universität
wien

Diplomarbeit

Titel der Diplomarbeit

"Typing the Untypeable:
Paralinguistic Features in IRC (Internet Relay Chat)"

Verfasser

Peter Waldner

Angestrebter akademischer Grad

Magister der Philosophie (Mag.phil.)

Wien, Oktober 2009

Studienkennzahl lt. Studienblatt: E 190 344 884

Studienrichtung lt.Studienblatt: Unterrichtsfach Englisch / Informatik

Matrikelnummer: 0216767

Betreuerin: Ao.Univ.-Prof. Mag. Dr. Ute Smit

Acknowledgements

I would like to thank my supervisor Professor Ute Smit for all her kind support, time and countless professional advice that made the creation of this thesis possible.
Thanks also to all my colleagues that accompanied me on the way through my study and special thanks to Heidrun Lunger for proof-reading this thesis.
Especially I would like to thank my wife Friederike for her never ending emotional support and help throughout my whole study.

Finally I wish to show my heartfelt gratitude to my parents who always supported me and were there for me whenever I needed them.

Abstract

Given that internet chat discourse as a 'new form' of communication is closely related to oral communication it includes many unique paralinguistic features, which presents an interesting field of research. This thesis examines paralinguistic behaviour in IRC (Internet Relay Chat) and focuses on emoticons and the performance of non-verbal actions. The basis for this research was an extensive corpus of IRC logfiles assembled for this study. With over 3 million words and ca. 250 logfiles that include 120 different IRC channels from 5 IRC networks, this corpus was carefully selected to provide an ideal basis to gather representative and unbiased results for language behaviour in IRC. The aim of this thesis was to provide an overview over the most important types of paralinguistic cues in internet chat and to describe how they are employed in IRC. It was also necessary to discuss the influences of oral and written communication on IRC. Furthermore it was discussed what unites the large IRC chat system to a coherent group suitable for research and whether it can be called a community. Emoticons and the transcription of non-verbal actions by chat users were considered the most interesting fields of research and various research questions have been addressed for both language features. The most frequently used emoticons and verbs for non-verbal actions were identified and reasons for their popularity discussed. A comparative analysis suggested that emoticon use might have changed in the last ten years. Furthermore an investigation of different humour markers revealed that they do not appear to be true synonyms and vary in intensity and range of application. It could be shown that the most frequently transcribed actions tend to represent cues of communicative closeness and that they can also be used to transcribe thoughts. As indicated multiple times in this thesis the style and amount of employed paralinguistic cues can vary from user to user, especially with some features like pause markers via three dots (...) or various different humour markers. Additional thoughts were given to the influences of the speech situation and other contextual features in IRC and it appears that the location within the IRC network can have an impact on the user behaviour. Some features like emoticons and actionmarkers are immensely popular and the use of paralinguistic cues seems to be essential for internet chat communication. With the findings presented in this thesis it should be easier to understand how and why paralinguistic cues are used in the medium of IRC.

Zusammenfassung

Um seiner Nähe zur mündlichen Kommunikation gerecht zu werden, beinhaltet Internet Chat Diskurs viele Möglichkeiten, um paralinguistisches Verhalten, wie man es aus traditioneller mündlicher Kommunikation kennt, einzubinden. Da dies ein interessantes Forschungsgebiet darstellt, untersucht diese Arbeit paralinguistisches Verhalten in IRC (Internet Relay Chat) und konzentriert sich dabei insbesondere auf die beiden Gebiete Emoticons und die Möglichkeit, eigene non-verbale Handlungen darzustellen. Grundlage dieser Studie war dabei ein umfangreicher selbst erstellter Korpus von IRC logfiles, der ein repräsentatives Bild von IRC Kommunikation ermöglichen sollte. Das Ziel war es, einen Überblick über die wichtigsten paralinguistischen Ausdrucksformen zu verschaffen und zu beschreiben, wie diese in IRC eingesetzt werden. Dafür wurde auch auf die grundsätzlichen Einflüsse von mündlicher und schriftlicher Kommunikation auf Internet Chat eingegangen. Emoticons und die Selbstzuschreibung von non-verbalen Äußerungen wurden als besonders interessant erachtet und viele Fragen konnten im Zuge dieser Studie in beiden Gebieten beantwortet werden. Mit Hilfe des Korpus wurden die wichtigsten Emoticons und für Selbstzuschreibungen verwendete Verben ausfindig gemacht, sodass Gründe für diese Verteilung diskutiert werden konnten. Eine vergleichende Untersuchung legte nahe, dass sich die Verteilung von Emoticons in den letzten zehn Jahren verändert hat. Außerdem ergab eine Untersuchung unterschiedlicher Humor Marker, dass es sich bei diesen um keine echten Synonyme handelt, sondern sich, unter anderem in der Intensität, Abweichungen ergeben. Es konnte gezeigt werden, dass die am häufigsten verwendeten Selbstzuschreibungen eine Tendenz haben, kommunikative Nähe herzustellen und dass ihr Aufgabengebiet, unter anderem, die Verbalisierung von eigenen Gedanken einschließt. Es scheint, dass das paralinguistische Verhalten von einzelnen Usern oft sehr unterschiedlich sein kann, was insbesondere bei einer Untersuchung von Pause Markern (…) gezeigt werden konnte. Des Weiteren scheint das sprachliche Verhalten auch vom virtuellen Ort abzuhängen, in dem sich die IRC Nutzer befinden. Viele der untersuchten paralinguistischen Ausdrucksformen werden sehr häufig genutzt und ihre Verwendung scheint für Internet Chat Kommunikation von größter Wichtigkeit zu sein. Mit den Ergebnissen dieser Studie sollte es leichter fallen, zu verstehen, wie und warum diese in IRC eingesetzt werden.

Table of contents

1. Introduction

Since the boom of the internet, computer mediated communication (CMC) has received much attention in the literature. An increasing number of people interact online via e-mails, blogs, web-forums, chats or instant messengers and these forms of communication have become part of every day life. Many fields of research have taken interest in CMC, among them social, cultural, political or economic studies. These "new" forms of communication also provide a very resourceful basis for linguistic research, because new communicative needs lead to the emergence of new language features that differ considerably from "old" forms encountered before the rise of the internet.

It is difficult to treat all aspects of CMC as one single type of communication. Blogs for instance do not only have a different purpose and communicative structure than internet chats but furthermore tend to include a different amount of oral influences. Despite being written, internet chat is strongly influenced by features of spoken discourse, which raises the question where this form of communication can be placed in the traditional dichotomy between written and spoken language. An essential feature that defines spoken discourse is the presence of paralinguistic cues that convey a considerable amount of meaning in a conversation. Noticing features like facial expressions, gestures, intonation, pitch or stress is often important to grasp the full meaning of an utterance, yet in internet chat discourse all these features become unavailable. The resulting lack of communicative tools threatens to limit communicative capabilities and provokes ambiguity. This need to replace important mechanisms of face-to-face conversation has led to the creation of many substitutes so that paralinguistic and contextual features can remain present in internet chat discourse.

Previous studies have already introduced the most relevant paralinguistic substitutes in internet chat, even though the interpretation of their worth varies from an "innovative set of linguistic devices" (Werry 1999: 57) to "somewhat desperate efforts" (Crystal 2006: 37) to replace linguistic features such as voice, gesture and tone. Even though these paralinguistic features of internet communication have been touched upon, they still await a more detailed analysis that provides more than just superficial results or

general statements. This study aims to investigate paralinguistic cues that appear in IRC (Internet Relay Chat) in unprecedented detail, as it presents representative results based on an extensive corpus of IRC logfiles. With the findings of this study it should be easier to understand the importance of some of these "new" language features, how they work and are used in IRC.

In order to investigate paralinguistic discourse it is important to establish where internet chat stands between the written and spoken medium and therefore the first chapter will discuss the position of internet chat discourse by addressing previous findings in the literature. Once this position is established a closer look will be taken at IRC by briefly describing what it is and how it works.

Like other studies in this area the present one will treat IRC as a single type of internet chat, despite the fact that it consists of many different channels and networks. Therefore it should also be discussed whether something like an "IRC community" exists or if other concepts than "community" would be better suited to describe IRC. I will also address why it makes sense to research paralinguistic behaviour on IRC in general as well as which characteristics define IRC and make it different from other CMC environments.

Chapter five introduces the research method applied in this study and describes how the corpus was assembled and processed. In chapters 6-8 I will present and explain the findings of this research project. The paralinguistic cues found in IRC are described and analysed with the support of various examples from the corpus. I will further discuss what possibilities IRC users have to replace paralinguistic cues of spoken discourse and examine how these possibilities are used. Apart from paralanguage other contextual features such as the speech situation also play an important role in IRC and thus the final chapter of this thesis will focus on contextual influences as well as provide a short analysis of virtual plays and their relevance in regular IRC conversation.

While all paralinguistic features such as loudness, stress or pauses will be discussed, the focus of this thesis is on the two aspects that I considered most interesting and important: emoticons and the performance of non-verbal actions. Especially in these two areas the present study aims to present new findings not yet discussed in the literature.

One of the research objectives is to produce a comprehensive list of all relevant emoticons in IRC communication and separate this group from all other smiley variations that exist. For this purpose it is essential to discuss the communicative functions of emoticons as well as what can be considered a 'relevant emoticon' for internet chat discourse. Furthermore the influence of loans from Asian CMC culture will be addressed. Another research question is whether emoticons in IRC are subject to language change and in the course of this investigation I will examine a possible change in the use of these paralinguistic signs. Finally it is interesting to observe the limited gradability of emoticons as well as their ambiguity, vagueness and flexibility in meaning. The key feelings happiness and sadness can for instance be expressed via multiple paralinguistic signs. Chapter 6.7 investigates this more closely by looking at three different humour markers and their occurrences in full detail. By analysing 150 random samples and individual user behaviour from the corpus this study tries to find out in which ways the most prominent humour markers in IRC differ from each other, or whether they work the same way and could be called true synonyms.

The research on action markers introduces *actionstrips* and *actionlines,* two popular ways of describing actions, feelings or interactive behaviour and will highlight some of their similarities and differences. One interesting question to ask is which purposes these two features serve in communication and in what situations they are used most. Through research on the IRC corpus this study aims to identify the most popular verbs used as collocations for these actions. I will address whether one can divide these collocations into groups serving different roles in communication. Of special interest in this context is the possibility to express feelings, thoughts and emotions with indirect statements in third person form via action commands. Possible reasons and implications for this usage of indirect phrases will be discussed as well.

2. Between Written and Spoken

The question whether internet chat can be classified as written or spoken discourse has been of interest to many researchers so far. This interest stems from the fact that CMC communication shares important features with both but is at the same time fundamentally different from any form of written or spoken communication that has existed in the past. Reid (1991: Preface) points out that "[c]hat programs deal in a form of synchronous communication that defies conventional understandings of the differences between spoken and written language", which makes this question difficult to answer.

The answer is undoubtedly not the same for every form of CMC communication and it is important to distinguish between synchronous forms of CMC such as internet chat or asynchronous forms such as e-mail, message boards or newsgroups. Synchronous CMC communication is definitely closer to the spoken medium which is reflected by the users' self-perception: "We 'write' e-mails not 'speak' them" (Crystal 2006: 32). Chats on the other hand – as the name already indicates – are perceived as a place where we 'talk', 'speak' and simply 'chat', which are all concepts closely associated with orality. It would be very unlikely for a chatter to state "as I have written before", instead they would most likely say "as I have said before". This fact is interesting in itself because we obviously write and do not speak in an internet chat.

The most quoted concept used for interpreting the position of synchronous CMC is that of conceptual and medial speech and writing by Koch and Oesterreicher (1994). Medial refers to the medium a text is realised in and is therefore an either/or distinction between written (graphical realisation) or spoken (phonic realisation). Conceptual however refers to the concept behind the language that is used and allows for much more grading. A text can be anywhere between the two poles on a conceptual scale.

Popular examples for this grading scale are a family conversation, which is clearly conceptually and medially spoken, or on the opposite a legal text, which would be both conceptually and medially written. A lecture or a speech however is medially spoken but would rather be conceptually written, especially if it is scripted in advance. On the other hand a transcript of a court hearing, despite being medially written would be

conceptually spoken. A private letter, while clearly written, can have strong conceptually oral influences, whilst a business letter would rather be written in concept.

Koch and Oesterreicher (1994: 588) argue that physical, temporal, social and emotional closeness as well as spontaneity are associated with conceptual speech, while conceptual writing stands for distance. In addition to the relation between *communicative closeness* and *distance* it is typical for written texts that production and reception are separated and that they are monologic with no possibility for the recipients to interfere and cooperate. A lecture speech would feature considerable amounts of communicative distance because it lacks spontaneity. The topic of the lecture is already set, it is monologic and there is emotional and social distance, making it rather close to conceptual writing.

Researchers on CMC discourse have come to the conclusion that synchronous CMC is very close to conceptually spoken language, despite having influences from its written medial realisation (for examples see Haase et al 1997, Hess-Lüttich 2003, Schönfeldt 2001, Schulze 1999, Storrer 2001a/b, Werry 1996). The fact that it features synchronous many-to-many conversation is already an indication for conceptual orality (Schönfeldt 2001:52). However in many ways it is unique and different from other text-types that are medial writing but conceptual speech: Scripted texts for radio plays, theatre, movies or song texts are intended to be spoken out orally, while transcripts for linguistic analysis or court hearings are used for preserving originally spoken discourse. In both cases the conceptual orality is bound to the medial orality (Storrer 2001a: 462, Storrer 2001b: 4) as they are based on oral talk in one way or the other. In synchronous CMC discourse however conceptual and medial orality are independent from each other because utterances are never intended to be spoken out and are not meant for reproduction. In fact it would even be very difficult to realise a spoken version of a chat transcript due to the fact that chat discourse is organised very differently than regular face-to-face conversations.

One of the most important differences between face-to-face and synchronous CMC discourse is the way how turn taking works. While being a very important concept in face-to-face communication, chat discourse discards turn taking completely and instead works on a client-server principle: Messages are typed by the users and are sent to the

server after their completion by pressing the enter key. They appear on the screen in the order they are received by the server, which means that whoever sends the message first will be displayed first. This leads to multiple serious implications. Typical features of face-to-face communication like back-channelling, interruptions or overlaps are omitted. Instead multiple strands of conversation are intertwined and the resulting sequence forms a multidimensional text (Werry 1996: 51). As a result chat logs are often difficult to read, because different strands of conversation might be discussed at the same time. In a face-to-face setting this would likely result in utter chaos. Even with paralinguistic information that could be interpreted as a kind of 'stage directions' one might struggle to act out a chat dialogue due to the fact that messages with a thematic relation do not necessarily stand next to each other, and consequently utterances following each other might not have a discursive relation at all. The positive outcome of this non-correlating and unique structure is that it makes it possible for the users to take part in multiple discussions at the same time, which would be odd, impolite and obviously very difficult in a regular face-to-face conversation. Furthermore there are no physical limitations of the communicative range as there would be in face-to-face discourse. While in regular spoken discourse it is difficult for a person to overhear a conversation that is held out of ears' range, no such limitations apply for CMC and users can be at many 'virtual places' at the same time and participate in all of them.

Schönfeld (2001: 34) suggests that the overall-chat discourse can be divided into several individual-chats. These individual chats, or in other words individual strands of conversation, are a series of contextually related utterances with a dialogic structure. Schönfeld assumes that an overall-chat containing a larger number of participants usually consist of several intertwined individual-chats and that the amount of individual chats rises proportionally to the number of members in a chat room. The aim for the participants is to filter out the relevant utterances they are interested in from the stream of the overall-chat, which can be seen as a continuous flow of messages that never starts or ends and is therefore hard to structure. Internet chat users have compensated the problem of addressing certain persons within multiple strands of unrelated conversation by the addition of direct references to whom the utterance is addressed, which will be further discussed in chapter 8.4.

Another important difference to regular turn taking is the fact that text production is "invisible" (Storrer 2001a: 452). In regular oral conversation ideally only one speaker is talking, while the others wait for a suitable moment to take over the turn to speak, which is simply not possible in synchronous CMC. There is no possibility to know whether other participants of the chat are producing a message at a given time or not, because the act of typing is invisible for others and the message is only displayed upon completion. A user could have dediced to ignore a member's latest utterance, but might also currently be busy writing a reply. There is no way of finding out what is happening behind the screens of other chat participant's PCs. Crystal (2006: 35) argues that there are many reasons why someone might not respond to an utterance and unlike telephone conversations where a silence would be complemented with comments like "Hello?Are you there?" chat groups are much less reliable. In addition to typing a response or deciding not to respond at all, a user could be distracted by another conversation in real-life or online. Chat participants may fail to notice a message for different reasons, such as too many messages being sent at the same time or because they are currently not present at their workstation (Crystal 2006: 35, Troest 1999: 2.2). Therefore delays in reactions are possible – but much more accepted than in face-to-face conversations – and little can be done to find out what has caused the delay. The fact that production (typing) and reception (message sent, received by server and read by user) of a statement are separated causes a certain degree of asynchrony (Storrer 2001b: 7) even though chat discourse is generally referred to as synchronous CMC. Hentschel (1998: 3.2.2) observed that chat users tend to interrupt their own sentences, usually in places where it is obvious that their utterance is not yet complete. This process of breaking up an utterance into smaller pieces is also observable in spoken discourse. However, due to the differences in turn taking there are other reasons behind this disruption. In face-to-face conversations pauses in strategic moments within an utterance are most likely made to prevent others from taking the turn or to provide an opportunity to think ahead what to say next. In regular chat discourse on the other hand it is of great importance to reduce the gap between production and reception and small simplistic utterances help the users to keep up to date. If users take too much time to formulate a long utterance the conversation might already have moved on to another point. As a consequence their messages might already be obsolete, might more likely be ignored by others, or will

complicate the further conversation. Therefore shorter utterances ensure higher interactivity and keep other users interested, rather than letting them wait for a long period of time. The process of ensuring interest by breaking up utterances will be addressed from the viewpoint of a moderated chat in chapter 8.3. Keeping a turn in synchronous CMC is indeed very difficult if not impossible, since according to Hentschel (1998: 3.2.2), not even talking without pause can be employed as a last resort. Constant typing without sending does not help at all and pauses are inevitable because new chunks of an utterance have to be written before being transmitted.

In addition to differences due to turn taking and conversational structure it is important to keep in mind that CMC usually features neither visual nor acoustic contact. Therefore "non-verbal and paralinguistic signals, such as proxemics, facial expression, gestures, body position, eye contact, prosody, talking speed, pauses [and] voice pitch" (Hentschel 1998: 3.) are all unavailable, and while efforts are made to replace some of them CMC discourse can not count on many traditional paralinguistic features of face-to-face conversation.

The considerable disadvantages of the written medium are compensated by many factors in CMC, but one of the most important ones is the fact that any information is stored in the chat log (Storrer 2001a: 460). While a failure to notice or understand an utterance in face-to-face communication leads to an inevitable loss of information, internet chats are much more forgiving. A message that has been overlooked at the time of its completion is still on the screen and can be re-read anytime, so there is a considerable chance that a user might pick up this information later on. Therefore IRC discourse has a longer "half life" (Troest 1999: 2.3) than regular face-to-face interactions. Especially due to the chaotic structure caused by multiple unrelated strands of conversation, the possibility to take more time disentangling the relevant bits of a strand is invaluable. However, older messages scroll of the screen quickly, which stresses the aspect of time bound, synchronised real time conversation, as older messages become more and more obsolete and irrelevant. Nevertheless one still has the advantage of being able to re-read every single message that was uttered from the moment the chat room has been joined to the moment the user leaves. In addition to the possibility to take part in multiple discussions and the unlimited communicative distance discussed above one can see that chat discourse still has some advantages over

regular face-to-face conversation despite the fact that the production of medially written texts is more time consuming (Storrer 2001b: 8).

Regarding all these important differences to spoken discourse one might wonder why synchronous CMC is still regarded as being very close to the spoken medium. The only explanation given so far was that the participants understand themselves as speakers, not as authors, which implies that utterances will most likely be formulated to resemble spoken discourse rather than a written one. The register used in chat discourse is therefore rather colloquial (depending on the formality of the setting itself) and Storrer (2001a: 446) found several lexical and syntactical similarities with spoken discourse, such as a preference for simple and short utterances which are colloquially marked or a reluctance to use structuring devices.

A study by Ooi (2002) based on Schulze's (1999) corpus shows that IRC participants use a great variety of interjection markers which "express a range of feelings, including approval, surprise, disappointment, and pain". The results indicate

> not only the 'creativity' involved in the choice of non-standard interjections but more so the speaker's desire to signify his/her precise degree of felt emotion and intonational contours by means of the limited orthography that is available on the keyboard. (Ooi 2002: 98)

Some examples that were tagged as interjection markers in Ooi's study were *hmm, oops, eh, ahh, yum, ooh* or *ouch,* which are rather typical for spoken discourse and range to exotic interjections like *ermmmmm, ugh* or *uhmmmmmmmmm,* resulting in a total of 113 different items tagged as interjections in the corpus. The example of interjections shows that internet chat users aim to imitate spoken discourse and its flexibility rather than adapt to written norms.

