et al. (1999) and has been added for reasons outlined in 11.1.2. above.

Each of these different structural types will be addressed and further illustrated in more detail in section 11.2. Of main interest will be their distribution in connection with grammatical mood and their frequencies across the different text types. It will also be of significance whether the different structural types are affirmative or negated. The four types of clauses listed in Table 11.1 (declarative, interrogative, imperative, and exclamative clause) can all be, as illustrated in the examples, of affirmative or negated syntactic structure. In contrast, some of the types of inserts do not feature negated variants. As this will be discussed in more detail below, I would now like to draw attention to the scope of negation as understood in this study. According to Biber et al. (1999: 175), "the scope of negation is that part of a clause that is affected by the negative form" and "the scope may be restricted to a single word or phrase" (referred to as "local negation"). Thus, both clausal units (as in *I don't know*) and non-clausal units (such as *no idea*) can be negated. This will be considered in the discussion of the different structural types in connection with the empirical findings. However, before turning to the results that the empirical investigation generated, I would first like to draw attention to the notion of "illocutionary force", the properties of which are related to the discussion of grammatical mood and syntactic structures.

11.1.4. The notion of illocutionary force

The term "illocutionary force" goes back to Austin (1962), who, together with Searle (1969), developed and formulated a theoretical framework for the analysis of speech acts. A profound discussion of the notion of illocutionary force, a concept of a rather complex nature, would exceed the scope of this study and will thus be held brief. It should also be noted that both Austin and Searle refer to spoken language rather than written exchanges in their theories. However, the notion of illocutionary force, as used in this study, is also applicable to written language.

Most central for this study is Austin's (cf. 1962: 109ff.) classification of performatives into three types: the "locutionary act" (the act *of* saying something, i.e. the utterance itself with determinate sense and reference), the "illocutionary act" (the act performed *in* saying something, i.e. the intention of the speaker), and the "perlocutionary act" (the act performed *by* saying something, i.e. the effect on the hearer). Of those three types the illocution is of most interest in that it reflects the intention of the speaker. It should be noted that Austin (1962) distinguishes between performatives (utterances used to do things, subject to felicity conditions), and constatives (utterances used to report things, subject to truth conditions). However, this distinction will not be taken into consideration in this study. This decision is based on the belief that, at their base, all utterances are performative in the sense that they are intentional and take effect on both the hearer as well as the communicative context established. Hence, the notion of illocutionary force applies to all types of utterances for the purposes of this study.

The illocutionary force of an utterance is not only expressed by its semantic content, but is also connected to its syntactic structure so that, for example, an interrogative (functioning as a question, e.g. asking for information) differs considerably in its illocution from a declarative (functioning as a statement, e.g. providing information). Consequently, the illocutionary force of an utterance is tied to its performance in a certain speech act situation. For the purposes of this study, the properties of illocutionary force, due to the absence of readership, can only be investigated in connection with the author's intention as expressed through his/her writing. However, as Fitzmaurice (cf. 2002: 61) aptly points out, even in the presence of the hearer, or reader in the case of this study, the speaker (author), cannot guarantee that the addressee will apprehend the addressor's intention in the way in which it was intended.

11.2. Grammatical mood, syntactic structures, and illocutionary force

After having briefly introduced all relevant terminology as used in this chapter, I would now like to turn to the findings that the empirical investigation into grammatical mood and different syntactic structures generated. Please note that due to very low numbers, passive voice constructions will not be treated separately. This means that in the following discussion, all frequencies and percentages include both active and passive voice constructions. Furthermore, the classifications presented in the upcoming Tables 11.2 to 11.6 are based on distinctive syntactic structures and it should be borne in mind that the functions and illocutionary force of similar syntactic structures may vary considerably. In any event, this concern will be met in the discussion of the findings.

11.2.1. Declaratives

Declarative clauses, which typically express statements, are marked with SV (subject-verb) structure, although they may have VS (verb-subject) order under special circumstances, also referred to as "subject-verb inversion" (cf. Biber et al. 1999: 152, 203). It emerged that the declarative clause is by far the most common syntactic structure in all of the text types. In addition, a lot of the non-clausal material, inserts excluded, was also categorised to function as statements. Compare examples (143) and (144):

(143) Hey sweetheart! My arms are wide open and waiting for you. <name>
(SMS text / author: male, 24 yrs)

(144) (…) They selected one of my photos to use for the Year 2005 Anniversary poster for the Market Festival, so that's a good thing. Good publicity.
(E-mail excerpt / author: male, 40 yrs)

In example (143), the SMS message consists of two inserts (the greeting *Hey sweetheart!* and the farewell <*name*>, see 11.2.5.) and the syntactically complete declarative clause *My arms are wide open and waiting for you*. In example (144), an e-mail excerpt, the underlined text segment shows a typical non-clausal unit of a declarative structure. The author includes the noun phrase *Good publicity* to further comment on the previous clause *(so) that's a good thing* (see 11.2.5. on the special status of discourse markers, such as *so*). He could have integrated the declarative noun phrase *Good publicity* somewhere in the preceding clauses, or he could have formed a clause out of the noun phrase itself (for example, *This will be good publicity*). Instead, the author decided to separate the non-clausal material by means of punctuation which, due to its curtness, draws attention to it.

Table 11.2 gives an overview of the distribution of different types of such declarative clauses and non-clauses across the five text corpora. As far as the reading of Table 11.2 is concerned, it should be noted that, next to clausal vs. non-clausal units, two further distinctions are made: first, between indicative mood (IM) and hypothetical mood (HM) and second, between affirmative and negated constructions. Of course, it is debatable whether it is apt to speak of "hypothetical declaratives" in connection with truth conditions. It raises questions such as 'how much of a declaration is a statement that is hypothetical?', because hypothetical mood is oriented towards expressing what is counterfactual, but otherwise possible. However, as has been pointed out in the beginning of this section, classifications are based on the syntactic structure of the clausal and non-clausal units. Differences in their functions and illocutionary force will be discussed shortly. Percentages as presented in Table 11.2 are given in relation to the total number of occurrences of declarative clauses and non-clauses in each of the text types.

Table 11.2: Declarative clauses and non-clauses in personal written communication (corpora-based results).

Types of declaratives		**SMS discourse** (18'426)		**E-mail** (25'733)		**Web Chat** (28'404)		**Personal HP** (16'030)		**EEC (letter)** (31'077)	
		IM	**HM**	**IM**	**HM**	**IM**	**HM**	**IM**	**HM**	**IM**	**HM**
AFF.	clauses	* 74.8	4.4	75.6	5.5	60.4	2.1	57.8	2.7	76.7	6.5
	non-clauses	12.1	0	5.7	0	26.4	0	36.0	0	0.7	0
NEG.	clauses	6.8	0.3	12.2	0.5	9.6	0.5	3.2	0.1	14.3	1.5
	non-clauses	1.6	0	0.5	0	1.0	0	0.2	0	0.3	0
TOTAL		** **1:9.7**		**1:12.2**		**1:10.9**		**1:12.9**		**1:13.0**	

Key: AFF. = affirmative; NEG. = negated; IM = indicative mood; HM = hypothetical mood; HP = homepage; * = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of declarative clauses and non-clauses in each of the text types; ** = declarative-to-word ratio (including all types, rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

Declarative structures (in the majority of cases subject-verb structures, followed by a full stop or comma), are overall the most frequent structural type found in all five text corpora. With respect to frequency, it is the text types SMS discourse and Web Chat that feature declaratives most numerously. A similar observation was made in connection with the frequency of textual units and verb phrases. Since the textual unit is related to the notion of clausal and non-clausal unit, it is a logical consequence that SMS discourse and Web Chat also feature more clauses and non-clauses. More interesting than the absolute frequency is, however, the distribution of the different types of declaratives. There are two main trends in Table 11.2. First, all media favour the affirmative declarative clausal unit over all other

types. Second, the use of hypothetical declarative clausal units does not exceed 7% for any of the text types, and no non-clausal units in hypothetical mood could be determined.

Apart from those two main trends it was found that both Web Chat and the personal homepage feature the declarative and affirmative non-clausal unit in indicative mood more frequently than all other text types. This is connected to particular idiosyncratic discourse structures, compare examples (145) and (146) below:

(145) <mary306girl>: where are u
<milkybytheway>: canada
(…)
<mary306girl>: canada's biggggggggg … ???
<milkybytheway>: ontario
(Chat excerpt / Room VI, session 5)

(146) **Political Resources**
Women's resources
GLB Resources
Progressive Resources
Electoral Politics
Environmental Resources
Right Resources
(Personal homepage excerpt, emphasis original / author: Ff_1)

In the Web Chat excerpt shown in example (145), a Chatter nicknamed <mary306girl> asks <milkybytheway> about his/her whereabouts. While this is nothing unusual in Web Chat (see also discussion on spatial deixis in chapter 10), it is special compared to all other text types, where such interactivity is not possible. It is this interactivity (made possible by the (near-)synchrony of the medium) that has Web Chat feature by far the highest frequency of interrogatives (see 11.2.2. below), which are in turn connected to non-clausal responses of declarative structure as shown in example (145). The Chatter <mary306girl> asks two questions, of which one lacks punctuation, but is in its syntactic structure definitely an interrogative (*where are u*), and the other one, based on the previous statement *canada's biggggggggg*, consists of question marks only (*???*). However, both interrogatives generate a non-clausal response (*canada* and *ontario*, respectively). This is most probably connected to the fact that the communicative context is clear, as well as to the pressure to type quickly. Example (146) on the other hand, is idiosyncratic to the personal homepage in that it is the only medium to feature headings and hyperlinks on a frequent basis—both of which turned out to have a tendency to be of a declarative, non-clausal structure. As illustrated in the above example, these non-clausal declaratives

are in most cases noun phrases. Both the heading *Political Resources* as well as the six hyperlinks from the personal homepage excerpt shown in example (146) classify as noun phrases.

Negated declarative clausal and non-clausal units were observed to be a lot rarer than affirmative structures in all text types, albeit with different frequencies. Of all declarative clausal and non-clausal units, percentages for negated structures are highest in EEC and e-mail (16.1% and 13.2%, respectively), followed by Web Chat (11.1%), and both SMS discourse and the personal homepages show lower frequencies (8.7% and 3.5%, respectively). It seems as if EEC and e-mail, which have shown similarities in their organisation of text and use of the tenses, are also similar in that authors are more likely to include negated statements as illustrated in examples (147) and (148):

(147) (…) Thanks for the call last night.. I wasn't expecting to hear from you … (…)
(E-mail excerpt / author: male, 22 yrs)

(148) (…) I am sorry that your busynes doth carye you another way that I shall not see you at Culford in your jorney towards Broome. (…)
(EEC excerpt, 1614 / letter collection Cornwall, author: female)

Examples (147) and (148) show another tendency in that most negations concern events in past time or future time. Negated declarative clausal or non-clausal units that refer to present time, on the other hand, turned out to be rare. Statements about present time are thus much more likely to be made by means of affirmative syntactic structures.

Although not included in Table 11.2, the subjunctive, if it is used, is also likely to be of declarative and affirmative structure. As has been pointed out in the theoretical discussion, the subjunctive occurs more frequently in EEC than in all other text types. If the subjunctive occurs in the modern text types, then it is most likely to be the past subjunctive (or "were-subjunctive") as illustrated in example (149):

(149) (…) I wish you were still here, we'd have such a giggle, you were a lot of fun to be around. (…)
(E-mail excerpt / author: female, age unknown)

All instances of the subjunctive found in the modern text corpora are comparable to the past subjunctive illustrated in example (149). Since only seven clausal units in subjunctive mood could be determined between the four modern text types, rendering thoughts in this mood is obviously not very popular among 21st century authors. In EEC, on the other hand,

a total of 64 clausal units were classified as subjunctives, of which 53 are affirmative and another seven negated declaratives (the remaining four subjunctives are of exclamative structure). In the majority of instances, these uses of the subjunctive are related to the religious figure God and are of declarative and affirmative structure, expressing the urgency of some matter. Also, they are very likely to occur towards the end of the letter, before the actual farewell section. Consider example (150) that illustrates such a prototypical use of the present subjunctive in EEC, the farewell section being indicated by square brackets:

(150) My sister is much better in health than when she came up. I pray God be with her. [I am your most affectinat brother, Edw. Harley.]
(EEC excerpt, 1661 / letter collection Harley, author: male)

Although the use of the subjunctive as illustrated in example (150) is more frequent in EEC than in any other text type, its occurrence, if compared to declaratives in indicative and hypothetical mood, is still rather low (one declarative subjunctive in every 575.5 words).

Thus, if these findings are brought into relation with function and illocutionary force, it appears that the authors of different types of personal written communication all tend to favour declarative clausal units in indicative mood and of affirmative structure. As both hypothetical mood and subjunctive mood have been shown to be rare, this implies that authorial intention is geared towards making statements that set forth particulars or facts. In most cases, these reports are concerned with present time (at writing time) and about issues to do with the author, readership and/or any other third party. Although the frequencies of declarative clausal and non-clausal units vary slightly across the five text types, they are nevertheless comparable (declarative-to-word ratios range between 1:9.7 and 1:13.0). It is a different picture with interrogatives which will be of interest next.

11.2.2. Interrogatives

According to Biber et al. (1999: 203), "interrogative clauses tend to occur in dialogue situations" and they come to the conclusion that out of the four main registers they investigated (conversation, fiction, news, and academic writing), "they are frequent only in conversation." It can thus be hypothesised that the more synchronous text types contain more interrogatives than the more asynchronous ones. As far as the structure of interrogative clauses is concerned, there are two main types of questions. The *wh*-question opens with a question word beginning with *wh* (such as *who*, *what*, *where* and so forth) or

how. This *wh*-word "indicates an element to be specified by the addressee" whereas "the rest is taken to be already known" (Biber et al. 1999: 204). The other type of question is the *yes/no*-question[78], which opens with the auxiliary (*be*, *have*, or *do*), and "all elements are taken to be already specified, and the addressee is expected to supply a truth value by answering *yes* or *no*" or other possible answers that indicate various degrees of certainty, such as *definitely*, *certainly*, *perhaps* and so forth (Biber et al. 1999: 206).

Example (151), taken from Web Chat, illustrates both the *wh*-question and the *yes/no*-question together with the answers these questions generated:

(151) <Bluesea>: what's your job <India>
(…)
<India28034>: office admin <Bluesea>
(…)
<Bluesea>: have you ever been in Europa
<India28034>: no <Bluesea>
(Chat excerpt / Room II, session 1)

Although both questions illustrated in example (151) lack punctuation (which is a prevalent feature in Web Chat, most probably connected to the ambition to economise on writing time), they are complete in their syntactic structures and thus identifiable as questions. The first question-answer exchange in example (151) concerns a typical *wh*-question with which one Chatter nicknamed <Bluesea> enquires after <India28034's> occupation (*what's your job <India>*), and the *wh*-word *what* indicates the element that needs specifying. The answer provided by <India28034> is *office admin*, and he/she thereby specifies the *wh*-word in <Bluesea's> question. The second question-answer exchange between those two Chatters is prototypical for a *yes/no*-question in that <Bluesea's> question *have you ever been in Europa* is responded to by <India28034> with *no*. On a side note, this is by no means always the case in Web Chat where many questions are left unanswered if they do not catch the interest of any of the other Chatters, even if they are repeated after a while.

The investigation into the frequencies of interrogative structures in personal written communication yielded results that confirm the above made assumption that the more synchronous text types are also more likely to feature interrogatives. The results are

[78] Next to the *wh*- and *yes/no*-question, Biber et al. (1999: 207) also mention the "alternative question" which is structurally similar to a yes/no-question (opening with the auxiliary verb which is followed by the subject). But "rather than expecting an answer in terms of yes or no it presents alternatives for the addressee to choose between," as in *Do you want one or two?*. However, this need not concern us any further as this type of question turned out to be non-existent in all of the five text corpora.

summarised in Table 11.3 below. Similar to Table 11.2 above, the results are presented in a threefold distinction: clauses vs. non-clauses, affirmative vs. negated structures, and indicative vs. hypothetical mood. Percentages are given in relation to the total number of occurrences of interrogative clauses and non-clauses in each of the text types.

Table 11.3: Interrogative clauses and non-clauses in personal written communication (corpora-based results).

Types of interrogatives		**SMS discourse** (18'426)		**E-mail** (25'733)		**Web Chat** (28'404)		**Personal HP** (16'030)		**EEC (letter)** (31'077)	
		IM	**HM**	**IM**	**HM**	**IM**	**IM**	**HM**	**IM**	**HM**	**IM**
AFF.	clauses	* 71.1	5.3	78.3	4.4	54.0	1.0	83.6	1.6	100	0
	non-clauses	21.1	0	13.9	0	42.5	0	8.3	0	0	0
NEG.	clauses	2.5	0	2.2	0	1.8	0	4.9	1.6	0	0
	non-clauses	0	0	1.2	0	0.7	0	0	0	0	0
TOTAL		** **1:57.9**		**1:136.9**		**1:31.1**		**1:262.8**		**1:3884.6**	

Key: AFF. = affirmative; NEG. = negated; IM = indicative mood; HM = hypothetical mood; HP = homepage; * = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of interrogative clauses and non-clauses in each of the text types; ** = interrogative-to-word ratio (including all types, rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

Web Chat not only features the highest number of interrogatives, it also features interrogative non-clauses on a more frequent basis than all other text types. According to Biber et al. (cf. 1999: 206), *yes/no*-questions in spoken conversation are frequently elliptic. It was observed that Web Chat shows similarities in this respect. There are, however, also frequent one-word-questions that can only be identified as questions because they are made up of *wh*-words and/or are accompanied by a question mark. Compare examples (152) – (154), where the suggested missing element(s) are supplemented in square brackets:

(152) *<machine-generated turn>: <peachesNcream^318> has joined Room III*
(…)
<peachesNcream^318>: [does] <u>any1 wana chat?</u>
(Chat excerpt / Room III, session 4)

(153) <MYNAMEIS>: how r u <Tinker>?
(...)
<Miss_TinkerBum>: im good
<Miss_TinkerBum>: [how are] <u>u?</u>
(Chat excerpt / Room VI, session 1)

(154) <kittos>: pokes <euqinu>
(…)
<euqinu>: why [do you poke me]?
(…)
<kittos>: <euqinu> cuz u said u are bored
(Chat excerpt / Room V, session 5)

Example (152) illustrates a *yes/no*-question where the auxiliary verb *do* is missing. This is frequent in both Web Chat (presumably connected to typing speed) and SMS discourse (presumably connected to the brevity of the texts). E-mail also features this type of non-clausal question occasionally, but it does not occur in the personal homepage corpus or in EEC.

However, examples (152) – (154) differ with regard to the "reconstruction potential" of the elliptic questions. Even without context it is easy to reconstruct the elliptical *any1 wanna chat?* in example (152) to *does any1 wanna chat?*. One can even make an educated guess in what sort of communicative context this question is most likely to be used. It is a different matter with the non-clausal interrogatives shown in examples (153) and (154), both illustrating elliptic *wh*-questions. In example (153), the suggested reconstruction of the non-clausal interrogative *u?* to the interrogative clause *how are u?* is only possible because the preceding turns are known. The same is true for example (154), where the exchange between the Chatters <kittos> and <euqinu> about "poking" helps to reconstruct a non-elliptic version of *why?*. Elliptic *wh*-questions of this kind only occur in the Web Chat corpus and rarely in SMS discourse. However, since the SMS text corpus is compiled of isolated messages, rather than ongoing discussions, it is most cases impossible to reconstruct a non-elliptic version with any certainty (cf. Biber 1999: 207).

Although not strictly independent clauses (and thus classified as non-clausal in this investigation) are question tags. The tag question (such as *isn't it* or *does she?*) is constructed of an auxiliary and a personal pronoun, of which the auxiliary is identical to the one of the clause to which it is appended (if there is no auxiliary in the clause the tag question is appended to, a form of do is inserted), and the personal pronoun is co-referent with the subject of the preceding clause (cf. Biber et al. 1999: 208). Although the question tag is according to Biber et al.'s (cf. 1999: 212) observations most frequent in conversation (analogue to the independent and elliptic questions discussed above), they are rare in the more synchronous text types Web Chat and SMS discourse. In fact, they are infrequent in all text types and only occur occasionally in e-mail correspondence.

As illustrated in Table 11.3, interrogatives in hypothetical mood have low frequencies in all corpora, SMS discourse being the only text type where the percentage of hypothetical interrogatives exceeds 5%. Interestingly, the hypothetical question is in many cases used to perform one of the main functions of SMS discourse: organising dates. Compare example (155):

(155) Hi my dear, can meet you guys at 2pm only, would that be ok or too late? (…)
(SMS text excerpt / author: female, 27 yrs)

Of course, as will be seen in the upcoming discussion of imperatives as well, the hypothetical mood often serves as a kind of hedging. In example (155), the preceding phrase *can meet you guys at 2pm only* sounds like this fact is set in stone. However, by adding a hypothetical question whether to meet at this particular time is convenient or not, it takes off the edge of the preceding statement.

Thus, overall it can be asserted that only the more synchronous text types, Web Chat and SMS discourse, include both clausal as well as non-clausal interrogative structures (in the majority of cases in indicative mood) on a frequent basis, which is reminiscent of spoken discourse structures. E-mail, personal homepages, and EEC, on the other hand, are less likely to feature interrogatives, and if it is the case, then clausal structures are more popular than non-clausal ones. With respect to illocutionary force, it emerged that the authorial intention behind questions in Web Chat and SMS discourse is predominantly aimed at eliciting information from the addressee, whereas authors of e-mail, the personal homepage, and EEC have a greater tendency to include either rhetorical questions or questions they answer themselves. Example (156) illustrates such a question, the corresponding answer is italicised.

(156) (…) Well I suppose the big news here is that as form 23/08/02 I have been made redundant. Yep I won't have a job. (…) What am I going to do? *God only knows! It will be difficult getting another job at my age, but we will see what happens.*
(E-mail excerpt, italics my emphasis / author: male, 50 yrs)

Although e-mail is quick in its transmission speed, the tendency to include questions (if at all) is low, and those questions are likely to concern matters that the author can respond to in his/her own right. This implies that compared to Web Chat and SMS discourse, e-mail is perceived as more of an asynchronous media with potentially longer time lags between two messages.

11.2.3. Imperatives

Formally, imperative clauses are characterised by the lack of the subject (although they may also contain one, see below), use of the base of the verb, as well as the absence of modals (cf. Biber et al. 1999: 219). As imperatives always contain a verb phrase (otherwise they could not be classified as imperatives), and may lack a subject (*do it*) or not (*don't you dare*), they are clausal by definition. It is for this reason that the investigation into imperatives will be focused on affirmative and negated structures. Also, because hypothetical imperatives are not possible, frequencies and percentages are only given for indicative mood. However, in addition to the formal imperative as described above, further attention will be paid to the notion of "softened imperative", that is all imperatives that are coupled with an expression that softens the sharpness of the command or request (such as *Please do me this favour* as opposed to *Do me this favour*.)

According to Biber et al. (1999: 219), "imperatives are typically used in contexts where the addressee is apparent; the subject is usually omitted but understood to refer to the addressee." Based on their corpus investigation, they come to the conclusion that "imperatives are many times more common in conversation than in writing" (Biber et al. 1999: 221). Thus, similar to the interrogative structures discussed above, it can be hypothesised that the more synchronous text types also tend to contain more imperatives. However, the investigation into the five text corpora produced results that only partly confirm this assumption. The most synchronous text type Web Chat features only the third highest frequency, both SMS discourse and the personal homepage show considerably higher frequencies than Web Chat. Authors of e-mail and EEC tend to include fewer imperatives in their messages. Table 11.4 summarises the distribution of imperatives and softened imperatives across the five text corpora. Percentages for imperatives and softened imperatives, affirmative or negated, are given in relation to the total number of occurrences of imperatives and softened imperatives found in each of the text types.

Table 11.4: Imperative and softened imperative clauses in personal written communication (corpora-based results).

Types of imperatives		**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
AFF.	imperative	* 64.6	69.6	84	82.4	47.4
	softened imp.	31.9	23.4	8.9	13.4	40.4
NEG.	imperative	2.3	4.7	4.9	2.8	4.4
	softened imp.	1.2	2.3	2.2	1.4	7.8
TOTAL		** **1:73.4**	**1:201.0**	**1:126.2**	**1:74.2**	**1:272.6**

Key: AFF. = affirmative; NEG. = negated; imp. = imperative; HP = homepage; * = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of imperative and softened imperative clauses in each of the text types; ** = imperative-to-word ratio (including all types, rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

As has been pointed out above, imperatives are more frequent in SMS discourse and the personal homepages. SMS discourse, however, shows a greater tendency to make use of the softened imperative than the personal homepage. Furthermore, there exists a special type of imperative formed with the verb *let* in combination with the first person plural pronoun *us* (usually contracted to *let's*) to express a suggestion involving both addressor and addressee (as in *let's do this*). As this type of imperative is felt to be less of a command and more of a motivation or "shared responsibility" between the correspondents, it was decided to classify it as a softened imperative. They contribute to the high frequency of softened imperatives in SMS discourse. Consider example (157) below:

(157) (…) meeting up next week's fine, maybe even sunday evening? let's phone! (…)
(SMS text excerpt / author: female, 23 yrs)

Again, similar to interrogatives, the softened imperative is recurrently used in connection with organising dates. The same is true for the formal imperative, affirmative and negated, as illustrated in examples (158) – (159):

(158) See ya later, bring meat! Hugs! <Nickname>
(SMS text / author: male, 33 yrs)

(159) Hi <nickname>, don't forget lunch with <nickname> tomorrow! (…)
(SMS text excerpt / author: female, 27 yrs)

In both examples (158) and (159), the authors make use of the imperative to point out something connected to an upcoming meeting. In example (158) the recipient is reminded

to bring along meat, whereas the negated imperative in example (159) is a reminder of the meeting itself.

E-mail is comparable in this respect to SMS discourse, as many of the formal affirmative imperatives are used to inform the addressee that he/she is required to get in touch about a certain issue. Opposed to SMS discourse, which often specifically refers to face-to-face meetings by means of imperatives, e-mail correspondence is not as explicit or definite as to how the addressee is supposed to establish contact next:

(160) I do hope you can come to the day away… <Name> would love to see you.
Let me know.
<Name>
(E-mail excerpt / author: female, 50 yrs)

In comparison to example (157) above, where the *let*-imperative is softened by the use of *us*, example (160) shows a formal imperative with *let* as the responsibility for action lies fully with the addressee (as opposed to being a shared responsibility between addressor and addressee). Just as how the addressee is expected to contact the person who sent the e-mail is not specified. However, people often have a strong tendency to answer within the same medium they are approached with, so it can be assumed that the addressee of example (160) reacted by sending an e-mail. If a change of medium is desired, then the author is most likely to communicate this in his/her message. This was most often observed in SMS discourse, in particular with issues of certain urgency, as illustrated in example (161) below.

(161) Don't touch word, I might have a way to find it if you haven't opened too many documents… Call me
(SMS text / author: male, 35 yrs)

As can be gathered from the content of the SMS message shown in example (161), the recipient is assumed to have experienced a loss of a word document. As this is perceived to be a delicate matter by the person who sends this message, he advises the addressee to call him and thereby change to a medium that is more synchronous than SMS discourse. Note as well the use of another formal imperative (negated) at the beginning of the message, which further emphasises the urgency of the matter (softened imperatives, as in *Please don't touch word*, are less urgent in their illocutionary force).

Participants in Web Chat, on the other hand, use imperatives predominantly to reflect on how other Chatters take part in the ongoing discussion. This is in most cases not very flattering, as illustrated in example (162):

(162) <FLAMING_GOAT!!>: <tom>, cool it with the language
(…)
<Tom>: i do wut i want
(…)
<Eliwood>: first tom
<Eliwood>: learn to spell
(Chat excerpt / Room VI, session 3)

The excerpt shown in example (162) is taken from a Chat session of quite offensive content. A Chatter nicknamed <Tom> makes contributions that contain a lot of swearwords and this offends other Chatters. When a Chatter named <FLAMING_GOAT!!> steps in and advises <Tom> to *cool it with the language*, <Tom> responds that he/she does as he/she pleases by stating *i do wut i want*. This comment, specifically its spelling, prompts <Eliwood> to contribute the formal imperative *(first <Tom>) learn to spell*. Needless to say, <Tom> continues to offend the other Chatters in the room until he/she is eventually kicked out and banned from the Chat room by an operator. Swearwords are also referred to as "expletives" and although they are not extremely numerous in Web Chat, they are more frequent in this text type compared to the other four. This issue will be re-addressed in the discussion of inserts below (11.2.5.). In any event, if a Chatter's behaviour, as expressed through his/her writing, is perceived as indecent by other Chatters, then they are likely to reprimand this Chatter by means of imperatives.

With respect to the personal homepage corpus, which features the second highest frequency of imperatives, predominantly of formal and affirmative structure, the imperative is used as a means to yet another different end. Authors of this medium tend to use the imperative as an instrument of navigation within the text, or to documents located outside the current page:

(163) Welcome to the <Full Name> Hompage
… use the bar below to navigate.
(Personal homepage excerpt / author: Nm_1)

(164) <URL of <Full Name> / external website> - check out his films, and be very glad you did.
(Personal homepage excerpt / author: Uf_1)

In example (163), the author instructs the reader with an imperative how to navigate through the page, whereas the author of example (164) advises the visitor (by means of two imperatives) to check out a certain person's films and by doing that in fact encourages her reader to leave the homepage. Another two types of formal imperatives that are frequent in personal homepages are *click here* and *send me (an) e-mail* (or the shortened versions: *e-mail me* and *mail me*). Although Loehr (2002: 29) states that "hypertext style guides prescribe against the use of *click here*" it is nevertheless numerously used. Presumably because of the simplicity of the term that unambiguously tells the reader what to do (*click*), and where to do it (*here*).

Imperative structures are least popular in EEC compared to the other text types. However, within the medium they are more frequent than interrogatives (see 11.2.2. above) and exclamatives (see 11.2.4. below). Authors of EEC show a slight preference for formal over softened imperatives and generally tend to favour affirmative structures. In contrast to SMS discourse in which imperatives are used for the organisation of dates, and the homepage where they are employed to navigate the reader, authors of EEC tend to employ imperatives to express emotional issues. Compare example (165):

(165) Sir, God hath taken away your eldest sonn by a cannon shott. (…) Hee is a glorious sainct in heaven, wherein you ought exceedingly to rejoice. Lett this drinke up your sorrowe. Seeinge theise are not fayned words to comfort you; but the thing is soe real and undoubted a truth, You may doe all thinges by the strength of Christ. Seeke that, and you shall easily beare your tryall. (…)
(EEC excerpt, 1644 / letter collection Charles, author: male)

The author of the letter excerpt shown example (165) has a sad task in reporting the death of the addressee's son in his letter. In trying to comfort his reader, the author also uses two imperatives in connection with this emotional issue. The first imperative, *Lett this drinke up your sorrowe*, refers back to the statement that the addressee's son is now in heaven. Another formal imperative is used in advising to seek comfort in the strength of Christ (*Seeke that*). In connection with the 18th-century letter, Fitzmaurice (2002: 87) observed that the familiar letter "was used extensively as a vehicle for giving people advice on all subjects, from their private life, their children's education, and their legal affairs to their health, their horticulture, and their dress." It turned out that the 17th-century letters in this corpus are more oriented towards narrating episodes or reporting sad news (as is the case in example (165) above). Although some of the letters contain pieces of advice, this is in most cases not the main thrive behind the epistolary interaction. It would thus be

interesting to compare a corpus of 18th-century letters with EEC from the 17th century and analyse the use of imperatives.

From the quantitative and qualitative investigation into the uses of formal and softened imperatives, it can be stated that they are in all text types mostly of affirmative structure, but used to achieve different communicative goals. While authors of SMS discourse and e-mail are likely to employ both formal as well as softened imperatives in connection with organising meetings or establishing contact, participants in Web Chat tend to reflect on other Chatters' behaviour, or contributions, by means of predominantly formal affirmative imperatives. In the personal homepages, which also shows preferences for the formal affirmative imperative, the majority of imperatives occur in connection with trying to navigate the reader through the page. Further, the number of imperatives in a particular page can be seen as an indicator of how interactive the page is with respect to reader involvement. EEC, on the other hand, shows tendencies towards using the imperative (both formal and softened) in relation to emotional issues, such as comforting the addressee. Thus, all five text types show tendencies of different authorial intentions behind using imperatives.

11.2.4. Exclamatives

Exclamations can be realised in a range of structures, both clausal and non-clausal (cf. Biber et al. 1999: 219). What unifies all exclamations is that they are followed (and identified) by an exclamation mark. Thus, both imperatives and interjections, which may or may not be followed by an exclamation mark, could also be classified as exclamations. So as to not bias the frequencies of the syntactic structures investigated in this study, it was decided to assign clausal and non-clausal units only one syntactic structure and therewith connected main function. This means that imperatives and interjections that are followed by exclamation marks are not included in the frequencies of exclamations.

Consequently, the structure of exclamatives is similar to clausal and non-clausal declaratives—albeit coupled with an exclamation mark as opposed to a period or comma (or no punctuation at all). According to Biber et al. (1999: 219), "exclamations occur chiefly in conversation" and it would thus be reasonable to hypothesise that the more synchronous media feature more exclamations. But, similar to the investigation into interrogatives, this is not the case. The results illustrated in Table 11.5 indicate that while SMS discourse has the highest frequency of exclamations, the more synchronous Web Chat is ranked fourth behind e-mail, and the considerably more asynchronous personal

homepage. Only EEC features fewer exclamations than Web Chat. The results from the investigation into exclamatives are presented, again, in a threefold distinction: clauses vs. non-clauses, affirmative vs. negated structures, and indicative vs. hypothetical mood. Percentages are given in relation to the total number of occurrences of exclamative clauses and non-clauses in each of the text types.

Table 11.5: Exclamative clauses and non-clauses in personal written communication (corpora-based results).

Types of exclamations		**SMS discourse** (18'426)		**E-mail** (25'733)		**Web Chat** (28'404)		**Personal HP** (16'030)		**EEC (letter)** (31'077)	
		IM	**HM**	**IM**	**HM**	**IM**	**IM**	**HM**	**IM**	**HM**	**HM**
AFF.	clauses	* 60.6	4.9	68.5	8.4	44.7	5.7	60.7	1.9	0	0
	non-clauses	30.0	0	16.3	0	39.0	0	34.6	0	100	0
NEG.	clauses	3.5	0.5	5.5	0	9.9	0	1.9	0	0	0
	non-clauses	0.5	0	1.3	0	0.7	0	0.9	0	0	0
TOTAL		** **1:42.1**		**1:156.0**		**1:201.4**		**1:149.8**		**1:10'359**	

Key: AFF. = affirmative; NEG. = negated; IM = indicative mood; HM = hypothetical mood; HP = homepage; * = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of exclamative clauses and non-clauses in each of the text types; ** = exclamation-to-word ratio (including all types, rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

As illustrated in Table 11.5, the exclamative is considerably more often non-clausal than any other structural type discussed so far. In EEC, even 100% of all exclamations are non-clausal in their structure. However, it needs to be taken into consideration that EEC has an extremely low exclamative-to-word ratio (1:10'359, or three exclamatives in the whole corpus). With regard to the other four text types, SMS discourse, as pointed out above, features the highest frequencies of exclamatives that are mainly both indicative and affirmative. While the clausal exclamatives (including clauses with first person pronoun ellipsis, cf. 11.1.3.) are, once more, predominantly related to upcoming meetings, many of the non-clausal exclamatives are primarily used to either give thanks or say sorry. Compare examples (166) and (167):

(166) 2.15 is fine for me. I'll be there! You bet :-)
(SMS text / author: male, 34 yrs)

(167) I'm happy & grateful you're part of my life & a big part,indeed,you're my best mate.thanks4everything!
(SMS text / author: female, 28 yrs)

While the author in example (166) refers to an upcoming face-to-face meeting and includes an exclamative guarantee of his physical presence therein (*I'll be there!*, further supported by *You bet*), the author of the SMS text shown in example (167) is giving thanks to the recipient by means of an exclamation (*thanks4everything!*).

In the personal homepage corpus, which features the second highest frequency of exclamations, the exclamation is not tied to specific contents, but serves primarily as a means to attract attention. Similar to SMS discourse, the personal homepages contain mainly affirmative exclamatives of both clausal as well as non-clausal structure. The majority of non-clausal exclamatives were observed to be noun phrases and, similar to declarative non-clausal units (see also example (146) above), they are predominantly incorporated into hyperlinks, as illustrated in example (168):

(168) Best freeware software!
My favourite Movies!
My Chess Page!
(Personal homepage excerpt / author: Of_1)

What is unusual about example (168) is that the hyperlinks are not fronted or headed by a title or any other introductory comment. In any event, with regard to exclamative noun phrases that are assumed to attract the reader's attention, example (168) is prototypical.

E-mail and Web Chat show comparable frequencies, although Web Chat is much more likely to make use of non-clausal exclamatives than e-mail. Interestingly, e-mail has a tendency to include exclamatives (both clausal and non-clausal) into the subject line. This is presumably done to attract attention to the message in the inbox of the recipient. Example (169) gives a selection of different exclamative subject lines found in the e-mail corpus:

(169)

Here I am with a reply!!!	(E-mail excerpt, subject line / author: female, 25 yrs)
Happy Birthday!!	(E-mail excerpt, subject line / author: male, 37 yrs)
Missed your birthday!!!!	(E-mail excerpt, subject line / author: female, 28 yrs)
I am coming to Europe!!	(E-mail excerpt, subject line / author: male, 30 yrs)

Example (169) illustrates another trend of the subject line in e-mail, namely that if punctuation is included, then it is likely to be of a repetitive nature. This is definitely aimed at attracting attention from the reader. With respect to exclamatives that are incorporated into the main body of the e-mail message, no particular trends could be determined.

The same is true of Web Chat which shows a tendency towards the employment of non-clausal exclamatives. This is assumed to be connected to typing speed where, similar to the declarative and the interrogative, syntactic structures are stripped off elements to the extent where some of the contributions can only be reconstructed if the previous textual units are available. The tendency to omit "superfluous" elements also includes punctuation. It can thus be hypothesised that quite a few "potential exclamatives" are not identifiable as such because the exclamation mark is missing. Since declarative structures without punctuation were classified as declaratives, this is assumed to contribute to declaratives being more frequent than exclamatives in Web Chat.

All in all, exclamatives, with the main function of drawing attention to the communicated content, emerged to be the least frequent of all the structural types investigated in this study. However, this is certainly connected to the fact that the frequencies given for exclamative structures exclude imperatives, inserts (see below), and all units of declarative structure that lack punctuation. As regards illocutionary force and authorial intention, it was observed that SMS discourse makes use of the exclamation in connection with the organisation of meetings and to say thanks or sorry. E-mail, on the other hand, features exclamations primarily in the subject line. It is strongly assumed that the authorial intention behind this is to attract attention to the e-mail in the receiver's inbox, which may be full with other e-mails. The personal homepages incorporate exclamatives predominantly into hyperlinks. Similar to e-mail, this is presumably aimed at attracting the reader's attention. With respect to EEC, however, only three non-clausal exclamatives could be determined. This may indicate that this medium is generally not felt to be the apt platform to communicate by means of exclamations.

11.2.5. Inserts

The last structural type that will be discussed is at the same time also the most diverse. As has been pointed out above, inserts are thought to be non-clausal material typical for the grammar of spoken language. However, five types of selected inserts turned out to be perfectly applicable to written language, these include: interjections, expletives, greetings and farewells, response forms and hesitators. As inserts are often single word-units, sometimes coupled with a modifier (cf. 11.1.3.), they are by definition non-clausal and can thus not be marked for grammatical mood. Furthermore, while some inserts can be locally negated, others cannot. Thus, frequencies given in relation to inserts do not consider distinctions related to mood and/or negation.

There are two preliminary remarks to the discussion of inserts that presuppose one another: first, inserts are (after declarative structures) the most frequently used structural type and second, this is mainly due to the frequency of greetings and farewells. Also, it should be noted that while greetings in personal written communication are comparable to greetings in face-to-face situations (as in *Hi Peter*), farewells may be comparable (as in *bye bye my dear*, where *my dear* would be the addressee), or they may differ from farewells in a spoken conversation (as is the case in *Love, Alain*, where *Alain* would be the addressor, but see hereto also 7.1.1.). Furthermore, although pre-closing sequences can also contain clausal units, particularly in EEC, it was observed that all farewells contain at least one non-clausal element, either a name or variation thereof (nickname, initials, and so forth), or some sort of prefabricated non-clausal farewell formula (such as*, bye bye*, *much love*, *yours*) or simply *x* for 'kiss'. It is for this reason that all greetings and farewells found in the text corpora qualify as inserts, both structurally and functionally. In any event, the distribution of the five different types of inserts are summarised in Table 11.6 below. Percentages are given in relation to the total number of occurrences of inserts in each of the text types.

Table 11.6: Inserts in personal written communication (corpora-based results).

Types of inserts	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (28'404)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
interjections (e.g., *wow!*)	* 5.7	13.3	32.5	20.9	0.5
expletives (e.g., *damn!*)	0.6	1.4	9.2	2.4	0
greetings (e.g., *hi darling*)	32.9	39.1	27.7	48.8	48.9
farewells (e.g., *bye for now*)	49.7	42.9	9.2	27.9	50.6
response forms (e.g., *ok*)	11.1	2.7	20.4	0	0
hesitators (e.g, *erm*, *uhm*)	< 0.1	0.6	1.0	0	0
TOTAL	** **1:17.7**	**1:87.5**	**1:16.2**	**1:372.8**	**1:174.6**

Key: HP = homepage; * = percentages (rounded to 1 decimal) are given in relation to the total number of occurrences of inserts in each of the text types; ** = insert-to-word ratio (including all types, rounded to 1 decimal) based on the total word counts (in brackets) for each of the text types.

Frequencies for the type of insert greetings/farewells are highest for all text types. Since greetings and farewells have been profoundly discussed with respect to their structural properties (see 6.3.3.), and also in connection with personal reference (see chapter 7), they will receive minimal attention in this section. The most frequent type of insert, after greetings/farewells, is the interjection. Interjections "have an exclamatory function" (they are, however, not always followed by an exclamation mark) and "are expressive of

the speaker's emotion" (Biber et al. 1999: 1083). Also, many interjections have onomatopoeic properties, such as *ouch!* (to express 'pain') or *urgh* (to convey 'disgust'). Others combine onomatopoeic properties with the functions of discourse markers, such as *oh*. It emerges that interjections can thus be of very heterogeneous form and function and "it is, indeed, to some extent an open-ended class" (Biber et al. 1999: 1085). What distinguishes them, however, from non-clausal exclamatives or non-clausal declaratives is their inability to enter into syntactic relations with other structures. Interjections are, so to say, separate non-clausal units.

As illustrated in Table 11.6, interjections are, compared to the other text types, most frequent in Web Chat where they are in the vast majority of the cases used to express positive feelings. By far most frequent is the imitation of laughter (as in *hahaha*, or *he he he*, see hereto also 6.1.2.), and the tendency to contribute such interjections as individual turns supports their status as separate non-clausal units. Contributing to the high frequency of interjections connected with laugher is the acronym *lol*. Although the acronym stands for a verb phrase (*laughing out loud* or *laugh out loud*), it is used and dispersed like an interjection. Also, in its function to express the emotion of the author, it classifies as an interjection from a functional perspective, too. It is for this reason that *lol* was assigned the status of interjection. Next to the imitation or expression of laughter, demonstrations of emotional involvement, albeit less frequently than laughter, were also observed:

(170) <divainross>: my lil 1 [little one] is screaming at her toy i think she thinks if she screams it will sing to her
<feline_hates_wierd_pms>: awwww <diva> how old babe? :)
(Chat excerpt / Room III, session 1)

In example (170), the Chatter nicknamed <divainross> makes a contribution that prompts the Chatter <feline_hates_wierd_pms> to start her turn with the onomatopoeic interjection *awwww*, expressing that he/she is touched by <divaninross'> comment. Interjections in connection with laughter and emotional involvement, such as sympathy, are not frequent in any of the other text types and are thus idiosyncratic to Web Chat.

Interjections are in general not very numerous in the other text types, but if they are included, then it was found that SMS discourse has a preference for expressing that something is pleasantly impressive (as in *wow!*) or, quite the opposite, is displeasing (such as *eek* or *yuck)*. Other than that no tendencies could be determined for SMS discourse.

With respect to interjections in e-mail, it was observed that they are predominantly associated with positive emotions. Among the more frequent interjections in e-mail are *yippee* and *yeebiee* that are both used to express joy. In the personal homepage corpus, on the other hand, no such preferences could be confirmed in the small number of nine interjections. The same is true for EEC with only two occurrences of non-clausal units that were classified as interjections (*alas!* and *ah*, for the latter see also example (22) in chapter 6).

Expletives, a term used for taboo expressions (swearwords), are related to interjections in that they are frequently coupled with exclamation marks. They are, however, distinctive from interjections in terms of meaning and pragmatic function. According to Biber et al. (1999: 1094), "expletives can be usefully subdivided into taboo expletives, which make reference to one of the taboo domains of religion, sex, or bodily excretion, and moderated (or euphemistic) expletives, which camouflage their taboo origin." Expletives are infrequent in all text types except Web Chat, where they make up just below 10% of all inserts. Interestingly, text types that identify their authors (by means of telephone number (SMS discourse), name and/or e-mail address in e-mail and the personal homepage) are likely to favour moderated expletives over taboo expletives. If they do contain taboo expletives, then they tend to be toned down by means of one or more "censor asterisks", as illustrated in example (172) below. In anonymous Web Chat, on the other hand, taboo expletives are by far more numerous than moderated ones. Consider examples (171) – (173), of which (173) is taken from the same Chat session as example (162) above (which was discussed in connection with imperatives, prompted by <Tom's> vulgarisms):

(171) oh shoot, but no worries. (…)
(SMS text excerpt / author: male, 27 yrs)

(172) Match ticket = f*ck all… (…)
(E-mail excerpt / author: male, 30)

(173) <Tom>: fuck authority
(…)
<funfunfun>: how about……. FUCK YOU
<Matthiasangel>: mother fucker
(Chat excerpt, emphasis original / Room VI, session 3, emphasis original)

Both the moderated expletive *oh shoot* (a euphemism for *oh shit*) and the censored expletive *f*ck* are softened in their illocutionary force. Another variant of moderated

expletive was found in the personal homepage, *for goodness sake!* (*goodness* being a euphemism for *God*), it is however, the only expletive in the whole homepage corpus.

In Web Chat the picture is very different. Although example (173) is taken from a particularly offensive Chat session, taboo expletives occur fairly frequent in Web Chat. They are usually connected to one Chatter who initiates swearing and thereby instigates others to join in. In the Chat session of which example (173) is an excerpt from, the discussion was free of swearwords until <Tom> logs in and starts swearing. All of a sudden, everybody is swearing back at <Tom> until he/she is banned from the Chat room and everybody returns to discussing issues without using expletives. This was observed in several Chat rooms and perceived to be an interesting phenomenon. I termed it the "scapegoat phenomenon". Most public Chat rooms have quite strict regulations about language policies, and it seems that participants who otherwise obey those rules, are quick to react to other Chatters who swear by swearing back themselves—banking on the fact that it was not them who started the swearing, but someone else (the "scapegoat"), and are thus less likely to be kicked out of the Chat room. However, besides the use of expletives, I agree with Bays (1998: OD) that participants also observe forms of politeness and "try to avoid conflict" in order to keep the discussion going. Bays (1998: OD) also points out that the Chat community "is a place for situational interaction, essentially, a social environment where the interactants consciously choose to be." And although swearing and insulting are part of human interaction, it is definitely not the main drive behind communicative exchanges, may it be among friends or strangers, as they tend to be geared towards saving each other's face rather than threatening it.

Another type of inserts is interesting from an interactive point of view, namely the response forms. They are inserts "used as brief responses to a previous remark by a different speaker" (Biber et al. 1999: 1089). These include affirmative responses to questions, such as *yes* or negating ones, such as *no* (as will be seen below, variant forms of *yes* and *no* frequently occur as well). Applied to written correspondence, this means that response forms refer back to a question or statement made in a previous message by another person. It is thus reasonable to hypothesise that they are more frequent in the more synchronous media, which is in fact the case. Of all inserts, Web Chat features over 20% response forms, SMS discourse just above 11% and e-mail just below 3%. Both the personal homepage corpus and EEC do not feature response forms at all.

In all three text types that contain response forms, affirmative response forms outweigh negating ones by far. The most popular affirmative response form is *yes* along

with its variants forms (as found in the corpora) *yeah*, *yep*, *yay* and *yesh* (reminiscent of a lisped *yes*). Other affirmative response forms to questions, as well as statements, include *okay* (with its derived forms *okey*, *oki-doki*, *okey dokey*, and the abbreviations *ok* and *kk*), as well as *sure*, *good*, *easy*, *fine* and *cool*. Response forms in written correspondence refer back to a previous message, this means that they have a strong tendency to front the message (or turn in Web Chat). Consider examples (174) – (176):

(174) ok i won't be long
(SMS text / author: male, 27 yrs)

(175) yeah please come!
(E-mail excerpt, subject line / author: female, 29)

(176) <**MooDY**>: my mums got cows
(…)
<**MooDY**>: lil calfs to
(…)
<**MooDY**>: and horses
(…)
<FLAMING_GOAT!!>: cool , <moody>
(Chat excerpt / Room VI, session 3)

Similar to truncated interrogatives, response forms can only be allocated in Web Chat because the preceding contribution, to which the response form refers to, can be identified, as shown in example (176) with *cool*. In both SMS discourse and e-mail this is not possible, as both corpora consist of isolated messages. The corresponding question or statement of the two response forms shown above, *ok* in example (174) and *yeah* in example (175), can thus not be ascertained.

The last type of inserts to be discussed are hesitators, filled pauses that "are occupied not by silence, but by a vowel sound" (Biber et al. 1999: 1053) and such hesitators are usually transcribed as *uh* and *um* (AmE) or *er* and *erm* (BrE). In spoken conversation, hesitators are pause fillers "whose main function is to enable the speaker to hesitate, i.e. to pause in the middle of a message, while signalling the wish to continue speaking" (Biber et al. 1999: 1092). As written correspondence is generally composed without the addressee in proximity, pause fillers are rare in Web Chat and e-mail, extremely rare in SMS discourse and the personal homepage corpus (only one hesitator, *umm*, was found, illustrated in example (21) in chapter 6), and in EEC they are non-existent. Even (near-)synchronous Web Chat features few instances of hesitators, probably connected to the wish of the Chatters to communicate rather than hesitate. It was also

observed that in the three text types that feature hesitators, they are used as a kind of hedging for the subsequent comment. Consider the fronted *um* in example (177) below:

(177) Um no I have to work on my presentation for tomorrow still
(SMS text / author: female, 30 yrs)

If the message is read without the hesitator, the sound is considerably sharper. However, these sound-oriented onomatopoeic expressions in the function of hesitators turned out to be infrequent in all text types. It can thus be asserted that they are in fact not a typical feature of personal written communication.

11.3. Chapter summary

Overall, it was found that all text types most frequently feature declarative structures that are affirmative and in indicative mood. In fact, the combination affirmative/indicative always overrules other possibilities regarding grammatical mood (hypothetical or rarely subjunctive mood) and negation. This means that authorial intention is geared towards communicating affirmative content that can be taken at face-value (rather than being hypothetical). However, while frequencies for declarative structures are more or less evenly distributed across the text corpora, other syntactic structures show differences in their distribution. Interrogative structures that are reminiscent of spoken discourse structures are predominantly found in the modern text types, which confirms H1 (albeit only partly, as questions in e-mail and the personal homepages are not frequent enough to be classified as a typical feature) and H2, claiming SMS discourse to be rich in contextual features typical of orality.

Furthermore, the investigation into imperatives (formal and softened) indicates that their use is generally connected to the communicative goal and not to the immediacy of the media, or whether the readership is known or not. Imperative structures tend to be slightly more frequent in SMS discourse and the personal homepages, which supports the assumption that their distribution is not connected to degree of synchrony of the text types. Thus in connection with imperatives, H1 – H5, all hypothesising idiosyncrasies due to differences in immediacy or type of readership, could not be confirmed. With respect to exclamations, it can be asserted that they are the least frequent of the syntactic structures investigated. However, as they are reminiscent of spoken discourse and considerably more frequently used in SMS discourse than all other text types, H2 could be confirmed.

Concerning inserts, a concept borrowed from the grammar formulated for spoken language, it was observed that out of all the syntactic structures investigated, they are tied the most to the degree of synchrony of the communicative exchanges. It was found that Web Chat features the highest frequency of inserts, followed by SMS discourse and e-mail. One type of insert, the response form, is highly reminiscent of spoken discourse as it refers back to preceding discourse units. They could be determined in the modern text types only and are particularly frequent in Web Chat. Thus, H1 (once more only partly because of the lower frequency of inserts in the personal homepages) and H2 could be verified. Also, EEC is distinctive in that it does not feature any inserts apart from a few interjections and greetings/farewells, which supports H3 (claiming EEC to be rich in features typical for literacy) and H4 (hypothesising differences between the modern text types and the older form of correspondence). However, overall, H5 could not be verified as it was observed that grammatical mood and syntactic structures in connection with authorial intention show tendencies that can be tied to the different communicative settings, but not to the type of readership.

12. Contextual effects in personal written communication

Several means of contextualisation that are available to authors of personal written communication, and how they are employed, have been discussed so far. More precisely, it has been analysed how media-related features and constraints influence the contextuality of personal messaging (chapter 6), or how authors deal with different kinds of readers in terms of address (chapter 7). It has also been of interest how the text itself is used as a means of contextualisation concerning textual networking (chapter 8), and how the notion of space is dealt with in text (chapter 9). Furthermore, it has been investigated in how far authors of personal written communication attempt to contextualise their writing by means of embedding it in a timeframe (chapter 10). And while the previous chapter has looked into authorial intention as expressed through grammatical mood and syntactic structures (chapter 11), this last chapter of Part III is aimed at another aspect of authorial intention and therewith connected questions of relevance: how the different elements of a text as contributed by the author may yield contextual effects in the comprehension process of the reader, which in turn modify the context established.

As has been pointed out in the theoretical part, the notion of contextual effects is taken from the larger framework of Relevance Theory (RT) by Sperber and Wilson (1995), an inferential approach to pragmatics. It has also been addressed that the concept of contextual effects was adapted considerably for the purposes of this study. Before re-addressing those adaptations in more detail, and discussing the findings that the empirical investigation into contextual effects (based on these adaptations) generated, I would first like to draw attention to some key terminology that is important in connection with the discussion and evaluation of textual units in terms of contextual effects.

12.1. Terminology

12.1.1. Explicatures and implicatures

Definitions of explicature (explicitly communicated assumptions against a certain communicative context) and implicature (implicitly communicated assumptions against a certain communicative context) are helpful concepts in connection with both the production of information (by the speaker/writer) as well as the comprehension process of information (by the hearer/reader). However, the definitions brought forward for these concepts vary somewhat. For reasons of consistency I will work with Sperber and Wilson's

(1995) interpretation of the two concepts, which differs from Grice's (1989) insofar as they assume a classificatory concept of explicitness.[79] According to Sperber and Wilson (1995), an explicature is a combination of linguistically encoded and contextually inferred conceptual features. The smaller the relative contribution of the contextual features, the more explicit the explicature will be: "explicitness, so understood, is both classificatory and comparative: a communicated assumption is either an explicature or an implicature, but an explicature is explicit to a greater or lesser extent" (Sperber & Wilson 1995: 182). Thus, although uttered in context (and to a certain extent drawing on this context), explicatures can in general be decoded by the hearer (or reader) based on the linguistic evidence provided, while implicatures feed from the context (and encyclopaedic knowledge) to a greater extent, and successful decoding (comprehension) is not possible from linguistic evidence alone.

In their more recent contribution to the *Handbook of Pragmatics*, Wilson and Sperber (2006: 613) formulate the comprehension process, in terms of relevance, as follows: "the hearer [or reader] should take the decoded linguistic meaning; following a path of least effort, he [or she] should enrich it at the explicit level and complement it at the implicit level until the resulting interpretation meets his [or her] interpretation of relevance:"

> **Relevance-theoretic comprehension procedure**
> a. Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguations, reference resolutions implicatures, etc.) in order of accessibility.
> b. Stop when your expectations of relevance are satisfied (or abandoned).

Hence, a speaker/writer is expected to make his/her utterance as easy as possible to understand for the hearer/reader, so that it is feasible for the latter to follow a path of least effort in decoding the utterance—in other words, the speaker/writer should formulate his/her utterance so that the first interpretation that satisfies the hearer/reader's expectation of relevance is in fact the one he/she intended to convey (cf. Wilson & Sperber 2006: 613, 614). Further, in RT, "the identification of explicit content [explicatures] is seen as equally inferential, and equally guided by the Communicative Principle of Relevance, as the recovery of implicatures" (Wilson & Sperber 2006: 615). This overall task, according

[79] It cannot be the purpose of this study to give a comprehensive account of explicatures and implicatures. Grice (1989), Carston (1988, 1998), Levinson (2000), Sperber and Wilson (1995, 2004), and Wilson and Sperber (1993, 2006), offer insightful reflection on the communication of linguistic content by means of explicatures and implicatures.

to Wilson and Sperber (2006: 615, emphasis original), "can be broken down into a number of subtasks:"

> **Subtasks[80] in the overall comprehension process**
> a. Constructing an appropriate hypothesis about explicit content (EXPLICATURES) via decoding, disambiguation, reference resolution, and other pragmatic enrichment processes.
> b. Constructing an appropriate hypothesis about the intended contextual assumptions (IMPLICATED PREMISES).
> c. Constructing an appropriate hypothesis about the intended contextual implications (IMPLICATED CONCLUSIONS).

The notion of "explicature", "implicated premise" and "implicated conclusion" are interrelated concepts in the overall comprehension process of a given utterance. For example, if a speaker utters *It is cold in here* in a living room with an open window, then the hearer may construct the hypothesis that the explicitly communicated content (i.e. the explicature) does not satisfy his/her expectations of relevance. The hearer may then continue to construct the hypothesis that the intended contextual assumption (i.e. the implicated premise) points to the circumstance that the speaker not only perceives the room temperature to be cold, but that he/she is freezing because of the low room temperature. Based on this implicated premise, the hearer may then construct the hypothesis that the intended contextual implication (i.e. implicated conclusion) is in fact the speaker's desire—provided that the outside temperature is lower than the room temperature—to have the window closed. It follows that the implicature of the utterance *It is cold in here*, formulated in a communicative context as outlined above, is something along the lines of 'the temperature of this room would rise if the window was closed', in other words, 'close the window'.[81]

Most important to this study, the inferring of explicatures and implicatures is connected to the concept of contextual effects (as simplified for the purposes of this study). In relevance-theoretic terms, utterances of explicit or implicit content are decoded by the hearer/reader against a certain communicative background (see above). The cognitive processing of these utterances then results in contextual effects that interrelate with the

[80] Wilson and Sperber (2006: 615) point out that these subtasks should not be seen as sequentially ordered, because "comprehension is an online-process, and hypotheses about explicatures, implicated premises, and implicated conclusions are developed in parallel against a background of expectations."