One can say that chat messages are not intended to represent sentences but rather oral utterances. This is partly reflected in differences regarding typology: full stops are generally avoided and commas are rare. Exclamation- and question marks are notably less effected (Dittmann 2001: 90) even though Gelleri (1998: 33) argued that they are sometimes left out when not necessary. Because of the simplistic structure of utterances most punctuation marks lose their functionality and capitalisation is very unlikely, even for proper names (Gelleri 1998: 33). The most important factor however is that chat

messages are generally grouped as communicative chunks rather than sentences. As discussed above chat users tend to divide longer utterances into smaller chunks and send them individually. Even in a moderated chat where only one person has the right to speak and takes longer turns, this speaker will usually group his utterance into small chunks which are transmitted one at a time. These chunks often do not represent grammatically complete sentences (Storrer 2001a: 454), instead the whole utterance is interrupted in strategic places much like pauses in face-to-face communication would be. (For an example of clustered chunks in moderated chat monologues and further discussion on that topic see chapter 8.3.) A study by Dittmann (2001: 66) indicates that ca. 75% of all IRC sentences are grammatically incomplete while only one quarter matches the minimum requirements. Additionally more than 90% of all IRC utterances only contain a single sentence and between 30%-40% of all utterances consist of only one or two words (Dittmann 2001: 65). Some deviations from written norms of typology like reduplication of letters (for example *hiiiiiii*) are often used as an intensifying tool or as a substitute for paralinguistic cues (see chapter 8.1), while others like constant misspelling and shortening of words (for example tho = though) are caused by the apparent need to save time in chat discourse (Gelleri 1998: 38).

An essential point why synchronous CMC is associated with conceptually spoken discourse is the communicative closeness (as defined by Koch/Oesterreicher) in chat conversations. Beisswenger (2001: 103) and Storrer (2001a: 451) argue that there is a metaphoric physical closeness because chat participants believe to be in the same "chat room". Crystal (2006: 45) investigated how typical criteria for spoken discourse are represented in various forms of CMC. He came to the conclusion that like spoken discourse internet chat is *loosely structured* and with some restrictions also *socially interactive*, *spontaneous* and *time-governed*, which are all important indicators for communicative closeness. The fact that synchronous CMC is time-governed means that a "message delivered in a chatgroup demands immediate response" (Haralampieva 2004: 17). This is a key feature of chat discourse, because unlike in asynchronous CMC and other genres more associated with the written medium production and reception are not separated (apart from small delays in message production), and discourse is created through cooperation and interaction. In synchronous CMC there is only little time

available for the production and reception of messages and a busy chat room can set high demands on the participants (Storrer 2001a: 447).

The high degree of interactivity ensures that users have the possibility to give constant feedback or ask for clarification, even though such interruptions are only possible message by message. Furthermore users can correct their mistakes made in earlier utterances. It is interesting that the option to correct one's spelling or lexical mistakes is used regularly (Storrer 2001a: 449) even though users could have already corrected their mistakes during the typing process by proof reading their messages. This underlines that a quick reply is much more important than its careful formulation and that participants in a chat room try hard to keep up with the pace of the conversation. Kramer (2008: 40) adds that mistakes are generally not seen as a sign of ignorance and have little negative connotation.

An important factor that complicates the participation in synchronous CMC is the fact that the users have to concentrate on production and reception at the same time (Storrer 2001a: 451). While they work on a message of their own, they have to keep track of what is currently being said in the chat room and sometimes a message even has to be adjusted during the typing process due to the constant inflow of new utterances (Troest 1999: 2.4). This divided attention leads to further reduced care in text production. Additionally a reaction has to be formulated before the message one wants to relate to becomes obsolete and is replaced by newer utterances, which further increases the time pressure to produce messages as quickly as possible.

The discussed time-governed and spontaneous features of synchronous CMC lead to a deviation from written norms and to an orientation towards simplistic utterances. The ever present need to save time becomes apparent in the common and wide spread use of abbreviations and acronyms: every keystroke saved is an advantage. Storrer (2001a: 453) suggests that new communicative skills are necessary to cope with this form of discourse: concentration and quick reactions are required when multiple topics are discussed at the same time and additionally some experience is needed to decode multiple strands of conversation in the chaotic stream of possibly unrelated chat messages – which are arguably difficult to read for new and inexperienced users. Troest (1999: 2.3) compares this necessity of filtering out relevant messages from irrelevant

ones to “communicative noise” which has no equivalent in any other written form of communication, but can be found in face-to-face interaction in a similar way: Large amounts of background noise can make it difficult to understand other parties. In IRC there are no possibilities to escape this noise as one cannot step aside or split up into smaller groups. Internet chat communication will always lack some of the speed of face-to-face discourse due to the simple fact that even a very fast and skilled writer cannot match the speed of a spoken utterance. However, the time constraints on synchronous CMC are very evident and Crystal (2006: 32) points out that as soon as a message scrolls off the screen it might not receive attention anymore, which emulates the urgency typical for face-to-face conversation. Gelleri (1998: 38) and Reid (1991: Part Two) both suggest that slow typists have considerable disadvantages and fast and skilled keyboard-writers have more time to think about witty replies and deliver them first, which allows them to make a far better impression and even seem more intelligent.

Despite the fact that traditional turn taking is omitted a chat conversation still follows typical patterns of spoken discourse such as question-answer or request-fulfilment-gratitude (Storrer 2001a: 453). Schönfeldt (2001: 52) adds that empirical studies show that chat users organise their interactions similar to face-to-face conversation, even though these interactions are less tightly structured due to the influences of the medial realisation discussed above. While the sequences remain the same as in regular conversation their interactional relation is much more flexible and the transition between them is more fluent and less divided into typical phases (Schönfeldt 2001:52).

Generally it makes sense to distinguish between two types of chat: moderated or expert chats and small-talk/leisure chats. According to Kramer (2008: 44) expert chats are considerably closer to written media and the importance of how utterances are formulated increases: “words are chosen with great care” and the text production is well planned, resulting in a higher text coherence and syntactic correctness (Kramer 2008: 44). The moderated chat takes an exceptional position regarding turn taking as well, because it will often include mechanisms to simulate this concept. Usually the right to speak is granted or taken by a moderator who leads the chat and therefore turn taking as well as passing on or keeping the right to speak remain relevant concepts in this form of chat. It is possible to hold the floor without worrying that others suddenly take over and

the time-pressure of reacting quickly is removed or at least weakened, ensuring that more time can be spent on formulation and accuracy of the statements.

Summarising one can say that synchronous CMC contains many features of communicative closeness: it is spontaneous, time-bound, loosely structured, interactive and furthermore contains a register very typical for spoken discourse. This puts synchronous CMC close to the spoken pole of the conceptual scale and therefore it can be labelled as a kind of "written speech" (Schönfeld 2001: 52, Schulze 1999: 71, Storrer 2001a: 461). Differences to spoken discourse are given through the medial realisation, whereas compensations are made for the obvious disadvantages caused by the transference into the written medium and the lack of face-to-face contact. Schulze (1999: 72) suggests that

> [p]rotocols of IRC sessions do […] look more like transcriptions of spoken communication than anything one would consider a typical piece of written language

and indeed internet chat users seem to go to great lengths to imitate spoken discourse in many ways.

This paper is especially concerned about the attempts which are made to replace some of the missing paralinguistic tools that are so essential to face-to-face communication. Although some features are inevitably lost others can successfully be replaced in internet chat discourse and it will be discussed how the written medium can be used in IRC to integrate paralinguistic features typical for spoken discourse. It should be noted that the amount of paralinguistic substitutes might be related to the level of formality of the chat discourse. A less formal setting is more likely to contain rich paralinguistic cues, while more serious conversations like business meetings or moderated chats display a reduced amount. This is caused by the fact that many of these cues - like emoticons or reduplication of letters - have a rather "playful" connotation and are considered less appropriate for formal settings.

3. What is IRC?

IRC (Internet Relay Chat) is an internet based chat system which was invented in 1988 by Jarakko Oikarinen. Communication usually takes place in public or private *channels*, which is the IRC term for chat rooms. IRC consists of a number of independent computer networks which are each hosted by individual servers. Nowadays IRC is still very popular and the five most important networks together can boast an average of over 3.300.000[1] users online at the same time. As of now there are not only five but over 830 networks, even though many of them are not very popular and some want to be hidden and remain undisturbed. Due to its immense popularity Gelleri (1998: 3) and other researchers have referred to IRC as "undoubtedly one of the most popular conferencing systems in the world", even though today newer programs like Skype and other instant messengers have caught up with IRC. Skype for instance claims to have between 7 and 15 million[2] users online depending on the time of day. Despite the fact that IRC has nowadays found strong competition in instant messengers and telephone conferencing, it is still undoubtedly the most popular system, primarily used for synchronous many-to-many conversation.

It is not always easy to imagine the structure of purely digital 'places' and Gelleri (1998: 11) suggests drawing an analogy to objects and places from the real world in order to better understand the structure of IRC networks. If we compare IRC to big company offices, then a network would be one office building. Inside this building there are many rooms, called channels, which the users can join in order to communicate with all the other people currently in that room. While they may join and leave channels freely, they are not able to enter rooms from other office buildings. Each network is a separate entity and even though some popular channels (*such as #allnitecafe*) might exist in multiple networks, it is important to remember that they are actually not the same channel. Speaking in our analogy: even though many office buildings may contain a room called 'cafeteria', you will meet different people depending on which building

[1] Source: http://irc.netsplit.de/

[2] Sources: http://www.skypestats.com/, the program Skype (which shows the number of online users) and Wikipedia

you go to. Therefore if people want to advertise their channel or arrange a meeting from outside IRC, the instructions need to include the channel name as well as the network it is on.

In order to connect to one of the IRC networks users have to choose personal nicknames and after they have logged on they may join channels by typing */join #channelname*. Of course the analogy to office buildings has its flaws and unlike in real-life, a user might be in many channels at the same time, which is very popular among IRC users. An infinite number of channels can be created in each network and as soon as a user joins a channel that does not already exist, a new one is created under this name. After the last user leaves the channel completely disappears, which causes IRC to be very flexible and dynamic (Gelleri 1998: 13). The first person in a channel has operator rights, which means that they can throw other people out of the channel or even permanently ban them from joining. All other members need to get assigned operator status by one of the other operators. To avoid empty channels from being removed or taken over by someone else users can employ a *bot*, which is an artificial program that simulates a user, stays on the channel indefinitely and can be assigned with specific commands (such as ban certain users or automatically assign them operator status). Bots have many other functions in IRC and are a common sight. Another popular type of bot for example is the trivia bot, which can be activated in a channel to host a 'quiz show'. The trivia bot was used multiple times in the corpus of this study.

Channels enable many-to-many communication, look very similar to regular internet chats and if multiple channels are active each one is represented in an individual window. See *Appendix 1* for screenshots as well as a short description of channel windows. In addition to the public many-to-many communication there is also the possibility to address people privately. This can be done by double-clicking on the target username or typing */msg username text.* As a result a new window containing only the two participants will be opened in order for both users to chat privately. Private conversations are popular in IRC and many users join a certain channel in order to find and contact users they can chat to privately. It is not unusual to contact completely unknown people and initiate a chat with them and despite being completely passive all the time many people tried to establish contact with me while I was online to assemble logfiles for my corpus. This initiating procedure often follows a typical pattern of

establishing contact: The first utterance usually contains a greeting such as *hi*, while the second very likely contains the acronym *asl,* which asks for age, sex and location of the potential chat partner (see Döring 2001: 163).

My study shows that even crowded channels with 100-1400 people online did sometimes not include any other utterances than automatically produced *join* and *leave* messages (displayed every time a user joins or leaves a channel), over observation spans that lasted more than 24 hours. Even though the lack of public conversations in these channels seems to indicate that many users prefer private conversations, there is also another reason for this almost characteristic silence in many IRC channels: Unlike many other internet chat visitors, IRC users tend to stay online and logged in without having much desire to actively participate in a conversation. A study by Geers (1999: 96) indicates that about 28% of the IRC users investigated in his study joined a channel without ever participating in a conversation, while 32% of the participants made less than 11 utterances and only 5% of the chat members made more than 61. Schulze (1999: 70) points out that IRC users call to the behaviour of being on a channel without actively participating *to idle,* which apparently is an important enough concept to be integrated into the German language as a loan word. Gelleri (1998: 26) and Hess-Lüttich (2003: 2.) also add the term *lurking* as a synonym for *idle* and according to Gelleri both terms have no negative connotation on IRC. As a result of this traditional silence some of the observed channels in this study contained little or no relevant communication even though they were often well attended.

A study by Döring (1995: 2.4.5) claims that IRC sessions are generally more time consuming and last longer that other internet related activities such as e-mailing, reading articles or participating in graphical chats (MUDs). It is possible that IRC users stay online for days without actually being at the PC most of the time and some even seem to have IRC activated 24 hours a day. Despite this distinct lack of communication in some channels, generally most of them are very busy and due to the sheer amount of channels and users online IRC is a very interesting medium for internet research and can provide large amounts of valuable data in very short time.

IRC is purely text based and there are no graphical representations for emoticons, or other possibilities to include pictures directly, which influences some paralinguistic

features such as emoticons (see chapter 6.2.). There are many additional features provided by IRC, which cannot be found in most other internet chats. Examples include performing actions via the */me* command, which will be discussed in chapter 7.2, or changing your nickname via */name* anytime during a session.

4. IRC and Community

In order to investigate language behaviour on IRC all IRC networks, channels and users are treated as one coherent group with enough shared characteristics to gain meaningful insights into the use of paralinguistic cues in general. It is very tempting to refer to this large group of people investigated as the "IRC community". However, this is problematic because the term is very slippery just like the concept of community itself. Thus, before using the concept of "community" it has to be discussed if and how IRC users form a single community and how this community can be defined. It is of interest which definitions of community can successfully be applied to the IRC system, and popular concepts like communities of practice or virtual community come to mind. However, Herring (2004: Introduction) warns not to apply the term community too light-heartedly when doing CMC research because not all groups of people interacting online can be considered to be a community and the concept becomes meaningless if applied indiscriminately.

4.1 Approaching IRC community

Wenger (1998: 73) defines communities of practice through three essential features: There is mutual engagement between members, they have a shared repertoire and a jointly negotiated enterprise (Corder 2007: 444). Corder makes it clear that all three criteria must be met to be able to speak of a community of practice Furthermore these factors are not fully independent for example a mutual engagement will lead to the development of a shared repertoire. When looking at IRC as a whole it becomes evident that the definition of community of practice can not be applied, as IRC completely lacks

a jointly negotiated enterprise, and mutual engagement between all members is more than unlikely because networks tend to be rather isolated from each other. Moreover users within one network of IRC are distributed among many different channels and might never meet or participate in anything together. Therefore the concept of community of practice is not applicable to IRC and other concepts might be more useful.

One other and more promising concept is that of the virtual community, however Lui (1999: Introduction) mentions that "researchers have difficulty coming up with a generally acceptable definition" and that it is not clarified "what exactly constitutes a virtual community". Jones (1997: Section 1) discusses this problem in more detail and argues that one popular use of the term is to simply call various forms of group-CMC, such as IRC or email-list forums, a community without ever investigating if they really constitute one. In order to avoid labelling everything as a community it needs to be clarified what characterises and distinguishes one form of group CMC from others. In other words one has to find out

> to what extent […] participation in these […] environments in fact constitute[s] "community" as opposed to being simply "people interacting online" (Herring 2004: Learning Environments)

Jones mentions that virtual communities cannot simply be reduced to a series of CMC messages - they are also a sociological phenomenon. This is underlined by a popular definition of the virtual community by Rheingold, who argues that

> virtual communities are social aggregations that emerge from the Net when enough people carry on those public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace. (Rheingold 1993: 5).

The problem of investigating the existence of a virtual community is addressed by Jones (1997) in his *virtual settlement* theory: It is necessary to distinguish between the communities and the place where they reside. These "cyber-places" where the groups meet online need to meet a minimum set of conditions defined by Jones to be called *virtual settlements* and thus qualify for hosting a virtual community. He proposes to conduct "*cyber-archaeology*" on these sites of virtual life, because the methods employed by the researcher are similar to regular archaeology: they have to investigate

and dig up “cultural artefacts” left behind by the CMC group, such as website structure or content, posts on message boards, log files of conversation and numerous other physical traces that group interaction might leave behind. With the help of those artefacts it can be evaluated whether a virtual settlement exists or not. The four necessary criteria as defined by Jones are:

> - a minimum level of interactivity
> - a variety of communicators
> - a virtual common-public-space where a significant portion of interactive group-CMCs occur
> - a minimum level of sustained stable membership
> (Jones 1997: Section 2)

Following the theory of *virtual settlement*, Jones argues that an IRC network or IRC as a whole does not indicate the existence of a single virtual community because its sheer size and structure causes it to be thousands of individual places rather than “one symbolically delineated place” (Jones 1997: Section 2). However, a single channel or a small group of channels could according to Jones very well fulfil these criteria.

Lui (1999: Virtual Common Public Space) agrees that single channels are most likely to fulfil the conditions of constituting a virtual community and has applied the four criteria mentioned by Jones to create a model for investigating the existence of community in individual IRC channels. He argues that IRC channels easily qualify to provide the essential common public space where most of the group interaction takes place and adds that just like in real-life communities, IRC users have the potential to exclude and ban unwanted users from their community or talk privately in addition to using their public platform. Jones’ condition of a *variety of communicators* is expanded by Lui to adapt to the flexible nature of IRC: firstly channels need to have a stable existence, which means that short-lived channels do not lend themselves as basis for a community. Additionally the number of participants is increased from “more than two” (Jones 1997: Section 2) to a not clearly defined but definitely larger number, as Lui found that only a large number of participants seems sufficient to create enough interaction for a community and small channels containing 3-5 people are most likely occupied only by artificial bots or people who are carry out their interaction somewhere else. Therefore channels like these should be excluded as well. With similar adaptations to the other two of Jones’ conditions Lui

proposed the following list of criteria to empirically test for the existence of a virtual community within an IRC channel:

> - Stability of Membership: The channel has a significantly large number of members whose participation is relatively stable throughout a reasonably long period of time.
> - Interactivity: Messages (including both verbal and action-simulating messages) posted in the channel demonstrate a significant level of interactivity
> - Stability of identity: Participants on the channel maintain stable identities as represented by nicknames (Lui 1999: Conceptual Hypotheses)

While all three criteria present a sensible framework to investigate whether a community exists within an IRC channel, they remain rather vague which makes it difficult for researchers to apply them in praxis. Lui's conceptual hypotheses rely on definitions such as "significant amounts" and it has yet to be investigated what qualifies as sufficient for a community. Lui especially mentioned the difficulty in measuring a sufficient level of co-appearance of channel members and stated his desire to implement software instruments that could perform such an investigation based on the methodology discussed in Lui's article from 1999.

Following Lui and Jones it can be said that when looking at IRC from the perspective of the *virtual settlement* theory the conclusion lies at hand that it consists of many different and distinct virtual communities rather than a single one. Even though not every channel is or has a community of its own, many of them might. It should be noted, that some of the channels that qualify for a virtual community might equally qualify to be a community of practice if they are based on a commercial or non-commercial enterprise like the administration of a website, or one of the various computer game 'clans' that are dedicated to participate in tournaments of a particular game. Apart from the *virtual settlement* theory there are other ways to approach virtual community and Herring (2004: Operationalization of Key Concepts) has formulated six key criteria for virtual communities based on various literature on this topic:

> 1) active, self-sustaining participation; a core of regular participants
> 2) shared history, purpose, culture, norms and values
> 3) solidarity, support, reciprocity
> 4) criticism, conflict, means of conflict resolution
> 5) self-awareness of group as an entity distinct from other groups

6) emergence of roles, hierarchy, governance, rituals

Arguably all six points could successfully be applied to IRC to some extent: There are obviously huge amounts of *participation* (1) throughout the networks and IRC can count on a large body of *regular core participants* (1). According to Herring (ibid) *history* (2) can be measured by the availability and existence of archives. These are usually uncommon for internet chats but exist for IRC for example as graphs depicting the popularity of the largest networks since IRC was first founded. *Culture* (2) is indicated by the use of "group-specific abbreviations, jargon, and language routines" (ibid) which arguably exist for IRC as well (see chapter 9). An illustrative example would be the popular "trout slap" phrase discussed in chapter 7.2.2. *Norms and values* (2) are revealed through the existence of netiquette statements or FAQs, again something that exists for most networks. Furthermore "reactions to violations of appropriate conduct" (ibid) can be observed (see chapter 9.1). *Solidarity* (3) "can be measured through the use of verbal humor" which is a prominent feature in IRC, while *support* (3) is provided through various help channels. According to Herring (ibid) *support* can also be identified by "acts of positive politeness", something that is advised and stressed in many of IRCs netiquette guidelines. An example of c*riticism* and *conflict* (4) can be observed in the log file example discussed in chapter 9.1. *Self-awareness* (5) is more typical of individual networks or channels: according to Herring (ibid) a "we do things this way here"-awareness indicates that self awareness exists, implying "that they might be done differently elsewhere". The Freenode network for example states: "The goals of freenode differ from those of most general chat IRC networks. For that reason, the design of our servers differs…" (http://freenode.net/freenode_and_irc.shtml) As mentioned above networks and channels may also have unique netiquette guidelines, stating a desire to handle things differently from other networks or channels. Finally *roles* and *hierarchy* (6) become apparent through the existing power structure of channel- and network operators.

Even though most of these criteria can be validated for IRC as a whole, they apply better to individual networks and even better to individual channels due to the fact that rules, customs, self-awareness and other factors may vary a great deal between the different virtual places within IRC. Kramer (2008) investigates whether the whole

internet could be classified as a single community and comes to the conclusion that concepts like community of practice or discourse community do not apply, encountering very similar problems to IRC. Instead she proposes to apply the concept of *semiotic domains* as introduced by Gee (2003: 17) (see Kramer: 2008: 95). It refers to "a set of practices, including language (written and spoken), images, symbols, gestures, artefacts, and similar sets of meanings" (Kramer: 2008: 95). IRC could, similar to the internet as a whole, be interpreted as a semiotic domain with several sub-domains: "they differ through goals and sets of practices but have some sets of practices in common with the semiotic mother domain". Every sub-domain uses some special rules or vocabulary even though everyone shares a core set: certain emoticons or paralinguistic cues for example will be used and understood in the same way throughout IRC.