[81] Notably, the utterance *It is cold in here* may not carry an implicature (and hence be explicit in its content) in a different communicative context. Consider, for example, the same utterance uttered as a factual statement by a speaker who is standing in a cold storage room.

context so far established in the comprehension process of the hearer/reader. As has been pointed out in chapter 5 (*Methodology*), several adaptations to the concept of contextual effects were necessary for the empirically oriented purposes of this study. The nature of these adaptations will be addressed shortly (see 12.2. below). First, I would like to draw attention to the notion of "formulaic sequence" as brought forward by Wray and Perkins (2000). The formulaic sequence emerged to be a helpful concept in connection with the classification of contextual effects.

12.1.2. Formulaic sequence

The evaluation of a textual unit and the contextual effect it may perform is connected to how a particular textual unit interacts with prior textual units in a given communicative context. Thus, in a spoken discourse that unfolds linearly, new information is added to a context of old information by the interlocutors. Applied to written correspondence, where the addressee is not in physical proximity, the result is that each textual unit of a particular message can be seen as a piece of new information that is appended to a stock of previous textual units. However, this would in theory have the consequence that the first unit of a given message would be difficult to classify in terms of contextual effects—because of a lack of linguistic context as the prior textual units are not available. Yet, a conflict arises with regard to interpersonal written correspondence that starts with a conventional greeting section. A textual unit of the type greeting section does not, owing to its traditional form and function, require a lot of processing effort on behalf of the reader. Also, a greeting section is typically explicit in its content and thus not highly dependent on context in terms of decoding—hence, it can be deduced from the textual evidence provided. It is, in other words, a typical effect of the type contextual strengthening in the comprehension process of the addressee. This conflict is abated by the notion of formulaic language as brought forward by Wray and Perkins (Wray & Perkins 2000, Wray 2002).

According to Wray and Perkins (2000: 1), the notion of a "formulaic sequence"[82] describes a phenomenon that "encompasses various types of word string which appear to be stored and retrieved whole from memory," rather than being subject to generation or analysis by language grammar at the moment they are uttered. From the larger framework

[82] There are various terminologies used in literature to describe formulaic sequences and formulaicity in language, amongst others, "amalgams", "prefabricated routines", "ready-made utterances", "idioms", "fixed expressions", "fossilised forms" and so forth (cf. Wray & Perkins 2000: 3). Similar to the discussion of DMs above, this linguistic phenomenon is difficult to classify due to the heterogeneity of its class members. There is also no general agreement on terminology.

of formulaic language, there is one aspect that is of particular importance to this study: the common purpose of formulaic language. Wray and Perkins (2000: 15) assign formulaic language several purposes, of which the purpose of formulaicity "as a short-cut in processing" is most interesting for the investigation into contextual effects. According to Wray and Perkins (2000: 18), many of the formulaic sequences that interlocutors use have a socio-interactional function and "just as the processing short-cuts are a means of ensuring that the speaker [or writer] achieves successful production, so the socio-interactional formulae are a means of ensuring that the hearer [or reader] achieves successful comprehension."

With respect to social interaction, Wray and Perkins (cf. 2000: 15) outline three main areas where formulaic sequences may be used as devices of social interaction: social manipulation of others, asserting separate identity, and asserting group identity. The function of "asserting group identity" is most useful to this study. Personal written communication is generally less concerned with manipulation and the assertion of separate identity, but is more aimed towards functioning as connector between author and readership (and is in this sense an adapted version of "asserting group identity"). The types of formulaic sequences subsumed under the function of asserting group identity are, amongst others, "forms of address", "hedges" and "institutionalised forms of words" such as *happy birthday* (an example for an institutionalised form of words) and the address form *dearly beloved* (Wray & Perkins 2000: 14). Conventional greeting sections (and farewells) qualify as forms of address and can thus be classified as formulaic sequences. Hence, a greeting section that opens a message and is, because of the discourse structure, not related to prior text in the same message, can under these circumstances still be classified as a textual unit processed as an effect of the type contextual strengthening. This is not only connected to the function of the greeting, but also to savings in processing effort.

12.2. Evaluation and classification of textual units in terms of contextual effects

12.2.1. Contextual effects: adaptations to the concept as formulated by Sperber and Wilson (1995) for the purposes of this study

As pointed out in chapter 4 (*Theoretical background*), Sperber and Wilson first formulated RT, of which the concept of contextual effects is a component, in 1986, and both Sperber and Wilson themselves as well as other scholars (see hereto section 4.4.) have further commented on the theory since it was first published. And while there have been more recent publications in the field of RT than the revised edition of Sperber and Wilson's

Relevance published in 1995 (for example, Wilson & Sperber's (2006) contribution to the *Handbook of Pragmatics*, see explicature/implicature discussion above), in none of these later publications is the concept of contextual effects as elaborately discussed as in the 1995 edition of *Relevance*, including additional postface and notes. It is for this reason that the upcoming discussion on the adaptations to the concept of contextual effects will take Sperber and Wilson's 1995 edition of *Relevance* as point of reference.

The analysis of personal written communication in terms of contextual effects, if, as is the case in this study, the data was made anonymous prior to investigation, faces one major constraint: neither author nor readership can be consulted to find out more about the authorial intention behind the encoding of the textual units of a message, and the decoding processes of those units on behalf of the reader. This means that the classification of textual units in terms of contextual effects is a) based on the textual evidence provided in a given communicative context and b) connected to the comprehension process of the researcher(s). While point b) might potentially lead to biased results if coded by one person only, this bias was averted by having large samples of each of the five text corpora test-coded by three independent coders until they reached above 70% agreement between them (see 5.2.1.). This means that although the coders acted as a kind of stand-in for the originally intended readers, it was still possible to be in agreement as to how personal written communication can be evaluated in terms of contextual effects.

This, however, deviates from how Sperber and Wilson (1995) discuss the evaluation of contextual effects as they not only analyse contextual effects in terms of speaker/hearer interaction (i.e. spoken interchanges as opposed to written communication), but also take the production cost on behalf of the speaker as well as the processing cost on behalf of the hearer into account. While Sperber and Wilson (1995: 116) illustrate their reasoning with examples from spoken exchanges, I believe that analogue to a speaker/hearer interaction, writer and reader also aim at maximising relevance:

> In verbal communication, the hearer [or reader] is generally led to accept an assumption as true or probably true on the basis of a guarantee given by the speaker [or writer]. Part of the hearer's [or reader's] task is to find out which assumptions the speaker [or writer] is guaranteeing as true. Our hypothesis is that the hearer [or reader] is guided by the principle of relevance in carrying out this task. He [or she] expects the information the speaker [or writer] intended to convey, when processed in the context the speaker [or writer] expected it to be contextualised in, to be relevant: that is, to have a substantial contextual effect, at a low processing cost.

Hence, both speaker/writer and hearer/reader are geared towards the maximisation of relevance in their communicative interactions. And while "an assumption is relevant in a context if and only if it has some contextual effect in that context" (Sperber & Wilson 1995: 122), the concept of relevance also involves "some form of const-benefit analysis" (Sperber & Wilson 1995: 123). In other words, "the contextual effects of an assumption in a given context are not the only factor to be taken into account in assessing its degree of relevance" (Sperber & Wilson 1995: 124)—the effort that goes into production and processing of an assumption are also crucial factors to be considered in assessing relevance. Both production cost and processing effort are negative factors in assessing relevance: "other things being equal, the higher the production cost, the lower the productivity" and "other things being equal, the greater the processing effort, the lower the relevance" (Sperber & Wilson 1995: 124). And this is where the framework underwent another adaptation, or rather simplification, for the purposes of this study. For reasons of confidentiality, neither author nor readership could be consulted and consequently, no analysis into the production cost and processing effort on behalf of author and originally intended readership, respectively, could be undertaken.

However, as pointed out above, the evaluation of contextual effects is a combination of how new information in the form of a textual unit (i.e. a new assumption) interrelates with old information in the form of previous textual units (i.e. old assumptions) in a given communicative context—and how these are processed by the reader. For this reason it is essential that the processing of a given textual unit, and more importantly, the result of this comprehension process in terms of contextual effects, be taken into account. This is the reason why it was decided to resort to the comprehension process of the test-coders and the researcher as a substitute for the originally intended readership in order to be able to analyse the five text corpora in terms of contextual effects—albeit without classification whether the processing effort was deemed high or low. Furthermore, it was also decided to disregard the production cost on behalf of the author. This means, basically, that while this study looks into what kinds of contextual effects can be traced in personal written communication, it ignores the "cost-benefit factor" which would be required for assessing the degree of relevance of these contextual effects. In other words, this study is less concerned with how relevant the individual contextual effects are, and more concerned with the different kinds of contextual effects that occur in the comprehension process of the reader. Speaking of which, four of the five types of contextual effects (contextual implication, strengthening, contradiction/elimination, and

weakening) investigated in this analysis are based on Sperber and Wilson's theory (1995, including postface), whereas the fifth type of contextual effect (non-contextual effect) was formulated specifically for the purposes of this study (see 4.4. for a more detailed discussion).

12.2.2. Contextual effects: evaluation/classification procedure

Adding new information to a discussion is crucial in keeping a discourse going: without new information the communicative interaction will come to a halt. Obviously, not just any new information will do: the contributions have to be relevant with regard to the communicative context established. Relevance, as we have seen, is highly context-dependent. The question is: what qualifies as new information and how can it be classified in terms of contextual effects? For example, the analysis in the last chapter showed that the notions of clausal (independent and dependent) and non-clausal material is useful in order to capture grammatical mood and distinguish between different syntactic structures. With respect to the notion of contextual effects, however, these research entities are too rigid. I will therefore return to the textual unit ('a string of words that form a coherent chunk of language as produced by a human being in written fashion', see also 4.3.1.) as research entity. In addition, symbols (such as emoticons or "x" for 'kiss') can, depending on their placement and function, either support a textual unit and be processed as a combined unit in the sense of "textual unit + symbol = assumption", or they can be assigned their own individual status as a piece of new information in the sense of "symbol = assumption". Since discourse unfolds linearly (this is true even of the potentially non-linear personal homepage)[83], new information can be seen in terms of such textual units (or symbols) that are appended within a particular message to the previous textual units or, in the case of Web Chat, are contributed as the discourse proceeds (analogue to a spoken interaction).

Thus, textual units (and/or symbols) in written discourse can be evaluated according to the following parameters[84]: (a) in how far a textual unit relates to prior textual evidence of the discourse, (b) in how far a textual unit introduces information that is new in relation to the prior textual evidence of the discourse (i.e. retrievable old information),

[83] As has been discussed previously (see 6.1.1.), the text arrangement of the personal homepages follows linear structures. It was shown that the non-linearity of the text type personal homepage is mainly a consequence of the (potential) reading process (rather than the text structure itself), which ceases to be linear when the reader, for example, decides to follow up on one of the hyperlinks in the page.

[84] See hereto also Figure 5.2, illustrating the coding procedure of contextual effects.

and (c) the way in which such a textual unit is deducible in relation to the prior textual evidence of the discourse and the larger (out-of-text) context. Therefore, a textual unit that (a) relates to the prior textual evidence of the discourse (i.e. relates to retrievable information in the text) classifies as new information processed as a contextual effect, whereas a textual unit that does not relate in any way to the prior textual evidence (i.e. does not relate to retrievable information in the text) classifies as new information that is not processable, hence a non-contextual effect. If (b_1) a textual unit relates to the textual evidence of the discourse but cannot be deduced by the content of the contribution alone, then this textual unit classifies as new information processed as a (c_1) contextual implication as the conclusion is dependent on the input (available) as well as the larger out-of-text context (not available). Consequently, a textual unit that (b_2) relates to the textual evidence of the discourse, and is deducible from the input alone, classifies as new information processed as a (c_2) contextual strengthening, contradiction/elimination or weakening, depending on how new and old information interrelate. However, there is one exception to this procedure. If a textual unit is assigned the status of a formulaic sequence (cf. 12.1.2.), then this textual unit classifies as new information processed as contextual strengthening even if it opens a message (and can thus not be related to prior textual units in the same message), because processing effort on the reader's behalf is minimal. As noted above, symbols (such as emoticons) can be classified, depending on their placement and function in the message, as either resulting in an individual contextual effect in the comprehension process, or supporting a textual unit and its contextual effect.

12.2.3. Illustration

I will in the following sections discuss each of the effects separately. However, it is necessary at this point to illustrate the classification of contextual effects, based on the discussion above, with empirical examples. In the e-mail message shown in example (178), the different textual units are marked with square brackets and each unit is labelled with the respective contextual effect it performed in the comprehension process of the researcher (contextual implication (CI), contextual strengthening (CS), contextual weakening (CW)):

(178) Subject: RE: [Dartmoor][CI]

[no greeting section]

[No jets, no people.][CI] [Didn't see another living soul 4 days apart from the odd sheep.][CS] [Couldn't see a single trace of human civilisation either,][CS] [no pylons no distant villages or cars][CS] – [nothing.][CS] [The silence at night was deafening too][CS] – [you know that weird ringing in your ears when there is not a single other sound on which to focus.][CS] [Truly wonderful.][CS]

[Gosh how poetic!][CW]

[no farwell section]

(E-mail / author: female, 32)

With regard to the larger discourse structure, one can assume that the e-mail shown in example (178) is a reply, based on the "RE" (short for 'reply') in the subject line, to a previous message. It is for this reason that the subject line *Dartmoor* is classified as relating to prior text, but the previous textual evidence is not available. Thus the textual unit *Dartmoor* is only deducible from input and context together[85], hence, it results in a contextual implication.

The message lacks a greeting section (which would have been classified as contextual strengthening in terms of processing effort) and starts with the negated declarative non clausal textual unit *No jets, no people*. Since it is not clear how this textual unit is related to the subject line, and it is also unclear as to what this unit refers to, it is not possible to process this textual unit from textual evidence alone, it is hence classified as a contextual implication. The remaining units of the paragraph refer back to this first unit of the main body in a strengthening fashion. Furthermore, each of them can be deduced by the textual evidence provided. Thus, they are all classified as contextual strengthening effects as their processing is possible from the textual evidence provided and results in supporting (strengthening) the first assumption *No jets, no people*. However, the message does not close with a conventional farewell (which would also have been classified as contextual strengthening in terms of processing effort), but ends with the exclamation *Gosh how poetic!*. This weakens the previously made assumptions, because it is understood to be a slightly sarcastic comment, or an expression of surprise at how poetical the message turned out to be. In any event, it is believed to indicate that the message is

[85] On a side note, *Dartmoor* is a National Park in Devon, England.

maybe more poetical than it was intended. Therefore, *Gosh how poetic!* is processed and classified as a contextual weakening.

There are two more effects that could not be illustrated, because they do not occur in the decoding of the above example (178), and they will thus be illustrated separately. Example (179) below illustrates the processing of a textual unit that results in a contextual contradiction/elimination (CC/E). Again, textual units are identified by means of brackets and labelled with the respective contextual effect they performed in the comprehension process of the researcher:

(179) <tracyt>: [he said I was a slag before]^{CI}
(…)
<tracyt: [am soooooooo not]$^{CC/E}$
(Chat excerpt / Room V, session 3)

In example (179), a Chatter nicknamed <tracyt> makes a backward reference to something another Chatter previously wrote. As it could not be determined to what kind of textual unit <tracyt> is referring to (it must have been made before recording of the Chat session started), the contribution cannot be successfully deduced by means of the textual input alone. Hence, the comprehension process involves reasoning beyond the textual evidence as to what exactly <tracyt> could mean with his/her contribution and thus results in a contextual implication (see also discussion of example (78) of the same textual unit in connection with exophoric backward reference). A couple of turns later, <tracyt> contributes another turn that makes a direct reference to *he said I was a slag before*. The textual unit *am soooooooo not*, however, contradicts the textual unit it refers to and the processing of *am soooooooo not* involves the cancellation of the old assumption *he said I was a slag before* in order to make sense. Thus, *am soooooooo not* results in a contextual effect that contradicts and eliminates an old assumption (hence, a contextual contradiction/elimination).

Example (180), on the other hand, illustrates, in terms of processing success (rather than effort), a typical non-contextual effect (NON-CE) as observed in another Web Chat session. Consider, in particular, the contribution by the Chatter nicknamed <DanTheAutomator>:

(180) <SunnyGurl> : [any1 wanna start a conversation ?][CS]
<machine-generated turn>: Cool_Rick entered room
<king_silky_B>: [wut about sunnygurl][CS]
<machine-generated turn>: Cool_Rick left room
<Maxim>: [Wazap!!][CS]
<machine-generated turn>: _freedom left room
<machine-generated turn>: sahin_25 entered room
<machine-generated turn>: Cool_Rick entered room
<machine-generated turn>: donab left room
<king_silky_B>: [hi sahin!!][CS]
<machine-generated turn>: killer_weed entered room
<machine-generated turn>: reddog10two54 entered room
<DanTheAutomator>: [All heil Plankton..!!][NON-CE]
(Chat excerpt / Room V, session 1)

There is nothing "unusual" in the unfolding of the Chat before the turn contributed by <DanTheAutomator>. Two Chatters, <SunnyGurl> and <king_silky_B>, have a short exchange on the question what to converse about (both turns processed as contextual strengthenings). Two other contributions belong to the category greetings or greeting formulae (<Maxim's> *Wazap!!* and <king_silky_B's> *hi sahin!!*), both of which perform effects of the type contextual strengthening. <DanTheAutomator's> turn, however, does not seem to make any sense at all. No conclusion is deducible from the input alone and neither is it possible to relate *All heil Plankton..!!* in any way to the communicative context established. Whatever the intention behind <DanTheAutomator's> contribution might be, none of the other Chatters reacts to this comment in the continuation of the discussion (not shown above), and <DanTheAutomator> does not follow up on it either. This textual unit, therefore, is a typical case of a non-contextual effect in terms of processing. Whichever way one puts it, it is impossible to successfully decode the textual unit *All heil Plankton..!!* so as to make sense in this particular communicative situation.

I will now turn to the findings that the empirical investigation into contextual effects yielded. The different effects will each receive the main focus in separate sections, but since different effects may be linked to the same example, there will also be cross-references. In order to be able to compare the frequency of the five types of contextual effects, Tables 12.1 – 12.5 will not only provide the number of textual units processed as a particular effect and the effect-to-word ratios for each of the text types, but also indicate the percentage of a particular effect in relation to the total number of all five effects. It should be noted that the "other, borderline cases" in Web Chat (i.e. the first turns of the Chat sessions that could not be processed due to the session being *in medias res* when data recording started, see also section 5.1.) are not included in the frequencies and

distributions shown in Tables 12.1 – 12.5. A total of 148 textual units (which translates into an average of 5 units per Chat session) were classified as "other, borderline case". These textual units accumulated to a total of 718 words, which were then deducted from the original total word count of Web Chat (resulting in 27'686 words), so as to not bias the results for the frequencies of the remaining effects.

12.2.4. Contextual implication

> Contextual implications are contextual effects: they result from a crucial interaction between new and old information as premises in a synthetic implication.
>
> (Sperber & Wilson 1995: 109)

Of all contextual effects, Sperber and Wilson (1995) elaborate on the contextual implication (henceforth also referred to as CI) most extensively. With regard to its impact on the context established, Sperber and Wilson (1995: 112) come to the conclusion that the contextual implication can be seen as a kind of "*dependent* strengthening" of old assumptions (as opposed to the "*independent* strengthening" of old assumptions, see 12.2.5. below), because the strength of the conclusion depends not only on the newly added information, but also on the context so far established.

It was observed that overall, the contextual implication (or dependent strengthening of old assumptions) and the contextual strengthening (or independent strengthening of old assumptions) are the contextual effects that occur most frequently in the processing of textual units that originate from personal written communication. This seems logical insofar as people in general communicate with each other with the aim of conversational success (and thus strive to understand each other, in other words, tend to contribute new information that is relevant and easy to process). Further, personal written communication can also be assumed to be geared towards the (re-)confirmation of old assumptions against a shared context, rather than contradicting or weakening them. In comparison to the occurrences of all five types of contextual effects that were identified for each of the text types, the comprehension efforts connected to personal written communication frequently involve the processing of contextual implications.

This is, of course, also related to the writing genre examined in this study.[86] Interpersonal communication creates (or, in the case of Web Chat and the personal homepages, may pretend to create) a communicative history between the correspondents. Every message that is exchanged adds to this history, and authors are also likely to draw on this history and imply it in their new messages—thereby fostering contextual implications in the comprehension process of the readership. Table 12.1 summarises the frequency and distribution of textual units processed as effects of the type contextual implication by the researcher across the five text corpora.

Table 12.1: Contextual implication in personal written communication (corpora-based results).

Contextual implication	**SMS discourse** (18 426)	**E-mail** (25'733)	**Web Chat** (27'686)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
Effects per text type	* 2417 ** 1:7.6	1521 1:16.9	1912 1:14.5	1163 1:13.8	1283 1:24.2
Percentage of all effects	*** 53.5	52.3	34.1	49.7	47.0

Key: HP = homepage; * = total number of occurrences (per text type) of textual units processed as effects of the type contextual implication; ** = contextual implication-to-word ratio (rounded to 1 decimal) based on the total number of word counts (in brackets, word count for Web Chat adapted ("other, borderline case" effects not included)) for each of the text types; *** = percentages (rounded to 1 decimal) given for the occurrences of textual units processed as effects of the type contextual implication in relation to the total number of occurrences of all types of contextual effects per text type.

Of all the five text types, the textual units of SMS discourse yield the most contextual effects of the implicational kind. This is believed to be connected to the discourse structure of SMS messaging, as it fosters "mini discussions" of two or more messages between two correspondents (see also 6.1.3.).[87] Hence, the quicker an SMS message "A" is responded to with an SMS message "B", the more likely the author of message "B" will be to create an implicature by making implicit reference to message "A"—thereby producing a contextual implication.

[86] Other genres are hypothesised to foster different types of contextual effects. For example, newspaper articles, aimed at informing the reader about a given topic, are assumed to be less likely to contain contextual implications, because newspaper journalists are not in the position to presuppose a shared communicative context with their readers. Furthermore, newspaper articles should also be understandable for different types of readers with different kinds of background knowledge. Hence, it can be expected that contextual strengthening occurs more frequently than contextual implication in newspaper writing.

[87] One is quite right in thinking that SMS messages, sent back and forth quickly, can in fact be seen as SMS discourse *in medias res* (the same is true for e-mail discourse performed under similar circumstances). However, the difference to Web Chat discourse is that each and every SMS message (or e-mail) was sent off as a text entity. This is, of course, different with respect to an ongoing discussion in a Chat room, where data recording starts at a random point in the discussion and where turns that would equal a message are often broken up into one or more turns.

As has been pointed out before, the SMS corpus consists of messages that cannot be brought into relation to each other. Although textual evidence in the form of preceding SMS messages is missing, the SMS messages differ from "other, borderline" cases in Web Chat insofar that they can be interpreted as isolated messages (see also footnote 87). However, even in these isolated SMS messages, evidence of such "mini-discussions" was observed. Compare examples (181) and (182) below:

(181) [thanks 4 the information!]^{CI} [I neva heard of em]^{CI} [I hope it won't be 2 expensive!]^{CI}
(SMS text / author: female, 25 yrs)

(182) [:-)]^{CI} [cheers.]^{CI} [have fun]^{CI} [& pls say hi to <name> from me,]^{CI} [love, <name>]^{CS}
(SMS text / author: female, 31 yrs)

Both examples (181) and (182) start with a response to a previous message (SMS text or otherwise). Even though the previous textual evidence is missing, both openings can be assumed to relate back to something the recipients wrote beforehand. In example (181), the author thanks the addressee for some kind of information she was provided with. It is a similar case in the SMS text shown in example (182), where the author starts her message with a smiling emoticon which, based on the positioning of the emoticon and the content of the message, is meant as a feedback to the content of a previous message. The emoticon in this example is classified as yielding an individual contextual effect because it is believed to express meaning that can be distinguished from the communicative content of the subsequent textual unit *cheers*.

However, both openings are not deducible from the textual evidence alone and are thus classified as contextual implications in terms of processing efforts. The same is true for the remainder of the message shown in example (181), where it is not clear, based on the textual input alone, what *em* or *it* refers to in the two remaining textual units (and although the content suggests that the referents of *em* and *it* are to a certain extent related, it is not clear what referent(s) they refer to). It is a similar case in example (182), which continues with three more textual units that cannot be successfully processed based on textual evidence alone (*cheers* (what for?), *have fun* (doing what?), and *pls say hi to <name> from me* (who does <name> refer to?))—hence, all three units result in contextual implications as regards processing effort. The message closes with a traditional farewell section (*love, <name>*), a formulaic sequence that results in a contextual strengthening for reasons outlined in 12.1.2. and 12.2.2. Many SMS messages in the corpus are composed in

analogue patterns. This contributes to the circumstance that textual units are processed most frequently as contextual implications in this text type, making up just over 53% of all contextual effects.

E-mail and the personal homepages are the two other text types where the processing of textual units most frequently results in contextual implications. Both Web Chat and EEC (as will be seen below) are made up of textual units that are processed as contextual strengthenings on a more frequent basis than contextual implications. With respect to e-mail, it was observed that this text type has a tendency to feature subject lines that foster contextual implications (see also example (178) above). Furthermore, messages that lack a greeting section start in the majority of the cases (similar to SMS discourse) with some sort of backward reference to the (implied) content of a previous message. Consider examples (183) and (184):

(183) Subject: RE: [no..]^{CI}
(E-mail excerpt (subject line) / author: female, 28)

(184) [no greeting section]
[It sounded as if you had a lot on your plate the other day.]^{CI} (…)
(E-mail excerpt / author: male, age unknown)

Similar to the *Dartmoor* example (178), the RE in the subject line in example (183) indicates that the message is a reply. However, based on the subject line alone, it is unclear what *no..* refers to, which means it is not possible to successfully process it from textual input alone and thus results in a contextual implication. Subject lines qualify as potential contextual implications in terms of processing effort because they tend to be short and are often implicit in that their meaning only becomes clear once the whole message has been read.

In contrast, subject lines are sometimes used as part of the greeting section (as in *hello…* (subject line) and then, for example, *my darling* fronting the main body of the message), or they may substitute the greeting section altogether (for example, *hi there* as a subject line and then the message starts right away with the main body without another greeting). These types of subject lines would then qualify as greeting sections which in turn qualify as formulaic sequences, resulting in contextual strengthening. They are, however, rare in the e-mail corpus investigated for this study. This could be related to the authors trying to be creative and original in their subject lines (rather than inserting a conventional greeting). Or, to say it in Flynn and Flynn's (2003: 22, emphasis original)

words, authors might aim at writing "subject lines with a real *oomph*" that make their e-mail messages stand out. Whatever the motivation may be, over 52% of all textual units resulted in contextual implications.

The high frequency of contextual implications fostered by the textual units of the personal homepage corpus, on the other hand, is a result of the large number of hyperlinks that are not contextually embedded in the running text, or otherwise commented on. The meaning of such hyperlinks can only be deduced by following them and gaining access to the textual units they are intended to refer to. Consider example (185) below:

(185) [While you are here,][CS] [be sure to visit my:][CS]
- [Photo Album][CI]
- [Travel Page][CI]
- [Jet Ski Adventures][CI]

(Personal homepage excerpt / author: Gf_2)

In the context of the personal homepage of Gf_2, both the textual units *While you are here* and *be sure to visit my* strengthen assumptions that are made earlier in the page. However, it is not clear what exactly the three hyperlinks shown in example (185) refer to. Photo albums can vary greatly in their content, and the same is true for the vague descriptions *Travel Page* and *Jet Ski Adventures*. The only way to deduce their meaning is to follow the hyperlinks and see to what sort of information they lead. Without additional information provided by the hyperlinked destinations, the three textual units in the form of hyperlinks are only processable as contextual implications. This was observed to be particularly frequent in connection with hyperlinks. The textual units in the running texts of personal homepages, however, were found to be more likely to classify as contextual strengthenings rather than contextual implications.

The two remaining text types, Web Chat and EEC, also feature textual units that result in contextual implications on a regular basis. However, they are the only two media that feature another contextual effect (the strengthening) more often (see 12.2.5. below). With respect to the contextual implication, however, it was observed that the textual units of Web Chat are prone to be processed as contextual implications because participants have no idea who they are conversing with. They thus run a greater risk of not being explicit enough in their contributions. This is at the same time also the reason why textual units processed as contextual strengthenings are so frequent in Web Chat.

Misunderstandings or unclarities often result in requests for clarification, which in turn fosters contextual strengthening. Consider example (186):

(186) <Bulgaria1>: [how can I fix that?][CI]
<PunkyBrewster>: [fix what?][CS]
(…)
<Bulgaria1>: [me][CI]
<PunkyBrewster>: [what???][CS]
<PunkyBrewster>: [what do you mean?][CS]
(Chat excerpt / Room II, session 4)

In the course of the Chat discussion from which example (186) is an excerpt, a Chatter nicknamed <Bulgaria1> asks a question (*how can I fix that?*), the meaning of which is not quite clear from the input alone (hence resulting in a contextual implication). Another Chatter nicknamed <PunkyBrewster> enquires into the issue by asking the question *fix what?* (yielding a contextual strengthening). <Bulgaria1> responds by contributing *me*, the meaning of which is again unclear (another contextual implication in terms of comprehension on behalf of the reader), and <PunkyBrewster> asks another two questions (resulting in contextual strengthenings) to find out what <Bulgaria1> is hinting at. It eventually turns out that <Bulgaria1> was unsuccessful in trying to change his/her nickname. However, both turns contributed by <Bulgaria1> have something to do with the context, but they are not explicit enough in order to be processed successfully, as illustrated by <PunkyBrewster's> attempts to make sense of <Bulgaria1's> contributions. As can be gathered from Table 12.1, contextual implications make up just less than 35% of all contextual effects identified based on the textual units from the Web Chat corpus. This means that they are considerably rarer than in any of the other four text types.