Even though some aspects of a virtual community can be applied to IRC other concepts like that of a semiotic domain seem more useful. The term virtual community should therefore be reserved to individual groups within IRC, which can indeed form strong bonds between members. Danet et al (1998) state that many channels have their regular real life gatherings, such as the channel #gb whose members are mostly from Great Britain and meet regularly in a London pub, which illustrates the ties formed within such communities. They also mention that there are numerous homepages of community-like-channels, featuring reports on real-life meetings and news about upcoming events. Many distinct and tightly knit groups have formed on the IRC network and Danet et al add that even though "each channel has no doubt developed its own subculture, unique in some respects, it is very likely that artful communication style is valued in all of them." Indeed the playful and creative nature of IRC discourse is mentioned by many authors investigating IRC and could be a part of the semiotic mother domain.

Concluding one can say that even though some researchers (for example Reid: 1991) call IRC a community, it is hard to justify such claims. Instead it makes more sense to describe it as a semiotic mother domain that hosts multiple sub-domains. However, so far the discussed definitions have done little to justify research on IRC in general and it has yet to be addressed why it makes sense to investigate the mother domain, rather than smaller parts of it.

4.2 Research on IRC – Why and what can be gained?

We have seen that IRC can be divided into countless different groups. Whereas some might qualify for virtual communities others do not. Therefore research on IRC as a whole will result in the description of the language behaviour of a large cluster of individual groups rather than one coherent virtual community as illustrated in the definitions above. In other words several semiotic sub-domains will be investigated in order to make assumptions about the mother-domain of IRC. This is very much in the interest of the present study as it is not the aim to investigate how members of a specific tightly knit group with close social bonds interact with each other. The aim is to provide a comprehensive analysis of how essential paralinguistic features of communication can be substituted in a certain area of CMC communication. For an investigation like the present one IRC is a very interesting choice: Even though it consists of many different parts and communities, they have more than enough in common to clearly distinguish themselves from sources outside of IRC.

This is largely due to the shared technical framework, which encourages them to use and produce similar language behaviour. Emoticons for example are not represented by graphical symbols instead they only appear as plain text, which greatly affects the amount and type of emoticons that will be employed. Even though users of other chats or message boards will include emoticons as well, their selection will be influenced by the graphical symbols they have at their disposal. As a result the use of emoticons might differ between two online chats simply because they have another set of graphical symbols available, while the technical conditions remain the same throughout IRC. Lui (1999: Introduction) points out the importance of a shared technical framework:

> Different technical settings of group communication on the Internet have different supporting mechanisms and communication processes. They do not work the same way in channeling communication messages. Virtual communities in one technological domain may exhibit different characteristics from communities in other domains

Baym (1995: 145) also agrees that communities are highly influenced by the structure of the technical system which ranges from the level of synchrony to the possibilities and conventions of how to quote or navigate through messages, ignore users etc. IRC has a

rich and uniting technical framework such as the performance of actions via */me* command, which is widely spread (see chapter 7.2).

Rather than a drawback the separation of IRC into many sub-domains is an advantage, because it is possible to make more general statements. The findings are not limited to an isolated, small and globally rather unimportant group of people but to a large body of more or less independent groups that use a similar way of communicating due to a uniting technical framework. They possess the same 'lexicon' of symbols and conventions to express paralinguistic behaviour, which distinguishes them as a group from other groups in the internet. It can be assumed that people on IRC speak the same paralinguistic language which makes them different from other internet chats or forms of CMC such as e-mail or message boards.

This uniting set of linguistic conventions is the major reason why researchers like Reid (1991) argued for IRC being a community: According to Reid IRC has found and "developed" a common solution to the "medium's lack of regulating feedback and social context cues, its dramaturgical weakness, and the factor of anonymity" (Reid 1991: Part 2) which she described as markers of their "common culture". They also have a self-regulating power system where operators and administrators enforce behavioural guidelines and rules created by "the community" and punish or permanently exclude violators. While points like these are definitely important and valid they might on their own not qualify to use broad and slippery terms such as "culture" or "community" as discussed above. Therefore this study will refer to these criteria as important characteristics of the IRC mother domain rather than indicators of an IRC community. Nevertheless they certainly help to distinguish IRC from other sources of CMC and render it as a large coherent group providing a good basis for meaningful research.

It should however be kept in mind that even though the technical framework of IRC provides the basis for a shared "paralinguistic vocabulary", language might still differ among the various semiotic sub-domains within IRC. Therefore even though statements can be made about language on IRC in general they can never be as accurate for every domain found in IRC. The most obvious example are emoticons: While this study can show which emoticons are generally used in IRC, some of them will be more popular in

certain areas of IRC than in others and IRC users might adapt their language behaviour depending on the virtual place they are in, as this quote from a long term IRC user illustrates:

> Seit mehr als sieben Jahren bin ich mit den gleichen Leute in einem IRC-Channel […] Ich denke, es entwickelt sich je nach Gruppe ein eigener Chat-Slang. […]
> So nutze ich hier im Forum Emotions, die ich im IRC niemals tippen würde. Und umgekehrt hacke ich dort Sachen in die Tastatur, die niemals den Weg in ein Forum schaffen werden.

In addition to the technical framework there are other characteristics that tie IRC users together as a loose group. Some networks have an individual focus they are dedicated to. Quakenet for example hosts many channels dedicated to computer games while Freenode is specialised on open source software. Therefore they both attract a crowd of users with special background and/or interest in these fields. This might lead to the use of slightly different expressions or emoticons, as some of them (^^ for example) are believed to be more popular with people that have a background influenced by computer games because their usage partly derived from that area (see chapter 6.3). Thus one can say that Freenode represents a sub-domain of IRC with a distinct focus and interest and that its channels are again sub-domains of Freenode. The linguistic differences between these sub-domains might in most cases hardly be noticeable but it is important to keep in mind that Freenode might attract a different crowd of average users than an average web-chat will.

Additionally IRC is not as easily accessible as a web-chat: you need to download a client, need to know which network you want to go, need to find a server for that network and need to know how to operate the client system, join channels and other commands such as the possibility to "auth" yourself in order to be recognised by the IRC network as a permanent user. For that reason it can be assumed that IRC has a high percentage of users that are experienced with computers and internet, while new or inexperienced users might prefer web-chats or other programs such as the windows messenger for their much easier accessibility. For the purpose of this study this fact has a considerable advantage: Firstly it is much more likely to encounter regular participants of the semiotic domain rather than brief visitors. Furthermore experienced

users will be much more skilled in using substitutes for paralinguistic cues and thus should be able to provide more valuable information on how these language features can successfully be employed than new and inexperienced users could.

Gelleri (1998: 6) conducted a so called *newbie experiment,* where new and inexperienced users were introduced to IRC and simply had to communicate with no further limitation. He found that IRC requires new users to "improve and expand their communicative competence". Many of the paralinguistic cues described in this paper were unknown to the users and they had to familiarise themselves with these language concepts in the course of the experiment to be able to blend in with the "community" and become an insider instead of an obvious outsider. Reid (1991: Part 2) agrees that successful communication in IRC depends on the use of the conventions that can be found in IRC to substitute paralinguistic cues. This is however not the only aspect non-CMC users will have to get used to. The results of Gelleri's experiment allow the conclusion that the test subjects' use of conceptual orality in the written medium (as defined by Koch/Oesterreicher) was less advanced and their utterances turned out to be over-formulated and too much like traditional writing, making their status as new and completely inexperienced users very obvious.

This distinction between in-group and out-group based on specific communicative abilities and knowledge is something that IRC has in common with communities of practice. They also make a distinction between core and peripheral group membership. Furthermore communities of practice encourage gradual learning of norms, which enables shifting from the periphery to the core. (Corder 2007: 445). Note that the in-group in question is the whole IRC domain rather than only one sub-domain within it. The essential features of communication necessary to adopt for blending in with the IRC crowd are basically the same for every place in the network - they are shared by the semiotic mother domain of IRC. Of course adopting the communicative practices of IRC does not imply that one might already be part of the in-group of specific semiotic sub-domains within IRC, since they might include special rituals and conventions or a more detailed set of rules. It should however allow the user to communicate and move through IRC without striking out as a "newbie" who does not know how to "talk properly".

In summary, IRC can hardly be seen as a coherent virtual community in the sense of the *virtual settlement* theory or other popular approaches to this slippery term. On the other hand it is more than just a huge unrelated body of "people interacting online". The uniting technical framework and the shared linguistic conventions provided by the mother domain are equally present and relevant in all sub-domains and it can be assumed that despite minor differences conclusions made about language in certain parts of IRC can be extended to the whole mother domain. With reasonable sampling and a large enough corpus the results will not be limited to one small and specific group of people, but to a whole area of CMC communication: Public many-to-many communication through the medium of IRC. Regarding this study the following characteristics of IRC communication are interesting:

- A common way of substituting paralinguistic cues leads to a similar "CMC language" throughout IRC.

- Regular members can be expected to be experienced users of CMC and should therefore possess the skills to express themselves with all necessary tools available to them.

- IRC users of various sub-domains are tied together by a framework of guidelines and norms given by the network they use to communicate.

So far researchers have always treated IRC as one big internet chat, whether they used the term community or avoided it. Like all previous studies the present one aims to make claims about the IRC mother domain. To do this careful sampling is necessary to provide a balanced and unbiased picture of the use of paralinguistic cues, which will be addressed in the next chapter.

5. Research method

In order to investigate IRC communication I have assembled a corpus of logfiles, which can be created automatically after enabling the option on an IRC client (in the case of this study the client *mIRC* was used). For each channel being joined, a new file will be created on the computer, consisting of the channel name, network and date of observation. It will log everything down that happens while being in the observed channel. As already mentioned it is common for IRC users to remain passive and 'lurk' in a channel without participating. Therefore a quiet and unknown participant is usually ignored and does not receive any attention. This makes IRC a good place for observing as the *observer's paradox* does not apply (see Gelleri 1998: 27) and the behaviour of active participants will not be influenced by passive observers.

I did not seek consent from the people observed to use these files, following Cameron's (2002: 27) argumentation that all regular IRC-users are aware of the possibility to create logs and thus know that their conversation is not really private. Additionally since every participant uses a nickname their real identity remains concealed. Most importantly however it would be impracticable to seek formal consent from every person involved, since people are joining and leaving channels continuously and even if a researcher would try to ask permission from all the thousands of users ever stepping foot in an observed channel, there would be no possibility to contact them or track them down once they leave. Lui (1999: Ethical issues) agrees that conversations in public IRC channels are "public acts deliberately intended for public consumption", similar to messages posted on newsgroups and concludes that "recording, analyzing and reporting" of such content can be done without explicit consent of channel users or other precautions by the researcher.

Out of the more than 300 log files created for this study some were not included in the final corpus for being irrelevant or inappropriate as discussed later. After the selection process 237 files remained in the corpus and 120 different channels were observed, containing a total number of 3 million words. Observation times differed considerably and were sometimes longer than 24 hours to take participants from all time zones into account. The focus of this study was on small-talk and general conversations even

though a small amount of gaming related and technical channels were added because they were characteristic for two of the major observed IRC networks (Quakenet and Freenode). A pilot study was conducted in May 2008 and the main study in February and March 2009. The tool *Word Smith* by Mike Scott was used to process and work with the corpus material. The corpus was intended as a project-related corpus of raw data and in the course of my research it was used to search for relevant paralinguistic language patterns, provide illustrative examples and gather numerical data.

Logfiles from five different IRC networks were included in this study. This serves not only to increase the target population, but is also an attempt to make the samples more representative for general IRC communication rather than for only one part of it. Networks can differ considerably; not only in their size, but also in their range of average users and have different emphases. Therefore it was necessary to investigate more than one to be able to make an unbiased and general statement. The following networks were included in this study:

- Quakenet (average users: 80,000-135,000; channels: 85,000-100,000[1])

Currently the largest network, originally created for a computer game and still attracting a large crowd of gamers. In addition to gaming communities numerous social chat channels exist and users from all over the world can be found talking about almost anything. The maximum amount of users online at the same time was 242,125 recorded in 2005.

- EFNet (average users: 50,000-55,000; channels: 25,000-26,000)

The modern-day descendant of the oldest and original IRC network. Many splits have decimated the size of this network, most notably the split into an American part (EFNet today) and a European part (IRCNet) also known as the "great split". However it firmly remains one of the four largest IRC networks today.

- Undernet (average users: 90,000-110,000; channels: 26,000-27,000)

Currently the second largest network. It was founded in 1993 and counts as one of the oldest. Today it is probably the most promising contender for the title "most popular network", which Quakenet has occupied since October 2002.

[1] Statistics taken from: http://irc.netsplit.de/networks/, a site that provides online statistics on IRC. Date: 25.5.2009

- **DALnet** (average users: 25,000-33,000; channels: 14,000-15,000)

  DALnet quickly rose to being one of the most popular networks but suffered from severe attacks and breakdowns in 2002. After that its user count dropped considerably. It still attracts users from all over the world and was included to represent a medium sized network.

- **Freenode** (average users: 45,000-57,000; channels: 21,000-24,000)

  A popular network with a major focus on free software and technical support. Also hosts a small number of social channels.

Much more decisive than the selection of networks was the decision which channels to observe. As previously discussed some language features might be more prominent in certain parts of IRC than in others. Therefore it can be assumed that the selection of certain channels will alter the numeric results of the study and careful sampling is necessary to provide a balanced and unbiased picture of the use of emoticons and other paralinguistic cues. This study has generally tried to broaden the range of samples in order to provide a representative picture of IRC. Despite all efforts, one might still expect that a completely different selection of channels could change the numeric outcome considerably. In order to investigate the extent of this thread I compared the results of the data gathered in the pilot study with the results that the major corpus provided. Both datasets were from a different timeframe, varied greatly in size and included largely different channels. Despite all these differences the end results are convincingly similar. When comparing the occurrence of emoticons the percentage for each smiley does not differ more than 4.2% and the average deviation is around 2%. A notable exception was the single German channel included in the corpus. It appears that this channel had a completely different distribution of emoticons than the majority of the logfiles. While minor differences exist between many channels the unrepresentatively high amount of emoticons in this single channel would have seriously altered the results. Therefore this channel was disregarded for the study of overall emoticon distribution in order not to distort an otherwise coherent picture. Possible reasons and implications for this deviation will be discussed in chapter 6.4. In all other aspects the mentioned channel behaved just like regular channels and was

included in the study for all other purposes. Therefore one can argue that features like emoticons may not always evenly distributed among IRC channels, however it seems that in most cases the differences are relatively small and easily balanced by the size of a representative and extensive corpus. The end results do not seem to be overtly affected by the selection of IRC channels, as long as the selection of samples is broad and diverse enough, which makes general assumptions about language use in IRC valid.

In the assembling process of the corpus I encountered two potential problems: logfiles without any relevant communication and logfiles containing massive amounts of spam messages. Due to the tradition of 'lurking' it was difficult to estimate beforehand whether a potential channel would feature relevant communication or not. The user count alone was often not sufficient and many well visited channels remained passive. The channel *#failures* on EFnet for example hosted over 60 persons but not one single utterance was made within more than 22 hours of observation. Because these channels are not very valuable in a corpus I decided to exclude every logfile that did not feature 'relevant' communication and discarded every logfile as 'irrelevant' that either contained only system messages or only single isolated utterances without any response by other participants. In the pilot study this procedure reduced the amount of channels from 62 to 44. In the main study this procedure followed similar lines.

Channels containing large amounts of spam messages were more problematic, since massive amounts of spam could seriously affect the numerical data. When I counted all the instances of emoticons with *Word Smith* in the pilot study I used the *plot* feature that clusters all instances of the investigated item. This makes it possible to see where a certain feature appears especially often in the corpus. Figure 1 shows an excerpt of the *plot* of the smiley :) and it is striking that the logfiles 25-27 contain an unrealistically high amount of this smiley. After looking at the suspicious files in more detail I discovered that the channel #search.pl contained massive amounts of spam messages, which means that the same utterance, advertisement or link was posted multiple times in rapid succession. Therefore this channel did in most parts consist of 'dumped' messages, rather than constructive communication. I decided that the repeated broadcasting of advertisements or requests would not be the kind of discourse I am interested in, but would seriously influence the results of the study, so I excluded all massive-spam logfiles from my corpus. After excluding only 3 out of 62 logfiles, the :)

smiley rate dropped from 815 to 460 in the pilot study and the :D rate from 614 to 307. It was interesting to see that the :(smiley rate dropped only from 158 to 156. While positive smileys like :) and :D where almost halved, the sad smiley :(was hardly affected at all. This underlines that the inclusion of spam-friendly channels in a research corpus does have a considerable impact on the distribution of paralinguistic features and that their exclusion has to be considered when doing research in that area. It is more difficult to discover and deal with massive spam when a channel becomes victim to a spam attack, as discussed in chapter 9.3. This happened more than once in the present corpus and I tried to cut out occurrences of spam attacks, while leaving the rest of the channel in the corpus. *Appendix 5* shows an excerpt of such a spam attack on the channel *#facebook*.

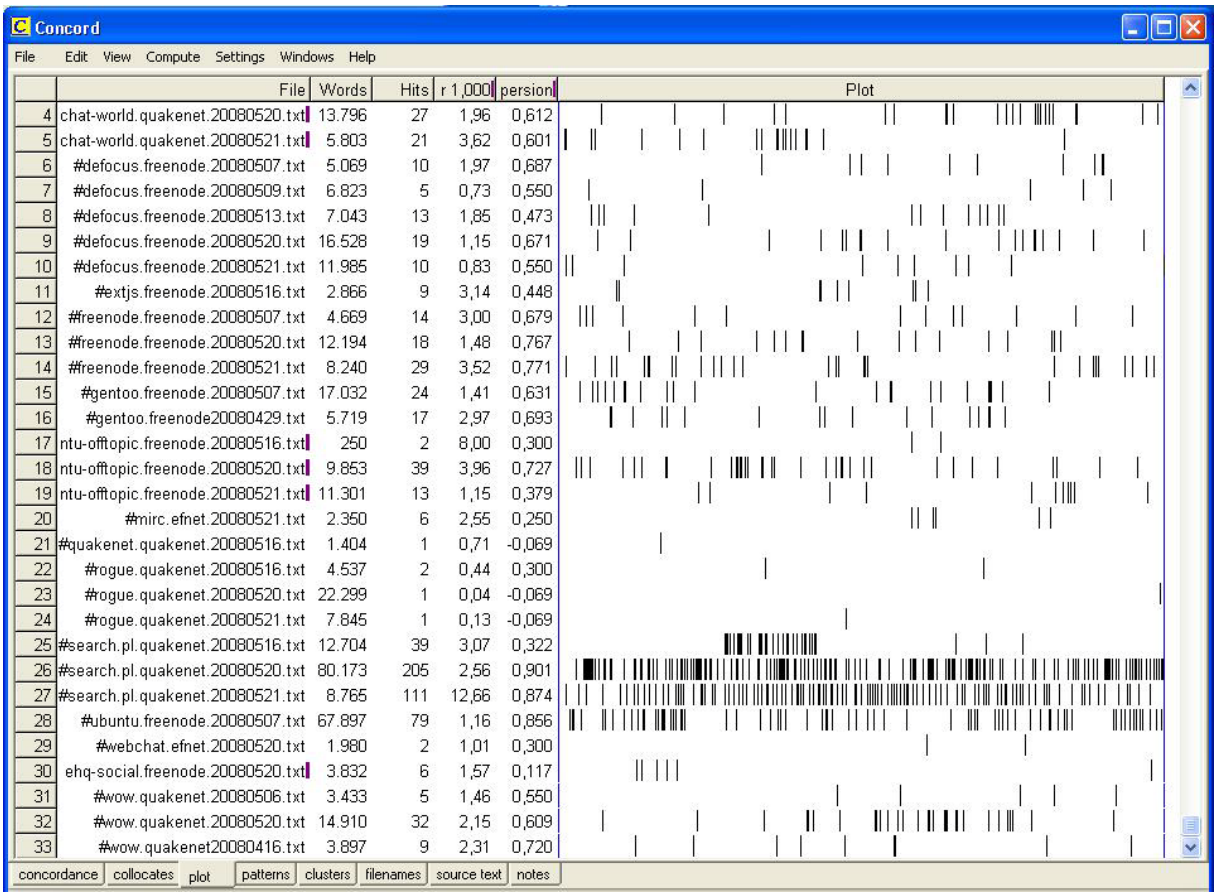


Figure 1: Plot of the smiley :) in Word Smith revealed mass spam in three logfiles

A further problem that occurred when working with the corpus data were the large amounts of system messages that made many search queries impossible. Several interesting language features made use of the asterisks sign, however so did all the

system messages. Every time a user joins or leaves a channel a message like this is created:

* tarantino (~moro@196.218.222.122) has joined #allnitecafe

Considering the fact that many channels contained a very high percentage of these system lines a sensible search for similar commands that were linguistically interesting (actionlines, actionstrips or intensifications) was impossible with *Word Smith* since every query would be flooded with system-message results. Since neither version 4 nor 5 of *Word Smith* could exclude the unwanted results another way had to be found and a regular expression has been executed on all files to remove all disturbing lines from the corpus and thus create a second and cropped version. Even though some information got lost in the process, most search queries were much more successful in the reduced corpus. With the following regular expression

> .*(has joined|Quit|has left|was kicked|Topic|set by|now talking in|
> disconnected|.**+).*

the following lines were permanently removed from the corpus:

1. * *user X* has joined *channel Y*
2. * *user X* Quit IRC
3. * *user X* has left *channel Y*
4. * *user X* was kicked by *user Y*
5. * Topic is... (displays channel topic)
6. * set by *user X* (displayed when a topic is changed)
7. * now talking in *channel X* (displayed when joining a channel)
8. * disconnected (displayed when disconnected from server/leaving the network)

Additionally every line containing more than two asterisks in a row was removed with the above regular expression. This was done to exclude any lines containing hints from trivia bots. Every question asked by a trivia bot will produce several hints until answered, for example:

<+triviabot> Question 4: Lyrics: When I'm lost at sea I hear your voice and it carries me?
<+triviabot> Hint: **a*** ** * P**** ** E*r** Beli*** **r*****
<+triviabot> Hint: **a*e* i* a Pl*** ** E*r** Beli*d* **rl****
<+triviabot> Time's up! The answer was: Heaven is a Place On Earth Belinda Carlisle

As one can see they often contain a large number of asterisks and made gathering numerical data for occurrences of language features containing asterisks problematic.

The original corpus was still used for many search queries because some of the cropped information might prove relevant - especially topic changes or trivia, but in some situations also leave-/join or kick messages. After removing all the lines mentioned above the total amount of lines in the corpus was halved from 400.000 to 200.000 as were the words from 3 million to 1.5 million. Many channels lost about two thirds of their lines, while from some channels almost all content was removed. Others were notably less affected by the procedure. Channels that turned out to be redundant after this process were removed from the corpus, causing 24 more channels to disappear even though they were originally large in size. The immense amount of cropped lines illustrates how many system messages flooded the corpus and it is not without reason that the most common words of the original corpus were *has / joined / left / IRC*, while apart from *has* none of them can be found anywhere near the top in the reduced one. Due to this procedure almost every unwanted asterisk was removed from the corpus and search queries on items containing this symbol could be made without further complications.

In addition to the main corpus I have occasionally used other corpora such as the Dortmunder Chat Corpus (http://www.chatkorpus.tu-dortmund.de/) to investigate the occurrence of emoticons. In the course of this study two smaller corpora were assembled for comparative analysis with the main corpus, one with a number of German logfiles used in section 6.4 and one consisting of dated logfiles from 2000 or older used in section 6.5.

With over 3 million words the size of the main corpus is more than satisfactory and I have tried to keep the samples as representative and unbiased as possible. Therefore this corpus should provide an excellent basis for an in-depth study of paralinguistic features in IRC.

6. Emotions

6.1 Communicative Functions of Emoticons

One of the most common markers of paralinguistic cues are emoticons or smileys. The word *emoticon* is a blend of emotion and icon (Haase et al 1997: 64). More or less iconographic symbols typically representing faces are used to convey facial expressions and basic emotions. Thus an addition of a smiley to an utterance ensures that an emotional drift is attached to it.

However one has to keep in mind that emoticons are not the same as the emotions they seem to represent. While in face-to-face conversation paralinguistic cues and facial expressions are often unintentional, emoticons always have to be set deliberately (Troest 1999: 2.3). Additionally they frequently serve as conversational turns in their own right as opposed to paralinguistic cues in face-to-face discourse, which are usually a parallel channel of communication (Troest 1999: 2.5). Marccocia et al (2007: Facial expression and emotion) note that

> the absence of smileys does not signal the absence of an emotion, whereas the absence of nonverbal expression raises questions about the presence of an emotion. At the same time, the presence of a smiley does not necessarily signal an experienced emotion, whereas most facial expressions are linked with an emotional experience.

Emoticons can fulfil multiple roles in CMC conversation: They have an evaluative function (Dittmann 2001: 74) which means that the addition of an emoticon allows others to interpret the meaning of an utterance. This can for example be used to mark otherwise offensive remarks as not meant seriously thus avoiding misunderstandings and preserving harmony. Secondly they can be used to express politeness and friendliness which can be seen most easily in greetings where positive emoticons are added to indicate goodwill and approachability (Dittmann 2001: 74). Dittmann further argues that positive emoticons (smiling faces) heavily outweigh negative ones, a fact that can be confirmed by the corpus of the present study. He assumes that positive emoticons serve to improve the prevailing mood and cohesion of a group. Furthermore emoticons can be used as a backchannel device (Troest 1999: 2.5; Kramer 2008: 61). In

example 01 the addition of the emoticon o.O (indicating suspicion) by *nazgjunk* shows not only that he is following the conversation but also allows him to give feedback ("I find it hard to believe that"). This might have encouraged *ivan_w* to elaborate on his claim to know a programming language that consists of only one statement.

```
01.   <ivan_w> nazgjunk: I saw there is even a language that has only
      *ONE* statement (and is yet Turing complete)
      <nazgjunk> o.O
      <ivan_w> The instruction is something like 'substract and
      multiply then jump on condition' (or something similar)
```

Many meanings conveyed by emoticons could also be described verbally (for example "I find it hard to believe that" from the example above), however the use of an emoticon is considerably shorter. As discussed earlier saving time and simplicity are of great importance in internet chat, which means that a very practical use for emoticons is the advantage to save many keystrokes compared to verbal constructions. Additionally one has to keep in mind that emoticons are often vaguely defined and leave room for interpretation (see chapter 6.5). This allows indicating a vague emotional drift without the necessity to provide a clear verbal description of the actual feeling or emotion involved, which makes describing feelings (like joy and happiness) less awkward and difficult. Haase et al (1997: 81) mention that these ideograms are difficult to verbalise (for example when reading a chatlog out loud), have no medially spoken pendant and no apparent translation convention exists. Additionally Haase et all (1997: 81) argue that it is not typical for written discourse that exact function and meaning has to be deduced from the co-text, as it is the case with emoticons. This pragmatic complexity can according to them rather be found in spoken communication.

A study by Witmer (1998) suggests that emoticons and other graphical accents (such as examples 5-8 in chapter 2.2) are more likely to be used by female users than male. The study was conducted with asynchronous CMC discourse and had to struggle with a low overall occurrence of graphical accents (only 13.2% of the total sample) and a low amount of female posters (16.4%), limiting the validity of the study. However the graphical accents that were used turned out to come predominantly from women

(Witmer 1998: 9). It remains questionable whether this result could also be applied to synchronous environments such as IRC. A reproduction of this study on IRC would be very difficult since there is no real possibility to deduce gender from nicknames, apart from some obviously gendered names (which still would rely on guesswork). Therefore the results of this study can only remain a vague suggestion that women tend to embrace graphical accents more easily and the assumption made by Witmer (1998: 9) that the acclaimed female emotionality might "translate to the computer-mediated environment" is rather difficult to support.

As discussed in chapter 2 acronyms and abbreviations like *cu* (see you) are very popular in internet communication and frequently used. Since they are simply shortened versions of longer utterances little paralinguistic worth can be attributed to them and they will not be discussed any further in this paper. However, there are some noticeable exceptions: The acronyms *lol* (laughing out loud) and *rofl* (rolling on the floor laughing) are commonly used to express laughter and mark humour. Especially *lol* is very popular with 6379 occurrences in the corpus. Although they are acronyms, one can argue that they are related to emoticons like :D as they have a very similar purpose in communication. Unlike abbreviations like *cu* they are not used to abbreviate phrases meant to be read out in order to save time. Instead they are attached to utterances (or constitute utterances of their own) in order to mark humour/laughter. Troest (1999: 2.5) speaks of *lol* and related language features (like *g*) as "semantic equivalents to emoticons" and Marccocia et al (2007: Facial Expression and Emotion) also suggest that these "emotional acronyms" are used to compensate the lack of nonverbal cues. Indeed IRC users always have the option to choose between using an emoticon or the acronym *lol* to express humour. Chapter 6.7 will provide a close analysis of these humour markers and compare their overlap in functions and their differences. Examples 02-04 illustrate how instances of *lol* and :D are used in a very similar way to mark humour. Also note that 04 includes the ^^ emoticon as another variant to mark laughter as discussed in chapter 6.7.

```
02.   <MidnightCommando> i'm mildly traumatised by this lol

03.   <Samanth0r> *sigh* so much for toner. i have 28 dollars to my
      name till friday.
      <Cann0n> Samanth0r, thats 25 more than i have! lol

04.   <Treval> Computer Science isn't that hard, I'm doing it. :D
      <Treval> I even have a book on logic. ^^
      <Treval> With triangle symbols and all.
```

6.2 What are relevant emoticons?

An interesting question to ask is which emoticons are relevant for communication and on a more general basis what can be considered an emoticon at all. Even though many smiley lists are available from literature or the internet, they usually do not manage to provide a suitable comprehensive list of all important emoticons in internet chat. Many sources claim that only the "basic" emoticons :) ;) and :((and their variations) are relevant (for example Beisswenger 2000: 99; Dittmann, 2001: 73; Hentschel 1998: 3.1.2) while others feature extensive lists of exotic and rarely used emoticons that hardly bear communicative relevance. While these lists tend to include a large amount of unnecessary smileys they still fail to include others that are very commonly used. Based on the extensive IRC corpus this study tries to provide a detailed list of all relevant emoticons in IRC conversation and discuss what separates them from thousands of other smiley variants.

Smiley collections including over 2000 items like Marshall's *canonical smiley (and 1-line symbol) list* might lead to the impression that internet chat discourse features a broad variety of different emoticons. This impression is however quite misleading and I agree with Schulze (1999: 76) that only a small number of emoticons are relevant for IRC discourse and used regularly in internet chat. Most emoticons collected in smiley lists appear to be 'joke smileys' like :-F "user is a bucktoothed vampire with one tooth missing" (Crystal 2006: 40). Even relatively small lists like the one from Crystal (2006: 40) mention only few 'basic smileys' compared to the large amount of 'entertainment' emoticons and while some very important items (like :D with over 5000 corpus

occurrences) are missing, other emoticons from Crystal's small basic category seem to be largely irrelevant (e.g. the emoticon %-) standing for "confused" was only used 3 times in the whole corpus, all instances produced by the same user).

It is claimed that the invention of new smileys is popular in IRC (Haase et al 1997: 64). However such rare variations and unique inventions like +-:-) ("user holds a Christian religious office" (Crystal 2006: 40)) mostly serve as riddles or as a humorous play with symbols and language. It is important that one can distinguish between these smiley 'wordplays', stories (see examples *07-08*) and riddles (along with more elaborate pictures constructed by ASCII characters, such as in *example 05*) and emoticons with a real communicative function. Only emoticons from the latter group are able to fulfil the important roles that smileys have in IRC discourse as discussed in chapter 6.1. Emoticons from this group act as important paralinguistic cues even though they can have slightly different functions than a pure translation of facial expressions. Therefore they can claim to be "icons transporting non-verbal information" (Schulze 1999: 76). The other group of emoticons does no such thing and rather represents the playfulness and creativity of the medium. In other words they are employed primarily or solely for entertainment, which is definitely not the case for basic emoticons like :). When talking about emoticons the two groups of communicative and entertainment emoticons should therefore be separated and discussed individually.

If one wants to define what 'relevant emoticons' in IRC are it should be based on the two following criteria:

- the emoticon is used regularly in IRC

- the emoticon transports non-verbal information and thus fulfils communicative roles

It appears that these two criteria overlap and all regularly used emoticons also transport non-verbal information. Emoticons that fail the second criteria are rare - probably due to their creative and individualistic nature - and largely insignificant when investigating signs of paralinguistic discourse. Thus it is possible to create a comprehensive list of all relevant emoticons in IRC based on the frequency of their occurrence, which will be discussed in chapter 6.4.

05. <hola> (___/)
<hola> (o'.'o)
<hola> (")_(")
- an animal 'painted' in IRC with ASCII characters. Not an emoticon but ASCII art.

The second question addressed in this section is which iconographic expressions can be called emoticons. A simple and valid argumentation is that every combination of characters that does not focus on the display of faces should not be called an emoticon but simply ASCII art. Many of the 2227 items from Marshall's list look like this:

```
06. `'~,.,~'`'~,.,~'`'~,.,~'`'~,.,~'  Wave or banner décor
07. -+#:|-|-<  <*,^^,--        Dragon vs. wizard
08. -+#:)->-< <* `vv`--        The wizard won
```

They use ASCII characters to tell a story, provide a riddle or present simple graphical illustrations and thus qualify as creative ASCII art just like the animal from *example 05*. However even though they might include a face somewhere (like *07* and *08* with characters 4-5) they are clearly more complex constructions and should therefore not be considered an emoticon. According to the present study every relevant emoticon in IRC represents a stylised a face. By defining that only character combinations that focus on a face are considered emoticons, no relevant emoticon would be excluded.

As discussed in chapter 4.2 IRC is an interesting place to observe the use of emoticons because it is a purely text-based medium. Usually online message boards and most other web-chats offer a broad variety of graphical representations for smileys. IRC however lacks any graphical representation or any buttons that could be used to automatically create entertaining but otherwise communicatively rather unnecessary smileys. Arguably only emoticons that carry enough necessary paralinguistic content are commonly used in such an environment and all smileys that do not express 'basic emotions' are more or less reduced to non-existence. Additionally only emoticons that are easy enough to produce and decode survive. Note that this might make smiley-behaviour in solely text-based CMC slightly different from graphically supported forms.

See *appendix 7* for an extensive list of smileys used in a message board (www.justfindout.de). While many of these smileys are used frequently in this board, only 9 out of the 97 have any form of textual representation in IRC. It appears that all 9

emoticons shared by IRC and the message board represent basic emotions. Therefore it seems that despite some deviations due to a different 'vocabulary' of paralinguistic signs the same basic smileys remain the most prominent emoticons in message boards and web-chats alike. This is illustrated by an excerpt from the statistics section provided by the message board mentioned above (see table 1): Six emoticons that have equivalents in IRC can be found in the top ten (assuming the :misstrau: emoticon is a direct equivalent to the suspicion emoticon o_O). While all six of these emoticons express basic emotions, the remaining four smileys in the top ten can not be found in IRC and belong to the 'individual vocabulary' of the board. Examples of these are the *:jump:* smiley on the fourth place and the *:ugly:* smiley on the fifth. The graphical representations of the 'basic' emoticons ;D :) and ;) are by far the most frequently used in the message board. Furthermore the emoticon ^^ (see 6.3 and 6.4) was found to be very popular in this board as well with ca. 4000 occurrences (not included in the table below because it does not have a graphical representation). Compared to table 2 (see 6.4) one can say that these emoticons enjoy a similar popularity in IRC. Thus one can say that graphically supported CMC environments will always be influenced greatly by the paralinguistic signs they have available. However the most basic emoticons, which are exactly the ones that can be found in IRC, should appear in almost every environment and remain the most frequently used.

;D
:)
;)
:jump:
:ugly:
:misstrau:
:uglyspitze:
:P
:D
::)

Top10 Smilies	
	7487
	7052
	5849
	2758
	2533
	2186
	2034
	1879
	1654
	1453

Table 1: This table shows the ten most frequently used emoticons from the message board www.justfindout.de. The box to the left shows the combination of characters that will produce each emoticon.

6.3 Influences from the East

Most discussions of emoticons in the literature, including the ones from Crystal, Schulze, Haase et al and Geers, only pay attention to the traditional 'Western' type of smileys. However there is a second influential type of emoticons that originates from Asia, but is becoming increasingly popular in the Western world of internet communication. These emoticons are loans from the Asian emoticon culture where smileys are displayed differently than in Western CMC communication.

The following examples show typical Eastern emoticons in their original form:

(-_-)	(^_^)	(o.O)	(T_T)	(^_~)	(>_<)
sighing, sad	smiling	suspicious	crying	winking	frustrated, angry

They represent a face (brackets) with two eyes and a mouth. As opposed to their Western counterpart they have the advantage that it is not necessary to tilt your head sideways to deduce their meaning. It is also very interesting to observe that while Western emoticons convey their whole emotional and paralinguistic content with the mouth, the meaning of Asian emoticons is deduced from the eyes, while the mouth remains expressionless. This difference in displaying emotion and facial expressions can be traced back to cultural differences between East and West:

> Given that the muscles around the eyes are more difficult to control than those around the mouth when a person conveys emotions [...], the eyes of others may be most diagnostic of their true emotional state for individuals in cultures where emotional restraint is the norm, such as Japan. By contrast, in cultures where overt emotional expression is the norm, such as in the United States, the more dynamically expressive mouth may be considered a better cue to another's emotional state. Thus, typical cultural practices in expressing or subduing emotions may also be manifested in the different cues that people use to interpret others' emotions. (Yuki 2007: 18)

Yuki showed American and Japanese people emoticons that displayed emotions both in the eyes and mouth. The subjects had to rate the faces from happy to sad and it turned out that

> Japanese weighted the cues in the eyes more heavily than Americans, whereas Americans weighted the cues in the mouth more heavily than Japanese (Yuki 2007: 10)

which provides an explanation "why stylized facial icons seem to differ" (Yuki 2007: 21) between East and West.

When used in Western internet communication as a loan the original brackets surrounding the emoticon are dropped and the emoticon is reduced to its eyes, which carry the whole paralinguistic content, even though many of them keep the 'mouth' for aesthetic reasons. Therefore the adopted versions of these emoticons look like ^^, T.T, -_-, o_O. It is important to keep in mind that these are loans rather than a simple adaption of the original emoticons and thus might be used differently in Eastern contexts than in Western discourse. This means that an Asian person could use and understand the emoticon (^_^) differently than an English chat user would use ^^.

Through contact between Eastern and Western culture over the internet these emoticons were adopted by Western internet chat users and while at first they were mainly used in Manga/Anime and internet gaming communities they became increasingly widespread in English and German CMC. Nowadays they can be found in many diverse places, for example the message board of legal studies of Vienna or popular services such as *studivz*.

6.4 Emoticons in IRC

Table 2 shows the occurrence of the most frequently used emoticons in the corpus, together with a vague description of their basic meaning. For this purpose variations like :p :-p ;p :ppp were added to the occurrence of the basic form. It is notable that it was only popular to include 'noses' with the basic smiley :-), while all other forms were almost always used without them. All Eastern emoticons were used in their Western form. A corpus search was performed with many different smiley combinations and it appears that apart from the 17 emoticons mentioned in this table there were only a few other items that occurred about 50 times. These few and rather rare emoticons (like :* or :c) were not included in the further observation because they were considered not

general, widely spread or relevant enough for this study. All other smiley combinations entered produced either no results or only very few (usually between 1 and 5). It seems that altogether not more than 20-25 emoticons have more than 10 occurrences in the assembled corpus. Therefore this study can agree with Schulze (1999: 76) that the group of relevant communicative emoticons in IRC is relatively small and clear-cut as it appears that the 10 most popular emoticons (with their minor variations) make up almost the whole range of emoticons used in IRC. It is also interesting that all of them express basic emotions.

:D or xD	5776	laughter	o_O	650	suspicion
:)	5720	'smiley'	-_-	539	sadness
:p	3193	pointing out tongue / annoyed / sarcasm	:/	454	unhappy…(see chapter 6.6)
^^	1480	happiness / smile-laughter	:>	426	very happy / vicious smile
:o	1404	surprise, shock	:X	261	awkwardness / reluctance to speak
:(	1234	sadness ('frowney')	:S	176	confused
;)	1102	wink	others, e.g: T_T , :< , :3 , :|		ca. 480

Table 2: Overview of emoticon occurrence in the corpus

One can see in this table that :) and :D are almost equally important, with :D even occurring slightly more often than the basic smiley. In the pilot study :) remained number one closely followed by :D. Next in importance is the emoticon :p often indicating sarcasm or annoyance. It occupies a stable third position between the two leading emoticons and had the same position in the pilot study. The next four emoticons are relatively similar in popularity and include the most widely spread Eastern emoticon ^^, an emoticon for displaying shock and surprise (:o), as well as the two other 'basic' forms :(and ;). The seven emoticons mentioned so far can be said to be the most influential emoticons in IRC. However the next two items which are the Eastern loans o_O and -_- and represent suspicion and sadness are also worth mentioning.

It is interesting that suspicion is a feeling not transported by any traditional Western emoticon and therefore o_O is the only popular Eastern loan that adds a completely new

feeling to the Western vocabulary. A possible reason for the lack of a suspicious Western emoticon is that no convenient way could be found to convey this feeling with the mouth, while eyes work much better. The other two Eastern loans (^^ and -_-) add another option to represent the feelings of happiness and sadness, while T_T as a third and less frequent loan (with ca. 80 occurrences) adds an intensified version of sadness (representing tears running down from the eyes). One should keep in mind that the fact that T.T is less frequent than other loans does not necessarily mean that it is a less known emoticon - there simply might be fewer occasions to display intensified sadness in IRC.

Altogether two basic feelings seem to dominate emoticons: happiness and sadness. 10 out of the 17 emoticons mentioned in the table directly represent either of the two. Furthermore positive/happy smileys are by far the most popular emoticons and heavily outweigh their negative cousins. :) and :D alone provide 50% of all emoticon occurrences in IRC. If one adds the over 6000 instances of *lol* the dominance of happy emotions becomes even more extreme. This underlines that emoticons are used frequently in their function to establish a positive climate as discussed in chapter 6.1.

It seems that emoticons are immensely popular which is shown by the fact that according to *Word Smith* the most common emoticons :) and :D are the 16th and 17th most popular 'words' used in the corpus with only the following words being more popular (in correct order)[3]:

the, I, to, a, is, you, and, it, in, of, for, that, on, was, lol

Note that *lol* occupies number 15 and also serves a similar purpose than :D as discussed above. The emoticon :p is number 44 between *if* and *now* while :o, ^^, :(and ;) are still within the 100-125 most frequently used words. With only a few basic words being higher in popularity one can argue that emoticons are one of the most important communicative tools in IRC communication. The observation that emoticons are very common was also made by Dittmann (2001: 73) who claimed that 24.2% of all his corpus utterances contained an emoticon, while 4.6% of all words were emoticons.

[3] The reduced corpus was used for this query to avoid that the words "left" "joined" and "channel" would be the most popular words due to the amount of system messages.