The last text type to be discussed in connection with contextual implication is the letter. EEC, similar to Web Chat, also contains more textual units that yield contextual strengthenings than contextual implications. Nonetheless, textual units that are processed as contextual implications are quite frequent in the letters, too. They are in the majority of cases concerned with accounts of other third parties, whose identities are treated as presupposed knowledge in the reader (see also 7.1.3.). EEC is at the same time also the text type most likely to embed a textual unit (and corresponding contextual effect) into another one:

(187) [For newes,]^{CS} [Sir Joh. Radclyfe is deade,]^{CI} [and his sonne hath [his company]^{CS} ;]^{CI} [the Counte Mansfeldes bysnis goes not forward as it aught;]^{CI} [the French kinge hath refused [him]^{CS} landinge and passage in France,]^{CI} [soe that [hee]^{CS} is now to shypp his troopes for Hamborough.]^{CI} (…)
(EEC excerpt, 1625 / letter collection Cornwall, author: male)

In a passage of 46 words, no less than five third parties are introduced in the discourse without much background information on the different personae. In consistence with the adapted classification system for contextual effects, a backward reference to a prior textual unit that can be deduced from the input alone, counts as a textual unit that result in strengthening the context. Thus, the first mentioning of the man *Sir Joh. Radclyfe* in example (187) is processed as an implication. The backward reference *his company*, however, is classified as a contextual strengthening because it is clear from the content of the input what it refers to. In any event, example (187) illustrates in a prototypical manner why textual units in EEC are frequently processed as contextual implications.

Hence, it was observed that the different text types show different idiosyncrasies in connection with the occurrence of textual units processed as contextual implications. SMS discourse and e-mail messages are likely to imply the contents of previous messages. In addition, the e-mail corpus shows tendencies for the textual units in the subject line to be processed as contextual implications. The textual units resulting in contextual implications found in the personal homepages, on the other hand, are most likely connected to the hyperlinks that repeatedly do not provide enough information so as to be deducible from their input alone. It was further observed that the textual units of Web Chat that yield contextual implications can be brought into relation with the tendency that Chatters are often not explicit enough in their contributions. It seems as though they sometimes presume too much common knowledge, where in fact there is little or none. It is quite the opposite in EEC, where the textual units that are processed as contextual implications can be attributed to the correspondents' shared (pre-)communicative context.

Including a question like "how often do you think you include textual units that are likely to be processed as contextual implications by the readership in your personal written communication?" in a questionnaire may confuse informants without a linguistic background. Nevertheless, I still wanted to find out whether the informants felt that there are differences in how they author and read messages transmitted by the five different media in connection with contextual effects. The notions of implicatures and explicatures proved to be helpful in the matter. Thus two questions concerned with the decoding of messages in connection with their degree of explicitness and implicitness were

incorporated into the survey. The first was aimed at the informant as an author. The question was how often they thought it happened that a recipient of one of their messages did not understand something they wrote (i.e. too implicit, successful decoding not possible) and asked for clarification. The survey showed that the majority of informants felt that this is sometimes the case in SMS discourse and Web Chat, but rarely in all other types of correspondence. The reverse question (how often informants thought it happened that they read a message from someone and found they did not understand something and needed clarification) generated similar results. The majority of informants answered that this is sometimes the case with SMS discourse, e-mail, and Web Chat, but rarely happens in the personal homepages they visit, or the letters they read. Hence, if the personal correspondence of the informants also tended to contain textual units that resulted in contextual implications in the comprehension process of the reader, then the outcome from the online survey suggest that they can generally be processed without great difficulty, above all with respect to e-mail messages, the personal homepages, and letters.

12.2.5. Contextual strengthening

> New information may provide further evidence for, and therefore strengthen, old assumptions.
>
> (Sperber & Wilson 1995: 109)

New information that strengthens and supports old assumptions and is thus processed as contextual strengthening (henceforth also referred to as CS) was observed to be frequent in all the text types, and it is even the most popular type of effect in Web Chat and EEC. Considering that the two text types are quite different in their discourse structure and in the kind of readership they address, this might seem at first glance a little odd. However, it emerged that they take different roads but arrive at a similar end. Participants in Web Chat make short contributions (with an average turn length of 5.2 words), and while such short contributions are predestined to be not explicit enough and thus processed as contextual implications (which, as has been discussed above, is also frequently the case), they also often stand in clear relation to previous textual units and/or consist of conventional greetings and farewells. Furthermore, it was observed that the anonymous communicative setting of Web Chat motivates those Chatters that are interested in having a discussion (as opposed to those who just want to make "noise", also referred to as "flooding" a Chat room by repeating the same comment or similar comments again and again), to ask for

clarifications if something is unclear and thereby fostering contextual strengthening (as illustrated in example (186) above).

The strengthening of old assumptions is crucial in a setting where participants do not know each other well (or not at all) in order to establish a common ground (Web Chat), or identify and clarify one's viewpoints (personal homepages). In a setting where correspondents know each other well, on the other hand, the strengthening of old assumptions may primarily serve to address (and emphasise) issues of certain importance (to sender and/or receiver) or urgency (SMS discourse). Furthermore, textual units that result in contextual strengthening frequently occur in text types that also tend to contain longer narrations, dealing with a selected number of topics (e-mail and EEC). Table 12.2 summarises the frequency and distribution of the textual units processed as contextual strengthening as observed in the five media.

Table 12.2: Contextual strengthening in personal written communication (corpora-based results).

Contextual strengthening	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (27'686)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
Effects per text type	* 1967 ** 1:9.4	1264 1:20.4	3138 1:8.8	1079 1:14.9	1338 1:23.2
Percentage of all effects	*** 43.6	43.4	55.9	46.0	49.1

Key: HP = homepage; * = total number of occurrences (per text type) of textual units processed as effects of the type contextual strengthening; ** = contextual strengthening-to-word ratio (rounded to 1 decimal) based on the total number of word counts (in brackets, word count for Web Chat adapted ("other, borderline case" effects not included)) for each of the text types; *** = percentages (rounded to 1 decimal) given for the occurrences of textual units processed as effects of the type contextual strengthening in relation to the total number of occurrences of all types of contextual effects per text type.

The percentage of textual units processed as contextual strengthening amounts to 55.9% in Web Chat, which means it is by far the most frequent type of contextual effect in this medium. This is assumed to be connected to the anonymous communicative setting of Web Chat as it runs an increased risk of participants misunderstanding each other. Or, in other words, participants are often not able to successfully decode messages because they often do not have a big enough common background knowledge. This would also mean that participants are quite frequently in need of clarification.

The results from the corpus investigation support these assumptions. Furthermore, because participants often express this need for clarification, they, at the same time, not only strengthen the context established, but motivate other participants to provide the clarification they need (leading to more strengthening of old assumptions). Or, as in the

example (188) below, other Chatters might help each other in order to clarify something that causes problems in comprehension:

(188) <Turkey17804>: [hi everybody]^{CS} [it is a question about inquiry:]^{CI} [what does it the most precious something in life?]^{CI}
<Oggimator>: [huh <Turkey17804>?]^{CS} [i don't get it…]^{CS}
<Sweden4251>: [it's life itself, <turkey> :-)]^{CS}
<Anonymous>: [<Sweden4251>, agree! :D]^{CS}
<Turkey17804>: [another thing]^{CI}
<India16971>: [what do u mean by that quest does it the most precious something in life?]^{CS} [<Turkey17804>?]^{CS}
<Sweden4251>: [yeah it's a bit funny…]^{CS} [I think he means what is the most precious thing in life]^{CS}
(Chat excerpt / Room II, session 2)

The Web Chat excerpt shown in example (188) is prototypical in that one Chatter makes a contribution, about which fellow Chatters debate. In this particular case <Turkey17804> contributes a somewhat strangely formulated enquiry about what the most precious thing in life could be. While two Chatters (<Oggimator> and <India16971>) ask <Turkey17804> what he/she means by that, <Sweden4251> and <Anonymous> contribute answers. Further, <Sweden4251> tries to help <India16971> understand, because <Turkey17804> withholds this information and responds with the elusive answer *another thing* instead. Thus, one Chatter contributes new pieces of information that cannot be successfully decoded by means of input alone (resulting in contextual implications). The four Chatters that try to process these new pieces of information by discussing it with each other thereby strengthen old assumptions, hence producing textual units that yield effects of the type contextual strengthening. In the last chapter it was determined that the Web Chat corpus features by far the highest frequency of interrogative structures. Here we see another reason why this is the case.

With respect to EEC, the other text type that features textual units processed as contextual strengthenings more often than implications, the motivation for strengthening old assumptions is different to Web Chat. Letter writers are generally not under such a time pressure to compose their correspondence as is the case with Chatters. Furthermore, they are also acquainted with their readership, so they know what could be of interest to them. This quite frequently results in story-like letters, which narrate in careful detail certain events that the author has heard of or witnessed. This detailed writing style means that new information that is added provides further evidence for, and therefore strengthens, the content of previous textual units. Furthermore, the letters tend to include elaborate

demonstrations of affection, resulting in numerous textual units that serve to strengthen previously made assumptions. Consider the letter in example (189) below:

(189) [Worthy Prince and my dearest brother:]^{CS}
[I received your most welcom Letter and kynd token [by Mr. Hopkins]^{CI},]^{CI} [highly esteeming them as delightfull memorialls of your brotherly love.]^{CS} [In which assuredly [(whatsoever ells may fayle)]^{CS} I will ever endevor to equall you,]^{CS} [esteeming that time happiest when I enjoyed your company,]^{CI} [and desiring nothing more than the fruition of it again:]^{CS} [that as nature hath made us neerest in our love together,]^{CS} [so accident might not separate us from living together.]^{CI}
[Neither do I account yt the leste part of my present comfort, [that though I am deprived of your happy presence,]^{CS} yet I can make these lines deliver this true message,]^{CS} [that I will ever bee during my lyfe,
[Your most kinde
and loving syster
Elizabeth]^{CS}.]^{CS}
[To my most dear brother the Prince.]^{CS}
(EEC 1600 / letter collection Original_3, author: female)

The whole letter is aimed at expressing the sisterly love that the author feels for her brother, the recipient of the letter. Of the fifteen textual units determined in the above letter, eleven are classified as yielding effects of the type contextual strengthening in terms of comprehension, and only four textual units are not deducible from the input alone (and thus classified as generating contextual implications). Letters like the one shown above, as well as letters explicitly expressing comfort, are numerous in the corpus. They contribute considerably to the strengthening of old assumptions being so frequent in EEC.

Similar to EEC, the e-mail corpus contains numerous messages that narrate one or more events in some detail. It has been pointed out earlier in this study that most e-mail messages have total word counts between 50 and 300 words (see 6.1.1.). Many of these messages are dedicated to addressing issues that are of concern to the author, who tries to convey a coherent picture of events to the reader:

(190) Subject: [London][CI]

[Hi sweetie!][CS]
[How are you doing?][CS]
[I have been in London for nearly a week,][CS] [my make-up course started on monday][CI] [and I am mainly exhausted.][CI] [I have learnt so much][CS] [and it's quite a bit to take all in][CS] [(in such a quick way we're supposed to).][CS] [There are [30 students][CI] in our class][CS] [and it's all a bit hectic.][CS] [We don't have time enough to practice all the things [they][CI] teach us][CS] [and it's going in a very quick pace.][CS] [While you do your make-up on a fellow student,][CS] [the 3 teachers walk around and give you some tips][CS] [(well correct you mainly).][CW] [There are just too many of us][CS] [and everyone needs the teacher][CS] [so you have to wait][CS] [but after some time you have to sit down and be the model][CS] [and your partner gets to work on you.][CS] [So while you wait you are running out of time.][CS] (…)
(E-mail excerpt / author: female, 30)

The e-mail excerpt shown in example (190) is an account of a make-up course that the author was attending in London at the time. Both the locality (*London* in the subject line) and the main topic (*my make-up course*) are introduced by means of textual units fostering contextual implications and this suggests that the author presupposed this information in the reader. The majority of textual units describing the make-up course, however, result in effects of the type contextual strengthening. In the whole excerpt shown in example (190), there are only four units that are processed as contextual implications (the two outlined above and the reference of both *they* and *30 students* is unclear) and one unit that results in a contextual weakening (the textual unit *well correct you mainly* weakens the preceding unit *the 3 teachers walk around and give tips*). This means out of 24 textual units determined in the excerpt above, 19 result in contextual effects of the strengthening type. This not only renders a coherent picture of events, it also indicates that this matter is of importance (i.e. stressful) for the author, and that she wants to share it with her reader.

SMS discourse features textual units processed as the contextual strengthening (next to greetings and farewells) predominantly in matters that are of importance to sender and/or receiver. However, opposed to e-mail (or EEC), SMS discourse offers considerably less space to communicate such issues. In order to accommodate the space restriction, authors of SMS texts frequently resort to using formulaic sequences:

(191) [*HAPPY*BIRTHDAY*][CS] [send you all the best wishes :-)][CS] [love you lots!][CS] [yours <nickname>][CS]
(SMS text / author: female, 27 yrs)

The content of the SMS text shown in example (191) is, judged by its content, assumed to be of emotional importance to both sender and receiver. The message opens with a textual

unit classified as a prototypical formulaic sequence and is thus processed as a contextual strengthening. This is followed by three more formulaic sequences that strengthen the first textual unit **HAPPY*BIRTHDAY**. In contrast to example (182) above, where the emoticon performs its own contextual effect, the emoticon in example (191) is classified as supporting the textual unit *send you all the best wishes* in its effect of the type contextual strengthening. SMS messages like the one shown in (191) are frequent in the corpus, although the event may be substituted, as required, by *merry Christmas*, *happy New Year*, *happy Valentine's Day*, *happy Easter* and so forth.

In the personal homepage corpus another trend was observed. The vast majority of textual units classified as fostering contextual strengthening served to both identify and clarify the author's viewpoints. This is connected to the circumstance that the author does not know who will be reading the contents of the page. Therefore, he/she has no idea what kind of background knowledge the reader may have. Hence, political views or special hobbies tend to be described by means of new information that strengthens old assumptions. Another interesting observation is that many of the textual units classified as strengthening old assumptions are given in parentheses:

(192) [Here is my hypothetical FAQ]^{CI} [*(answers to questions that I think you might ask).*]^{CS}
(Personal homepage excerpt, italics my emphasis / author: Em_2)

(193) (…) [I you've got any questions about anything]^{CS} [*(I do my best to answer all questions),*]^{CS} [click on the email me.]^{CS}
(Personal homepage excerpt, italics my emphasis / author: Cf_2)

In both the examples (192) and (193), the textual units in parentheses provide additional information that strengthens the preceding textual units. This is so common in the personal homepage corpus (and not very frequent in the other text types) that it can be seen as idiosyncratic of this text type.

Overall, textual units that result in contextual strengthening, similar to the contextual implication, are frequent in all text types but show different areas of use. Participants in Web Chat are highly likely to contribute new information that results in contextual strengthening in communicative situations that require the resolving of a misunderstanding between two or more Chatters. Also, the high frequency of interrogatives found in Web Chat fosters contextual strengthening, as the answering of questions in most cases entails the strengthening of old assumptions. Both e-mail and EEC

have shown trends towards containing textual units that strengthen old assumptions in longer narrations that are often of emotional value to either sender and/or receiver. This was not observed in SMS discourse, and it is believed that this is related to the limitation in text length. However, this text type features formulaic sequences on a more frequent basis than any of the other text types, and this too, is believed to be related to the restriction in text length. The personal homepages show yet another trend in that a great number of textual units, aimed at strengthening old assumptions, are placed in parentheses. Hence, the investigation into the two most frequent types of contextual effects exposes tendencies that can be tied to either discourse structure or the communicative setting of the text types.

As illustrated in Tables 12.1 and 12.2, textual units resulting in contextual effects of the types implication and strengthening make up more than 90% of all contextual effects determined for each of the text types. Although Web Chat features the lowest frequency (90% as opposed to all other text types, where frequencies amount to over 95%), this is nevertheless astonishing, considering the circumstance that the participants, in most cases, do not know each other. It seems as though authors of personal written communication have a general tendency to modify the context by means of textual units that are processed as contextual implications (dependent strengthening) and contextual strengthenings (independent strengthening). I will now look into three more types of contextual effects that have shown to be considerably less numerous, but still position themselves as a means for context modification.

12.2.6. Contextual contradiction/elimination

> [New information] may provide evidence against, and perhaps lead to the abandonment of, old assumptions.
>
> (Sperber & Wilson 1995: 109)

A textual unit that yields a contextual effect of the type contradiction/elimination (henceforth also referred to as CC/E) provides evidence against old assumptions. In the case of personal written communication that is authored by one individual, CC/E would consequently mean that the author either contradicts him-/herself within the same message or, in order for the contradiction to be apparent, would have to explicitly refer back to a text (segment), produced by the author or another person, and contradict that. Web Chat is insofar an exception as its (near-)synchrony and discourse structure allow several participants to interact at the same time (see also example (179) above, where the textual

unit *am soooooooo not* contradicts *he said I was a slag b4*, the third person personal pronoun *he* indicating that this comment was made by another Chatter). Hence, it can be assumed that textual units processed as effects of the type contextual contradiction/elimination occur more frequently in communicative settings where different people with different viewpoints come together. As illustrated in Table 12.3, the percentages for textual units that result in contextual contradiction/elimination are, however, lowest for Web Chat and highest for EEC:

Table 12.3: Contextual contradiction/elimination in personal written communication (corpora-based results).

Contextual contradiction/elimination	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (27'686)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
Effects per text type	* 51 ** 1:361.3	41 1:627.6	57 1:485.7	31 1:517.1	52 1:597.6
Percentage of all effects	*** 1.1	1.4	1.0	1.3	1.9

Key: HP = homepage; * = total number of occurrences (per text type) of textual units processed as effects of the type contextual contradiction/elimination; ** = contextual contradiction/elimination-to-word ratio (rounded to 1 decimal) based on the total number of word counts (in brackets, word count for Web Chat adapted ("other, borderline case" effects not included)) for each of the text types; *** = percentages (rounded to 1 decimal) given for the occurrences of textual units processed as effects of the type contextual contradiction/elimination in relation to the total number of occurrences of all types of contextual effects per text type.

The observation that EEC features almost twice as many textual units processed as contextual effects of the type contradiction/elimination than all other text types can be explained in terms of politeness. Consider the letter excerpt shown in example (194):

(194) [Mr Parr, -]^{CS} [I thanck you for part of your letter,]^{CI} [as farre as you ware indifferent in this busines;]^{CI} [but I must tell you]^{CS} [that I did never expect that you would have ben a persuader of me [to a gave awaye the increase of my owne estate]^{CI},]^{CI} [being you have eaver heard me earnestli to protest that I would not,]^{CS} (…) [This is no new thinge to you]^{CS} [for you know that before I eaver saw [Mr Bacon]^{CI} that this was my mind,]^{CS} [and from which I neaver will be removed.]^{CS} [Besides, you know how often you have brought me word from Mr Bacon]^{CS} [that I should do with my own estate, [besides my child's,]^{CW} what I would;]^{CS} [which, if it should now apear to me that all this was but done to entice my affection,]^{CI} [would be a grate reson to direct it another waye,]^{CS} [and I fere I shall find such hard mesure;]^{CS} [[but, howsoever, [(yet)]^{CW} [yf you carry yourself justly,]^{CI} you shall eaver find me your frend to my power,]]$^{CC/E}$ [Ja. Cornewaleys.]^{CS}
(EEC excerpt, 1613 / letter collection Cornwall, author: female)

The author of the letter excerpt shown in example (194) addresses a delicate issue between her and the recipient of the letter. Her viewpoint in the matter, namely that she feels she is not only entitled to do what she wants with her own estate (in consideration of her child, the insertion *besides my child's* being a textual unit fostering a contextual weakening), is

not only demonstrated in the content of the letter, but it also emerges that she feels that she has made this clear to the recipient on several other occasions. Nonetheless, she suspects that the addressee acts against her wishes in trying to persuade her to give away some of her property. The detailed outline of the issue, of which some has been substituted by (…) in the excerpt above, is realised predominantly by means of textual units that strengthen old assumptions, resulting in effects of the type contextual strengthening in terms of comprehension process. The same is true for the most important claim of the letter directed at the addressee: *if it should now apear to me that all this was but done to entice my affection, would be a grate reson to direct it another waye, and I fere I shall find such hard mesure*. The statement *and I fere I shall find such hard mesure* is particularly strong in its illocutionary force. This is why it is assumed that the author decides, for reasons of politeness, to include an additional textual unit that in fact contradicts the previous one: *but, howsoever, (yet) yf you carry yourself justly, you shall eaver find me your frend to my power*. Note that the contradicting textual unit is conditioned (and thereby) softened by the insertion of *(yet) yf you carry yourself justly*.

Further, because of this weakening of old assumptions, the textual unit only contradicts but not completely abandons the previous textual evidence. Hence, the larger textual unit *but, howsoever, (yet) yf you carry yourself justly, you shall eaver find me your frend to my power* contradicts previously made assumptions and is therefore processed as a contextual contradiction/elimination. However, embedded in this contradiction are another two textual units yielding different types of effects: *but, howsoever, (yet)* conditions and thereby softens the overall effect of the larger textual unit (resulting in a contextual weakening), whereas the textual unit *yf you carry yourself justly* is not explicit enough in its content to be deduced by textual evidence alone and although it is conditional (*yf*) it is at the same time also implicit as to how exactly the reader ought to behave (resulting in a contextual implication).

In the case of Web Chat, as has been illustrated in example (179) above, participants of this type of discourse are more likely to use contradictory statements for the occasional negation of a contribution made by another Chatter (resulting in an average of one contradiction/elimination in every 500 words). However, both e-mail and SMS discourse show tendencies for including textual units of the contradicting kind in connection with face-saving strategies[88]:

[88] See hereto also the discussion of the DM *well* in the function of face-threat mitigator in section 8.1.

(195) Subject: [boring lunchtime?][CI]

[Dear <Forename>][CS]

[I thought you might be having [a boring Wednesday lunchtime][CS] at work][CI] [so might be checking your e-mails,][CS] [so I thought I'd send you a message…][CS] [Hope your day is not too bad,][CI] [hope [they][CI] are giving you enough work to do.][CS] [Actually, I don't suppose you ever have boring lunchtimes at work.][CC/E] (…)
(E-mail excerpt / author: male, 41)

The above e-mail does not only have the issue *boring lunchtime* as its subject line, the whole of the first paragraph is dedicated to what could be a boring lunchtime for the addressee. However, it seems as if the author at some point realises that statements such as *I thought you might be having a boring Wednesday lunchtime at work* and *hope they are giving you enough work to do* are potentially face-threatening for the recipient, because they could also be understood along the lines of the recipient being lazy during lunchtimes, or, and probably even worse, is seen as not being qualified enough to be given more work. Whatever the reason, the author includes at some point the textual unit *Actually, I don't suppose you ever have boring lunchtimes at work* which results in a contextual contradiction/elimination of previously made assumptions—to the assumed effect that the reader does not feel offended by the prior statements. As to why the author chose to structure the content in this fashion is unclear. Had he deleted the previous units before sending the e-mail, he could have avoided this contradiction and the reader would have been oblivious to it (something that is not possible in a spoken conversation). However, this kind of communicative strategy was observed in both e-mail and SMS discourse.

The personal homepage corpus, on the other hand, incorporates all of the few textual units processed as contextual contradiction/elimination in narrative passages. This may be aimed at creating suspense. Consider example (196):

(196) [I moved into my current apartment over a year ago.][CI] [Shortly after getting settled,][CS] [I noticed that one of the light switch tabs [-next to the back door from the garage into the lower level-][CS] pointed the wrong direction.][CI] [Typically, tab down equals light off.][CS] [The switch controls an overhead bulb in the garage.][CS] [The light fixture has a chain.][CS] [I was perplexed as to why the tab in the up position turns the light off.][CS] [I thought maybe the chain on the fixture needed to be at a different click.][CS] [Not.][CC/E] (…). [I gave up.][CS] [I was bothered by the switch every time I passed.][CS] [I thought I was being anal about wanting the tab in the traditional position.][CS]
(Personal homepage excerpt / author: Af_2)

The author of the above homepage excerpt narrates a little story about a light switch. She looks at the issue from a technical side (*Typically, tab down equals light off*), but also expresses her concerns from a psychological perspective (*I was perplexed as to why the tab in the up position turns the light off*). However, most interesting is the twist (realised by the negation *Not*) that after all this reasoning, it is in fact something else that makes the light switch point in the wrong direction (it turns out later in the narration that it is a partner switch). Similar to the letter excerpt shown in example (194), the outline of the issue consists predominantly of textual units that strengthen old assumptions (hence, resulting in contextual strengthening) up until the point where they are contradicted by a textual unit that contradicts and to a certain extent negates old assumptions (resulting in contextual contradiction/elimination). The use of contradictory statements to create suspense in narrative text passages is more frequent in the personal homepages compared to the other types of correspondence, and it is thus idiosyncratic to a certain extent.

12.2.7. Contextual weakening

In the postface to the second edition of *Relevance* (1995), Sperber and Wilson include "a fourth type of contextual effect, namely weakening of existing assumptions" (1995: 294). It has been pointed out in the theoretical discussion that the status of the contextual weakening (henceforth also referred to as CW) within RT is uncertain. It was not treated as an independent type of contextual effect in the first edition of *Relevance* (1986), and it has also more recently been stated that the weakening of assumptions is not treated as a distinct type of contextual effect because it only contributes indirectly to relevance (cf. Higomashimory & Wilson 1996, see also 4.4.). In any event, as illustrated in example (178), and see also examples (191) and (194), the notion of the independent weakening of old assumptions has shown to be a useful concept for the purposes of this study.

Textual units processed as contextual weakening are most frequent in the text type Web Chat, and this is believed to be related to the anonymity of the communicative setting. Since Chatters do not know who they are corresponding with, they are likely to soften the illocutionary force of certain textual units with the aim of keeping offence at bay (hence, a motivation comparable to the incorporation of contradictory textual units which are then processed as contextual contradiction/elimination in EEC and e-mail in connection with saving face). Consider example (197):

(197) <NBK>: [need a cigarette!][CI]
(…)
<NBK>: [right, I'm gonna go inhale some cancer now][CS]
(…)
<SicklySweet>: [blah smoker][CS]
(…)
<NBK>: [I'll do as I wish?][CS]
(…)
<NBK>: [Why does when I die matter to you?][CS]
<NBK>: [lol][CW]
(Chat excerpt / Room I, session 1)

A Chatter nicknamed <NBK> is in need of a cigarette and expresses this need in his/her first turn shown in the excerpt above (*need a cigarette!*), which is classified as a textual unit resulting in a contextual implication because the relevance of this textual unit is not quite clear as it cannot be tied to the context established. He/she re-addresses this issue again by announcing that he/she is going to *inhale some cancer now* (meaning 'smoke a cigarette') and thereby strengthening the old assumption *I need a cigarette!*, hence resulting in a contextual strengthening. Another Chatter, <SicklySweet>, comments on <NBK's> project with *blah smoker* and thus also contributes to the cigarette issue with a textual unit that strengthens old assumptions (fostering a contextual strengthening). <NBK> reacts to this comment with a statement disguised as a question (*I'll do as I wish?*), as well as the bold question *Why does when I die matter to you?*.

However, <NBK> seems to realise that this is quite a harsh comment, and he/she softens the illocutionary force of this textual unit by adding *lol* ('to laugh' or 'laughing out loud') in the next turn (hence, a contextual weakening in terms of comprehension process). The difference in the use of *lol* in this example to, for example, a contradictory textual unit such as *Actually, I don't suppose you ever have boring lunchtimes at work* in example (195) above, is the following: while *lol* does not contradict or eliminate the previous textual unit *Why does when I die matter to you?*, but merely weakens it in its illocutionary force, i.e. reduces the potential face-threat this comment might have on the reader, the use of a contradictory textual unit like *Actually, I don't suppose (…)* eliminates old assumptions in that they are cancelled. Analogue communicative scenarios were observed with the acronyms *lmao* ('to laugh my arse off' or 'laughing my arse off'), onomatopoeic expressions of laughter (*hihihi*, *hehehe*), and emoticons.

Textual units resulting in contextual weakening occur more frequently than textual units yielding contextual contradiction/elimination in all of the text types, except for the

personal homepage corpus. Table 12.4 illustrates the frequencies and distribution of the textual units processed as contextual weakening across the five text types.

Table 12.4: Contextual weakening in personal written communication (corpora-based results).

Contextual weakening	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (27'686)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
Effects per text type	* 79 ** 1:233.2	84 1:306.3	228 1:121.4	13 1:1233.1	54 1:575.5
Percentage of all effects	*** 1.8	2.9	4.0	0.6	2

Key: HP = homepage; * = total number of occurrences (per text type) of textual units processed as effects of the type contextual weakening; ** = contextual weakening-to-word ratio (rounded to 1 decimal) based on the total number of word counts (in brackets, word count for Web Chat adapted ("other, borderline case" effects not included)) for each of the text types; *** = percentages (rounded to 1 decimal) given for the occurrences of textual units processed as effects of the type contextual weakening in relation to the total number of occurrences of all types of contextual effects per text type.

The contextual weakening as illustrated in the Web Chat excerpt in example (197) above is typical for the occurrence of textual units generating this effect in all text types. Both SMS discourse and e-mail show a similar tendency in their use of textual units processed as contextual weakening, although this is achieved by means of emoticons and the imitation of laughter (rather than with acronyms), as illustrated by the SMS message shown in example (198). Of course, contextual weakening is not restricted to emoticons or onomatopoeia, but may also be brought about by textual units as illustrated in the e-mail excerpt in example (199). Furthermore, example (199) also illustrates that, next to the function of face-saving, the weakening of old assumptions is also sometimes unconnected to politeness:

(198) [fine]^{CI} [your the boss]^{CI} [;-)]^{CW}
(SMS text / author: male, 27 yrs)

(199) [You're right,]^{CI} [New Orleans is one of the most dangerous places in the world]^{CI}
[But I guess you shouldn't worry too much about it]^{CW}
(E-mail excerpt / author: male, 30)

In example (198), the emoticon ;-) ('smile + wink') weakens the illocutionary force of the previous textual unit *your [you're] the boss*. The first textual unit contains the response form *fine* (see 11.2.5.), which implies that the SMS message is a response to a previous message, and that the author of the current message agrees to something as proposed in the previous message (indicated by *your the boss*). Thus, the winking smiley does not contradict the previous textual units, because it can still be assumed that the addressee gets

his/her way. It weakens, however, the strong illocutionary force of *your the boss* (which could be understood as 'you are bossy') by means of the emoticon, which indicates that *your the boss* is not meant in a negative sense.

In the e-mail excerpt shown in example (199), the communicative goal of textual units yielding contextual weakening differs from example (198). While example (198) is concerned with face-saving, the textual unit *But I guess you shouldn't worry too much about it* (where the *too much* indicates that one should worry, but just not too much) weakens the previous unit *New Orleans is one of the most dangerous places in the world* for other reasons. From the whole of the e-mail message, of which example (199) is an excerpt, it can be understood that the addressee intends to travel to New Orleans. The contextual weakening by means of *But I guess you shouldn't worry too much about it* may have been included in the message so as not to scare the recipient, and maybe risk putting him/her off travelling to New Orleans.

EEC also shows tendencies to use include textual units processed as contextual weakening for reasons of politeness, although this is exclusively achieved by means of textual units consisting of letters—emoticons, onomatopoeic imitations of laughter, and acronyms are non-existent in this corpus. The majority of textual units processed as weakening old assumptions are embedded in longer sentences in EEC. These insertions typically include textual units such as *mistake me not altogether* or *if it may stand your liking*. Textual units resulting in contextual weakening occur on average every 550 to 600 words in EEC. The personal homepage corpus, on the other hand, hardly features any textual units resulting in effects of the type contextual weakening and features the lowest weakening-to-word ratio of all the text types (1:1233.1). This may be connected to the circumstance that the authors do not know their readership. While the use of textual units that contradict or eliminate old assumptions (hence, contextual contradiction/elimination) in the personal homepages was observed to create suspense, the weakening of old assumptions might be rare because authors do not wish to generate this effect on the readership. Also, in connection with politeness, they are under no pressure to hedge their content because they do not directly address a particular reader, but an implied readership. This stands in contrast to the other four text types, where the author runs the risk of offending his/her correspondent as they are directly addressed.

12.2.8. The notion of non-contextual effect

The non-contextual effect (henceforth also referred to as NON-CE) was formulated specifically for the purpose of this study, and it has shown to be a useful concept for both Web Chat and the personal homepage data. The text corpora containing correspondence directed at a known readership (SMS discourse, e-mail, and EEC) do not feature textual units resulting in non-contextual effects in any of the messages. Table 12.5 illustrates the frequencies and distributions of textual units processed as non-contextual effects in the two corpora in which they were observed.

Table 12.5: Non-contextual effects in personal written communication (corpora-based results).

Non-contextual effect	**SMS discourse** (18'426)	**E-mail** (25'733)	**Web Chat** (27'686)	**Personal HP** (16'030)	**EEC (letter)** (31'077)
Effects per text type	0	0	* 280 ** 1:98.9	56 1:286.3	0
Percentage of all effects	0	0	5.0	2.4	0

Key: HP = homepage; * = total number of occurrences (per text type) of textual units processed as effects of the type non-contextual effect; ** = non-contextual effect-to-word ratio (rounded to 1 decimal) based on the total number of word counts (in brackets, word count for Web Chat adapted ("other, borderline case" effects not included)) for each of the text types; *** = percentages (rounded to 1 decimal) given for the occurrences of textual units processed as effects of the type non-contextual effect in relation to the total number of occurrences of all types of contextual effects per text type.

With respect to Web Chat and the personal homepages, it turned out that the messages composed on those two media contain textual units that result in different types of non-contextual effects. While Web Chat has a tendency to feature textual units that do not make sense in the context established (also referred to as "noise"), the personal homepages are often equipped with dysfunctional hyperlinks. If the reader decides to follow up on such a dysfunctional hyperlink he/she will be taken to a dead-end (see also Figure 6.1.), and since the originally intended destination of such a dead-end hyperlink cannot be consulted, i.e. the content is not available, this results in a non-contextual effect in terms of comprehension—for the simple reason that because no "new information" is available, it can also not be processed against a background of old assumptions. This is true even if the text of the hyperlink (i.e. the link itself) is related to the context within the current homepage, but because one does not know to what sort of other site the hyperlink would have led, it is in fact not possible to deduce its full meaning.

As it is not possible to include a dysfunctional hyperlink in this study, I will concentrate on two examples from the Web Chat corpus to further illustrate the non-contextual effect. Web Chat is time and again subject to "noise" (contributions that

make no sense within the context established, see also example (180) above), "flooding" or "spamming" (the repetitive contribution of non-relevant material). Although most Chat room operators advise their Chatters to refrain from flooding and spamming, it happens quite frequently. Chatters who nevertheless flood the room are often ignored by the other participants. They are, however, sometimes advised by fellow Chatters to cease flooding. Example (200) is a case where the Chatter who is accused of flooding responds to these allegations. Normally, "flooders" do not interact with other Chatters, but see the disturbance of the discussion as their main aim.

(200) <gothic_babe>: [ppppppppppppppaaaaaaaaaa][NON-CE]
<gothic_babe>: [ppppppppppppppaaaaaaaaaa][NON-CE]
<gothic_babe>: [ppppppppppppppaaaaaaaaaa][NON-CE]
(…)
<gothic_babe>: [kkkkkkkkkkkkllllllllllllllllliiiiiiiii][NON-CE]
<gothic_babe>: [kkkkkkkkkkkkllllllllllllllllliiiiiiiii][NON-CE]
(…)
<keat>: [<gothic>: can u pls stop flooding!][CS]
<gothic_babe>: [ppppppppppppppaaaaaaaaaa][NON-CE]
<gothic_babe>: [ppppppppppppppaaaaaaaaaa][NON-CE]
(…)
<newcomer>: [stop it <gothic_babe>][CS] [you are redicoules][CS]
(…)
<gothic_babe>: [opps][CI]
<gothic_babe>: [ok][CS]
(Chat excerpt / Room IV, session 5)

In the Web Chat excerpt shown in example (200), <gothic_babe> is flooding the Chat room with his/her contributions, all of them classified as textual units resulting in non-contextual effects. At some point two other participants, <keat> and <newcomer>, step in and they both tell <gothic_babe> to stop flooding. In general, Chatters who flood the room leave it again at some point (by choice, or because they are kicked out) and tend to avoid interaction with other participants. In this example, however, <gothic_babe> first reacts by contributing *opps* (a textual unit resulting in a contextual implication because it is not quite clear what is meant by it), to then assent by means of *ok* to discontinue flooding. The remaining data of this particular Chat session shows that <gothic_babe> keeps his/her word, as no more flooding is recorded.

Another trend that was observed is that sometimes, contributions are made in another language than English. If only one participant chats in this language, and no-one else understands him/her, then these textual units classify as resulting in non-contextual

effects.[89] Turns in a foreign tongue are similar to flooding in that none of the remaining Chatters comprehend the contributions. They thus perceive these contributions to be a hindrance in the unfolding of the discussion. Consider example (201):

(201) <aditya>: [ge hodiyo]$^{NON-CE}$
(...)
<aditya>: [sathodana]$^{NON-CE}$
(…)
<aditya>: [le ille iddini kane]$^{NON-CE}$
(…)
<aditya>: [yallargu aussi ne baku]$^{NON-CE}$
(…)
<ribs>: [hey mate]^{CS} [what's this language?]^{CS}
<NeGaRiSh~AsS~*>: [i dont know <ribs>]^{CS}
<aditya>: [this is gali language]^{CS}
<aditya>: [u will not understand it <ribs>...]^{CS}
(Chat excerpt / Room V, session 4)

In this Web Chat excerpt, <aditya> writes some of his/her contributions in what is later identified as Gali language (spoken in Indonesia). However, since none of the fellow Chatters respond in the same language, it can be assumed that they do not understand it. They are thus not in a position to process any of <aditya's> contributions, which therefore result in non-contextual effects. Also, one of the Chatters, <ribs>, enquires what sort of language <aditya> is writing in (*what's this language?*), to which <NeGaRiSh~AsS~*> responds *i don't know <ribs>*. Both these contributions are textual units resulting in contextual strengthening, as they make a clear reference to previous textual units and can be deduced from textual evidence alone. Finally, <aditya> solves the riddle by identifying the language of his/her previous contributions. However, he/she then points out that the Chatter who enquired after the language (<ribs>), is not able to understand it. The question as to what sort of communicative goal <aditya> has in mind in contributing turns in a language that none of the other Chatters understand remains, however, unanswered.

In terms of distribution, textual units resulting in non-contextual effects only occur in Web Chat and the personal homepages on a regular basis, albeit, as illustrated above, under different circumstances. Nevertheless, both these communicative settings are relatively new to the field of personal correspondence, and it seems as though difficulties

[89] If two or more participants chat in a language with each other that the remaining Chatters in the room do not understand, then the situation would be different. In this case, the textual units would only be processed as non-contextual effects by those Chatters that are not in the position to comprehend what the Chatters are conversing about in a foreign tongue. However, this particular scenario was not observed in the data analysed for this study. This is assumed to be connected to the wide array of Chat rooms available for every imaginable need.

arise in communicating with strangers in terms of conversational success. Also, with respect to the trend in the personal homepages to contain dysfunctional hyperlinks, it appears that this type of modern communication is more prone to textual units that are likely to be processed as non-contextual effects because of media-related constraints. It is frustrating to follow up a hyperlink of interest, and then be sidelined because the linked website no longer exists. Hence, hyperlinks raise several issues in connection with contextuality and conversational success.

12.3. Chapter summary

The investigation into textual units evaluated in terms of contextual effects showed several trends. For example, of all the textual units analysed, the vast majority (more than 90%) resulted in either contextual implications or contextual strengthenings with comparable frequencies across the five text types—apart from Web Chat which clearly contains more textual units processed as contextual strengthenings than as implications. However, with respect to the total frequencies, the modern text types show higher contextual implication-to-word ratios than EEC. As the decoding of new information that is context-dependent cannot be achieved by means of textual evidence alone, this is reminiscent of spoken conversation (where, for example, deictics are also often only interpretable together with the context established). Hence, H1 and H2 (hypothesising more features of orality in the modern text types than in EEC) and in reverse also H3 (predicting more features of literacy in EEC than in the modern media) and H4 (assuming differences in the contextualisation between the modern media and EEC) could all be confirmed.

Textual units resulting in contextual contradiction/elimination on the other hand, are more or less evenly distributed across the five media and only slightly more frequent in the more synchronous text types (SMS discourse and Web Chat). Also, in comparison to the total number of contextual effects, percentages for the contradiction/elimination are below 2% for all the media. The same is true for textual units processed as contextual weakening (predominantly softening the illocutionary force of preceding textual units) that do not occur more than 4% in any of the text types. This suggests that both the contradiction/elimination and the weakening are not typical contextual effects as brought about by the textual units of personal written communication, which may be related to the circumstance that in general (apart from Web Chat) only one author is responsible for the content of a particular message. Since contradiction/elimination and weakening of the context established do not help to strengthen it, the employment of both effects could be

seen as acting against one's interest (in establishing a context that serves as common ground for author and receiver) and thus tends to be avoided.

With respect to textual units processed as non-contextual effects, it was observed that they only occur in Web Chat and the personal homepage data. In Web Chat this is connected to the phenomenon of "flooding" or "spamming" a Chat room with contributions that are (or at least seem to be) completely unrelated to previous textual units and the context established—and thus not processable. In the personal homepages this is tied to numerous dysfunctional hyperlinks that make it impossible for the reader to deduce the input the hyperlinks were supposed to provide. Both trends are believed to be connected to the unknown readership in Web Chat and the personal homepages. In comparison, acquainted correspondents refrain from spamming each other and also tend to make sure that the hyperlinks they send each other are functional. As this type of effect is non-existent in the processing of the textual units from SMS discourse, email, and EEC, this supports H5 that hypothesises differences in the contextualisation of messages directed at an acquainted vs. unknown readership.

PART IV: CONCLUSION

13. Conclusion and outlook

> The characteristics of language as an interactive phenomenon have challenged traditional notions of linguistic structure and linguistic rules, suggesting a view of the relationship between language and context as a process that emerges and changes through time and space.
>
> (Goodwin & Duranti 1992: 31)

Recent work in a number of different fields in linguistics (intersecting with psychology and Internet studies) have called into question the adequacy of earlier definitions of context in favour of a more dynamic view of the relationship between linguistic and non-linguistic dimensions of communicative events (cf. Goodwin & Duranti 1992: 31). One prime example that takes the dynamics of language in use into account is Sperber and Wilson's (1995) Relevance Theory that not only accommodates the interpretation of utterances as highly context-dependent, but also considers that the context itself is modified by the effects that these utterances have on the hearer/reader in terms of comprehension.

This study has drawn on the concept of contextual effects from the larger framework of Relevance Theory (1995) in an attempt to incorporate the interrelation between utterance and context into an empirical investigation. Also, various other parameters in the reciprocal relationship of text and context have been considered in the analytical framework designed for this study. Applying this analytical framework to different types of communication, modern electronic messages vs. Early English correspondence, was aimed at shedding light onto the contextuality of personal communication. It is in this sense that this study seeks to support, first, the dynamic view of the relationship between language and communicative context, and second, capture the idiosyncrasies that new media like the Internet and the mobile phone have introduced into the field of personal written communication.

In the following pages, a conclusive summary will be provided that briefly highlights the main findings that the empirical investigation generated, relates them to each other and discusses them in the context of the main research hypotheses H1 – H5 (13.1.). I will then evaluate the analytical framework that was designed for the purposes of this

study and consider the possibility of a computerised coding system (13.2.). The chapter will close with reflections on the future of personal written communication in terms of its use in private communication and provide perspectives for the analysis of personal correspondence from an academic point of view (13.3.).

13.1. Conclusive summary

Previous work in the field of modern written high speed telecommunication, in particular CMC and SMS discourse, has suggested that these forms of correspondence are reminiscent of speech. In the case of CMC, it has been argued that it contains features that go beyond orality and literacy, also referred to as cyberdiscursivity. In the terminology of Koch and Oesterreicher (1994), this means that in a communicative immediacy-distance continuum, CMC and SMS discourse are placed closer to the communicative immediacy pole (and in the case of CMC also go beyond the continuum). On the other hand, more traditional forms of correspondence, such as the handwritten letter, are perceived as closer to the pole of communicative distance. With respect to personal written communication, which is directed to a (specific or implied) readership, modern telecommunication has also made it possible for strangers to converse on a personal level. Hence, the field of personal written communication has experienced some fundamental changes since the public gained access to these types of correspondence in early 1990. The five main research hypotheses (H1 – H5) were formulated in consideration of these changes. In the conclusive summary that is to follow, I will discuss the main findings of the empirical investigation in relation to H1 – H5 and highlight which features of the different text types confirm the research hypotheses, and which ones have shown to be contradictory.

Chapter 6 looked into structural aspects of the five types of correspondence and the uses of different communication channels. Some of which are reminiscent of spoken discourse (for example, the use of emoticons and onomatopoeia) and others are subject to media-related features and constraints (such as the incorporation of hyperlinks and moving text segments). The analysis of the structural aspects of the different text types placed a focus onto the insertion of paragraphs as well as how greetings and farewells are incorporated into the messages. With respect to how authors structure their messages, no striking differences between the text types were observed (excluding SMS discourse, which is constrained by media-related features in connection with paragraph insertion). However, EEC was found to be the most conventional of all text types by including greetings and farewells in almost every message. Regarding the available communication channels, it

was observed that the authors of messages produced on more synchronous media make more use of onomatopoeic language, emoticons, and different means of emphasis (such as capitals). Authors of SMS discourse, one of the more synchronous text types, tend to compensate for lack of cue-bandwidth by more frequently employing those communication channels that are available to them (in particular emoticons and means of emphasis, such as capitals and the repetition of individual letters). Interestingly, the uses of emoticons and onomatopoeia were found to be less frequent than is sometimes suggested. Furthermore, the incorporation of signs and hyperlinks was only observed to be frequent in the personal homepages. Concerning the structural aspects and uses of available communication channels, it was found that overall, H1 (assuming features of orality and cyberdiscursivity for CMC), H2 (hypothesising features of orality for SMS discourse), H3 (claiming EEC to contain more features of literacy) and H4 (conjecturing differences in the contextualisation of modern communication vs. letter writing) could be largely confirmed. With respect to H5 (predicting differences in the contextualisation of correspondence that is aimed at an acquainted vs. unknown readership), no tendencies could be determined.

The focus of chapter 7 was placed on how authors contextualise their correspondence in terms of personal references. This area of research yielded results that can be tied to the type of readership that the messages are aimed at. It was observed that SMS discourse, e-mail, and EEC all feature authorial self-address more frequently than Web Chat and the personal homepages. With respect to readership address, however, Web Chat (together with SMS discourse) features the highest and the personal homepage the lowest frequency, whereas the other three text types show comparable frequencies. The most typical readership address in Web Chat was observed to be by means of (anonymous) nicknames, which was classified as a feature of cyberdiscursivity (confirming the second part of H1). Authorial address of a third party is most recurrent in e-mail and EEC. This is believed to be tied to text length, allowing the introduction of a third party, and if no introduction accompanies such third party addresses, then this implies that the reader is expected to be familiar with these third parties. Also quite frequent are third party addresses in the personal homepages, predominantly coupled with a hyperlink providing more information on these people. Again, this was classified as a feature of cyberdiscursivity (and thereby supporting the second part of H1). All in all, however, H1 – H4 (all of which assume differences between the modern text types and EEC) could not be entirely confirmed in connection with forms of address. It turned out that the main

findings of chapter 7 indicate that personal reference is less media-related and more tied to the kind of readership that is addressed, verifying H5.

In chapter 8, the investigation was focused on different types of textual reference. It was found that with respect to text-internal (endophoric) reference, SMS discourse, e-mail, and EEC feature fewer references than Web Chat and the personal homepages. Concerning text-external (exophoric) references, the main trends include a preference for forward over backward references in all text types except EEC. However, with regard to different types of exophoric forward references, a higher frequency of references to upcoming telephone calls and face-to-face meetings was observed in media directed at an acquainted readership. If frequencies and distributions of endophoric and exophoric references are compared, then a clear trend towards contextualising personal written communication in terms of text-internal structuring devices could be determined. Not many features distinctive of either modern telecommunication or the letters were found. The findings of chapter 8 thus do not (or only partly) confirm H1 – H4. It emerged, however, that textual references have a strong tendency to be tied to the type of readership, which means that, again, only H5 could be fully confirmed.

The notion of spatial deixis was dealt with in chapter 9. It turned out that the deictic origo in SMS discourse, e-mail, and EEC is predominantly oriented towards the location of the author (at writing time), while the Web Chat data and the personal homepage corpus both show a tendency for the text as deictic origo. In addition, both Web Chat and the personal homepages are idiosyncratic in their references to virtual space, which was classified as a feature of cyberdiscursivity and thereby confirms the second part of H1. It was once more not possible to determine distinctive differences between the modern text types and EEC, which means H2 – H4 could not be verified. In comparison, authors who direct their correspondence to an unacquainted readership tend to refrain from using their location as deictic origo because it is not essential to the communicative goal. While the participants in Web Chat and the authors of the personal homepages have shown tendencies to discuss or reveal their locations in the "real world", they have also shown an even greater interest to deal with their readership on a textual level. This is assumed to be related to the circumstance that the connection between authors and readers in Web Chat and the personal homepages is in most cases restricted to virtual space and thus remains textual. Communicative partners of SMS discourse, e-mail, and EEC, on the other hand, are all more likely to meet in person, hence, the communicative goal of these media is often related to reflect on past meetings or organise future dates. This confirms, once more, H5.

The focus of chapter 10 was placed on how authors of personal written communication contextualise their messages from a temporal point of view. The investigation into the use of the tenses showed that the more synchronous media favour present over past tense constructions. Furthermore, if authors of the more synchronous media include temporal expressions (fixed or unfixed to the calendar), then they are most likely to refer to present time at writing time. EEC and the personal homepages, on the other hand, feature lower frequencies for present tense and higher frequencies for past tense constructions in comparison to the other three text types. With respect to the overall frequency of temporal expressions, both Web Chat and EEC feature them to a lesser extent. This means that the degree of synchrony does not seem to play a role in the use of temporal expressions. However, the tendency of CMC and SMS discourse to feature more present tense constructions, and to favour temporal expressions that refer to present time at writing time, is reminiscent of spoken discourse, which confirms H1 (except for the personal homepages) and H2. In reverse, the tendency of EEC to contain more past tense constructions can be classified as a feature of literacy, which supports H3. In contrast, no general tendencies distinguishing the modern text types from the Early English letter could be determined, and the same is true for the type of readership. Consequently, H4 and H5, respectively, could not be verified.

The investigation into grammatical mood and syntactic structures was the main focus of chapter 11. The findings exposed the trend that the combination indicative mood and affirmative syntactic structures overrule hypothetical mood and negation. By far the most frequent syntactic structure in all text types is the affirmative declarative. This means that authorial intention as expressed in grammatical mood and syntactic structures is geared towards the reporting of affirmative statements. Since affirmative declarative clausal and non-clausal units distribute evenly across the five text types, no distinctive tendencies could be determined. It is a different picture with interrogatives and inserts, which are clearly more frequent in the more synchronous text types Web Chat, SMS discourse, and e-mail. As both these syntactic types are reminiscent of spoken discourse structure, it confirms H1 with respect to Web Chat and e-mail, and it fully confirms H2. Regarding exclamative clauses and non-clauses, it was observed that they are only frequent in the SMS text corpus, which supports H2. Formal and softened imperatives, on the other hand, are most frequent in SMS discourse and the personal homepages, indicating that their distribution is not connected to the immediacy of the correspondence. In reverse, interrogatives, inserts, and exclamations all occur rarely in EEC. As these three types of

syntactic structures are all reminiscent of spoken discourse, this verifies H3 (assuming EEC to be rich in features of literacy and, hence, show few features of orality). Furthermore, inserts are, apart from greetings and farewells, practically non-existent in the letter corpus, distinguishing EEC from all other text types in this respect and thus supporting H4. However, overall, H5 could not be validated because grammatical mood and syntactic structures show tendencies that can be tied to the different communicative settings, but not to the type of readership.

The last chapter of Part III, chapter 12, was concerned with the analysis of textual units and the different types of contextual effects they yield in the comprehension process of the readership (as discussed elaborately in 12.2., the concept of contextual effects, taken from the larger framework of Relevance Theory (Sperber & Wilson 1995), underwent considerable adaptations for the purposes of this study). It was found that the contextual implication and the contextual strengthening are the effects that occur most frequently in the processing of the textual units from all the five types of correspondence (amounting to more than 90% of all effects for each medium). Yet while the distribution of the contextual strengthening-to-word ratios is comparable across the five text types, the modern media have shown to be considerably higher in their contextual implication-to-word ratios than EEC. As the decoding of new information that is highly context-dependent cannot be achieved by means of textual evidence alone (and hence, results in a contextual implication), this is reminiscent of spoken conversation (where, for example, deictic expressions are often only successfully processable together with the context established) and distinguishes EEC in this respect from the modern text types. Thus, in connection with textual units processed as contextual effects of the type implication, H1 – H4 could be verified. Similar to the contextual strengthening, the frequencies for textual units processed as contextual effects of the type contradiction/elimination and weakening are comparable across the five media. However, the percentages in relation to the total number of contextual effects yielded, with 2% for the contextual contradiction/elimination and 4% for the contextual weakening, are a lot lower. This means that analogue to the contextual strengthening, no distinctive patterns in the distribution of textual units processed as contextual contradiction/elimination and weakening could be determined. The notion of non-contextual effect, on the other hand, has shown to be selective in its distribution and only apparent in the processing of textual units from Web Chat and the personal homepages. This was observed to be connected to spamming in Web Chat and recurrent dysfunctional hyperlinks in the homepages, both of which can be related to the

circumstance that authors do not know their readership. Authors that write to an acquainted readership normally refrain from spamming and also usually do not include dysfunctional hyperlinks in their correspondence. Hence, the idiosyncrasies connected to the notion of non-contextual effect can be tied to the readership and thus confirms H5.

Conclusively, the empirical investigation yielded two main tendencies: first, electronic communication is not collectively different in its contextualisation compared to Early English correspondence, and second, whether written communication is directed at an acquainted or unknown readership has considerable influence on how personal messages are contextualised. H1 hypothesised that CMC, being immediate in nature and multifaceted in its communication channels, is both rich in contextual features typical for orality and shows features of cyberdiscursivity. However, while the investigation into different aspects of Web Chat tended to confirm the first part of H1 (features of orality) but not the second, personal homepages have shown to be more supportive of the second part of H1 (features of cyberdiscursivity) but not the first. With e-mail, on the other hand, it was in many areas not possible to confirm either part of H1. Also, e-mail was found to be in many aspects closer to the letter than to the other types of CMC.

Regarding SMS discourse, being immediate in nature but limited in its available communication channels, H2 hypothesised that regardless of the leanness of the medium, it ought to be rich in features of orality. It turned out that H2 could be confirmed in the majority of the aspects that were investigated. Only in the uses of textual references and spatial deictics was it not possible to find distinctive patterns that support H2. In connection with H3 and EEC, being asynchronous in nature and limited in its available communication channels, it was assumed that EEC is rich in features of literacy. While the structural and syntactical aspects of EEC are indeed reminiscent of literacy, it was with several other investigated areas not possible to fully confirm H3. For example, high frequencies of pronominal readership address, textual references and use of spatial deictics all did not confirm H3. It was further not possible to entirely support H4, which predicted considerable differences in how modern messages are contextualised compared to 17th-century letters—this is, however, as argued above, in several of the investigated aspects not the case. The main research hypothesis H5, predicting differences in connection with acquainted vs. unknown readership, was verified in all areas except use of communication channels (no tendencies could be determined), temporal deixis, and distribution of syntactic structures. Hence, the influence on the contextuality of personal

written communication by the medium is in many cases overruled by the type of readership that is addressed.

13.2. Critical evaluation of analytical framework (code system)

The analytical framework that was designed for the purposes of this study has been shown to be applicable to personal written communication in order to investigate many relevant aspects in connection with its contextuality. In addition, the empirical investigation, the coding and the analysis of the data, yielded results that expose clear tendencies in connection with the contextuality of personal correspondence as influenced by media-related features and/or the type of readership that is addressed. The code system also accommodates the case when interpersonal messaging takes place in a larger context of communication (in particular in parent codes 03, "intra-/inter-textuality" and 10, "contextual effects") and is not restricted to the individual messages. However, owing to constraints in connection with data collection, it was often not possible to investigate those larger structures in detail because all natural language corpora (except the Web Chat data) were comprised of isolated messages. Also, in connection with the Web Chat data, it was necessary to adapt the code system with respect to parent code 03 ("personal deixis") in order to accommodate the circumstance that several authors are involved in the production of the discussions. Other than that the code system was perfectly applicable on all of the five text types. It is in this sense that the analytical framework is believed to be a useful contribution to the field of discourse analysis.

There is, however, one problematic issue in connection with the analytical framework. I cannot see, at this point, how the code system as a whole could be automated to run through a set of data without a researcher closely surveying the coding process. Particularly in connection with the parent codes that deal with different kinds of deixis (03 – 08). Furthermore, it seems difficult, if not impossible, to investigate authorial intention as expressed through mood and syntactic structures (parent code 09), or issues connected to relevance (parent code 10) by means of a computer program doing the coding. The problems with an automated coding of language in connection with deixis are obvious. How can a computer be programmed to identify whether backward and forward references are endophoric or exophoric, or distinguish between the author's location and the text as deictic origo? With respect to personal reference, it appears highly problematic to teach a computer to differentiate, for example, between all the potential referents of the second person personal pronoun *you*, particularly in view of the fact that it can also be used

as an impersonal pronoun. The same is true for the identification of the time reference of verb phrase constructions that, depending on the context, may index different points in time. Furthermore, while a computer program may be successful in coding punctuated clausal units, unpunctuated clauses and all non-clausal material would pose problems. Also, the notion of a contextual effect does not lend itself easily to automated coding.

Conclusively, the code system was classified as applicable, but with the constraint that it does not appear to be possible to conduct an automated coding of the data. In order to ensure the reliability of the code system as applied by humans, samples of all five text corpora were test-coded by three independent coders until agreement between coders reached above 70% for all codes in each of the five text types (see also 5.2.1.). However, this means that with current standards of information technology, the investigation into the contextuality of personal written communication relies on the human mind to conduct the coding of the data. This is a disadvantage of the code system because it involves many hours of careful coding by trained person(s). It is to be hoped that the future will bring forward computer programming that facilitates empirical research into topics as complex as the contextuality of language in use.