Even though the percentage for this study was lower with 13% of all utterances containing an emoticon, the fact that more than every tenth line contains an emoticon and that the two most frequently used smileys are amongst the 20 most popular words shows that they are truly very common. *Examples 09* and *10* were randomly taken from the corpus and illustrate the high density of emoticons that can be found in some extracts of IRC dialogue.

```
09.   * @kitty pokes alvari with a sauna
      <+alvari> sauna? :DD
      <@kitty> yep ;>
      <+alvari> the hot steaming room? :p
      <@kitty> coz ur hot, right? ;p
      <+alvari> lol
      <+alvari> okay I will poke you with a livingroom
      * evil^angel frowns
      <+alvari> coz your living, rite?
      <+alvari> :D
      <evil^angel> how bout bedroom?
      <+alvari> a7x
      <evil^angel> cozim sleeping
      <+alvari> never heard of that :<
      <evil^angel> yeah
      <DanQnA> lulz

10.   <@Waller> Tuut tuuut
      <@Luna> wall e
      <@Luna> :P
      <@Waller> Now I stay :P
      <@Luna> uh?
      <@Luna> stay where
      <@Waller> Here :o
      <@Luna> oh
      <@Luna> :S
      <@Waller> :D
      * @Luna gives u blanket
      <@Waller> Thanks :>
      <@Luna> :D
      <@Luna> i wanna pm ya for a bit
```

As discussed in chapter 5 the single German channel observed in this study had a notably different distribution of emoticons and was therefore disregarded so far. It appeared that the concentration of :D, xD and ^^ was much higher than on average and the occurrence of :) was much lower. Since differences between channels are normal in IRC this would not have been a reason to exclude the channel on its own, but the amount of emoticons used in this particular channel was so high that it would have

seriously distorted the general results. :D and xD outnumbered :) 10:1 while ^^ turned out to be three times as popular as :).

It was therefore interesting to find out what caused this deviation from the norm. One implication that lies at hand is that German IRC users might have a different habit of using emoticons than English ones. According to Dittmann (2001: 73) emoticons can be used to a different extent in each language and he found differences in German and French emoticon use. For example in Dittmann's study 33.1% of the German emoticons, but only 6.8% of the French emoticons appeared to be of the type ;) Language differences regarding paralinguistic signs can also be seen by the fact that German chat users regularly employ the abbreviation *g* for *grin*, which is hardly used in English other languages.

In order to investigate whether the observed deviations could be language based I observed further German channels that were not included in the main corpus. It would go beyond the scope of this study to investigate cross-language behaviour in full detail and therefore only limited results can be provided at this point. It turned out that the additional German logfiles were similar to the channel in the main corpus, which means that :D, xD as well as ^^ were all more popular than :), even though at a much less extreme rate. To investigate further I worked with the Dortmunder Chat Corpus that included a large number of German IRC logfiles and was compiled in 2005. While ^^ was significantly more common than in the corpus of the present study, it turned out that :D was used less often. Altogether it seems very plausible that the reasons for this deviation is at least partly language based and that ^^ is more popular in German than English at the present time.

Apart from probably being influenced by the target language, there appear to be several other factors that have an impact on the distribution of emoticons. A close study via *Word Smith*'s plot feature revealed noticeable differences between many corpus channels regarding the frequency of some emoticons. It is possible that the tendency to produce large amounts of a certain emoticon can be influenced by both the network a channel is located on and the expected chat-expertise of the channel users (are they likely to have beginner status?). However from the perspective of this study they can only increase the likelihood of a certain emoticon pattern rather than being a definite

factor influencing the result. For instance the main corpus revealed that especially Quakenet channels appear to have an affinity for high ^^ and :D concentrations (as opposed to :)). However Quakenet channels do not generally behave differently than channels from other networks, as many channels on Quakenet are very similar to the ones on others. Likewise channels from other networks can also have the same proportion of :D and ^^ as the most extreme Quakenet channels, but it seems slightly less likely that they do so. At the same time channels where new users are expected (e.g. *#beginner*) seem to be more likely to feature a higher amount of :), even though not all of them do. Because such deviations are usually not extreme one can still produce stable and meaningful results regarding emoticon distribution in IRC as long as the samples are broad and diverse enough, the corpus is big enough and the most extreme channels are watched. Even though most additional German channels observed in this study were from the network Quakenet it still seems that German chat users tend to include more ^^ emoticons than English ones.

Concluding one can say that emoticons are an extremely popular tool in IRC and despite minor variations regarding popularity in individual networks and channels it is possible to determine all influential emoticons in IRC discourse. There are ca. 10-15 widely spread items that make up almost all emoticons used in IRC. More than 50% of all emoticons seem to display happiness, while other basic feelings like annoyance, sarcasm or surprise are also of considerable importance.

6.5 A Change in Emoticon Use?

Despite some similarities between the results presented by Schulze (1999: 76) and this study there are also considerable differences. It is very striking that the distribution of smiley occurrences had little similarities. While Schulze found that more than 90% of the emoticons were variations of the basic smileys :) and :(, the corpus of the present study implies a very different conclusion. In table 3 the major differences of the two corpora are shown. Most impressing are the huge differences in popularity of the emoticons :D from 0.6% to 25.2% as well as :) from 80.1% to 25.0%.

Smiley	Waldner		Schulze		Difference
:)	5720	25.0%	1308	80.1%	-55.1%
:(	1234	5.4%	192	11.8%	-6.4%
:p	3193	14.0%	84	5.1%	+8.9%
:/	454	2.0%	11	0.7%	+1.3%
:D	5776	25.2%	10	0.6%	+24.6%
^^	1480	6,5%	not mentioned		n/a
:o	1404	6.1%	not mentioned		n/a
;)	1102	4.8%	not mentioned		n/a

Table 3: comparison with Schulze's results

While :) has clearly dominated in Schulze's corpus as the single most important emoticon, the present one indicates that it has lost much of its former position. :D on the other hand has massively increased in popularity and seems to have gained equal importance. It can be implied that the 'basic' emoticons :) and :(have decreased in popularity while all other forms of emoticons mentioned by Schulze have increased at least slightly. Because the third traditional emoticon ;) is missing in Schulze's study no immediate claims can be made about this item. Schulze (1999: 76) mentioned that the seven smileys identified in his study are the only ones that appear in IRC. This exclusiveness was not the case in the present IRC corpus and even though there are still only about 15 popular and widely used emoticons, various rare items and variations exist that occurred less than 50 times.

These observations in the difference of emoticon use suggest that paralinguistic language in the internet, just like regular language, might currently be subject to change, even though at a much faster pace. Schulze's corpus was created in 1998 and in only ten years the use of emoticons in IRC has supposedly changed considerably. The basic forms have lost some of their former importance while many other forms of emoticons have rapidly spread and increased immensely in popularity.

To support this claim further evidence is needed than a numeric difference between two independent corpora queries can provide. It would have been very interesting for this study to investigate Schulze's corpus more closely, however it is no longer publicly

available. Further research on his corpus could have revealed whether other factors such as target population, samples or different search queries might have influenced the results. Especially the emoticon ;) would have needed further investigation as Schulze's table does not mention it at all, even though is very often treated as one of the three most important emoticons in the literature.

Without the possibility of a closer investigation it is even more important to find other factors that support such a claim and further sources that enable a comparison. Dittmann's corpus from 2001 lends further credibility to the claim of emoticon change as it shows similar results to Schulze with a total of ca. 70% of all emoticons estimated to be :) and a further 20% being ;). Another significant indicator that smileys are subject to language change is the increasing appearance of Eastern emoticons in Western internet chat. While older studies seem to disregard this variant of emoticons completely, they receive increasing attention in newer studies such as Kramer (2008). Wikipedia states that Eastern emoticons are commonly used in IRC since the year 2000 and are especially popular with younger users. Studies from the late nineties like the ones by Schulze (1999) and Geers (1999) might have had little opportunity to observe these emoticons, which would explain the complete lack of Eastern loans in their research. Crystal (2006) also does not mention any of them in his study (nor does he mention :D). He compiled his emoticon list "after Sanderson, 1993" and has apparently not updated it to the year 2006. It should be added that Crystal like many other authors investigated 'Netspeak' in a more general sense (rather than a particular area of it, like IRC) and Eastern emoticons today are still not believed to be used or understood in all areas of the internet, even though they seem to become increasingly widespread as mentioned in chapter 6.3.

To provide a further base of comparison I assembled a second corpus that consists of old IRC logfiles from around 2000 or older. The queries performed on the main corpus were repeated on this corpus, which will further be referred to as the Comparative Corpus. Claims about language change could be supported depending on whether the results of the Comparative Corpus would be similar to Schulze or the present study. Because IRC logfiles about casual conversation that are ten or more years old appear to be very difficult to find, the Comparative Corpus could only be rather limited in size. It includes files from a study by Gelleri (1998) as well as the small corpus from the

website “Investigating the Language of New Communication Technologies” (http://www.demo.inty.net/), which includes files from 1999-2000.

As a result the emoticon distribution is strikingly similar to Schulze, with :) providing a large percentage of the overall smiley use. Table 4 shows all the emoticons found in the Comparative Corpus:

:)	375	81.3%
;)	36	7.8%
:(	16	3.5%
:p	22	4.8%
:/	6	1.3%

Table 4: Emoticon occurrence in the Comparative Corpus

It was interesting to observe that :D (as well as ;D or xD) did not occur even once, and neither did :o or any other emoticon that is considered slightly popular in the main corpus of this study. Furthermore not a single Eastern emoticon could be found in the corpus. The fact that both old corpora (the one from Schulze and the Comparative Corpus) indicate :) as 80% of all emoticons lends credibility to the argument that :) has once been the dominant smiley but lost much of its importance in the last ten years.

Despite the small size of the Comparative Corpus, the results are so convincingly similar to other studies of that time, such as Dittmann or Schulze, that increasing the corpus size would only unlikely change the outcome considerably. Together with the fact that Eastern emoticon loans seem to have risen in popularity over the last decade there is much evidence that the variety, use and distribution of emoticons has changed. It seems that the necessity of expressing more than just the basic emotions of :) and :(has led to the development of a broader basis of widely used smileys.

6.6 Variety and Ambiguity

As discussed in chapter 6.1 even though emoticons attempt to imitate facial expressions and emotions, it has to be considered that they are certainly not the same. The key characteristic that Cook (2006: 45) attributes to paralinguistic features is that they are "graded rather than discrete". One example would be laughter which

> has many different forms and languages have many different words for types of laugh. Thus in English we have among others: titter, giggle, snigger, cackle, chuckle , guffaw. (Cook 2006: 46)

This represents a major problem that emoticons have to face: their ability to be graded is very limited. While it is possible to emphasise and intensify a smiley by adding additional characters, like :)))))))), xxxxD or -_____- this is a very unsatisfying way to distinguish between a hearty laugh and a giggle.

Some emoticons do however have certain overlaps, especially in the area of happiness/sadness. For example :(and :/ are both used to express sadness, but one could expect a slightly different connotation. Furthermore :D and ^^ both express humour and an IRC user does have the choice how he wants to express a laugh (see chapter 6.7). Despite some overlaps it is without question, that emoticons will never be able to match the variety and gradability of all the different emotions we show in traditional face-to-face communication.

Another drawback of emoticons is their potential ambiguity. Crystal (1999: 39) points out that "an individual smiley allows a huge number of readings (happiness, joke, sympathy, good mood, delight, amusement etc.)" and the intended meaning of an emoticon is not always obvious. Additionally the 'meaning' of a smiley is often not very well defined and for some items it might be difficult to grasp their exact connotation. A good example is the emoticon :/ which might be read in various ways. Schulze (1999: 76) describes it as "slight frown", English Wikipedia states "uncertainty", "bored", "annoyed" or "awkward" (05.2009 and 05.2008) Haralampieva (2004: 38) mentions it being "expressionless" and German Wikipedia includes: "das finde ich nicht lustig, unzufrieden, sprachlos" (05.2008) (That's not funny, unsatisfied, speechless). Unsatisfied with this variety of explanations, I also asked seven people

what this smiley indicates for them and the results were: afflicted, sorrowful, unhappy, uncomfortable, disappointed (but not really sad) and "a bit pissed off". A user that reads up the English Wikipedia explanation of 'uncertain' and does not know about the other possible meanings of this smiley might easily misinterpret the emoticon. To investigate the possible meanings of this emoticon more closely, random samples were taken form the corpus. The aim was to show that the emoticon can indeed be used in various situations. *Examples 11-20* illustrate some of these situations and include instances where users admit they do not know something (*ex.11-12*), show sympathy (*ex.15*), state an unpleasant fact (*ex.16-17*), employ sarcasm, or realise that they might have caused some harm (like 'scaring someone away' in *ex.14*). In *example 19 black_rose* emphasised her negative feelings about piracy with the emoticon :/ while *Zuu* used it in *ex.20* to underline the fact that he likes something that other people seemingly do not. Note that *example 13* shows a combination of :(and :/ emoticons when talking about being drunk. It would be interesting to find out existing differences between those two emoticons by conducting a close study of random samples similar to the one on humour markers in the following sub-chapter.

11. <starcannon> f0rmat very odd, I've had this laptop up almost non stop since release day, nothing like that, I actually rebooted this morning after I downloaded some updates
<f0rmat> starcannon, :/ i dunno then

12. <MrTaVi> what's nj ? :/

13. <+PrincesS^Rinoa^> im headache :(
<evil^angel> thats coz u drank too much Princess
<evil^angel> u should stay sober like me:)))
<+PrincesS^Rinoa^> Yes :/

14. <ZiniN> that includes lucifer, satan, the devil, et
* +iustus (~iustus@iustus.users.quakenet.org) Quit (Read error: EOF from client)
[...]
<ZiniN> nothing further, iustus?
<Watr> you scared him away
<ZiniN> :/

15. <PillzE> man, tomorow I have 10 hours of college courses
<PillzE> do these people think we don't EAT?
<+sweet`pea> :/

16. <Ivan-Shekleburgerheimenstein> lol omg I just set myself up as a poo expert
<Ivan-Shekleburgerheimenstein> :/

17. <Nadezhda> I drink to omuch :/ lol

```
18.   <+triviabot> Question 9: What is the flower that stands for:
      aversion?
      <@DevilNith> mistletoe
      <@DevilNith> :/

19.   <+dazedman> i watched slumdog millionaire..it was really good
      <@Abrienda> durn! im still waiting for that to be released here
      [...]
      <biotech> get pirated copy
      [...]
      <@Abrienda> no biotech i suck at pirating
      <black_rose> Piracy :/

20.   <Treval> Zuu everyone hates discrete mathematics.
      <Zuu> i dont :/
```

After asking seven people and reading four descriptions already 15 different meanings exist and the corpus examples suggest that all of these meanings can be possible uses for this emoticon. Concluding one can say that the emoticon :/ can be employed in various situations and that its exact meaning is not easily grasped. It has to be deduced from each situation individually. It can range from empathy to aversion and can stand for sadness as well as sarcasm, which explains the diversity of descriptions this emoticon receives.

The same flexibility is true for other emoticons, for example :p ranges from annoyance to sarcasm. In the course of this study especially the meaning of :X was very difficult to narrow down. Various people had to be asked and corpus samples analysed just to find a few words able to represent this emoticon in *table 2*. (See *appendix 4* for a few random corpus samples exemplifying the diversity of the emoticon :X)

Despite the fact that emoticons are not as accurate or diverse as their paralinguistic equivalents in face-to-face communication, their linguistic worth should not be underestimated. The possibility to attach emotions to utterances, however crude they may be is very powerful and an essential tool in spoken discourse. Smileys are able to diffuse many ambiguous situations by getting across some of the intentions and attitudes of the speaker. They are not used to emulate the whole range of paralinguistic cues in all its graded variety, but to give certain messages an emotional drift. We will see later on that should there be the need to display more elaborate emotions or use a more graded approach - for example to diffuse a special ambiguity or provide a more

colourful picture of ones emotions - IRC users can make use of other tools like *actionmarkers.*

6.7 Humour Markers

The most important area of emoticons seems to be the expression of humour/happiness and sadness, which is reflected in the fact that one can choose between multiple items in these areas. While sadness can be expressed with :(:/ -_- or T.T, humour is added via :D, ;D, xD, ^^, *lol, rofl*, and to some extent also :). It is still arguable whether choosing one of these items is a possibility to grade an emotion or simply a stylistic decision based on personal preferences.

In this chapter we will take a close look at the most influential humour markers :D, ^^, and *lol* (which will be treated as an equivalent to an emoticon in this chapter due to its similar function). The aim is to investigate whether they are used as true synonyms expressing exactly the same emotion or not. When looking at examples from the corpus, one can see that a single humorous situation can lead to responses using different humour markers. It is also possible that the same user employs multiple humour markers as a reaction to one comment. This implies that they have an overlap in functions and can be used to express the same emotion.

```
21.   <@^^^fIzI^^^> fizi is my name
      <monit> monit.
      <hothothot> ur name make me dizzy fizi
      <@^^^fIzI^^^> lol
      <@^^^fIzI^^^> dont become dizzy
      <hothothot> lol
      <@^^^fIzI^^^> :D

22.   <darkzii> Are you man enough to cut your own balls with a
      hammer?
      <Mezmor> lol
      <zwiep`> :D

23.   <flowerboy> c-r-i-m-i-n-a-l
      <Cr|m|naL> Oh goodie! You can spell! :p
      * +Dareena- slaps a large trout around a bit with flowerboy
      <Cr|m|naL> f-l-o-u-r-b-o-y
      <flowerboy> ;)
      <Cr|m|naL>  I can't :\
      [...]
      <+Dareena-> lol flour :D
```

```
24.   <Dreamer`> im going to pick up some weed shortly
      <Dreamer`> and have a nice relaxing evening
      <@vGl-CoW> w33d
      <ShoCkeyy> cow youre back
      <@vGl-CoW> yes i am triggered by certain keywords
      <Dreamer`> weed?
      <@vGl-CoW> hi
      <Dreamer`> lol
      <Dreamer`> ^^
      <ShoCkeyy> lol

25.   <@Sunie> juusty: are you angry? :(
      <juusty> no
      <juusty> why would i be
      <@Sunie> who is then?
      * |BrainLess| sets mode: +v Bubaboba
      <@Sunie> !kick Bubaboba
      <+chris^> you are Sunie :D
      <+chris^> see?^^
```

Lol and :D can both be used to respond to a humorous comment or situation (see *example 22*) while in the last example :D and ^^ are both used by *chris* to add humour markers to his utterances (in *example 04* above the same pattern can be observed). In *example 23* two humour markers are used in conjunction by the same user. The co-occurrence of these items shows that it is possible to use more than one humour marker in the same situation. However it does not imply that they are true synonyms and after investigating these expressions more closely it seems that they do differ in important characteristics. I have taken 50 random samples from occurrences of :D*, lol* and ^^ and investigated them closely.

It appears that one of the major distinguishing factors is that they differ in intensity. To prove this assumption I added an intensity scale and assigned each sample a value from 0.5 to 3. A genuine and full hearted laugh would be represented by 3, while 1 would stand for a mild “hehe”. The minimum of 0.5 was used to indicate only light traces of humour – best translated into real life with a warm smile. It has to be kept in mind that this is a subjective approach that relies on interpretations rather than clear rules. However I tried to avoid taking into account how funny I found a particular situation and the value was rather based on an estimation of the perceived intensity of the user reaction. The aim of this scale was to apply the same subjective criteria to all samples in order to show a difference between the investigated items. See *appendix 3* for examples for each type of intensity from 3 to 0.5.

After adding up all humour points I calculated the average per sample. As a result *lol* turned out to be the most intense with an average of 1.6 humour points per sample. *:D* was second with a value of 1.3 and ^^ was the least intense with 0.9. It is interesting to see that the steps between the items are fairly regular with a gap of 0.3 and 0.4 and that the difference between them is very noticeable.

Figure 2 summarises how the three investigated items are distributed between high and very low intensity. *Lol* succeeds in being the most intense because of the fact that it is the most likely item to represent a laugh of intensity 3. Seven samples of *lol* were marked with this intensity, while only two items of :D and none of ^^ could match this criteria. The occurrences of intensity 2 on the other hand are evenly distributed between *lol* and :D with both having 14 samples classified as medium intensity. While :D had more instances of level 1 intensity, the differences between :D and *lol* do not seem to be very significant in this regard and both can be equally used to express mild amusement and "hehe" types of laughter.

^^ turned out to be the least intense humour marker which can be seen in the large amount of low intensity ratings. Many reactions received a value of 0.5 (e.g. *ex.26*), most others could not exceed an intensity level of 1. However there are also a small number of level 2 instances, indicating an overlap in function with :D and the range of application also includes sarcastic comments such as *ex.27*.

```
26.   <sANDAKER> pfft, thats not have you get your hand on cheap
      electronics
      <sANDAKER> you know some guy that knows some guy
      <Wilitus> that steals from some guy
      <sANDAKER> and you get it 50% without any warranty
      <Wilitus> =)
      <sANDAKER> I wish I knew a guy like that
      <sANDAKER> ^^

27.   <Ben_1> can anyone help me?
      <Big-Mama> sure
      <Big-Mama> suck me first ^^
```

Generally ^^ can be seen as a rough mixture between :) and :D regarding intensity, ranging from a warm smile to a "hehe" kind of laugh. :D seems to be the most flexible and ranges from a broad grin and mild laugh to a much more sophisticated one. *Lol* can

be used in the same way as :D but also features full hearted laughs and seems to be preferred in intense situations.