13.3. What the future holds

Correspondence via the computer and the mobile phone has only been available to the wide public since the beginning of the 1990's. This is a short time span if compared to hundreds of years of letter writing. The empirical investigation into the contextuality of personal written communication showed on many occasions that the modern text types cannot be classified as considerably different in how they are contextualised. However, the more synchronous text types did show tendencies to be more reminiscent of spoken conversation than the more asynchronous ones. It seems as though users bring implicit knowledge of other "communication templates" to the experience of composing electronic messages, including the handwritten genres of the personal letter or the post card, as well as the oral genres of face-to-face communication and telephone conversation (cf. Danet 1997: OD). The question is what happens to electronic correspondence when it eclipses handwritten genres, which can then no longer serve as templates?

Opinions differ on whether or not electronic correspondence will at some point replace handwritten messages. With respect to the online survey, in which one of the open questions addressed this issue, around two thirds of the participants stated that they do not think that handwritten correspondence will ever be replaced by electronic messaging.

However, another third felt that this process has already started and stated that they themselves rarely or never compose and/or receive handwritten messages. Many of the supporters of handwritten communication commented that it is more personal, shows more effort on the author's behalf, is more authentic, and that the handwriting itself is irreplaceable in its emotional effect on the reader. Furthermore, it was pointed out in several answers that handwriting will prevail because it is an essential skill, and that both handwriting and handwritten messages are cultural values that should not be lost. However, according to the other third of the informants, we are already in the process of losing handwritten correspondence. Many stated that electronic communication is easier, faster, and also "cooler" than handwritten messages. It was further emphasised that this is a sign of the times, and that communication has become more relaxed in general. In a similar vein, others stated that handwritten messages seem old-fashioned, and that a fast transmission of the messages is more important than the emotional value of handwriting. This includes traditional events and festivities such as birthdays and Christmas, where many refrain from sending handwritten cards but send their wishes via SMS text or e-mail.

The answers in the online survey deviate with regard to whether or not handwritten communication will be eclipsed by electronic messaging. Hence, it will have to be monitored closely what the future holds for personal written communication. Not only in terms of whether or not modern text types will replace handwritten texts at some point, but also how users adapt, or completely change, the way they communicate as the modern media undergo further technological developments. One illustrative example is SMS discourse, where messages were restricted to 160 characters at the time of data collection in 2003/2004. Four years later and this has evolved into SMS discourse with messages of unlimited length[90] and most current mobile phone software also features the possibility to send and receive e-mail. It was pointed out in the discussion of the findings that some of the results in connection with SMS discourse might have been different if I had worked with a newer corpus of SMS messages unlimited in length.

I therefore included a question in the online survey that addressed this issue and informants were asked if they tended to write SMS texts that are longer than 160 characters,

[90] Of course, in the days where SMS texts were limited to 160 characters, users could send more than one SMS text to communicate more content. This was, however, rarely done as it involved the repetition of the same procedure (activating the service, writing and sending the message) and in addition, every SMS text was charged separately. Nowadays, users may write SMS messages of unlimited length, although the software indicates at 160 characters that the space of the SMS message is used up and that a "new" message is being added. This means less effort on the user's behalf in terms of handling the text messaging, but the user is still charged separately for every 160 characters he/she uses.

and whether or not they send e-mail from their mobile phones. The vast majority answered that they rarely compose SMS texts of more than 160 characters and avoid sending e-mail via their mobile phones for reasons of impracticality (tedious typing due to small keys). There was a general agreement among the informants that if a message does not fit into one SMS text, then it is probably more appropriate to call or write an e-mail. However, mobile telephony is developing fast and it is assumed that user habits will change accordingly. This means that more research into the communicative exchange via the short message service will be needed.

The same is true of the three other types of electronic communication investigated in this study. Personal e-mail, Web Chat, and the personal homepage are all bound to undergo changes as technology advances. This, however, makes it at the same time difficult to investigate these media. For example, of the 60 personal homepages that were collected in November 2006 and monitored until January 2007, six had closed down as of February 2007, and another four had morphed into Web Blogs, more commonly referred to as "Weblogs"[91]. This means that 10% of the original personal homepage corpus underwent severe changes, whereas the other 90% were updated at varying frequencies. Web Chat, on the other hand, is problematic with respect to data collection. It has been pointed out that Chat discussions were not only *in medias res* when data recording started, but the content of the Chat discussion was only available in plain text. This means that the recorded Chat sessions are in general extracts rather than text entities, and that it is not possible to investigate features connected to format based on these Chat logs. Hence, different research tools are needed in order to be able to better capture the idiosyncrasies of the communicative exchanges that take place in Chat rooms.

This is, normally, where specific areas for future research would be suggested. However, this study has addressed such a wide array of contextual features in the field of personal written communication, with a particular focus on modern correspondence, that it does not seem appropriate to list each of those many features and recommend how exactly other researchers could be contributing to gain more insight. Also, it has been pointed out several times throughout the study that there is indeed ample room for future research in many areas. Therefore, it is suggested on a general note that more work is needed in the

[91] Weblogs are similar to personal homepages in that they are often aimed at portraying someone's personae and his/her life online. However, it is an even newer online phenomenon than the personal homepages, and not much is known about the idiosyncrasies of Weblogs. They differ from personal homepages in the arrangement of the content (newest contributions go on top, hence the reading process in reversed) as well as in several other features. See Bausch et al. (2002), a guide to creating and maintaining Weblogs, for a good overview of the (short) history of Weblogs.

field of how personal messages are contextualised. It is further assumed that it would be particularly insightful to work with text corpora that contain longer communicative exchanges (rather than isolated messages), so as to be able to track larger discourse structures more concretely. Above all in view of the circumstance that the modern media are undergoing changes right at this moment and that computer-mediated communication as we know it today (at the beginning of the 21st century) is indeed vastly different than what we saw in the early nineties (cf. Krikorian et al. 2000: OD). More important still, "new technologies are being developed even before old ones have had time to stabilise (and hence be examined)," and "what was new is rapidly becoming old and without looking ahead, one falls behind" (Krikorian et al. 2000: OD).

In a similar vein, Androutsopoulos and Ziegler (cf. 2003: 253) argue that regardless of the significant findings that linguistic research into the different aspects of electronic communication has yielded, we are still far away from an adequate description and explanation of the manifold varieties of electronic correspondence—both from a *use*- as well as *user*-related perspective. In an attempt to reduce this gap between praxis and theory, this investigation is aimed at contributing insight into how users approach (and make use of), from a contextual point of view, modern media in comparison to an older form of correspondence, the Early English letter. It is believed that this study is another step towards the goal of trying to describe and explain how changes in the communicative context, as brought about by the Internet and mobile telephony, influence the way we communicate.

14. Bibliography

Aijmer, K. 2006. "Introduction." In: K. Aijmer & A.-M. Simon-Vandenbergen (eds.), *Pragmatic Markers in Contrast*. In: *Studies in Pragmatics* 2. Amsterdam: Elsevier, 1-10.

Akman, V. & F. N. Alpaslan. 1999. "Strawson on Intended Meaning and Context." In: P. Bouquet, P. Brezillon & L. Serafini (eds.), *Modelling and Using Context*. Proceedings to 2nd International and Interdisciplinary Conference, Context '99. Heidelberg: Springer, 1-14.

Androutsopoulos, J. & G. Schmidt. 2001. "SMS-Kommunikation: Ethnografische Gattungsanalyse am Beispiel einer Kleingruppe." Online Document: PDF, 3-31. http://www.ids-mannheim.de/prag/sprachvariation/tp/tp7/SMS-Kommunikation.pdf [28.10.2008]

Androutsopoulos, J. & E. Ziegler. 2003. "Sprachvariation und Internet: Regionalismen in einer Chat-Gemeinschaft." In: *Standardfragen: Soziolinguistische Perspektiven auf Sprachgeschichte, Sprachkontakt und Sprachvariation*. Frankfurt am Main: Peter Lang, 251-279.

Arnold, J. & H. Miller. 2003. "Self in Web Home Pages: Gender, Identity and Power in Cyberspace." In: G. Riva & C. Galimberti (eds.), *Towards CyberPsychology: Mind, Cognitions and Society in the Internet Age*. Amsterdam: IOS Press, 73-94.

Asher, R. E. 1994. *The Encyclopaedia of Language and Linguistics*. Oxford: Pergamon Press.

Askehave, I. & A. Ellerup Nielsen. 2005. "What are the Characteristics of Digital Genres? – Genre Theory from a Multi-modal Perspective." Online Document: PDF, 1-8. http://csdl2.computer.org/comp/proceedings/hicss/2005/2268/04/22680098a.pdf [28.10.2008]

Austin, J. L. 1962. *How to Do Things with Words: The William James Lectures delivered at Harvard University in 1955*. J. O. Urmson (ed.). Oxford: Clarendon Press.

Baker, P. 2006. *Using Corpora in Discourse Analysis*. London: Continuum.

Barber, C. 1997. *The English Language: A Historical Introduction*. Cambridge: Cambridge University Press.

Baron, N. S. 2000. *Alphabeth to Email. How Written English Evolved and Where It's Heading*. London & New York: Routledge.

Baron, N. S. 2005. "The Future of Written Culture – Envisioning Language in the New Millennium." Online Document: PDF, 1-40. http://www.american.edu/lfs/tesol/In%20Press%20Paper--Future%20of%20Written%20Culture.pdf [28.10.2008]

Bartsch, R. 1995. *Situations, Tense, and Aspect. Dynamic Discourse Ontology and the Semantic Flexibility of Temporal System in German and English*. Berlin & New York: de Gruyter.

Bausch, P., M. Haughey & M. Hourihan. 2002. *We Blog : Publishing Online with Weblogs*. Indianapolis: Wiley Publishing.

Bays, H. 1998. "Framing and Face in Internet Exchanges: A Socio-cognitive Approach." In: *Linguistik Online*: *Computer-mediated Communication* 1.1. Online Document: HTML. http://www.linguistik-online.de/bays.htm [28.10.2008]

Beard, A. 2004. *Language Change*. London & New York: Routledge.

Bechar, H. 1995. "From <Bonehead> to <cLoNehEAd>: Nicknames, Play, and Identity on Internet Relay Chat." In: M. McLaughlin & S. Rafaeli (eds.), *Journal of Computer-Mediated Communication* 1.2. Online document: HTML. http://jcmc.indiana.edu/vol1/issue2/bechar.html [28.10.2008]

Bergs, A. T. 2004. "Letters: A new Approach to Text Typology." In: I. Taavitsainen & A. H. Jucker (eds.), *Journal of Historical Pragmatics* 5.2. Amsterdam: Benjamins, 207-227.

Bergs, A. T. 2006. "Analysing Online Communication from a Social Network Point of View: Questions, Problems, Perspectives." In: B. Ben-Israel, M. Bensoussan & D. Stein (eds.), *Language@Internet* 1.3. Online Document: PDF, 1-17. PDF located at: http://www.languageatinternet.de/articles/2006 [28.10.2008]

Berry, M. D. 2004. "Internet Research: Privacy, Ethics and Alienation – An Open Source Approach." In: *Studies in Social and Political Thought* 9. Online Document: PDF, 53-71. http://opensource.mit.edu/papers/berry2.pdf. [28.10.2008]

Bertacco, M. & B. Deponte. 2005. "E-mail as a Speed-Facilitating Device: A Contribution to the Reduced-Cues Perspective on Communication." In: S. C. Herring (ed.), *Journal of Computer-Mediated Communication* 10.3. Online document: HTML. http://jcmc.indiana.edu/vol10/issue3/sohn.html [28.10.2008]

Biber, D., S. Johansson, G. Leech, S. Conrad & E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Blakemore, D. 1987. *Semantic Constraints on Relevance*. Oxford: Blackwell.

Blakemore, D. 2001. "Discourse and Relevance Theory." In: D. Schiffrin, D. Tannen & H. E. Hamilton (eds.), *The Handbook of Discourse Analysis*. Oxford: Blackwell, 100-118.

Blakemore, D. 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.

Blakemore, D. 2006. "Discourse Markers." In: G. Ward & L. Horn (eds.), *The Handbook of Pragmatics*. Oxford: Blackwell, 221-240.

Boardman, M. 2005. *The Language of Websites*. London: Routledge.

Bolter, J. D. 2001. *Writing Space: Computers, Hypertext, and the Remediation of Print* (2nd edition). Mahwah (NJ): Lawrence Erlbaum Associates Inc.

Booher, D. 2001. *E-writing. 21st – Century Tools for Effective Communication*. New York: Pocket Books, Simon & Schuster, Inc.

Brown, B., N. Green & R. Harper (eds.). 2002. *Wireless World – Social and Interactional Aspects of the Mobile Age*. London: Springer-Verlag.

Brown, P. & S. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Bryant, J. A., A. Sanders-Jackson & A. M. K. Smallwood. 2006. "IMing, Text Messaging, and Adolescent Social Networks." In: S. C. Herring (ed.), *Journal of Computer-Mediated Communication* 11.2. Online document: HTML. http://jcmc.indiana.edu/vol11/issue2/bryant.html [28.10.2008]

Bühler, K. 1934. "The Deictic Field of Language and Deictic Words." Reprinted 1982 in: R. Jarvella & W. Klein (eds.), *Speech, Place and Action: Studies of Deixis and Related Topics*. New York: John Wiley, 9-30.

Carlson, L. 1984. *'Well' in Dialogue Games. A Discourse Analysis of the Interjection 'well' in Idealised Conversation*. Amsterdam: Benjamins.

Carston, R. 1988. "Implicature, Explicature, and Truth-Theoretic Semantics." In: R. Kempson (ed.), *Mental Representations: The Interface between Language and Reality*. Cambridge: Cambridge University Press, 155-181.

Carston, R. 1998. "The Semantics/pragmatics Distinction: A View from Relevance Theory." In: J. Harris & C. Iten (eds.), *UCL Working Papers in Linguistics* 10. Online Document: PDF, 1-53. http://www.phon.ucl.ac.uk/publications/WPL/98papers/carston.pdf [28.10.2008]

Carston, R. 2002. "Linguistic Meaning, Communicated Meaning, and Cognitive Pragmatics." In: *Mind and Language* 17.1, 127-148.

Chandler, D. 1998. "Personal Home Pages and the Construction of Identities on the Web." Online Document: HTML. http://www.aber.ac.uk/media/Documents/short/webident.html [28.10.2008]

Chandler, D. & D. Roberts-Young. 1998. "The Construction of Identity in the Personal Homepages of Adolescents." Online Document: HTML. http://www.aber.ac.uk/media/Documents/short/strasbourg.html [28.10.2008]

Chung, S. and A. Timberlake. 1985. "Tense, Aspect, and Mood." In: T. Shopen (ed.), *Language Typology and Syntactic Description: Grammatical Categories and the Lexicon* 3. Cambridge: Cambridge University Press.

Claridge, C. 2001. "Structuring Text: Discourse Deixis in Early Modern English Texts." In: *Journal of English Linguistics* 29. SAGE Publications. Online Document: PDF, 55-71. PDF located at: http://eng.sagepub.com [28.10.2008, subject to charges]

Coe, L. 1995. *The Telephone and Its Several Inventors: A History*. North Carolina: McFarland.

Conrad, B. 1982. *Referring and Non-referring Phrases: Studies in the Use of the Gerund and the Infinitive*. Copenhagen: Akademisk Forlag.

Crystal, D. 2001. *Language and the Internet*. Cambridge: Cambridge University Press.

Crystal, D. 2003. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.

Crystal, D. 2004. *The Language Revolution*. Cambridge: Polity Press.

Curwen, P. 2002. *The Future of Mobile Communications. Awaiting the Third Generation*. Houndmills: Palgrave Macmillan.

Cusack, B. (ed.). 1998. *Everyday English 1500-1700. A Reader*. Edinburgh: Edinburgh University Press.

Daft, R. L. & R. H. Lengel. 1984. "Information Richness: A New Approach to Managerial Behaviour and Organisation Design." In: L. L. Cummings & B. M. Staw (eds.), *Research in Organisational Behaviour* 6. Homewood: JAI Press, 191-233.

Danet, B. 1996. "The Ethnography of Writing." Online Document: HTML. http://pluto.mscc.huji.ac.il/~msdanet/ethno.html [28.10.2008]

Danet, B. 1997. "Language, Play and Performance in Computer-mediated Communication." Final Report submitted to the Israel Science Foundation. Online Document: HTML. http://pluto.mscc.huji.ac.il/~msdanet/report95.htm [28.10.2008]

Danet, B. 2002. "The Language of Email." Online Document: PDF, 1-27. http://pluto.mscc.huji.ac.il/~msdanet/papers/email.pdf [28.10.2008]

De Saint-Georges, I. 1998. "Click Here if *You* Want to Know Who *I* Am: Deixis in Personal Homepages." In: *System Sciences* 2. Proceedings of the Thirty-First Hawaii International Conference. Online Document: PDF, 68-79. PDF located at: http://ieeexplore.ieee.org/Xplore/guesthome.jsp [28.10.2008, subject to charges]

December, J. 1993. "Characteristics of Oral Culture in Discourse on the Net." Paper presented at the 12th annual Penn State Conference on Rhetoric and Composition. University Park, Pennsylvania, July 8th, 1993. Online Document: HTML. http://www.december.com/john/papers/pscrc93.txt [28.10.2008]

Declerck, R. 1991. *Tense in English. Its Structure and Use in Discourse*. London: Routledge.

Diller, H.J. & M. Görlach (eds.). 2001. *Towards a History of English as a History of Genres*. Heidelberg: Carl Winter.

Dittmann, M. 2001. *Sprachverwendung im Internet. Untersuchungen zu Sprache und Nutzung des Internet Relay Chat (IRC) in Deutschland und Frankreich*. Sarlat: Editions Indoles.

Döring, N. 2002a. "Kurzm. wird gesendet - Abkürzungen und Akronyme in der SMS-Kommunikation." In: *Muttersprache. Vierteljahresschrift für Deutsche Sprache* 112.2, 97-114.

Döring, N. 2002b. "1x Brot, Wurst, 5Sack Äpfel I.L.D. Kommunikative Funktionen von Kurzmitteilungen (SMS)." In: P. Vorderer (ed.), *Zeitschrift für Medienspsychologie* 3. Online Document: PDF, 118-128. http://nicola-doering.de/publications/sms-funktionen-doering-2002.pdf [28.10.2008]

Döring, N. 2002c. "Personal Home Pages on the Web: A Review of Research." In: M. McLaughlin & S. Rafaeli (eds.), *Journal of Computer-Mediated Communication* 7.3. Online Document: HTML. http://jcmc.indiana.edu/vol7/issue3/doering.html [28.10.2008]

Döring, N. 2003. *Sozialpsychologie des Internet* (2nd edition). Göttingen: Hogrefe, Verlag für Psychologie.

Duffley, P. J. 1992. *The English Infinitive*. In: R. Quirk (ed.), *English Language Series* 19. London & New York: Longman.

Dürscheid, C. 2002a. "E-mail und SMS – ein Vergleich." In: A. Ziegler & C. Dürscheid (eds.), *Kommunikationsform E-mail*. Tübingen: Stauffenburg Verlag, 94-114.

Dürscheid, C. 2002b. "SMS-Schreiben als Gegenstand der Sprachreflexion." In: *Networx. Die Online-Schriftenreihe des Projekts Sprache@Web* 28. Online Document: PDF, 1-26. http://www.mediensprache.net/networx/networx-28.pdf [28.10.2008]

Dürscheid, C. 2005. "E-mail – verändert sie das Schreiben?" In: T. Siever, P. Schlobinski & J. Runkehl (eds.), *Websprache.net. Sprache und Kommunikation im Internet*. In: *Linguistik - Impulse & Tendenzen* 10. Berlin & New York: de Gruyter, 85-97.

Dresner, E. & M. Dascal. 2001. "Semantics, Pragmatics, and the Digital Information Age." In: E. Rigotti et al. (eds.), *Studies in Communication Sciences* 1.2., 1-22.

Erickson, T. 1996. "The World Wide Web as Social Hypertext." Online Document: HTML. http://www.pliant.org/personal/Tom_Erickson/SocialHypertext.html [28.10.2008]

Faulstich, W. 2006. *Mediengeschichte von 1700 bis ins 3. Jahrtausend*. Göttingen: Vandenhoeck & Ruprecht.

Fennell, B. 2001. *A History of English: A Sociolinguistic Approach*. Massachusetts: Blackwell Publishers.

Ferguson, Ch. 1994. "Dialect, Register, and Genre: Working Assumptions about Conventionalisation." In: D. Biber & E. Finegan (eds.), *Sociolinguistic Perspectives on Register*. Oxford: Oxford University Press, 15-30.

Fetzer, A. 2004. *Recontextualising Context: Grammaticality Meets Appropriateness*. In: A. H. Jucker (ed.), *Pragmatics & Beyond New Series* 121. Amsterdam: Benjamins.

Fetzer, A. (ed.). 2007. *Context and Appropriateness: Micro Meets Macro*. In: A. H. Jucker (ed.), *Pragmatics & Beyond New Series* 162. Amsterdam: Benjamins.

Fillmore, C. J. 1997. *Lectures on Deixis*. Stanford: CSLI Publications.

Finnegan, R. 1988. *Literacy and Orality: Studies in the Technology of Communication*. Oxford: Blackwell.

Fitzmaurice, S. M. 2002. *The Familiar Letter in Early Modern English: A Pragmatic Approach*. Amsterdam: Benjamins.

Flynn, N. & T. Flynn. 2003. *Writing Effective E-Mail*. California: Crisp Learning.

Fox, K. 2001. "Evolution, Alienation and Gossip. The Role of Mobile Telecommunications in the 21st Century." Social Issues Research Centre, Oxford. Online Document: HTML. http://www.sirc.org/publik/gossip.shtml [28.10.2008]

Francis, W. N. 1982. "Problems of Assembling and Computerising Large Corpora." In: S. Johansson (ed.), *Computer Corpora in English Language Research*. Bergen: Norwegian Computing Centre for the Humanities, 7-24.

Fraser, B. 1996. "Pragmatic Markers." In: *Pragmatics* 6.2, 167-190.

Fraser, B. 1999. "What are Discourse Markers?" In: *Journal of Pragmatics* 31.7. Online Document: PDF, 931-952.
http://vaporiser.student.utwente.nl/~timmie/afstuderen/fraser.pdf [28.10.2008]

Furniss, G. 2004. *Orality: The Power of the Spoken Word*. New York: Palgrave Macmillan.

Gains, J. 1998. "Electronic Mail: A New Style of Communication or Just a New Medium? An Investigation into the Text Features of E-mail." In: *English for Specific Purposes* 18.1, 81-101.

Garton L., C. Haythornthwaite & B. Wellman. 1997. "Studying Online Social Networks." In: M. McLaughlin & S. Rafaeli (eds.), *Journal of Computer-Mediated Communication* 3.1. Online document: HTML.
http://jcmc.indiana.edu/vol3/issue1/garton.html [28.10.2008]

Giora, R. 1997. "Discourse Coherence and Theory of Relevance: Stumbling Blocks in Search of a Unified Theory." In: *Journal of Pragmatics* 27.1, 17-34.

Geser, H. 2004. "Towards a Sociological Theory of the Mobile Phone." Online Document: PDF, 1-47. http://socio.ch/mobile/t_geser1.pdf [28.10.2008]

Goffman, E. 1963. *Interaction Ritual. Essays on Face-to-Face Behaviour*. New York: Anchor Books.

Goodwin, C. & A. Duranti. 1992. "Rethinking Context: An Introduction." In: A. Duranti & C. Goodwin (eds.), *Rethinking Contex. Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press, 1-42.

Görlach, M. 2004. *Text Types and the History of English*. In: W. Bisang, H. H. Hock & W. Winter (eds.), *Trends in Linguistics, Studies and Monographs* 139. Berlin & New York: de Gruyter.

Goutsos D. 2005. "The Interaction of Generic Structure and Interpersonal Relations in Two-Party E-Chat Discourse." In: D. Stein (ed.), *Language@Internet* 3. Online Document: PDF, 1-39.
http://www.languageatinternet.de/articles/188/Goutsos0308_DOULOS.rtf.pdf
[28.10.2008]

Green, G. M. 2006. "Some Interactions of Pragmatics and Grammar." In: G. Ward & L. Horn (eds.), *The Handbook of Pragmatics*. Oxford: Blackwell, 411-426.

Grice, H. P. 1975. "Logic and Conversation." In: P. Cole & J. L. Morgan (eds.), *Syntax and Semantics: Speech Acts* 3. New York: Academics, 41-58.

Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge: The MIT Press.

Grunder, A. 2001. "Forming the Text, Performing the Work—Aspects of Media, Navigation, and Linking." In: H. Francke (ed.), *Human IT* 5, 81-206.

Haase, M., M. Huber, A. Krumeich & G. Rehm. 1997. "Internetkommunikation und Sprachwandel." In: R. Weingarten (ed.), *Sprachwandel durch Computer*. Opladen: Westdeutscher Verlag, 51-85.

Hanks, W. F. 1992. "The Indexical Ground of Deictic Reference." In: A. Duranti & C. Goodwin (eds.), *Rethinking Context. Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press, 43-76.

Hentschel, E. 1998. "Communication on IRC." In: *Linguistik Online*: *Computer-mediated Communication* 1.1. Online Document: HTML.
http://www.linguistik-online.de/irc.htm [28.10.2008]

Haarmann, P.-L. 1994. "Copyright." In: H. Günther & O. Ludwig (eds.), *Schrift und Schriftlichkeit* 10.1. Berlin & New York: de Gruyter, 898-902.

Heritage, J. 1984. *Garfinkel and Ethnomethodology*. Oxford: Basil Blackwell.

Herring, S. C. 1996. "Linguistic and Critical Research on Computer-mediated Communication: Some Ethical and Scholarly Considerations." In: R. Kling (ed.), *The Information Society* 12.2, 153-168.

Herring, S. C. 1999. "Interactional Coherence in CMC." In: M. McLaughlin & S. Rafaeli (eds.), *Journal of Computer-Mediated Communication* 4.4. Online Document: HTML. http://jcmc.indiana.edu/vol4/issue4/herring.html [28.10.2008]

Herring, S. C. 2001. "Computer-mediated Discourse." In: D. Schiffrin, D. Tannen & H. E. Hamilton (eds.), *The Handbook of Discourse Analysis*. Oxford: Blackwell, 612-634.

Herring, S. C. 2004a. "Computer-mediated Discourse Analysis: An Approach to Researching Online Behaviour." In: S. A. Barab, R. Kling & J. H. Gray (eds.), *Designing for Virtual Communities in the Service of Learning*. New York: Cambridge University Press.

Herring, S. C. 2004b. "Content Analysis for New Media: Rethinking the Paradigm." In: *New Research for New Media: Innovative Research Methodologies Symposium Working Papers and Readings*, 2004. Online Document: PDF, 1-16. http://ella.slis.indiana.edu/~herring/newmedia.pdf [28.10.2008]

Herring, S. C. 2007. "A Faceted Classification Scheme for Computer-Mediated Discourse." In: D. Stein (ed.), *Language@Internet* 1. Online Document: PDF, 1-37. http://www.languageatinternet.de/articles/761/Faceted_Classification_Scheme_for_CMD.pdf [28.10.2008]

Herring, S. C., K. Börner & M. B. Swan. 2003. "When Rich Media Are Opaque: Spatial Reference in a 3-D Virtual World." Online Document: PDF, 1-56. http://ivl.slis.indiana.edu/km/pub/2003-herring-whenrich.pdf [28.10.2008]

Higashimori, I. & D. Wilson. 1996. "Questions on Relevance." In: *UCL Working Papers in Linguistics* 8. Online Document: PDF, 1-14. http://www.phon.ucl.ac.uk/publications/WPL/96papers/higashi.pdf [28.10.2008]

Hiltz, S. R. & M. Turoff. 1993. *The Network Nation: Human Communication via Computer*. Cambridge (Mass.): MIT Press.

Höflich, J. R. 2001. "Das Handy als 'persönliches Medium' – Zur Aneignung des Short Message Service (SMS) durch Jugendliche." In: C. Stegbauer, J. Schmidt & K. Schönenberger (eds.), *kommunikation@gesellschaft* 2. Online document: PDF: 1-19. http://www.soz.uni-frankfurt.de/K.G/B1_2001_Hoeflich.pdf [28.10.2008]

Höflich, J. R.. 2003. "Einleitung: Mediatisierung des Alltags und der Wandel von Vermittlungskulturen." In: Höflich, J. R. & J. Gebhart (eds.), *Vermittlungskulturen im Wandel*. Frankfurt am Main: Peter Lang, 7-20.

Hörisch, J. 2004. *Von der Oblate zum Internet*. Frankfurt am Main: Suhrkamp Taschenbuch.

Hundt, M., N. Nesselhauf, & C. Biewer (eds.). 2007. *Corpus Linguistics and the Web*. Amsterdam: Rodopi.

Ifantidou, E. 2001. *Evidentials and Relevance*. In: A. H. Jucker (ed.), *Pragmatics & Beyond New Series* 86. Amsterdam: Benjamins.

Ishii, K. 2000. "A Comparative Study of Personal Web Pages." Paper presented at IAMCR (Internet Development in the Asia Pacific) 2000. Singapore, July 19th, 2000. Online Docment: PDF, 1-19.
http://infoshako.sk.tsukuba.ac.jp/~ishii/IAMCR2000.pdf [28.10.2008]

Jacobsen, M. M. 2002. *Transformations of Literacy in Computer-Mediated Communication*. New York: The Edwin Mellen Press.

Jakobs, E.-V. & K. Lehnen. 2005. "Hypertext – Klassifikation und Evaluation." In: T. Siever, P. Schlobinski & J. Runkehl (eds.), *Websprache.net. Sprache und Kommunikation im Internet*. In: *Linguistik - Impulse & Tendenzen* 10. Berlin & New York: de Gruyter, 159-184.

Jucker, A. H. 1993. "The Discourse Marker *well*: A Relevance-theoretic Account." In: *Journal of Pragmatics* 19, 435-52.

Jucker, A. H. & Y. Ziev. 1998. *Discourse Markers: Descriptions and Theory*. In: A. H. Jucker (ed.), *Pragmatics & Beyond New Series* 57. Amsterdam: Benjamins.

Kasesniemi, L. & P. Rautiainen. 2003. "Das Leben in 160 Zeichen: Zur SMS-Kultur finnischer Jugendlicher." In: J. R. Höflich, & J. Gebhart (eds.), *Vermittlungs-kulturen im Wandel*. Frankfurt am Main: Peter Lang, 291-311.

Kauffman, R. J. 2005. "Is There a Global Digital Divide for Digital Wireless Phone Technologies?" Online Document: PDF, 1-39.
http://www.misrc.umn.edu/workingpapers/fullPapers/2005/0501_011505.pdf
[28.10.2008]

Karlsson, A.-M. 2002. "Web Literacy, Web Literacies or just Literacies on the Web?—Reflections from a Study of Personal Homepages." In: *The Reading Matrix* 2.2. Online Document: PDF, 1-19.
http://www.readingmatrix.com/articles/karlsson/article.pdf [28.10.2008]

Kehler, A. 2006. "Discourse Coherence." In: G. Ward & L. Horn (eds.), *The Handbook of Pragmatics*. Oxford: Blackwell, 241-265.

Kilian, J. 2001. "Geschriebene Umgangssprache in computervermittelter Kommunikation. Historisch-kritische Ergänzungen zu einem neuen Feld der linguistischen Forschung." In: M. Beisswenger (ed.), *Chat Kommunikation*. Stuttgart: Ibidem, 55-78.

Kim, H., G. J. Kim, H. W. Park & R. E. Rice. 2007. "Configurations of Relationships in Different Media: FtF, Email, Instant Messenger, Mobile Phone, and SMS." In: S. C. Herring (ed.), *Journal of Computer-Mediated Communication* 12.4. Online document: HTML. http://jcmc.indiana.edu/vol12/issue4/kim.html [28.10.2008]

Koch, P. & W. Oesterreicher. 1990. *Gesprochene Sprache in der Romania*. Tübingen: Niemeyer.

Koch, P. & W. Oesterreicher. 1994. "Schriftlichkeit und Sprache." In: H. Günther & O. Ludwig (eds.), *Schrift und Schriftichkeit* 10.1. Berlin & New York: de Gruyter, 587-604.

Koch, P. 1999. "Court Records and Cartoons: Reflections of Spontaneous Dialogue in Early Romance Texts." In: A. H. Jucker, G. Fritz & F. Lebsanft (eds.), *Historical Dialogue Analysis*. In: A. H. Jucker (ed.), *Pragmatics & Beyond New Series* 66. Amsterdam: Benjamins, 399-430.

Krikorian, D., J.-S. Lee, T. Makana Chock & Ch. Harms. 2000. "Isn't that Spatial? Distance and Communication in a 2-D Virtual Environment." In: M. McLaughlin & S. Rafaeli (eds.), *Journal of Computer-Mediated Communication* 5.4. Online Document: HTML. http://jcmc.indiana.edu/vol5/issue4/index.html [28.10.2008]

Langacker, R. W. 2001. "The English Present Tense." In: R. Hogg (ed.), *English Language and Linguistics* 5. Cambridge Journals Online: 251-272 / PDF: 1-19. http://mind.ucsd.edu/syllabi/05-06/cogs200/English%20Present.pdf [28.10.2008]

Lakoff, R. 1973. "Questionable Answers and Answerable Questions." In: B. B. Kachru, R. B. Lees, Y. Malkiel, A. Pietrangelui & S. Saporta (eds.), *Issues in Linguistics. Papers in Honour of Henry and Renée Kahane*. Urbana (IL): University of Illinois Press, 453-467.

Leech, G. 2004. *Meaning and the English Verb* (3rd edition). London: Longman.

Lenk, U. 1998. *Marking Discourse Coherence: Functions of Discourse Markers in Spoken English*. Tübingen: Gunter Narr Verlag.

Levinson, S. C. 1983. *Pragmatics.* Cambridge & New York: Cambridge University Press.

Levinson, S. C. 2000. *Presumptive Meanings. The Theory of the Generalised Conversational Implicature*. Cambridge: The MIT Press.

Levinson, S. C. 2006. "Deixis." In: G. Ward & L. Horn (eds.), *The Handbook of Pragmatics*. Oxford: Blackwell, 97-121.

Ling, R. 2005. "The Socio-linguistics of SMS: An Analysis of SMS Use by a Random Sample of Norwegians." In: R. Ling & P. Pedersen (eds.), *Mobile Communications: Renegotiation of the Social Sphere*. London: Springer, 335-349.

Lynn Weber, B. 1988. "Tense as Discourse Anaphor." In: *Computational Linguistics* 14.2. Online Document: PDF, 61-73. http://ucrel.lancs.ac.uk/acl/J/J88/J88-2006.pdf [28.10.2008]

Loehr, D. 2002. "Hypertext and Deixis." Online Document: PDF, 29-38. http://acl.ldc.upenn.edu/W/W97/W97-1404.pdf [28.10.2008]

Mancini, C., C. Pietsch & D. Scott. 2007. "Visualising Discourse Structure in Interactive Documents." Online Document: PDF, 1-4. http://mcs.open.ac.uk/ds5473/publications/mancini-pietsch-scott-ewnlg07.pdf [28.10.2008]

Marshall, J. 2004. *Language Change and Sociolinguistics: Rethinking Social Networks*. New York: Palgrave Macmillan.

Mc Enery, T. 2006. *Corpus-Based Language Studies: An Advanced Resource Book*. London: Routledge.

Mejías, U. A. 2004a. "Online Discourse: Past, Present and Future." Online Document: PDF, 1-15. http://eduforge.org/docman/view.php/29/109/mejias_online_discourse.pdf [28.10.2008]

Mejías, U. A. 2004b. "Distributed Textual Discourse: A New Approach to Online Discourse." Online Document: PDF, 1-16. http://eduforge.org/docman/view.php/29/117/mejias_dtd.pdf [28.10.2008]

Mitra, A. 1999. "Characteristics of the WWW Text: Tracing Discursive Strategies." In: M. McLaughlin & S. Rafaeli (eds.), *Journal of Computer-Mediated Communication* 5.1. Online Document: HTML. http://jcmc.indiana.edu/vol5/issue1/mitra.html [28.10.2008]

Moessner, L. 2001. "Genre, Text Type, Style, Register: A Terminological Maze?" In: H. Grabes, C. Bernard & C. Connelly (eds.), *European Journal of English Studies* 5.2. London: Routledge, 131-138.

Mori & Lycos (Media Group). 2000. "I just Text to Say I Love You." Press Release 5.9.2000. Online Document: HTML. http://www.ipsos-mori.com/polls/2000/lycos.shtml [28.10.2008]

Nevala, M. 2004. "Inside and Out: Forms of Address in Seventeenth- and Eighteenth Century Letters." In: I. Taavitsainen & A. H. Jucker (eds.), *Journal of Historical Pragmatics* 5.2. Amsterdam: Benjamins, 271-296.

Nevalainen, T. 2004. "Letter Writing: Introduction". In: I. Taavitsainen & A. H. Jucker (eds.), *Journal of Historical Pragmatics* 5.2. Amsterdam: Benjamins, 181-191.

Nevalainen, T. & H. Raumolin-Brunberg (eds.). 1996. *Sociolinguistics and Language History. Studies based on the Corpus of Early English Correspondence*. Amsterdam: Rodopi.

Nevalainen, T. & H. Raumolin-Brunberg. 2003. *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. London: Longman.

Nickisch, R. M. G. 2003. "Der Brief – historische Betrachtungen." In: J. R. Höflich & J. Gebhart (eds.), *Vermittlungskulturen im Wandel*. Frankfurt am Main: Peter Lang, 63-73.

Noh, E.-J. 2000. *Metarepresentation: A Relevance-Theory Approach*. In: A. H. Jucker (ed.), *Pragmatics & Beyond New Series* 69. Amsterdam: Benjamins.

Olson, D. R. & N. Torrance (eds.). 1991. *Literacy and Orality*. Cambridge: Cambridge University Press.

Ong, W. J. 2002. *Orality and Literacy: The Technologising of the Word.* London & New York: Routledge.

Paolillo, J. 1999. "The Virtual Speech Community: Social Network and Language Variation on IRC." In: M. McLaughlin & S. Rafaeli (eds.), *Journal of Computer-Mediated Communication* 4.4. Online Document: HTML. http://jcmc.indiana.edu/vol4/issue4/paolillo.html [28.10.2008]

Paraboni, I. & K. van Deemter. 1999. "Issues for the Generation of Document Deixis." In: *Proceedings of the Workshop on Deixis, Demonstration, and Deictic Belief at ESSLLI XI*, Paper 7. University of Brighton, UK. Online Document: PDF, 44-48. http://folli.loria.fr/cds/1999/library/pdf/paraboni-etal.pdf [28.10.2008]

Plant, S. 2000. "On the Mobile. The Effects of Mobile Telephones on Social and Individual Life." Online Document: PDF, 18-88. http://www.motorola.com/mot/doc/0/234_MotDoc.pdf [28.10.2008]

Pons, S. 1997. "A Prototype Approach to the Concept of Connective." Unpublished Manuscript, University of Valencia.

Preece, J., D. Maloney-Krichmar & C. Abras. 2003. "History of Online Communities." Online Document: PDF, 1-11. http://www.ifsm.umbc.edu/~preece/Papers/Community_%20Encyclopedia_03.pdf [28.10.2008]

Quirk, R., S. Greenbaum, G. Leech & J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. New York: Longman.

Rehm, G. 2002. "Schriftliche Mündlichkeit im World Wide Web." In: A. Ziegler & C. Dürscheid (eds.), *Kommunikationsform E-mail*. Tübingen: Stauffenburg Verlag, 263-308. / Online Document: PDF: 1-32. http://www.uni-giessen.de/germanistik/ascl/dfg-projekt/pdfs/Muendlichkeit-im-Web.pdf [28.10.2008]

Raible, W. 1994. "Orality and Literacy." In: H. Günther & O. Ludwig (esd.), *Schrift und Schriftlichkeit* 10.1. Berlin & New York: de Gruyter, 1-17.

Rheingold, H. 2000. *The Virtual Community. Homesteading on the Electronical Frontier* (revised edition, originally published in 1993). Cambridge, Massachusetts & London: The MIT Press.

Ribera, J. 2007. "Text Deixis in Narrative Sequences." In: J. M. Hernández-Campoy (ed.), *International Journal of English Studies* 7.1. Online Document: PDF: 149-168. Document located at: http://dialnet.unirioja.es/servlet/dcfichero_articulo?codigo=2541006&orden=0 [28.10.2008]

Ridings, C. M. & D. Gefen. 2004. "Virtual Community Attraction: Why People Hang Out Online." In: S. C. Herring (ed.), *Journal of Computer-Mediated Communication* 10.1. Online Document: HTML. http://jcmc.indiana.edu/vol10/issue1/ridings_gefen.html [28.10.2008]

Riva, G. & C. Galimberti. 1998. "Computer-mediated Communication: Identity and Social Interaction in an Electronic Environment." Online Document: PDF, 1-46. http://www.cybertherapy.info/pages/cmc.pdf [28.10.2008]

Romaine, S. 1994. *Language in Society: An Introduction to Sociolinguistics*. Oxford: Oxford University Press.

Runkehl, J., P. Schlobinski & T. Siever. 1998. *Sprache und Kommunikation im Internet. Überblick und Analysen*. Opladen: Westdeutscher Verlag.

Schiffrin, D. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Schiffrin, D. 1994. *Approaches to Discourse: Language as Social Interaction*. Oxford: Blackwell.

Schlobinski, P. 2005. "Editorial: Sprache und Internetbasierte Kommunikation – Voraussetzungen und Perspektiven." In: T. Siever, P. Schlobinski & J. Runkehl (eds.), *Websprache.net. Sprache und Kommunikation im Internet*. In: *Linguistik - Impulse & Tendenzen* 10. Berlin & New York: de Gruyter, 1-14.

Schlobinski, P., N. Fortmann, O. Gross, F. Hogg, F. Horstmann & R. Theel. 2001. "Simsen. Eine Pilotstudie zu sprachlichen und kommunikativen Aspekten in der SMS-Kommunikation." In: *NETWORX-Online-Publikationen zum Thema Sprache und Kommunikation im Internet* 1.2. Online Document: HTML. http://www.mediensprache.net/networx/networx-22.pdf [28.10.2008]

Schmidt, G. & J. Androutsopoulos. 2004. "löbbe döch. Beziehungskommunikation mit SMS." In: A. Deppermann & M. Hartung (eds.), *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 5. Online Document: PDF, 50-71. http://www.gespraechsforschung-ozs.de/heft2004/ga-schmidt.pdf [28.10.2008]

Schulze, R. 1985. "A Selective Bibliography of Recent Writings on the English Tense and Aspect System." In: W. Hüllen & R. Schulze (eds.), *Tempus, Zeit und Text*. Heidelberg: Carl Winter, 11-38.

Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language*. New York: Cambridge University Press.

Short, J. E., E. Williams & B. Christie. 1976. *The Social Psychology of Telecommunication*. New York: Wiley.

Shortis, T. 2001. *The Language of ICT: Information and Communication Technology*. London: Routledge.

Smith, C. S. 2003. *Modes of Discourse: The Local Structure of Text*. Cambridge: Cambridge University Press.

Snowden, C. 2000. "Blinded by Text: Revaluing The Oral Imperative in Communication." Paper presented at the Communications Research Forum 2000. Department of Transport Economics and Communication, Canberra, 1-12.

Sperber, D. & D. Wilson. 1995. *Relevance: Communication and Cognition* (2nd edition with an additional postface and notes, originally published in 1986). Oxford: Basil Blackwell.

Sperber, D. & D. Wilson. 1997. "Remarks on Relevance Theory and the Social Sciences." In: R. J. Watts (ed.), *Multilingua* 16, 145-151.

Sperber, D. & D. Wilson. 2004. "Relevance Theory." In: *UCL Working Papers in Linguistics* 14. Online Document: PDF, 249-290. http://people.bu.edu/bfraser/Relevance%20Theory%20Oriented/Sperber%20&%20Wilson%20-%20RT%20Revisited.pdf [28.10.2008]

Stegbauer, C. 2003. "Form und Beziehung am Beispiel schriftlicher Kommunikation." In: J. R. Höflich & J. Gebhart (eds.), *Vermittlungskulturen im Wandel*. Frankfurt am Main: Peter Lang, 75-94.

Stein, D. 2006. "The Website as a Domain-Specific Genre." In: B. Ben-Israel, M. Bensoussan & D. Stein (eds.), *Language@Internet* 1.6. Online Document: PDF, 1-9. http://www.languageatinternet.de/articles/LangInternetSpecialVolume1/374/The_website_as_a_domain-specific_genre.pdf [28.10.2008]

Storrer, A. 2000. "Schriftverkehr auf der Datenautobahn: Besonderheiten der schriftlichen Kommunikation im Internet." In: G. Voss et al. (eds.), *Neue Medien im Alltag*. Leverkusen: Leske & Budrich, 153-177.

Storrer, A. 2001. "Getippte Gespräche oder dialogische Texte? Zur kommunikationstheoretischen Einordnung der Chat-Kommunikation." In: A. Lehr et al. (eds.), *Sprache im Alltag*. Berlin & New York: de Gruyter, 439-465.

Suler, J. 2003. "E-Mail Communication and Relationships." Online Document: HTML. http://www-usr.rider.edu/~suler/psycyber/emailrel.html [28.10.2008]

Svartvik, L. 1980. "'Well' in Conversation." In: S. Greenbaum, G. Leech & J. Svartvik (eds.*), Studies in English Linguistics for Randolph Quirk*. London: Longman, 167-177.

Tannen, D. (ed.). 1984. *Spoken and Written Language: Exploring Orality and Literacy*. In: *Advances in Discourse Progress* 9. New Jersey: ABLEX Publishing Cooperation.

Tanskanen, S. K. 2004. "Intertextual Networks in the Correspondence of Lady Katherine Paston." In: I. Taavitsainen & A. H. Jucker (eds.), *Journal of Historical Pragmatics* 5.2. Amsterdam: Benjamins, 255-269.

Tieken-Boon van Ostade, I. 2006. "'Disrespectful and too Familiar?' Abbreviations as an Index of Politeness in 18th-Century Letters." In: Ch. Dalton-Puffer, D. Kastovsky, N. Ritt & H. Schendl (eds.), *Syntax, Style, and Grammatical Norms: English from 1500-2000*. Berlin, New York: Peter Lang, 229-247.

Unger, Ch. 2006. *Genre, Relevance and Global Coherence: The Pragmatics of Discourse Type*. In: N. Burton-Roberst & R. Carston (eds.), *Palgrave Studies in Pragmatics, Language and Cognition Series*. New York: Palgrave MacMillan.

Van der Herst, J. B. & D. Sperber. 2004. "Testing the Cognitive and Communicative Principles of Relevance." In: I. A. Noveck & D. Sperber (eds.), *Experimental Pragmatics*. Basingstoke: Palgrave Macmillan, 229-279.

Wellman, B. 1997. "An Electronic Group is Virtually a Social Network." In: S. Kiesler (ed.), *Culture of the Internet: Research Milestones from the Social Sciences*. Mahwah, N.J.: Lawrence Erlbaum Associates, 179-208.

Wellman, B. 2001. "Physical Place and Cyber Place. The Rise of Personalised Networking." In: *International Journal of Urban and Regional Research* 25. Online Document: HTML. http://www.chass.utoronto.ca/~wellman/publications/individualism/ijurr3a1.htm [28.10.2008]

Wenz, H. 1998. "Formen der Mündlichkeit und Schriftlichkeit in digitalen Medien." In: *Linguistik Online*: *Computer-mediated Communication* 1.1. Online Document: HTML. http://www.linguistik-online.de/wenz.htm [28.10.2008]

Wilkinson, S. & C. Kitzinger (eds.). 1995. *Feminism and Discourse: Psychological Perspectives*. Thousand Oaks (CA): Sage.

Wilson, D. 1994. "Relevance and Understanding." In: G. Brown, K. Malmkjaer, A. Pollitt, & J. Williams (eds.), *Language and Understanding*. Oxford: Oxford University Press, 35-58.

Wilson, D. & D. Sperber. 1993. "Linguistic Form and Relevance." In: *Lingua* 90, 1-25.

Wilson, D. & D. Sperber. 2006. "Relevance Theory." In: G. Ward & L. Horn (eds.), *Handbook of Pragmatics*. Oxford: Blackwell, 607-632.

Wirth, U. 2005. "Chatten. Plaudern mit anderen Mitteln." In: T. Siever, P. Schlobinski & J. Runkehl (eds.), *Websprache.net. Sprache und Kommunikation im Internet*. In: *Linguistik - Impulse & Tendenzen* 10. Berlin & New York: de Gruyter, 67-84.

Wood, J. L. 2004. "Text in Context: A Critical Discourse Analysis Approach to Margaret Paston." In: I. Taavitsainen & A. H. Jucker (eds.), *Journal of Historical Pragmatics* 5.2. Amsterdam: Benjamins, 229-254.

Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Wray, A. & M. R. Perkins. 2000. "The Functions of Formulaic Language: An Integrated Model." In: *Language and Communication* 20, 1-28.

Online-sources without reference names

1. *The British Postal Museum and Archive*
   http://www.postalheritage.org.uk [28.10.2008]

2. *Internet World Stats: Usage and Population Statistics*
   http://www.internetworldstats.com/stats.htm [28.10.2008]

3. *Research and Markets: World's Largest Market and Research Resource*
   http://www.researchandmarkets.co.uk [28.10.2008]

4. *Universities of Helsinki and Jyväskylä (FI): Research Unit for Variation, Contacts, and Change in English*
   http://www.eng.helsinki.fi/varieng/ [28.10.2008]

5. *Journal of Computer-mediated Communication (JCMC)*
   http://jcmc.indiana.edu/ [28.10.2008]

6. *The Oxford Text Archive*
   http://ota.ahds.ac.uk/ [28.10.2008]

7. *CEECS* (Corpus of Early English Correspondence Sampler) *Manual*
   http://khnt.hit.uib.no/icame/manuals/ceecs/index.htm [28.10.2008]

8. *Yahoo Web Directory*
   http://dir.yahoo.com/Society_and_Culture/People/Personal_Homepages/
   [28.10.2008]

9. *John Benjamins Publishing Company: Pragmatics & Beyond New Series*
   http://www.benjamins.com/cgi-bin/t_seriesview.cgi?series=P%26bns [28.10.2008]

15. Appendix

15.1. Code system

Table 15.1: Complete code system.

01	**COMMUNICATION CHANNELS**
	1.1. written word(s) 1.2. onomatopoeia 1.3. emoticons 1.3.1. happy (all variations) 1.3.2. sad (all variations 1.3.3. winking (all variations) 1.3.4. kiss (all variations) 1.3.5. surprised (all variations) 1.3.6. wry (all variations) 1.3.7. tongue sticking out (all variations) 1.3.8. other 1.4. signs 1.4.1. symbol 1.4.2. icon 1.4.3. index 1.4.4. combinatory 1.5. hyperlink 1.5.1. to a destination inside text entity 1.5.2. to a destination outside text entity 1.5.3. to a non-identifiable destination 1.5.4. other 1.6. photograph 1.7. picture (drawing, physical or digital) 1.8. logo 1.9. graphics 1.10. attachment, enclosure (other than photograph, picture) 1.11. video sequence 1.12. sound (spoken or sung text segment) 1.13. other, borderline case 1.16.1. contributions in a language other than English
02	**MEANS OF EMPHASIS**
	2.1. capitalised 2.2. bigger font 2.3. bold font 2.4. different font 2.5. italicised 2.6. underline 2.7. moving (text) segment 2.8. sound (alert) 2.9. use of colour 2.10. repetition of letter

	2.11. repetition of word 2.12. repetition of punctuation 2.12.1. full stop 2.12.2. exclamation mark 2.12.3. question mark 2.12.4. dash 2.12.5. mixed 2.13. repetition of other 2.13.1. quotation marks 2.13.2. hash 2.13.3. asterisk 2.13.4. brackets 2.13.5. other 2.14. different combination of means of emphasis 2.15. no use of means of emphasis
03	**INTER- / INTRA-TEXTUALITY**
	3.1. acknowledgment of receipt 3.2. reference to previous writing specified 3.2.1. same medium 3.2.2. other medium 3.3. reference to current writing specified 3.4. reference to upcoming writing specified 3.4.1. same medium 3.4.2. other medium 3.5. reference to contact face-to-face specified 3.5.1. previous 3.5.2. future 3.6. reference to contact telephone specified 3.6.1. previous 3.6.2. upcoming 3.7. reference to writing unspecified 3.7.1. previous 3.7.2. upcoming 3.8. reference to contact oral unspecified 3.8.1. previous 3.8.2. upcoming 3.9. reference to contact unspecified 3.9.1. previous 3.9.2. upcoming 3.10. quotation 3.11. discourse markers 3.11.1. above all 3.11.2. additionally, in addition 3.11.3. alas 3.11.4. and, & 3.11.5. anyway, anyhow 3.11.6. as for 3.11.7. besides

	3.11.8. but 3.11.9. by the way 3.11.10. despite 3.11.11. finally 3.11.12. generally 3.11.13. hence 3.11.14. however, howsoever 3.11.15. I mean, you know what I mean 3.11.16. in any case 3.11.17. in other words 3.11.18. last but not least 3.11.19. like 3.11.20. meanwhile 3.11.21. now 3.11.22. obviously 3.11.23. of course 3.11.24. oh, o, ah (+ combinations) 3.11.10.1. single or + other 3.11.10.2. + shoot 3.11.10.3. + great 3.11.10.4. + sorry 3.11.10.5. + well 3.11.11. ok (then) 3.11.12. or 3.11.13. other (than that), on the other (…) 3.11.14. otherwise 3.11.15. personally (speaking) 3.11.16. so / *soe* 3.11.17. so far 3.11.18. still 3.11.19. then 3.11.20. therefore / *therfor*, *tharfore* 3.11.21. though 3.11.22. thus 3.11.23. well 3.11.24. what can I say 3.11.25. yet 3.11.26. you know what 3.12. contact postal address 3.13. contact e-mail address 3.14. contact information other 3.16. other, borderline case
04	**PERSONAL DEIXIS**
	4.1. greeting section 4.2. farewell section 4.3. author 4.3.1. references to author other than 1. person pronoun 4.3.1.1. name, surname, full name, initial(s), title 4.3.1.2. nickname – derived from name 4.3.1.3. pet name 4.3.1.4. nickname – anonymous (i.e. Chat) 4.3.1.5. kinship terminology 4.3.1.6. other 4.3.1.7. unknown 4.3.2. all text segments directed at author 4.3.3. text entities without authorial self-address

4.4. readership
- 4.4.1. address of readership (other than pronouns)
  - 4.4.1.1. name, surname, full name, initial(s), title
  - 4.4.1.2. nickname – derived from name
  - 4.4.1.3. pet name
  - 4.4.1.4. nickname – anonymous (i.e. Chat)
  - 4.4.1.5. kinship terminology
  - 4.4.1.6. other
  - 4.4.1.7. unknown
- 4.4.2. all text segments directed at readership (other than pronouns)
  - 4.4.2.1. to a specific chatter
  - 4.4.2.2. to the whole room
  - 4.4.2.3. other
    - 4.4.2.3.2. specific group of chatters
    - 4.4.2.3.1. females only
    - 4.4.2.3.2. males only
  - 4.4.2.4. other, borderline case
- 4.4.3. text entities without address of readership

4.5. impersonal second person pronoun *you* / third person pronoun *one*

4.6. any other 3. party (other than pronouns)

4.7. ellipsis
- 4.7.1. author pronoun
- 4.7.2. receiver/visitor pronoun
- 4.7.3. any other 3. party pronoun

4.8. selected terms of non-pronominal personal reference
- 4.8.1. all
- 4.8.2. anybody
- 4.8.3. anyone / *any one*
- 4.8.4. both
- 4.8.5. boy, boys
- 4.8.6. each (other)
- 4.8.7. everyone / *every one*
- 4.8.8. folks
- 4.8.9. girl, girls
- 4.8.10. guy, guys
- 4.8.11. no one
- 4.8.12. nobody
- 4.8.13. other(s)
- 4.8.14. man, men
- 4.8.15. people
- 4.8.16. person
- 4.8.17. somebody
- 4.8.18. someone
- 4.8.19. together
- 4.8.20. relative pronoun (personal)
  - 4.8.18.1. that
  - 4.8.18.2. which
  - 4.8.18.3. who, whose, whom / *hoo*, *whoe*, *whos*, *whome*
- 4.8.21. woman, women

	Pronominal references 4.9. 1. person personal pronouns, singular 4.9.1. I 4.9.2. me / *mee* 4.9.3. mine / *myn*, *myne* 4.9.4. (my) 4.9.5. myself / *my self*, *my selfe* 4.10. 1. person personal pronouns, plural 4.10.1. we / *wee* 4.10.1.1. addressee included 4.10.1.2. addressee excluded 4.10.2. us / *vs* 4.10.2.1. addressee included 4.10.2.2. addressee excluded 4.10.3. (our / *oure*) 4.10.3.1. addressee included 4.10.3.2. addresses excluded 4.10.4. ours 4.10.4.1. addressee included 4.10.4.2. addressee excluded 4.11.5. ourselves / *our selves*, *oure selves* 4.11.5.1. addressee included 4.11.5.2. addressee excluded 4.11. 2. person personal pronouns, singular 4.11.1. you, ya, u, y, yee, thou, thee / *youe*, *yow*, *ye* 4.11.2. (your / *yors*, *yr*) 4.11.3. yours 4.11.4. yourself / *your selfe* 4.12. 2. person personal pronouns, plural 4.12.1. you 4.12.2. you all, you guys 4.12.3. (your) 4.12.4. yours 4.12.5. yourselves 4.12. 2. person pronouns, not clear whether singular or plural 4.13. 3. person personal pronouns, singular / female 4.14.1. she / *shee* 4.14.2. (her / *hir*) 4.14.3. hers 4.14.4. herself, *her self*, *herselfe* 4.15. 3. person personal pronouns, singular / male 4.15.1. he / *hee* 4.15.2. him / *hym* 4.15.3. (his) 4.15.4. himself / *him self* 4.17. 3. person personal pronouns, plural 4.17.1. they / *thy*, *thay*, *thaye* 4.17.2. them, ‘em, em 4.17.3. (their / *there, thire*) 4.17.4. theirs 4.17.5. themselves / *them selves*

	4.18. other, borderline case 4.18.1. author action in 3. person (Chat)
05	**SPATIAL DEIXIS**
	5.1. spatial references to (specific) physical places 5.1.1. area, region, town, city, canton, county 5.1.2. island, state, country, continent 5.1.3. the world, universe 5.1.4. public transport, station, airport 5.1.5. somebody's place (flat, house, estate) 5.1.6. workplace 5.1.7. restaurant, bar, pub, club, party 5.1.8. public institution, building 5.1.9. shop, shopping 5.1.10. any other location 5.2. spatial references to virtual places 5.2.1. current homepage (chat room) 5.2.2. other personal homepage / weblog 5.2.3. public website (news, business, commercial) 5.2.4. discussion forum 5.2.5. (other) chat room 5.2.6. Internet, Web 5.2.7. other service 5.2.8. other 5.3. selected spatial expressions (in consideration of deictic origo author vs. text) 5.3.1. (at) home / *hom* 5.3.2. above 5.3.3. abroad / *abrode* 5.3.4. across 5.3.5. among, amongst 5.3.6. anywhere 5.3.7. around 5.3.8. at 5.3.9. away 5.3.10. back 5.3.11. below 5.3.12. beside 5.3.13. beyond 5.3.14. bottom / *bottome*, *botome* 5.3.15. by 5.3.16. centre, central 5.3.17. channel 5.3.18. close 5.3.19. compass 5.3.19.1. north, northern 5.3.19.2. east, eastern 5.3.19.3. south, southern 5.3.19.4. west, western 5.3.20 domain 5.3.21 down, download / *doun* 5.3.22 far / *farr* 5.3.23 forward / *forwarde* 5.3.24 from 5.3.25 front, in front of 5.3.26 global 5.3.27 here / *heer*, *heere*, *hear*, *heare*, *her* 5.3.28 horizon, horizontal, horizontally