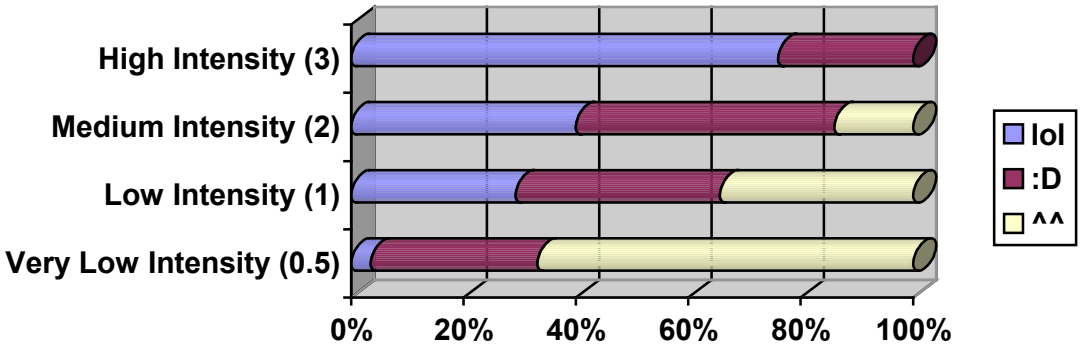


Figure 2: Humour markers and intensity

In addition to an intensity scale I have added a short description of the situation for each sample. The aim was to identify the reasons for each laugh. In the analysis of the various reasons a second difference between *lol* and :D could be identified. :D was more likely to be used as an addition to ones own utterances while *lol* was more often used as a reaction to comments made by other users. While more than three out of five *lol* samples were used as reactions, three out of five :D samples were referring to own utterances. When only looking at laughs made primarily for humorous (rather than social) reasons, it was almost three times as likely that :D was used to modify ones own utterances than *lol*. The following examples illustrate the typical pattern of reaction and modification:

```
28.   <+triviabot> Question 9 (biology): Every human has one of these
      on their tummies.
      <+Julies85> bellybutton
      <@HHC> beer
      <@^^^fIzI^^^> lol

29.   <+_HellMngr> lawyers are not cheap
      <+Zeratul> ya
      <+Zeratul> i know
      <+Zeratul> but my uncle is :D
      <+Zeratul> hehe
      <+Zeratul> he loves me :D
```

Example 28 shows an excerpt of a trivia quiz. The trivia bot asks a question and the answers "bellybutton" and/or "beer" are considered so funny by *fizi* that he reacts with a laugh. In *ex.29* on the other hand *Zeratul* modifies his own utterances with *:D* adding a humour marker to them. One can also see the relation to intensity: While *lol* in this case can be interpreted as a genuine laugh, *:D* on the other hand would most likely be a "hehe" or broad grin. Reactions to jokes and funny comments made by other people are more likely to produce full hearted laughs than additions to ones own utterances. *Lol* could therefore either be better suited for a reaction than an addition to own utterances, thus providing it with more opportunities to represent genuine laughter, or be better suited to represent intense laughter thus being used more often as a reaction.

Even though all three items are used primarily for humorous purposes, they are also employed for other reasons. :D and ^^ are both frequently used as gestures of politeness, for example when paying or receiving a compliment. Messages that show gratitude for receiving help, or suggest a topic change are also likely to contain these emoticons and it turns out that :D and ^^ can be used for various social reasons. Both are also used as an expression of general happiness or joy, for example when a user is coming back, when someone succeeds in doing something or when something cute happens. It seems that ^^ is more likely used for politeness while :D is more likely to express happiness/joy. Likewise in my samples ^^ was rather used to pay compliments while :D was more commonly employed as a reaction receiving them, even though more data would be necessary to prove this claim.

Another relevant category for laughter is embarrassment or awkwardness. This includes for example excuses for making a mistake or being made fun of. In *example 30* one can see that *LilMe* explains that her boyfriend is much older than she is, which is accompanied by multiple instances of :D. The reactions of the other users like the :o emoticon (surprise/shock) and statements like "ohmy" indicate the awkwardness of the situation and it seems very likely that these instances of :D were employed for reasons of awkwardness/embarrassment. All three investigated humour markers seem to play a role in expressing this type of laugh and even *lol* can be used to refer to ones own inability to do something right, as can be seen in e*x.33*: After *Dan45m* has been warned not to repeat his utterances he excuses himself by stating that the repetition was a mistake caused by his mouse, followed by the humour marker *lol*.

```
30.   <@OROZZZZZ> how old is he ?
      <LilMe> 36
      <+^^^fIzI^^^> :O
      <+^^^fIzI^^^> 36]
      <+^^^fIzI^^^> :O
      <LilMe> :D
      <@OROZZZZZ> how old are you
      <LilMe> 18 :D
      <@Urokhtor> :o
      <+juusty> matches
      <@OROZZZZZ> double ur age
      <@Urokhtor> ohmy
      <LilMe> :D

31.   <y0g0ss> im not english
      <+Omni> but you can speak english :)
      <y0g0ss> little
      <y0g0ss> ^^

32.   <Shadow_mil> bad-shadow_, your very bad
      <Shadow_mil> bad-shadow_, don't run IRC as root.
      <bad-shadow_> ok sorry ^^

33.   <Dan45m> [msg(jana11)] my face is ugly
      <Dan45m> [msg(jana11)] my face is ugly
      [...]
      <+sambuca`> pls dont repeat Dan45m thank u
      [...]
      <Dan45m> damn mouse lol
```

As mentioned above all three items investigated were used primarily in humorous situations, even though there is a slight decline from *lol* to ^^, as :D and ^^ increasingly take over non-humorous roles. While *lol* seems to be more common as a reaction to jokes, all three are frequently used for funny comments, statements, word plays and situational comic. :D has also been used as a reaction to popular action commands such as “trout slaps” or *licking* which will be discussed in chapter 7.2.2.

It can be seen from the 150 samples taken from the corpus that there seem to be several important differences in the use of the three humour markers discussed in this chapter. At the same time another important observation was that the use of these markers seems to be highly personalised, because different people appear to employ humour markers differently. This makes generalisations about their use and function very difficult. Therefore in addition to the random samples I have also observed the behaviour of five different people in full detail to see how individuals mark humour over time. Each of

them made more than 200 utterances in a channel and all of them seem to behave differently.

The most striking and extreme example was a user that exclusively made use of *lol* and did not include any other emoticon. In this case *lol* took over all possible rolls of humour, was frequently used to refer to his own utterances, mild cases of humour and other non-humorous roles of laughter. To intensify his laughter this user employed lengthening or capitalisation of *lol* (see chapter 8.1). Altogether he used *lol* ca. 70 times in about 240 utterances and additionally employed 7 intensified versions. Another user seemed to prefer the opposite and included 23 instances of :D, while only one *lol* was used as a reaction to a funny comment. The other users were more balanced in their use of humour markers, even though they displayed the same tendency to use :D more often than *lol*. One person used *lol* only as reaction to other people's utterances and never to modify his own. :D on the other hand was used for both purposes and they were frequently used in combination, e.g. "lol :D".

The person observed last had the most equally distributed use of humour markers with 20 *lol*, 15 xD and 9 :D. The user further had the interesting quality of never adding a humour marker to the end of his utterances. Instead he always sent them separately as a message only containing the humour marker. The instances of *lol* employed by this user were equally divided between reactions to other users and modifications of own utterances, while xD and :D were exclusively used as a modification. Because three different humour markers were used to accompany own utterances, one might expect to find a reason behind the individual choices and a pattern in the use of the three items – probably a different grade of intensity or another purpose of laughter. However after looking at the examples it was very difficult to observe any reasonable pattern in the use of this user's humour markers. It seems as if the choice was more random than planned and all three were used in varying intensities.

Note that none of the observed persons made use of the ^^ emoticon, which illustrates that ^^ is considerably less common and not as widely spread as the other two investigated items. It is an alternative many people do not seem to consider. When not employed by a user at all, the functions usually attributed to ^^ are taken over by the

other emoticons. Users that include this smiley into their vocabulary simply have additional possibilities to express humour that other people miss.

Even though the function of *lol* is often similar to the other two emoticons it is sometimes used as a verbal expression within utterances. Despite its major function as a humour marker one should not forget that *lol* is still an abbreviation and as such has become a kind of neologism. I have heard people using *lol* as a normal word in real-life face-to-face conversations and this use of *lol* as a word - rather than an indicator of laughter - enables it to be used in different situations. Therefore five samples were excluded form the intensity scale because they were not used as a humour marker, such as:

34. <Neur0ticism> actually saw a man in a car once..strung out on drugs LOL that was bad.

As one can see from this example *lol* can also be integrated into an utterance as a discourse marker or interjection. Despite the obvious difference that *lol* can also be read as a word rather than an indicator of humour, it is still in most cases used as a humour marker and often seems to be interchangeable with :*D* both in function and intensity.

To summarise this study of humour markers one can say that they are clearly not true synonyms, but could on the other hand be considered near synonyms as they overlap in many aspects and are often mutually interchangeable. Because every person can be expected to use humour markers differently it is almost impossible to assign clear meaningful roles to them. However it is possible to say that on a general scale each humour marker has its strengths and weaknesses and while *lol* seems to be preferred in intense situations, :D is more likely to add a hint of humour to one's own utterances. Nevertheless *lol* can easily be less intense than :D and is also frequently used as a modification of a speaker's own utterances, or as an expression of awkwardness making the borders more than blurry. In the end it is up to the individual user which humour marker they employ and the choice is not only influenced by the situation but also very much by personal preferences and style. This can vary considerably, ranging from rejecting emoticons altogether to adding humour markers exclusively via extra utterances.

7. Performing actions

"Speech is accompanied by an intricate set of vocal and gestural non-verbal signals, which affect meaning, emphasis and other aspects of utterances" (Argyle 1972: 243). In some situations non-verbal communication "is used to communicate attitudes and emotions and to manage the immediate social situation" (Argyle 1972: 268) while it "is also used to support and complement verbal communication." (ibid) Therefore non-verbal cues like yawning, "head-nods, shifts of gaze, fine hand-movements [or] bodily posture" (Argyle 1972: 243) can play an important role in a conversation and the inability to perform such actions in internet communication called for new devices to integrate these elemental features of spoken discourse into the written medium. In order to simulate actions typical of face-to-face conversations actionmarkers are introduced, which can be divided into actionstrips on the one hand and actionlines on the other (also referred to as the action command in IRC). They are often compared to stage directions in play scripts and while the original scripts help actors to play their role, their CMC equivalent enables users to describe what they are doing or rather what their "virtual being is doing or wishes to do, had it been given a physical body" (Lui 1999: Action-Simulating Messages). Actionmarkers add descriptive and narrative moments to the communicative spectrum (Storrer 2000: 13-14) and Haralampieva (2004: 30) sees similarities in their way of narration to role playing.

7.1 Actionstrips

Actionstrips brace verbs or verbal stems with asterisks as in the following examples:

```
35.   <Samanth0r> *sigh* nm, this sort of thing has been debated
      endlessly.....

36.   <evant_> yeah, sydney is really boring
      <pluesch0r> evant_: i liked sydney
      <evant_> *shrug* i guess im just used to it xD
```

```
37.   <+PrincesS^Rinoa^> red wine or white wine ?
      <Daemona> white of course
      <Daemona> so i don't get headache *chuckle*
```

In the literature this form of verbally describing nonverbal actions is often compared to comic strip language, where similar conventions exist to express a sigh or a moan. (see Schulze 1999: 75). Geers (1999: 92) defines actionstrips as short transcriptions of paralinguistic behaviour, which are sometimes used instead of emoticons because they allow expressing facial expressions and gestures more elaborately than smileys. They enable to express emotions, feelings and thoughts with more options for grading and detail and without the restriction to only a few basic forms. Furthermore actionstrips do not allow for interpretation like emoticons do, because clear verbal descriptions are used instead of vague symbols (Haase et al 1997: 65). This can be seen in *example 37* where *Daemona* used the actionstrip *chuckle* even though there would have been various humour markers and laughter smileys to choose from. In this example it was important for the user to describe and grade a certain type of laugh in full detail and instead of relying on a symbol that can be interpreted as various forms of humour or laughter *Daemona* wanted to convey that her laugh was a *chuckle.* There are only a handful of relevant emoticons in IRC which display only a very limited variety of paralinguistic cues and therefore actionstrips are a much appreciated and a relevant tool to include certain actions and emotions, like sighs, shrugs or yawns. However because IRC features a more common way of describing actions with the *action command* they are probably less common than in many other places of the internet.

Their exact number is hard to estimate because they share with *intensification* the convention of bracing words with asterisks (see chapter 8.1). There are ca. 850 instances of words or groups of words braced by asterisks in the corpus, but it has to be taken into account that about one quarter up to half of the occurrences could be *intensifications* rather than instances of actionstrips in the corpus.

The most common one is definitely *hug* with all its variations including *hugz* or *hugglesnuggles,* altogether featuring 67 occurrences. Hugging has a rich tradition in IRC as a greeting and is also one of the most common *action commands,* as we will later see. As table 5 shows it is followed in popularity by *nod* and *shrug.* Other common

actionstrips are waving in greeting or goodbye*, sigh, sniff, sing, cough,* and *yawn,* but generally very few actionstrips have more than 10 occurrences in the corpus.

hug	67	wave	9
shrug	17	sing	8
nod	17	cough	6
sigh	12	yawn	5
sniff	9	wink	4

Table 5: popular actionstrips

It is interesting to observe that some actionstrip verbs are almost exclusively used in the third person form: out of the 67 *hug* instances, 63 were in third form, for example **hugs** or **hugz*,* while only four were formulated in the first form. Most other actionstrips however were used more frequently in the first form and it was much more common to say *shrug* than *shrugs* (every fifth instance of *shrug* contained a third person s). It was three times as common to use *sigh, sing* and *sniff* in the first person form, while *yawn* was never used in third person at all. Out of the 9 most popular actionstrips investigated in this study only two were primarily used in third person which were *hugs* and *waves. Nod* also features considerably more instances of third person forms but variations like *nodnod* or *nodnodnod* even out the scale and make *nod* the only actionstrip where third and first person instances are evenly spread. These findings correspond to Kramer (2008: 51) who argues that English actionstrips allow the use of both third person and infinitives, while German for instance relies exclusively on infinitives.

Actionstrips are not limited to one word utterances, even though they are most commonly used this way. They can also include multiple words such as in the following examples:

```
38.   <+Ivory> I dont think "EvilShoe" is bot *even more confused*

39.   <McHearty> may cpl rip *moment of silence*
```

```
40.   <BonesolTeraDyne> dthacker: Don't worry, it wasn't me who
      removed your clothes while you slept. Uhh... *looks for someone
      to blame*
      <BonesolTeraDyne> It was Daisuke!
```

As one can see longer clusters can be used for various purposes as for example simulating a moment of silence or looking for something specific *(ex.40)*. The two-word clusters in English typically feature verb and object such as *rubs temples*, *rolls eyes*, *shakes/scratches head*. According to my corpus they are extremely rare in IRC and seem to be more common in German as most of the two-word instances came from the single German channel observed and include phrases such as "verliebt gugg", "hart bleib", "neidisch bin", "ohren reib", "zurück geh", "markus knuddel" or "tränen wegwisch". It appears that German use of two-word clusters is not limited to verb/object combinations and allows more variety. Beisswenger (2000: 106) claimed that German actionstrips are exclusively one word utterances and any spaces in between words are left out to form 'single-word-clusters' for example in *ganzgenauwissenwill* (**wanttoknowmoreaboutit**). This claim does not hold true in the IRC corpus of this study where none of these word-chain utterances could be found. According to Dittmann (2001: 75) and Beisswenger (2000: 108) this unusual use of verb clusters is a German CMC language characteristic. It was presumably introduced by the translator of Donald Duck comics and entered CMC later on (ibid). Utterances that are longer than two words appear ca. 100 times in the corpus and are limited in popularity by the much more common alternative of using the *action commands* which seems to satisfy the needs of expressing longer phrases much better.

7.2 Action Commands

Action commands allow the creation of whole utterances in the third person form: The command */me text* will produce * *username text*, for example if *Dr_Future* types */me yawns* it will produce the output of *example 41*. To distinguish these utterances from regular chat messages IRC clients like *mIRC* display action command lines in a purple

colour (see the screenshot in *appendix 1*). This way IRC users have the possibility to 'do something' by building a sentence with 'me…':

```
41.   * Dr_Future yawns
      * xteddy hugs sleepy Dr_Future

42. (line 35) * cal dies

43.   <War2> I am AGAIN MISUNDERSTOOD in this WORLD.
      * War2 sobs
      * PacMan85 hugz War2
      * Dave2 has NO SYMPATHY for the man who keeps ABUSING him :'(
      * xteddy offers a hug
      <PacMan85> i understand you
      <War2> :)
      * Bspec looks at xteddy then proceeds to look at starshine
```

Actionlines are very popular in IRC and over 8000 occurrences appear in the corpus. According to the present corpus (ca. 200.000 utterances) this would mean that 4% of all utterances are made in third person form. They are used to represent a broad variety of different types of actions. One can see in the examples above that they are used to substitute nonverbal cues like yawning or sobbing much like the actionstrips. Kramer (2008: 54) suggests that both types share the same features and characteristics and that they are mainly distinguished by their different technical realisation: while the actionstrip uses symbols to indicate an action, actionlines rely on a software dependant command available in IRC. Nevertheless the action command undoubtedly seems to be a much more popular way to realise actions in IRC which corresponds to the findings of Geers (1999: 90) and even in the area where actionstrips shine most – in descriptions of short one word actions – they are outperformed regarding popularity by the action command as will be discussed later on.

One major advantage of the action command is that it provides more elaborate possibilities to express oneself than the actionstrips, since it is not limited to only short verbal phrases, but may consist of whole utterances. It allows objects to be included into the action statement as can be seen in *example 41* and while simple subject-verb-object phrases are one of the most common uses of the action command, verb-object combinations are almost nonexistent in English IRC actionstrips as discussed above. Even for simple actions like hugging the action command offers more possibilities for

grading and detail than its simpler cousin and it is not only possible to hug other people but also to "offer a hug" as can be seen in *example 43*.

Enabling interactions with the physical or virtual environment as well as the conversation partners is a very important function of the action command and apart from hugging other channel members many other interactive actions are performed regularly. *Example 43* shows how users can 'transcribe' how they look at 'virtual objects' as *Bspec* looks at one user and then proceeds to look at another. By including these more or less elaborate transcription-like utterances, IRC users try to compensate the lack of face-to-face contact and enrich their speech situation with important non-verbal elements that would otherwise be missing in a purely written medium. According to Reid (1991: Part Two)

> users of the IRC system feel it important to create a physical context within which their peers can interpret their behaviour.

In other words chat users establish a *fictitious stage* (Beisswenger: 2001: 101) which allows them to virtually interact and act. Beisswenger (2001: 105) argues that the chat room as a *fictitious stage* has similar qualities to a real room, even though the position of objects is very open and vague. While it is possible to look at certain users from a specific angle (e.g. user X looks at user Y from the side) it is usually not possible to determine where someone is positioned in a room and the only reference points remain other virtual objects or persons.

Despite the fact that there is no possibility to see what the communication partners are doing or physically interact with them, there is still an option to let the others know 'what you are doing', or what you would be doing in a face-to-face situation, such as *Bspec* when virtually 'looking' at other channel members. It should be added that although many interactive actions might imitate behaviour that could also be found in real-life speech situations, others might be more surreal or imaginative due to the purely virtual environment. The fact that the users are in a virtual environment provides them with much more liberty in choosing and performing their actions and they can climb skyscrapers as easily as they can clap their hands. This creativity that actionlines can provide contributes greatly to the playfulness that is attributed to IRC conversations.

Gelleri (1998: 18-23) mentions that actionlines can have several different functions. While some actions for example refer to events that happen in the users' real life environment others clearly do not. Real-life events might provide additional information what users are doing (e.g. * Merlyn was on the phone; Gelleri 1998: 21) or they serve as a way to compensate the lack of physical and personal presence. Other actions express gestures or emotions and might provide "indirect references to the participant's state of mind" (Gelleri 1998: 22). Yawns and sighs for example are usually not interpreted as real life actions but as means of expressing boredom or dissatisfaction. Apart from the important function of bridging the lack of physical cues action commands are according to Gelleri often used simply for amusement. They furthermore help to

> break the first-person monotony of IRC, playfully transforming the interaction into a kind of script or play. In this manner, descriptions may be utilised to create the illusion of a physical world where non-verbal cues are also available" (Gelleri 1998: 21).

In 7.2.2 a possibility to group actionlines into different categories regarding their function will be discussed.

7.2.1 Sharing Thoughts

When looking at examples from the corpus it becomes obvious that the action command is also frequently used to paraphrase statements into third person form that do not imitate physical actions of any kind: It is possible to share thoughts and feelings such as in *example 44*. This process can also be seen in *example 43* above where *Dave2* states that he "has no sympathy for…" via action command. Likewise additional information might also be provided in a similar way via notes, such as in *example 45*. This way users can "think out loud [and] thoughts and feelings can be expressed in a way impossible in any other communication" (Kramer, 2008: 54) and a comparison can be drawn to thought bubbles found in cartoons.

```
44.   * unixSnob thinks it's a bit ethnocentric to assume the will to
      learn english is to pursue work the US
```

```
45.   <fred> war tends to drastically accelerate technology
      development
      * fred notes he isn't saying he thinks war is overall a good
      thing
```

The fact that IRC users can always choose whether they want to say things directly or paraphrase them into third person form is very interesting and further research could be done to find out what factors might influence their choice. Sometimes the decision seems arbitrary, in *example 43 Dave2* could easily have said *"I have no sympathy for..."* instead of using the action command, while *PacMan85* could have said *"* PacMan85 understands you"* just as well. However, the fact that the command is used for sharing feelings and thoughts at all, even though it consumes arguably more time to create it, indicates that it does have a stylistic or even communicative function.

Gelleri (1998: 20) suggests that by using this type of third-person utterances users aim to

> express background or extraneous information without creating a disrupting effect upon the actual conversation. Whereas it might be impolite to ignore a first-person utterance, in the case of action descriptions communicants can decide whether or not they want to respond.

It is difficult to validate this assumption because first person utterances may be ignored just as easily while many third person utterances in the corpus triggered a reaction, like the following example shows: *flipflops* indirect comment about beer receives both agreement and disagreement (from *tremere* and *franchise*) and is discussed further, while *flipflops* muttering also influences the discussion.

```
46.   <FRANCHISE> i'd rather drink Keystone then Bud or Bud Light
      <anna_s> Sup Team
      * flipflop thinks beer is beer is beer is bear. Especially
      after a few.
      <pornflakes> hey DiCeR whats up :)
      * flipflop whistles.
      <pornflakes> hey flipflop :)
      <@DiCeR> Working..
```

```
<`tremere> flipflop.. true.. once youre hammered. it dont
matter
<@DiCeR> or atleast pretending to work..
<FRANCHISE> i dont agree...
* flipflop mutters - no one noticed the last one!
<NailBunny> Beer is...bear?
<pornflakes> id hate to work on a pc
<flipflop> Yay, NailBunny!
```

It is unlikely that *flipflop* formulated his utterance in third person because he did not want to influence the conversation, more likely he expected a reaction similar to a first person utterance, which can be seen in his dissatisfaction by the fact that participants failed to notice the "bear is bear" part of the utterance. It could still be discussed whether the formulation in third person aimed to make the statement more neutral or not. *Flipflop's* whistling and muttering might indicate a wish to display a certain level of neutrality and secrecy rather than boasting his argument out loud, even though it was most likely done simply for playful reasons rather than a sincere wish to stay neutral. Because of the fact that third person utterances stand out as something 'special' in IRC due to their break in monotony and different colour scheme it would also be a valid argument that they attract more attention than regular first person utterances. Therefore a deliberate break in monotony sometimes even seems to be one of the main reasons for paraphrasing:

```
47.  <yaloki> tacit: true. but personally, I'd miss the contact with
     coworkers IRL
     [...]
     * Chrysantine does not miss people.
     <Beineri> Chrysantine: not even Germans? :-)
     <Chrysantine> Not even ze germans :p
```

Example 48 shows another use of the third person form and it seems that the choice in this case is made for a different reason: Instead of paraphrasing a personal opinion into third person form, the user wants to share what is going on in his head. After seeing *MrJump* entering, *Bless* has to think of a song and starts singing it. However she also has to let the other people know what she is doing and the simple utterance of "jump, go

ahead and jump…" might not have sufficed to convey that *Bless* is actually singing a song. Therefore she decided to use the option of sharing her thoughts with other people, making clear that she is reminded of this song.

```
48.   <@MrJump> yo
      <PatricQ> ello MrJump
      <@Jj`> Jump Jump
      <+Bless> jump
      * +Bless thinks of an old song
      <+Bless> JUmp  go ahead and Jump.....
```

7.2.2 Collocations for Actionlines

It has been discussed that the action command is very popular with over 8000 uses in the corpus. However it is interesting to investigate for what purposes it is used most frequently. Therefore I have investigated which verbs are typical collocations for the command and even though it collocates with many different verbs, there are some that are notably common. The most striking is probably the verb *slap* which turned out to be the most common collocation in the whole corpus. It occurs 546 times which is a large number compared to the total amount of action commands and illustrates that IRC users like to playfully slap their colleagues from time to time. However most of these occurrences turned out to be instances of a very popular phrase which is:

* *user x* slaps *user y* around a bit with a large trout

This line is often produced by the shortcut */slap user y* which is the so called "slapping feature" that is included in IRC and will create the entire phrase above. The use of this phrase is widespread and not restricted to individual communities, channels or networks, it can be found evenly distributed throughout IRC. Wikipedia states that its origin might be a reference to *The Fish-Slapping Dance* which is a popular sketch in *Monty Pythons Flying Circus*. In this sketch Michael Palin performs a dance that includes slapping John Cleese in the face with a small fish four times. After the dance music stops Cleese hits back with a very large fish, so hard that Palin falls into the nearby water. Even though according to the urban dictionary (http://www.urbandictionary.com/) the heritage of the slapping feature is frequently connected with the Monty Pythons, it adds that this connection remains a vague suggestion. The dictionary further defines a "fish slap" as a practice of slapping someone with a fish (usually in the face) that is rather uncommon in real-life even though other examples than the Monty Python sketch exist.

Illustration 1: Monty Pythons fish slapping dance, possibly the reason for the slapping feature (source: Wikipedia)

Many instances of this phrase might be created via the shortcut of the *slap command*, but the phrase has become so popular that it is often used with playful variations:

```
49.   * Flatterdings slaps dSi`Qwnage around a bit with a large
      fishbot
      * dSi`Qwnage slaps Flatterdings back with an unbelievably large
      trout
```

444 of the 546 occurrences of slap can be attributed to the fish slap or one of its many variations while the remaining 102 are regular appearances of slapping. However popular the tradition of slapping might be it is important to remember that it is a playful and humorous way of interaction and never used in an aggressive or humiliating way. Theoretically the collocation slap can also be used for other things than slapping other people for example in *ex.50* even though this can be expected to be very rare.

```
50.   * DyslexicGhost slaps a knee laughing
```

The second most frequently used collocation is *hug* which appears 344 times as an action command, including all its variations such as *hugz* or *huggels.* Because *hug* is also the most popular actionstrip one can see that the act of hugging truly is a common language feature in IRC. Table 6 shows the most popular collocations for action commands and their occurrences according to *Word Smith*. Compared to table 5 above one can immediately see that typical collocations for actionlines are vastly more common than the ones for actionstrips.

slaps	546	gives	82	goes	61
hugs	334	pokes	78	nods	59
looks	113	hands	75	hides	59
licks	89	wonders	63	throws	57
waves	83	gets	63	yawns	51

Table 6: the 15 most popular collocations for the action command

The fact that *look* is in third place with over 110 occurrences illustrates the importance of visual contact. IRC users often establish 'eye contact' by 'looking' at other people. Furthermore they can let their peers know what they can 'see' or what they are looking at *(ex.51*). This can be used for narrative effect as in the following scene:

```
51.   * @`sambuca looks at the time and taps her nails
      [...]
      * @`sambuca looks at juniperlee and Black_Dahlia
      <+sraet> whom you waiting for `sambuca
      <juniperlee> `sambuca...he is a dead man
      <@`sambuca> yes he is!!!!!!!!!! cause once he's back im sooo
      ignoring him!
```

Apart from the items already mentioned the verbs *lick* and *poke* stand out as being actions rarely performed in face-to-face conversation but being popular in IRC. The act

of playfully *licking* and *poking* other people seems yet to be another tradition that can be observed in IRC and *example 52* shows an excerpt where both appear:

```
52.   * Vexar wiggles out
      <scouty> wake up time o.o
      * KitKateGrr pokes Marc-O
      * pornflakes rocks up nerd n all :)
      * scouty licks KitKateGrr all over her forehead.
      <Marc-O> hehya :)
      <KitKateGrr> eek
      <scouty> salty..
      <scouty> g'morning ppls
```

As one can see they were both used as part of a playful greeting ritual, which illustrates nicely what many of the most popular collocations of the action command are used for: It seems that through interactions such as hugging, slapping, licking and poking IRC users try to overcome the *communicative distance* and emulate the physical, social and emotional closeness that is so important for conceptionally spoken discourse. All four interactions include touching other people, which leads to a metaphorical physical closeness between the participants. This friendly and playful way of interaction aims to build an emotional link. All of them are frequently used as greetings and while hugging is the most obvious interaction for greetings and goodbyes, even pokes and slaps can be used to establish contact, as can be seen in the example above where the poke from *KitKateGrr* is interpreted as a form of greeting and triggered the reply “hehya :)” by *Marc-O*. Therefore it is possible to divide common collocations into three different groups: utterances that are made primarily for their ability to establish *communicative closeness* and their intentional playfulness, utterances that largely serve the role of transcribing missing non-verbal cues or physical actions and the ones that enable to share thoughts and feelings:

Playfulness/closeness	Physical actions / non-verbal cues		Sharing thoughts and opinions
slap	look	nod	wonder
hug	give	yawn	think
lick	hand	put	
hide	get	sit	
poke	go	sighs	
wave	throw	takes	

Table 7: The 20 most popular connotations for the action command grouped according to function

Note that *hide* has been grouped somewhere between a playfulness and physical action cue because it is rather difficult to categorise due to the fact that it is most likely used in a playful way even though to a lesser extend than other words grouped in this category. A typical use for this collocation would be that someone tries to metaphorically 'hide' after he has said something that might annoy another user. This way he can either indicate humour and/or preserve harmony. *Wave* has been classified as a closeness cue because similar to hugging it is used as a form of greeting, establishing contact or saying goodbye. *Nod* on the other hand is an important feature to signal approval and is an ideal example for a direct substitute of a paralinguistic cue. Nods have a clear communicative function and represent an important non-verbal feedback tool in face-to-face conversation. They typically paraphrase "Yes, I agree" or "Yes this is true" in IRC as in the following examples:

```
53.   <Latios_Ex> why are asians small?
      [...]
      <AsianaGal> Good thing comes in small package
      * +Iseult nods to that

54.   <Ladylya> gopsy..arent u a girl?
      * gopsy nods

55.   <`tremere> 3 days off from work, wife and kids outta town for a
      week. didnt know i would be so bored. Lol
      <Ryannon> masturbate, when one properly, it can be very time
      consuming
      <`tremere> im sure i can wait til my wife gets home lol
      * Neur0ticism nods darkly
```

```
56.   <+Nisha> oh also, the word for pineapple in Finnish is similar
      to spanish
      <+Nisha> ananas
      <Black_Dahlia> cool
      * +Nisha nods
```

Words associated with *communicative closeness* can work together as means of establishing contact like in *example 57*: *Abrienda* gives another user an extensive hug as a greeting, followed by a question aimed to establish contact. After a short pause without reaction she decided to poke him, playfully expressing her annoyance of not being answered and further indicating her wish to communicate. After all means of getting attention have failed she sighs as *helenaxis* does not seem to be able to participate in this conversation. Even though pokes might be used as a means of establishing contact they are unlike waves or hugs not primarily used this way and often work very similar to slaps. Both might display mild annoyance or are used as a reaction towards certain behaviour of other participants such as in *example 58* where *anniebabeblues* provokes a poke because of her sarcastic comments on *Insideth's* complaints about having a long day. Note that both poke comments (*ex. 57-58*) include creative information with what objects and where the other person is poked, which further stresses the playfulness of these actionlines.

```
57.   [18:54] * @Abrienda hugglesnugglekissandwiggles helenaxis with
      dazedman's tree
      [18:55] <@Abrienda> how's it going helenaxis?
      [18:56] * @Abrienda pokes helenaxis with her new rhino's horn
      [18:58] * @Abrienda sighs


58.   <Insideth> Nothing to do, nowhere to go
      <Insideth> Man, it's gonna be a long day
      <+anniebabeblues> booo hooo
      <Insideth> I know eh?
      * +anniebabeblues gets the violin out
      [...]
      * Insideth pokes anniebabeblues in the knee with a broom handle
```

It cannot be stressed enough that this approach has only looked at the most popular collocations for the action command so far. As mentioned before actionlines make use of a broad variety of verbs and altogether over 150 verbs have been identified as collocations in the corpus. Taking all verbs into account it becomes obvious that the majority of action commands are used to substitute non-verbal cues and/or physical actions. Even though the verbs associated with *communicative closeness* are among the most popular, they are very few in number. The same is true for verbs associated with sharing thoughts: *think, wonder, feel and note*[4] seem to be one of the few verbs exclusively used for this function. However sentences with *be* and *have* can also be used to express thoughts and feelings such as "x has no sympathy for..." or "x is concerned that...".

Comparing collocations for actionstrips and actionlines it can be observed that the verbs typically associated with actionstrips are used as actionlines as well. Almost every verb is more commonly used as an actionline, even though to a different degree: while *yawn* and *wave* are considerably more common actionline commands, the use of *sigh* is more balanced and *shrug* is almost evenly spread. *Cough* and *sniff* on the other hand strike out as having more occurrences as an actionstrip than -line. Even though actionlines are generally much more popular it appears that some words collocate better with actionstrips. Words like *cough* and *sniff* seem to have a special characteristic that distinguishes them from other action verbs and makes them more attractive to be used as a strip. Possible reasons could be due to the immediate nature and disruptive effect of these actions.

With more than 150 different verbs being used for actionlines one can see that the action command is a very diverse tool that can express a multitude of non-verbal actions. It seems that most actionlines are used for transcribing physical actions such as *give, go, throw, take, sit* etc. However many of the most frequently used verbs for actionlines seem to concentrate on emulating physical- and *communicative closeness* and aim to establish a positive, light-hearted atmosphere, such as the widespread greeting rituals or the playful acts of poking and slapping. A final group of collocations that attracts special attention are the verbs used for sharing thoughts and opinions, and it

[4] Note that feel and note are not included in table 7 because they are not among the most popular collocations

is possible to divide verbs that collocate with the action command into these three major groups.

7.3 The use of Actionmarkers

Gelleri (1998: 19) reports that none of the participants of his *newbie* experiment made use of actionmarkers, even though they had previously all been familiarised with it. He suggests that a possible reason for the lack of actionmarkers would be that the utterances of the test subjects were based solely on experience gained from face-to-face interactions. Since real-life interactions do not feature the use of third-person action descriptions it is difficult for inexperienced users to integrate this new language feature they have never encountered before. It seems that emoticons, despite being completely new symbols, are more easily accessible than unfamiliar patterns of verbal expressions.

It has been mentioned in Haase et al (1999: 77) that action descriptors only seem to appear in some parts of the IRC network while others rarely use it. To investigate this claim I have used the plot feature of *Word Smith* to see whether the use of action descriptors is limited to a small number of channels. The 15 most common collocations for the action command have been investigated and results show that their range of use is not considerably different from emoticons of a similar popularity. Verbs with 60-90 occurrences could be found in 22-35 different channels, while the emoticon :3 with 87 occurrences appears in 27 channels. Some emoticons on the other hand seem to be much more limited than individual actionlines, the emoticon T_T (sadness) for example has 43 occurrences but only appears in 12 channels. *Slap* (546) as the most common collocation can be found in 86 channels. As a comparison the emoticon :o (577) shows a very similar range with 91 channels. Altogether action commands appear to be used in almost every channel, beating even the emoticon :) (165 channels) in its range of use and the claim that the performance of actions is limited to certain groups can not be supported by this study.

Action commands and strips allow IRC users to play with language in a very creative way and enable many different ways of expressing oneself. *Example 59* shows two interactive actions by *Bspec*, but more interesting are the reactions by *Cann0n*. Note that

he first chooses to use the action command to sniff and even makes use of indirect speech within the action to finish his utterance. The second time however he chooses to integrate *sniff* as an actionstrip into a direct verbal utterance and creates the effect of continuous sniffing. *CannOn* would have had many other options to express himself, for example he could easily have avoided the indirect speech by simply starting a new utterance after the actionline, or he could have used an actionstrip instead (*sniff* I love the smell of gun powder). It does not seem that there are any definite rules when to use which command or when to say something directly or indirectly and IRC encourages creative and playful use of language. Actionmarkers provide the users with a new set of communicative possibilities, help to make language figuratively speaking more colourful and break monotony.

```
59.   * Bspec smothers Cann0n in gun powder hugs
      [...]
      * Cann0n sniffs and says, "I love the smell of gun powder"
      [...]
      * Bspec gives Cann0n some Kleenex
      [...]
      <Cann0n> thanks, *sniff* i uh.. i got *sniff* gun powder got in
      my eye..
```

7.4 Actionmarkers and Emoticons

Geers (1999: 90) argues that even though emoticons are more common than actionmarkers, the use of actionmarkers is of special significance, since their creation takes considerably more time than adding a simple emoticon. As already discussed time saving is a very important factor in internet communication and the fact that the time consuming command is still used excessively underlines that it serves an important role in IRC discourse. According to Geers' study 23.9% of the IRC utterances of his corpus contained emoticons, while 12.3% consisted of action commands.

Regarding the relation between actionmarkers and emoticons it seems that whenever an utterance needs to convey crude basic emotions such as humour, shock or irony, smileys

are an adequate and efficient way to do so. As soon as a user wants to create more elaborate descriptions or more detail is required to diffuse a situation, IRC users have the possibility to spend some extra time and switch to *actionstrips* or the even more powerful *action command* to satisfy their communicative needs.

As opposed to the fast and simple emoticons, verbal constructions allow a very accurate emulation of graded paralinguistic cues by "falling back on literary expressions" (Crystal 2006: 38).

Haralampieva (2004: 37) notes that

> smileys appear to stand as a short form of […] verbalised descriptions of physical actions. This substitution was reinforced by the economic principle of language, or the principle of least effort.

In other words that the symbol ;) simply stands for the verbalised *wink* which is the shortened form of "I am winking with my eye". Because internet chat discourse is time-bound and typing takes more effort than speaking, the "principle of least effort" dictates that every keystroke is worth saving. The use of emoticons can thus be seen as a simple convention to save keystrokes. Common "verbalised descriptions" like *wink* are attributed to partly iconic symbols like ;) or acronyms like *lol*. This is why 'basic smileys' like 'laughter' or 'being annoyed' differ so much from emoticons like "Bill Clinton enjoying some French fries" (Crystal 2006: 40). The latter category simply does not have the same function of representing the most common transcriptions of basic actions or emotions (otherwise expressed by actionmarkers) in order to save time and still modify an utterance with paralinguistic information. It should be added that other factors apart from saving time can influence the choice of using an emoticon (see chapter 6.1) such as aesthetic reasons: symbols are a more elegant way to convey popular emotions than adding cumbersome verbal expressions to every comment that contains humorous intentions.

However the introduction of too many standard-emoticon symbols would inevitably make internet communication more complicated and there is a limit to how many smileys remain useful before the effort to decode and memorise them counters their advantages. It seems that not more than 15 of these iconic symbols are used regularly at this time in IRC and some even appear to cover the same basic feelings such as humour

or sadness. Any other feelings and emotions like boredom (*yawn*) or popular actions like hugging or greetings completely lack symbolised icons and rely solely on verbalised actions. This ensures that actionmarkers are and remain one of the two most important ways to substitute paralinguistic cues in IRC.

8. Other paralinguistic features

8.1 Loudness and intensity of speech, intonation

Intonation, loudness and stress are very important paralinguistic features and their absence limits communicative possibilities. Therefore the very common convention has evolved in Netspeak that capitalised letters stand for 'shouting':

```
60.   <Simeon_H> I said DON'T COPY THAT FLOPPY

61.   <Jaomi> Who's conan? Isent that the American gouvenour?
      <Windcape> yes
      <Jaomi> Forgot his name -.-
      <Windcape> ARNOLD!

62.   <Murkeli> i might just come and KILL YOU WITH MY BEAR HANDS
      <Murkeli> yes, i said BEAR
```

Capitalisation can often be received as yelling and an expression of negative feelings, anger or outrage, even though there are also positive examples such as "HELLO!" (Kramer 2008:47). Prolonged and extensive use of shouting will likely be interpreted as hostile behaviour and might trigger negative reactions and demands by other users to change back to normal letters. However when used strategically as in the examples above it is a very effective way of emulating emphasis, loudness and intensity.

Another very common convention is the imitation of a drawn-out or expressive intonation by reduplicating letters (*ex.63-71*). It can also be combined with capitalisation to further increase the effect (*ex.65-66*). Hentschel (1998: Section 3.2.4) distinguishes between two types of reduplications: those that mirror real life speech behaviour and those that do not. While it would be natural in real life to lengthen vowels such as the *o* in *noooooooooooooo* or consonants the *r* in *rrrrrrrrright* it would be unlikely to reduplicate *c* and *h* in as in *chumi* from *example 67*, the *g* in *argggggggg* or any other stops such as *p* in *example 68*. According to Hentschel (ibid) reduplication works different in IRC to real-life because it "does not take into account the phonetical and practical aspects of spoken language, but rather makes exclusive use of the graphemes." Onomatopoeic expressions such as snoring or roaring can also be found in IRC discourse such as in *examples 69-71*. A comparison can be drawn to comics were these expressions are used in a similar way.

63. <mestizah> heelllloooo

64. <Athan> it'll be loooooooooooooooooooooong since greyed out

65. <Cheez> "yay im passed the shit doldrum level... NOOOOOOOOOOOOOOOOOOOOooooooooooooooooooooooo"

66. <Ziggy_Sawdust> NOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOO OOOOOOOOOOOOOOOOOOOOOOOOOOO

<tetsuOS> YESSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSs

<Maddy_> MAAAAAAAAAAAAAYYYYYYYYBEEEEEEEEEEEEE

<Ziggy_Sawdust>

SOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOO

OOOOOOOOOOOOOOOOOOOOOOOO

67. <ghieauleztha> ccccccccccccccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhuuuuuuuuuuuuuuuuu uuummmmmmmmmmmmmmmmmmiiiiiiiiiiiiiii

68. * HiddenMickey shakes pepperspray: helppppppppppppppppppppppppp!

69. <+PrincesS^music^> Rawwwwwwwwwrrr

70. <TechnoPotato> ZzZzZzZzZzZzZz...

71. <Thiefcraft> "halp halp we r under at---*BRRRZZZZZPPPPPPPT* *sound of falling body parts*"

Furthermore reduplicating may be used on question- or exclamation marks to indicate increased puzzlement (*ex.72-73*) or intensify questions and exclamations (73-75). Čmejrkova (1999: 122) observed that this kind of emphasis may depict "the author's emotions, mood and state of mind".

72. <ubuntu_> ??????????????

73. <STIGY> what?????

74. <masaca> who will help me???????????????????

75. <dustybin> it works!!!!!!!!!!!!!!!!!!!!!!!! thanks for your help guys! :D

Another important paralinguistic feature substituted in IRC conversation is adding emphasis to single words. As a convention asterisks are commonly used to indicate stress such as in *examples 76-79*:

76. <tulcod> Lord_Deathscythe: if you expect not to get any more ram, then 32-bit *can* be a *little* bit faster

77. <NoorulIslaam> mynetdude thats true. there was only 1 bill gates, and his time is gone :P

<NoorulIslaam> *however*

<NoorulIslaam> there are places around the world

<NoorulIslaam> emerging markets

78. <@chookie> wish i could, wish i could play *any* instrument

79. <trivvy913> Come on! Play our Trivia! WE *DARE* YOU!