	5.3.27. house / *hous*, *hows*, *howse*, *howes* 5.3.28. in, inner, indoors, inside 5.3.29. international 5.3.30. km, m, cm / mile, foot, inch 5.3.31. left, left hand side, on the left 5.3.32. local, locally, locate, location 5.3.33. long, longer, longest 5.3.34. near, nearer, nearest, nearby / *neer*, *neere*, *nearer* 5.3.35. next to 5.3.36. on, onward(s) 5.3.37. online 5.3.38. onto 5.3.39. open 5.3.40. out, outdoors, outside / *owt* 5.3.41. over, come over / *ouer* 5.3.42. present (physically) 5.3.43. right, right hand side 5.3.44. room 5.3.45. short, shorter, shortest 5.3.46. somewhere 5.3.47. space, place / *plas* 5.3.48. that 5.3.49. there / *ther*, *thear*, *theare* 5.3.50. these / *theise* 5.3.51. this 5.3.52. those 5.3.53. through, throughout 5.3.54. top, on top of 5.3.55. towards 5.3.56. up, upon, upload / *vp*, *vppon* 5.3.57. verticality, vertical, vertically 5.3.58. way 5.3.59. where, whence / *wher*, *whear(e)* 5.3.60. wherever, wheresoever / *whersoever*, *whersoeuer* 5.3.61. whither 5.3.62. other, borderline case
06	**TEMPORAL DEIXIS – TENSE**
	6.1. present tense 6.1.1. present simple, present participle 6.1.1.1. present simple / present simple + infinitive 6.1.1.1.1. thereof referring to past time 6.1.1.1.2. thereof referring to present time 6.1.1.1.3. thereof referring to future time 6.1.1.2. present + modal 6.1.1.3. present + modal + infinitive 6.1.1.4. present + modal + present participle 6.1.1.5. present + present participle. 6.1.1.6. present + infinitive + present participle 6.1.1.7. present participle 6.1.1.8. present participle + present participle 6.1.1.9. present participle + infinitive 6.1.1.10. present + modal + present participle 6.1.1.11. present simple in passive voice 6.1.2. present continuous 6.1.2.1. present continuous + infinitive 6.1.2.2. present continuous + present participle 6.1.2.3. present continuous in passive voice

6.1.3. present perfect
6.1.3.1. present perfect + infinitive
6.1.3.2. present perfect + present participle
6.1.3.3. present perfect + modal
6.1.3.4. present perfect in passive voice

6.1.4. present perfect continuous
6.1.4.1. present perfect continuous + present participle
6.1.4.2. present perfect continuous + infinitive

6.2. past tense
6.2.1. past simple, past participle
6.2.1.1. past simple
6.2.1.2. past + modal
6.2.1.3. past + infinitive
6.2.1.4. past. + modal + infinitive
6.2.1.5. past + present participle
6.2.1.6. past participial phrases
6.2.1.7. past simple in passive voice

6.2.2. past continuous
6.2.2.1. past continuous + infinitive .
6.2.2.2. past continuous. + present participle
6.2.2.3. past continuous in passive voice

6.2.3. past perfect
6.2.3.1. past perfect + infinitive
6.2.3.2. past perfect + present participle
6.2.3.3. past perfect + modal
6.2.3.4. past perfect in passive voice

6.2.4. past perfect continuous

6.3. future tense
6.3.1. future – will
6.3.1.1. future simple / future simple + infinitive
6.3.1.1.1. thereof referring to present time
6.3.1.1.2. thereof referring to future time
6.3.1.2. future continuous
6.3.1.3. future perfect
6.3.1.4. future perfect continuous
6.3.1.5. future + present participle
6.3.1.6. future + modal + infinitive
6.3.1.7. future continuous + infinitive
6.3.1.8. future in passive voice

6.3.2. future – going to
6.3.2.1. future simple / future simple + infinitive
6.3.2.1.1. thereof referring to present time
6.3.2.1.2. thereof referring to future time
6.3.2.2. future continuous
6.3.2.3. future perfect
6.3.2.4. future perfect continuous
6.3.2.5. future + present participle
6.3.2.6. future + modal + infinitive
6.3.2.7. future continuous + infinitive
6.3.2.8. future in passive voice

	6.3.3. future – shall 6.3.3.1. future simple / future simple + infinitive 6.3.3.1.1. thereof referring to present time 6.3.3.1.2. thereof referring to future time 6.3.3.2. future continuous 6.3.3.3. future perfect 6.3.3.4. future perfect continuous 6.3.3.5. future + present participle 6.3.3.6. future + modal + infinitive 6.3.3.7. future continuous + infinitive 6.3.3.8. future in passive voice 6.4. infinitive constructions (without tensed verb) 6.5. constructions with "let" 6.6. constructions with "had better" 6.7. constructions with modal verb (no main verb) 6.8. constructions with periphrastic "do" / "did" 6.9. passives without tensed verb 6.10. other, borderline case 6.10.1. abbreviations – verb form not exclusively classifiable
07	**TEMPORAL DEIXIS – CALENDRIAL REFERENCES**
	7.1. (official) holidays 7.2. decade(s), century(ies), millennium 7.3. year 7.3.1. variations of "year(s)" (e.g., year, years, yr, yrs) / *yeer(s), yeare(s)*, *yeaur(s)* 7.3.2. in numbers e.g., 2006, '07) 7.3.3. specific date (d/m/y - or any 2 of the 3) 7.4. season 7.4.1. variations of "season(s)" / *sesone* 7.4.2. variations of "spring(s)" *7.4.3.* variations of "summer(s)" / *somer(s)*, *sommer(s)* 7.4.4. variations of "autumn(s)" 7.4.5. variations of "winter(s)" 7.5. month 7.5.1. proper noun (incl. abbr.) (e.g., January, march, Feb.) 7.5.2. variations of "month(s)" (e.g., month, months, m.) / *muthe*, *munth(s)*, *moneth(s), -(es)* 7.5.3. in numbers (1-12) 7.6. week *7.6.1.* variations of "week(s)" (e.g., week, weeks, w.) / *weeke(s)*, *weake(s)* 7.6.2. variations of "weekend(s)" (e.g., w-end, weekends) 7.6.3. in numbers (e.g., (in week) 4) 7.7. day 7.7.1. weekday: proper noun (incl. abbr.) (e.g., Friday, wednesday, Sat.) *7.7.2.* variations of "day(s)" (i.e. day, days, d.) / *daie(s)* 7.7.3. in numbers (e.g., 24th, 12., in 12 (days)) 7.8. time of day 7.8.1. clock 7.8.1.1. variations of "clock(s)" (e.g., (at 2) o'clock) 7.8.1.2. variations of "hour(s)" / *houre(s)*, *hower(s)* 7.8.1.3. variations of "minute(s)" 7.8.1.4. variations of "second(s)" 7.8.1.5. specific time of day (e.g., 2pm, 14:00h)

	7.8.2. particular sequence 7.8.2.1. variations of “morning(s)” 7.8.2.2. variations of “noon(s)”, “midday(s)” 7.8.2.3. variations of “afternoon(s)” 7.8.2.4. variations of “evening(s)” 7.8.2.5. variations of “midnight(s)” 7.8.2.6. variations of “night(s)"
08	**TEMPORAL DEIXIS – SELECTED TEMPORAL EXPRESSIONS**
	8.1. after 8.2. afterwards 8.3. age(s), for ages 8.4. ago / *agoe* 8.5. already / *alredy* 8.6. always 8.7. another time 8.8. anymore / *any more* 8.9. anytime / *any time* 8.10. at / *att* 8.11. at the moment (mo) 8.12. at the same time 8.13. before 8.14. beforehand 8.15. beginning, in/at the beginning / *begynynge* 8.16. birthday 8.17. breakfast (time) 8.18. brunch (time) 8.19. current, currently 8.20. date(s), update, updated 8.21. dinner, dindins (time) 8.22. due, overdue 8.23. early, earlier, earliest 8.24. end, in/at the end 8.25. estimated time of arrival (eta) 8.26. ever / *euer*, *eaver* 8.27. everyday, every day, daily / *dayly*, *daylie* 8.28. every time, moment, hour, minute, (…) 8.29. fast, faster, fastest 8.30. final, finally 8.31. first (time) 8.32. for good 8.33. forever / *for ever* 8.34. former, formerly 8.35. frequent, frequently 8.36. future 8.37. haste, in haste 8.38. holiday(s), vacation(s) (personal) 8.39. hurried, hurriedly 8.40. hurry, in a hurry 8.41. immediate, immediately / *imediate(ly)* 8.42. last, lastly 8.43. late, later, latest 8.44. lately 8.45. long, longer, longest 8.46. look/looking (…) forward to 8.47. lunch (time) 8.48. meanwhile, in the meantime / *mean whyle*, *in the meane tyme* 8.49. never / *neuer*, *neaver* 8.50. new

8.51. next
8.52. no more
8.53. now / *nowe* , *xxx*
8.54. often
8.55. old, older, oldest
8.56. once
8.57. original, originally
8.58. over / *ouer*
8.59. past
8.60. periodic, periodically
8.61. perpetual, perpetually
8.62. point, at any/that/some point
8.63. present, presently
8.64. prior, previous, previously
8.65. quick, quickly
8.66. rare, rarely
8.67. recent, recently
8.68. regular, regularly
8.69. right away
8.70. rush, in a rush
8.71. second (time)
8.72. short, shortly (time)
8.73. since / *sinc*, *sinse*, *sens*, *syns*
8.74. slow, slowly
8.75. some other time
8.76. sometime(s) / *some tyme*
8.77. soon + combinations / *soone* + combinations
  8.77.1. soon (+ other)
  8.77.2. as soon as (possible), asap
  8.77.3. sooner (or later)
  8.77.4. soonest
  8.77.5. until soon
  8.77.6. call (…) soon
  8.77.7. come (…) soon
  8.77.8. see (…) soon
  8.77.9. hear (…) soon
  8.77.10. speak/talk/chat (…) soon
  8.77.11. write/reply/respond (…) soon

8.78. still
8.79. the other day
8.80. then
8.81. through, throughout
8.82. time(s) / *tyme(s)*
8.83. today
8.84. tomorrow / *to-morowe*
8.85. tonight, tonite
8.86. until, till, til / *untill*, *vntill*
8.87. wait and see
8.88. when
8.89. whenever
8.90. while, whilst
8.91. yesterday / *yester night*
8.92. yet
8.93. young, younger, youngest / *younge*
8.94. other, borderline case

09	**MOOD & SYNTACTIC STRUCTURES**

9.1. indicative mood
- 9.1.1. declaratives
  - 9.1.1.1. affirmative
  - 9.1.1.2. negated
- 9.1.2. interrogatives
  - 9.1.2.1. affirmative
  - 9.1.2.2. negated
- 9.1.3. imperatives
  - 9.1.2.3. affirmative
  - 9.1.2.4. negated
- 9.1.4. softened imperatives
  - 9.1.4.1. affirmative
  - 9.1.4.2. negated
- 9.1.5. exclamatives
  - 9.1.5.1. affirmative
  - 9.1.5.2. negated
- 9.1.6. inserts
  - 9.1.6.1. interjection
  - 9.1.6.2. expletive
  - 9.1.6.3. greeting/farewell
  - 9.1.6.4. response forms
  - 9.1.6.5. hesitators
- 9.1.7. other non-clausal material
  - 9.1.7.1. declarative structure
  - 9.1.7.2. interrogative structure
  - 9.1.7.3. exclamative structure
- 9.1.8. mixed

9.2. hypothetical mood
- 9.2.1. declaratives
  - 9.2.1.1. affirmative
  - 9.2.1.2. negated
- 9.2.2. interrogatives
  - 9.2.2.1. affirmative
  - 9.2.2.2. negated
- 9.2.3. exclamatives
  - 9.2.3.1. affirmative
  - 9.2.3.2. negated
- 9.2.4. mixed

9.3. subjunctive mood
- 9.3.1. affirmative
- 9.3.2. negated

9.4. other, borderline case

10	**CONTEXTUAL EFFECTS**
	10.1. contextual implication 10.2. contextual strengthening 10.3. contextual contradiction/elimination 10.4. contextual weakening 10.5. non-contextual effect 10.6. other, borderline case
11	**ORGANISATION OF TEXT**
	11.1. greeting section 11.1.1. none 11.1.2. separated from main body by paragraph (or turn) 11.1.3. starts separated by paragraph (or turn) and continues into main body (or another turn) 11.1.4. not separated from main body by paragraph (another turn) 11.1.4.1. same paragraph as main body / separated by exclamation mark 11.1.4.2. same paragraph as main body / separated by period 11.1.4.3. same paragraph as main body / separated by comma 11.1.4.4. same paragraph as main body / separated by dash 11.1.4.5. same paragraph as main body / separated by colon 11.2.4.6. same paragraph as main body / separated by emoticon 11.2.4.7. same paragraph as main body / separated by other 11.1.4.8. same paragraph as main body / not separated 11.2. farewell section 11.2.1. none 11.2.2. separated from main body by paragraph (or turn) 11.2.3. starts in main body (or turn) and is separated at some point 11.2.4. not separated from main body by paragraph (or another turn) 11.2.4.1. same paragraph as main body / separated by exclamation mark 11.2.4.2. same paragraph as main body / separated by period 11.2.4.3. same paragraph as main body / separated by comma 11.2.4.4. same paragraph as main body / separated by dash 11.2.4.5. same paragraph as main body / separated by colon 11.2.4.6. same paragraph as main body / separated by emoticon 11.2.4.7. same paragraph as main body / separated by other 11.2.4.8. same paragraph as main body / not separated 11.3. number of paragraphs (or turns) 11.3.1. 01 paragraph (min. 1 word, symbol) 11.3.2. 02 paragraphs 11.3.3. 03 paragraphs 11.3.4. 04 paragraphs 11.3.5. 05 paragraphs 11.3.6. 06 paragraphs 11.3.7. 07 paragraphs 11.3.8. 08 paragraphs 11.3.9. 09 paragraphs 11.3.10. 10 paragraphs (…) 11.3.976. 976 paragraphs
12	**WORD COUNT**
	12.1. 0001-0010 words 12.2. 0011-0020 words 12.3. 0021-0030 words

	12.4. 0031-0040 words 12.5. 0041-0050 words 12.6. 0051-0100 words 12.7. 0101-0150 words 12.8. 0151-0200 words 12.9. 0201-0300 words 12.10. 0301-0400 words (…) 12.37. 3001+ words

Key: terms in italics = alternate spellings in EEC.

Table 15.2: Coding procedure "personal reference: forms of address".

	Forms of address	**Code**
A	authorial self-address: pronoun (e.g., *I*)	(+A, -R, -3P / +P, -O)
	authorial self-address: other (e.g., *Paul*)	(+A, -R, -3P / -P, +O)
	authorial self-address incl. reader(s): pronoun (e.g, *we* → *I + you*)	(+A, +R, -3P / +P, -O)
	authorial self-address incl. reader(s): other (e.g., *our group* → *I + you*)	(+A, +R, -3P / -P, +O)
	authorial self-address incl. third party: pronoun (e.g., *we* → *I + him*)	(+A, -R, +3P / +P, -O)
	authorial self-address incl. third party: other (e.g., *our group* → *I + him*)	(+A, -R, +3P / -P, +O)
R	reader(s)-address: pronoun (e.g., *you*)	(-A, +R, -3P / +P, -O)
	reader(s)-address: other (e.g., *Kaspar*)	(-A, +R, -3P / -P, +O)
	reader(s)-address incl. third party: pronoun (e.g. *you* → *you + him*)	(-A, +R, + 3P / +P, -O)
	reader(s)-address incl. third party: other (e.g., *your group* → *you + him*)	(-A, +R, + 3P / -P, +O)
3P	third party-address: pronoun (e.g., *him*)	(-A, -R, + 3P / +P, -O)
	third party-address: other (e.g., *Ursula*)	(-A, -R, + 3P /-P, +O)

Key: A = author; R = reader(s); 3P = third party; P = pronoun; O = other, incl. = including.

With respect to the coding procedure of personal references as illustrated in Table 15.2, it should be noted that there exist two more possibilities for authorial self-reference in personal written communication: (+A, +R, +3P / +P, -O) and (+A, +R, +3P / -P, +O), where the former could be any first person plural pronoun that, based on contextual information, refers to author, reader, and any other third party and the latter any non-pronominal reference that includes all three parties. These types of authorial self-address were, however, not observed in any of the text corpora and were therefore neglected.

15.2. Questionnaire (online survey)

The online survey was launched on 13.8.2007 and terminated on 3.9.2007. The questionnaire included below is an identical print version of the online questionnaire (including all explanatory remarks). Informants were sent an e-mail with the link to the online survey, which was located at http://freeonlinesurveys.com, and could participate anonymously.

QUESTIONNAIRE
DISSERTATION CLAUDIA AESCHBACHER, UNIVERSITY OF ZURICH (SWITZERLAND)

0) Introduction

Dear Participant,

Below you will find a questionnaire (consisting of 10 pages with a total of 31 questions) that focuses on **personal written communication.**

"Personal written communication" is understood as communicative interactions in written fashion with people that you are familiar with, such as your partner, your friends, or members of your family, and the nature of the communication is **informal** (as opposed to formal, business-like). In the case of settings where you as the author and/or your reader(s) may remain anonymous (on the Internet, for example), "personal written communication" is understood as communicative interactions in written fashion with people that you are in most cases not familiar with, but the nature of the communication is nevertheless informal.

For all questions that offer predefined answers, please indicate your choices by checking the respective circles. As regards the open questions, please enter your answers into the respective text boxes. All questions marked with an asterisk (*) need to be answered before you can go to the next page (this is the case for all questions except for question Nr. 29, which gives you the opportunity to make additional comments). However, once you have completed a page and switched to another one, you may always go back to make any changes.

Please note that examples given in brackets are for illustration only ("e.g." means "for example"). By no means are they intended to guide you in any direction with your answers.

I thank you very much in advance for your participation!

PAGE 1: FREQUENCY OF USE
This set of questions is concerned with how often you think you perform personal correspondence via the various media listed. Although this questionnaire is focused on written communication, I am also including here a question on diverse media for audio(-visual) communication to be able to compare the two modes with regard to frequency of use. Please indicate below how often you think you engage in communicative exchanges with the different types of media listed.

Please note: If you check "I don't use this medium" for any medium in the questions of Page 1, please continue to check "no answer" in the remaining questionnaire for questions concerned with that medium.

1) How often do you communicate in writing via the media listed below?

	Never	Several times per year	Several times per month	Several times per week	Several times per day	More than 5 times per day
SMS text	○	○	○	○	○	○
E-mail	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○

2) How often do you orally communicate via the media listed below?

	Never	Several times per year	Several times per month	Several times per week	Several times per day	More than 5 times per day
Mobile phone	○	○	○	○	○	○
Home phone	○	○	○	○	○	○
Face-to-face	○	○	○	○	○	○
Video message	○	○	○	○	○	○

PAGE 2: COMMUNICATION CHANNELS

There are several communication channels that can be employed in personal written communication, e.g. the use of emoticons ("smileys") to express emotions. In this set of questions you will find a few selected strategies and please indicate how often you think you use them in your writing.

3) How often do you draw something instead of writing it (e.g. a heart shape for "love")?

	no answer	never	rarely	sometimes	often	very often	always
Handwritten letter (or card)	○	○	○	○	○	○	○

4) How often do you use emoticons to express emotions (e.g. :-) for "happy")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

5) How often do you use a symbol instead of a word (e.g. # for "number")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

6) How often do you attach or enclose another document (e.g. a picture or file of any kind)?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

7) How often do you include hyperlinks (e.g. "check out this link: www.address.com")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

PAGE 3: MEANS OF EMPHASIS

In personal written communication we have several means of emphasis at hand to highlight certain points or aspects of a message. In this set of questions you will find a few selected strategies and please indicate below how often you think you use them in your writing.

8) How often do you use CAPITALS for reasons of emphasis?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

9) How often do you repeat punctuation (e.g. "what a great idea!!!") for reasons of emphasis?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

10) How often do you repeat certain words (e.g. "that's very very kind") for reasons of emphasis?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

PAGE 4: ORGANISATION OF TEXT

There are several ways to organise text in personal written communication, e.g. visually separate text into greeting section, main body and farewell section. The main body of a text message may further be divided into paragraphs for reasons of clarification. In this set of questions you will find a few selected strategies and please indicate below how often you think you use them in your writing.

11) How often do you include a greeting section (e.g. "hello Andreas") in your messages?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

12) How often do you include a farewell section (e.g. "bye bye, love Sarah") in your messages?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

13) If you decide to include a greeting and/or farewell section in your message, how likely are you to separate them (by paragraph or punctuation) from the main body of the text?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

14) How often do you include a list or enumeration for a clearer presentation in your messages?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

PAGE 5: PERSONAL DEIXIS

"Personal deixis" is concerned with how we address ourselves and others in our writing. In this set of questions you will find a few selected strategies and please indicate below how often you think you use them in your personal written communication.

15) How often do you refer to the addressee with a pronoun in the greeting section (e.g. "hey you!")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

16) How often do you refer to the addressee with his or her proper name in the greeting section (e.g. "hey Alexander" or "good morning Susan")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

17) How often do you refer to the addressee with his or her nickname in the greeting section (e.g. "hey Alex" or "hello Susie")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

18) How often do you refer to the addressee with a pet name in the greeting section (e.g. "hey sweetheart", "my darling")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

19) How often do you include your own name or nickname in the farewell section (e.g. "see you later, love Alexander" or "yours, Susie")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

PAGE 6: INTRA- / INTER-TEXTUALITY

"Intra-textuality" means referring to a text segment inside the current writing (e.g. “as I have mentioned above”) whereas "inter-textuality" refers to a text segment outside the current writing (e.g. “as I mentioned in my email last week”). In this set of questions you will find a few selected strategies and please indicate below how often you think you use them in your personal written communication.

20) How often do you refer to something you wrote in the past (e.g. "did you get my last e-mail?")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

21) How often do you refer to something you intend to write in the future (e.g. "I'll send you an SMS text tomorrow")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

22) How often do you refer to something you have written in the current message (e.g. "see below" or "as I have already mentioned")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

PAGE 7: TEMPORAL DEIXIS

"Temporal deixis" is concerned with how we refer to the notion of time in communication. In this set of questions you will find two selected strategies and please indicate below how often you think you use them in your personal written communication.

23) How often do you include dates (written by yourself, not computer- or mobile phone-generated dates) in your messages (e.g. "12.7.2007" or "July 2007")?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

24) How often do you use temporal expressions (such as "now", "yesterday", "today", "next week" etc.) to embed your messages in a timeframe?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

PAGE 8: DECODING OF MESSAGES

"Decoding" is concerned with how we process (or: strive to understand) language input in terms of meaning. The decoding of a written message is an interplay between author, reader, medium and the communicative context. A message that is easily understandable (with a straightforward, explicit, and clear content) requires little effort from the reader to decode its meaning. On the other hand, a message that is not as easily understandable (with an ambiguous or otherwise unclear content) requires a certain effort from the reader to decode its meaning - and in some cases the reader might not be able to successfully decode (understand) a message and is in need for clarification. In this set of questions you will find two selected scenarios concerned with this matter and please indicate below how often you think this applies when you communicate on a personal level.

25) How often do you think it applies when you communicate that the reader of a message of yours didn't understand something you wrote and asks you for clarification?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

26) How often do you think it applies when you communicate that you read a message from someone and you find something is unclear and needs clarification by the author?

	no answer	never	rarely	sometimes	often	very often	always
SMS text	○	○	○	○	○	○	○
E-mail	○	○	○	○	○	○	○
Web Chat	○	○	○	○	○	○	○
Personal HP / Blog	○	○	○	○	○	○	○
Letter / Card	○	○	○	○	○	○	○

PAGE 9: OPEN QUESTIONS

Below you will find three open questions (Nr. 25-27).
You need not write an essay (but of course you may if you wish to do so!), but if you could briefly formulate your thoughts in 2-5 sentences that would be great.

27) Message length in SMS texts used to be limited to 160 alphanumeric characters (including blanks). Nowadays, with the newer mobile phone models, it is possible to write longer texts or even log onto the Internet via mobile phone to write e-mails.
On average, are your SMS texts longer or shorter than 160 characters? And do you also write emails from your mobile phone? Can you give a few reasons why?

Please type your answer into the text box below.

28) Do you think that in personal written communication handwritten messages will at some point be replaced by electronic (digital) ones?
If your answer is “no”, what do you think makes handwritten messages so special?
If your answer is “yes”, what do you think are the consequences for human communication?

Please type your answer into the text box below.

29) I am interested in your opinion!

If there's anything you would like to let me know in connection with this questionnaire or personal written communication in general, then please don't hesitate to enter your comments (optional) into the text box below.

PAGE 10: DEMOGRAPHICS

30) Please indicate your sex for statistical reasons. Thank you.

○ Male
○ Female

31) Please indicate your age for statistical reasons. Thank you.

○ 10-20 yrs
○ 21-30 yrs
○ 31-40 yrs
○ 41-50 yrs
○ 51-60 yrs
○ 61+ yrs

15.3. Web Chat sessions

The 30 Web Chat sessions (15 minutes each) took place between 31.1.2006 and 23.3.2006. Tables 15.3 to 15.8 below provide details of the individual Web Chat sessions as recorded in Rooms I-VI. It should be noted that participant numbers are approximations as Chatters entered and left the rooms quite frequently throughout the data collection.

Table 15.3: Web Chat sessions recorded in Room I.

Session	Date	Time	Nr. of participants	Word count
1	10.02.2006	17:35 – 17:50	17-18	647
2	13.02.2006	17:47 – 18:02	14	362
3	22.02.2006	20:57 – 21:12	25	752
4	03.03.2006	11:18 – 11:33	14	580
5	14.03.2006	10:21 – 10:36	15-21	283

Table 15.4: Web Chat sessions recorded in Room II.

Session	Date	Time	Nr. of participants	Word count
1	22.02.2006	13:50 – 14:05	9-12	312
2	03.03.2006	12:53 – 13:08	6-12	324
3	10.03.2006	10:55 – 11:10	3-5	131
4	14.03.2006	22:00 – 22:15	5-7	193
5	23.03.2006	10:14 – 10:29	2-5	194

Table 15.5: Web Chat sessions recorded in Room III.

Session	Date	Time	Nr. of participants	Word count
1	21.02.2006	21:10 – 21:25	250	1836
2	03.03.2006	12:35 – 12:50	140	0720
3	10.03.2006	10:37 – 10:52	80-90	1126
4	14.03.2006	21:55 – 22:10	231	2655
5	23.03.2006	09:45 – 10:00	77-88	1167

Table 15.6: Web Chat sessions recorded in Room IV.

Session	Date	Time	Nr. of participants	Word count
1	11.02.2006	18:09 – 18:24	13-15	0807
2	20.02.2006	08:58 – 09:13	13-16	1021
3	03.03.2006	11:40 – 11:55	22	1431
4	10.03.2006	10:16 – 10:31	12	0778
5	14.03.2006	21:37 – 21:52	10-15	0740

Table 15.7: Web Chat sessions recorded in Room V.

Session	Date	Time	Nr. of participants	Word count
1	08.02.2006	10:38 – 10:53	16	0886
2	15.02.2006	17:30 – 17:45	80	0524
3	22.02.2006	17:13 – 17:28	77	1874
4	03.03.2006	10:45 – 11:00	50	1212
5	14.03.2006	10:05 – 10:20	70-80	2104

Table 15.8: Web Chat sessions recorded in Room VI.

Session	Date	Time	Nr. of participants	Word count
1	31.01.2006	18:50 – 19:05	42	1221
2	13.02.2006	17:21 – 17:36	54	1216
3	22.02.2006	16:53 – 17:08	46	0651
4	03.03.2006	11:01 – 11:16	61	1990
5	14.03.2006	09:48 – 10:03	10-16	0667

Curriculum Vitae

Claudia Aeschbacher was born 31st May 1976 in Wädenswil, Switzerland. She attended primary and secondary school in Wädenswil (1983-1992) and subsequently the *Gymnasium Kantonsschule Enge* in Zurich, Switzerland (1992-1998). During her time at *Kantonsschule Enge*, she spent a year abroad in Canterbury, GB (1995), attending the Sixth Form at Barton Court Grammar School. After completing the *Gymnasium (Typus D)*, she worked two years for *Swissair* (now known as *Swiss International Airlines*) in the position of Flight Attendant as well as Assistant to Management of Flight Attendants and Performance. In 2000, Claudia Aeschbacher started her academic education at the University of Zurich, Switzerland. She studied "English Language and Literature Studies" as main subject (as well as Geography and Sociology as minor subjects), graduating with a diploma thesis in English Linguistics in 2005. At the beginning of 2006, she began her research for the present PhD thesis, supervised by Prof. Dr. A. H. Jucker and Prof. Dr. Daniel Schreier, financially supported by the Research Committee of the University of Zurich, which she finished in 2008. Claudia Aeschbacher is currently working as a lecturer at the English Department of the University of Zurich, teaching the course "Introduction to Linguistics, Parts I & II".