Instead of asterisks underbars (e.g. _can_) may also be used for emphasis, although it is much less common. As discussed above capitalisation could also add emphasis and a user has a choice between 'shouting' or emphasising single words which provides a possibility to grade the intensity of intonation. *Example 80* shows that *ErrantEgo* decided to shout 'inappropriate', while he could have also emphasised it with asterisks. By choosing to shout, he lent extra emphasis to this word and made his utterance even more intense. *Appendix 2* contains a longer excerpt including these two lines and one can see that the intensity suits the heat of the discussion about using swear words well. Asterisks would also have emphasised *inappropriate*, but with a more neutral and less aggressive undertone than shouting. *Example 79* shows how both features can be

combined and illustrates that emphasis is gradable in a few steps as Haase et al (1997: 61) point out: *dare* < *DARE* < *D A R E*.

```
80. (line 27) <jhonijim> its just a word
    (line 28) <ErrantEgo> its an INAPPROPRIATE word.
```

8.2 Addressing recipients

In many-to-many conversation it is often necessary to address the recipient of an utterance directly. The lack of eye contact, gaze and body gestures poses a serious threat to gain the attention of the listener on the one hand and identify the addressee of an utterance on the other. The need to provide a direct reference to whom an utterance is directed is greatly enhanced by the chaotic structure of an internet chat, where different strands of conversation are merged into a single stream of utterances. Therefore it is very common to compensate the lack of these non-linguistic contact devices by writing the name of the addressee as the first word of an utterance, which is typically followed by a colon, but not necessarily so as the following example taken from *appendix 2* shows:

```
81. (line 49) <troubled> cal: which is why im smilin ;)
    (line 50) <mynetdude> cal or unserious depends on where you go
```

This form of directly addressing people is very popular and according to Dittmann the semicolon is the most frequently used punctuation mark next to the question mark because it is regularly employed in this function. The amount of direct references generally rises with the amount of people taking part simultaneously in a discussion as well as the amount of independent strands of conversations taking place at the same time. In my corpus it appeared that technical support channels, with their tendency to include short, specific requests and answers and immediate topic shifts show a special affinity for referencing. The following excerpt from *#gentoo* illustrates the amount of direct referencing quite well. It was randomly selected from one of the technical-

channel logfiles and out of the 25 utterances 15 contained direct references. I have added numbers in brackets to show the individual strands of conversation (2.1 and 2.2 are dialogues between different people about the same topic).

```
82.
(1)   <andip> jihsro, you could tail their history-files.
(2.1) <andip> _Sync_, i think you wanna use your mountpoints for that
(3)   <Moult> hey i emerged kde 4 using the overlay and unmasked a
      bunch of stuff. i seem to have removed most of the kde4
      files...however, some stuff like konqueror and konsole still
      remain...doing emerge --unmerge konsole or konqueror don't seem
      to detect that i have it installed...any ideas?
(2.1) <_Sync_> oh, yeah i forgot it's in the fstab
      *** pykid has left #gentoo ("Leaving")
(2.1) <_Sync_> so useradd --home-dir /home/stefan is it?
(3)   <bonsaikitten> Moult: kde-svn ?
(2.1) <andip> try with just useradd first, and you'll see what else to
      add ;)
(2.2) <tulcod> _Sync_: useradd -m Stefan
(2.2) <tulcod> i guess
(2.2) <_Sync_> tulcod, i don't want it to create a home
(3)   <Moult> bonsaikitten cannot remember, i followed a guide
(2.2) <tulcod> i can't seem to start jackd with realtime:
      ttp://rafb.net/p/wh3jqY77.html i think i followed
      http://www.gentoo.org/proj/en/desktop/sound/realtime.xml
      correctly
(2.2) <_Sync_> because i already have one
(2.2) <tulcod> _Sync_: i don't think it'll cause any damage...
(2.2) <_Sync_> mmm
(2.2) <tulcod> _Sync_: but read the man page & enjoy :)
(2.2) <_Sync_> :D
(3)   <bonsaikitten> Moult: hmm, so how do you know bits are still
      installed
(3)   <bonsaikitten> ?
(3)   <Moult> bonsaikitten i see it in my applications menu, and they
      run when i click on em
(3)   <bonsaikitten> Moult: are there entries for them in /var/db/pkg/
      ?
(3)   <bonsaikitten> Moult: category/package/ is the directory
      structure there
(1)   <jihsro> andip: ah yup  cheers
(3)   <Moult> bonsaikitten i've got a kde-base there
(3)   <bonsaikitten> Moult: and a konqueror-4.x.x subdirectory?
```

One can see the necessity of this device when multiple strands of conversation intertwine as the short question and answer sequences would have been very confusing without a clear structure indicating direct references. Note that there might be better examples for intertwining strands of conversation, as this excerpt was only a purely random sample taken from the *#gentoo* logfile. The aim was to demonstrate the necessity of adding direct references to a discussion containing only moderately intertwined strands of conversation, which can be found regularly in these channels. Furthermore *ex.82* indicates that short question and answer sequences promote a high use of direct referencing, which can be seen by the sheer amount of direct references included in this short excerpt.

8.3 Rhythm and pauses

In spoken discourse we divide our utterances into chunks that belong together and often make strategic pauses between them. The most obvious device to make pauses in an internet chat is to just type and send one chunk after the other. However it appears that there is still the need to mark pauses explicitly. Pauses are typically indicated with dots as the following examples show:

```
83.   <GaintSura> I noticed something when uploading my files to the
      server.... the file sizes are different.... the local copy is
      21914 Bytes, while the uploaded copy (no matter how many times I
      try to re-upload it) is 21667 Bytes...


84.   <ivan_w> I know what an FSM is.. that.. I can do


85.   <mynetdude> cc77, I'm listening... go on... :)
```

Dots can often be seen as a break in thought (Kramer 2008: 48) and can furthermore be understood as indicators of conceptual orality because of their function to simulate pauses or delays typically made in face-to-face discourse (Dittmann 2001: 90). Statements are also sometimes concluded with dots indicating a strategic pause at the end of an utterance (e.g. *ex.83*). The use of explicit pause markers is relatively common

and was found ca. 5000 times in my corpus. This popularity is also reflected in a study by Dittmann (2001: 90) who discovered that three dots constitute ca. 16% of all punctuation marks used in IRC.

The corpus further indicates that regular use of dots to indicate pauses depends very much on personal style and preferences. While some users will be very fond of including pauses, others might hardly use them. To support this claim I observed the pause-marker behaviour of various IRC users. One of the observed users made very excessive use of this feature. The 20 utterances he made in a channel already contained 31 pause markers and he did not formulate a single utterance without one. Eight of these utterances contained two pauses while two of them even contained three. It turned out that this user ended each of his 20 utterances with a strategic pause and that he was accountable for roughly half of all the pauses made in the observed channel.

This is definitely an extreme example but it illustrates that the explicit indication of pauses can be seen as a linguistic habit that characterises speech acts of certain users. I observed three other users that each made more than 100 utterances in a channel and appeared to use pause markers on a regular basis. I found that between 10-33% of their utterances contained pauses, which seems to be a rough average for users that make extensive use of this language feature.

In moderated chats pauses at the end of an utterance receive special importance. Usually one person has the right to speak and keeps it until the moderator hands it on to someone else. However utterances are still grouped as communicative chunks and sent individually rather than written as a whole (Storrer: 2001a: 454). The advantages of this method are obvious: if users would write their whole utterance as one line of text and send it upon completion, the other participants would have nothing to do because the text production is hidden. Smaller chunks enable the other users to remain active and continuously read small bits of what is being said. However this makes it difficult for the moderator to see whether the turn of the speaker is complete or if he is producing further messages. Therefore pauses are used at the end of utterances to indicate that the turn is not finished yet and the speaker wishes to continue. When there is no pause it is assumed that the speaker is done and the turn is passed on to the next candidate (Storrer

2001a: 454). An example from Storrer 2001 illustrates the use of pauses and clustering of utterances in moderated chats:

> leiter: „Ich übergebe an Rowo“
> Paul: „danke.“
> Rowo: „Bei der einfachsten Methode werden die Tokens einfach durch...“
> Rowo: „Leerzeichen getrennt...“
> Rowo: „also Token=Wort...“
> Rowo: „Das kann man mit einem deterministischen Transduzer machen, ...“
> Rowo: „bei dem jedes Leerzeichen durch eine Trennung ersetzt wird. ...“
> (...)

This use of explicit pause markers can also be applied in regular chat contexts and pauses can generally be used to indicate that the speaker has not finished yet. On the other hand they can also represent a strategic moment of silence and add a certain weight to an utterance. In other words dots at the end of regular chat utterances do not necessarily indicate that a user still wants to add something to this topic. This means that others have to deduce from the co-text what function pauses at the end of an utterance have. However since it is relatively easy to tell whether a message is incomplete and needs to be finished, pragmatic confusion should not be very common.

8.4 Referencing and quotes in IRC

Arrows (like <-- or <<) in IRC have the function of making a reference to something. They might be used as a self reference and are commonly used to paraphrase “I am from” as examples *86* and *87* show (Kramer 2008: 52). Additionally they may reflect the user’s mental or physical condition (Crystal 2006: 95). This use of reference is very similar to the action command *(ex.88-89)* and despite having a slightly different form they seem to have the same function.

In addition to self references arrows can also be used as a reference to quotes or external links. By quoting statements from other participants it is possible to comment on utterances made earlier in the conversation (*examples 90-91*). This possibility to ‘quote’ others and directly refer to what they have said is an advantage chat-discourse has over regular face-to-face conversation. It helps to compensate the difficulties that arise due to

multiple strands of conversation and time constraints: Even in a very complicated conversation it can be made sure which utterance the user is referring to, while older utterances can be brought up again at later points in time.

Ex.92 shows that this tool is not restricted to quotes: To explain the term *trolling ivan_w* has made up an example and uses reference arrows to point at it. Finally external links, pieces of programming code or anything else can be commented this way which appears to be very useful in a medium where the lack of hand gestures makes it difficult to point at something.

```
86.   <Kicher> <- german ;S
87.   * +GothicAngel <--- Scotland
88.   <+anniebabeblues> <--- eating
89.   <+Moi^> <-- tries to be happy always
90.   <+Sweets^> <esthero> she's gonna use you <--- i'm worried too
91.   <War2> <<Maddy_>> It will be hot in my bedroom   <--- I stopped
      reading after that. O_O O_O O_O
92.   <ivan_w> No.. Trolling would be: "Zuu is a n00b" << That's
      "trolling"
93.   <oiaohm> http://legacy.planetannihilation.gamespy.com/  << This
      is a mother of a map pack.
```

8.5 Conclusion

One can see that paralinguistic cues can have overlaps in form or function. Arrows for references can have the same function as actionmarkers and both may be used to transcribe physical actions or mental conditions. As another example intensifications can be performed by bracing words with asterisks or using capitalisation. On the other hand the same convention can have multiple meanings. Words braced by asterisks can either be intensifications or actionstrips, dots can either indicate a moment of silence or the wish to continue and arrows for references can be used for referring to quotes or links, paraphrasing statements like “I am from”, or transcribing actions. However since in most cases it is easy to deduce which of the diverse meanings of a convention is intended, there should be little pragmatic confusion.

Many of the features described in this section, like capitalisation, reduplication, pauses or references are not restricted to internet chat and can also be used in other forms of CMC like e-mails. Kramer (2008: 50) however notes that they tend to be more common in chats than e-mails, possibly due to the "inherently traditional function of writing a letter" that is associated with e-mails. The conceptual closeness of internet chat allows for a more playful way of expression and depends on these conventions to include paralinguistic cues, while other forms of CMC which are closer to the conceptually written pole have less to no need for the language features described above.

9. Contextual influences

So far this paper has discussed various paralinguistic features of IRC dialogues. However, apart from linguistic and paralinguistic features there are also several other factors that have an impact on spoken interactions. In addition to paralanguage Cook (1995: 37) identifies the following contextual features: situation, participant knowledge and participant attitudes. All of them have a considerable influence on spoken discourse and all of them can be found in IRC.

Many channels specialise on a certain topic or area which makes the influencing factor of participant knowledge easy to observe. For example most channels dedicated to computer games feature a large amount of special vocabulary only used in a particular game. References are frequently made to the world the game is set in and if a visitor has never played the game before it will often be very difficult, if not impossible to follow the conversation. Even if they are familiar with the game, special vocabulary that is used and often invented by the game community might pose a considerable threat to understand the content and be able to participate in conversation as many new phrases and abbreviations are invented such as in *example 94*, taken from *#wow*. A similar sort of specialist knowledge and vocabulary can be identified in technical channels that are dealing with computers. Because the speech situation seems to be the most interesting of the three aspects I will focus on this contextual feature.

```
94.   <DominoEffect> but I've got a bunch of nice meta designs
      <Xil0> Done the Consortium grind?
      <DominoEffect> and several gems for melee dps
      <DominoEffect> nah, not yet
      <DominoEffect> I just got done doing maghar rep runs
```

9.1 The Speech Situation

Cook (1995: 37) defines situation as the "features of the immediate physical surroundings including features of the participants." Even though in internet chat discourse every person is physically isolated from the others it could be argued that the physical surroundings of each individual member remain relevant parts of the speech situation. Since participants might from time to time tell their conversation partners what they are doing in the real world or make references to something that is happening around them, their physical environment sometimes influences the discourse. However, it would be a mistake to assume that this is the only type of speech situation relevant in CMC. There is also the virtual space the IRC user is located at and one can argue that the location on IRC is most likely much more relevant to the discourse than the physical surroundings of the user and that individual channels do indeed create a unique speech situation.

Firstly, many channels have a clearly defined topic, which makes it quite obvious what the subject of the discussion will look like. Sometimes the channel name might indicate what the focus and aim of a channel might be, for example *#EURO2008* (football EM) or *#linux.de*. It would be odd to join a linux channel and talk about football; most likely it will result in negative comments from other users or even a kick or ban from the channel. Other channel names remain more or less uninformative *(#eggdrop101, #88*) and it is up to the user to find out what is being discussed there. Given the dynamic nature of IRC, channels appear and disappear frequently, although some of them are fairly stable. In order to help users in finding the right places to talk many networks like Undernet have an in built search function that can list currently existing channels according to search criteria provided by the user. If the user for example would search for a place to talk about cats they could enter the command */list cats* in the Undernet

status window. The network will search all channel descriptions, topics and names for matches and one can join promising search results.

Most importantly however, a considerable influence on the speech situation is the fact that IRC channels often provide a framework of rules and behavioural guidelines. While on some channels swearing is considered inappropriate, others prohibit Off-Topic discussions. Sometimes it is allowed to dump considerable amounts of spam messages (for example sending the same link or advertisement several times in a row), while most channels outlaw such behaviour. Considering all these different rules and attitudes it becomes obvious that communication will differ according to the 'virtual place' someone is in. These rules are also important to know for a researcher, because one has to consider carefully which channels are appropriate when observing certain language patterns and which channels will almost never or unrealistically often produce the desired behaviour due to their individual rules and attitudes.

The excerpt in *appendix 2* was taken from the channel *#defocus* on Freenode and shows some of these attitudes very well. One can see that the repeated utterance of "fuck" by a channel user immediately caused various negative comments stating that "#defocus is a family friendly channel" (line 15) and that "spamming and using bad language is a no no" (line 53).

The statement:

```
This channel defines it as inapprorpiate. By joining this
channel, you agree to abide by our rules. (line 40)
```

illustrates that *#defocus* defines what is appropriate and that any user on this channel has to follow the individual rules of the speech situation, while

```
<cal> troubled: sorry, but the internet _is_ serious business :)
(line 46)
[...]
<mynetdude> cal or unserious depends on where you go (line 50)
```

illustrates, that these rules are highly dependant on "where you go" and each channel might have a very different attitude towards swearing and bad language.

Since it is so important for the participant to know these rules, the question remains how to find out what is expected in a channel and what is inappropriate. The most useful methods to broadcast channel rules are to include them into the *channel topic*, which is a description that can be edited by users with appropriate rights; or create an automated message by an IRC bot. Irrespective of which option is chosen, every time a user joins a channel the message will be displayed and they can stay informed what this channel is about and how users should behave. The following examples from the corpus illustrate some examples for channel topics and messages:

95. -Q- [#Chat-World] Welcome to Chat-World Please speak english. Advertising or spamming will result in a ban

96. *** Topic is 'Welcome to #freenode, official help channel | Available staff are voiced | Feel free to message us! | Looking for the social channel? Try #defocus | Guidelines found at http://freenode.net/poundfreenode.shtml | Check out our blog: http://blog.freenode.net'

97. -Q- [#rogue] #Rogue! The most helpful and enlightened bunch of rogues in the wor... oh who am I kidding, come in and see the freaks. [Likely to contain explicit images not fit for under 18's.]

98. *** Topic is '[## Socialites] Welcome to Freenode's most awesome social channel! Feel free to sit around and chat, drink beer, etc. xteddy will be your bartender for this evening. Warning: The bouncers have pistols and might 'accidentally' shoot you if you get out of line =) [< starshine> Raping the channel dÃ©cor is not cool.]'

As one can see these messages might include clear statements on what is expected and what is prohibited, as well as the response to violation (*ex.95*), links to other channels which might be more appropriate (as in *example 96*: “Looking for the social channel?”), or links to more extensive guidelines and many other hints. *Example 97* mentions that this channel will be likely to contain explicit images (which in this case seems to be more of a joke and theoretical warning than actual implication), while other channels might indicate that swearing and any bad language is decidedly unwelcome. The last description (*ex.98*) depicts the channel atmosphere in more detail and one might imagine this place as something cosy and informal. This is supported by expressions like “sit around”, “drink beer” and “bartender” mentioned in the invitation. By this short

atmospheric text, the channel differentiates itself from other supposedly less informal channels and one could also imagine that the attitude towards 'bad language' in *#rouge* or *##socialites* might be less strict than in *#defocus* (and indeed the word "fuck" appeared 50 times in *#rouge* and only 3 times in *#defocus* apart from the passage the appendix 2 dialogue is from). Altogether there is much support for the assumption that IRC channels form unique speech situations and that the channel location can be as important to the conversation as the location is to traditional face-to-face communication.

Behavioural rules may also be provided by the networks or servers a user joins. They are of course more general and primarily concerned with violation or disruption of communication and abuse of the system. Networks or servers might for instance deny the use of certain bots or prohibit illegal or harmful activities, such as channel takeovers (see 9.3), or the flooding of channels with spam and harmful scripts. The following quotes are excerpts from Quakenet and Undernet server policies that appear in the IRC status window every time one connects to the network. Both policies mention that the use of their servers is "a privilege not a right" which might be removed any time.

Excerpt from Quakenet policies:

```
- Please note that the network may not be used for the following:
-
-   o No Flooding
-   o No DoS bots / Virus / Cheats / Trojan distribution
-   o No Distribution of Warez / Pornography / Copyright material
-   o No Spamming / Advertising
-   o No Takeovers
-   o No Phishing
-   o No Racism / Nazism
```

Excerpt from Undernet server policies:

```
- ==>     Bot Policies:
-
-     It is allowed to run NON abusive bots on this server, all
      abusive bots will be killed on sight.
[...]
-     The use of this server is no right, but a privilege. The
      admin(s)and opers can revoke this priviledge without further
      notice and without a reason.
```

9.2 Metaphorical Rooms

As mentioned in chapters 2 and 7.2 channels can be seen as virtual rooms and simulate metaphorical physical closeness (Storrer 2001b: 18). Therefore comments can refer to the metaphorically constructed chat room or the real life settings of the user. "Being away" for example can mean that the user leaves the workstation, but very often simply means that the user is busy with another activity on the computer (Storrer 2001b: 18). In the statement "I have just read an article, so I was away for a while" for example the *away* refers to the metaphorical chat room, not the workstation and the user has never left the real-life setting. Sometimes these terms are ambiguous as the following example from Haase et al (1997: 67) shows:

> Karin: Gleich wird Theo herkommen.
> Horst: Hier in den IRC?
> Karin: Horst: Nein, er kommt mich besuchen für das Wochenende.

References like „here" can refer both to real-life or fictional setting which may cause misunderstandings.

Unlike in face-to-face conversation users can also make references to what has been said before due to the fact that the chat log saves every utterance. Therefore statements like "see above" can be found which would be impolite for oral communication, but are very tempting in CMC because it saves time (Storrer 2001b: 19).

9.3 Acts of Authority and Virtual Territory

It has been mentioned before that a failure to follow the channel or network rules may lead to kicks and bans, which both remove the user from the channel or network in question temporarily or permanently. Hentschel (1998: Section 3.1.1) describes both actions as "territorial activities", with the major difference to real-life that the territory in question is virtual rather than real. The effect and intention however is similar: a kick is supposed to have a pedagogical effect forcing the user to change their behaviour as soon as they get back on the channel. Bans are considered much more serious and are a more drastic and rather unfriendly action with a strong negative connotation. A kick can sometimes be performed "just for fun" by operators to taunt, mock or play with users

and friends - Hentschel (ibid) adds that it can even be used for flirting. Bans however are used to permanently remove unwanted guests. According to Hentschel (ibid) bans can even be employed in "wars" fought over virtual IRC territory, where certain malicious scripts try to ban all users occupying a channel, leading to a takeover. She mentions as an example that during the Balkan war the channels *#Croatia* and *#Serbia* attacked each other with so called "war scripts". In addition to ban-scripts, other scripts can be used to invade a channel with countless spam messages, thus flooding it and making meaningful conversation impossible. Even though open wars are certainly a rare exception and highly unwanted by the networks, it shows how important "virtual space" is to the IRC users. Despite the dynamic structure of IRC, considerable amounts of time are invested to shape the virtual home territory and various bots can enhance the capabilities of a channel. They might for example perform trivia or act as security and distributor of channel rights or news.

The "law enforcement" varies from channel to channel but some appear to be very fast to ban or kick suspicious users. They might even get banned from a channel simply for being on "bad channels": A script or operator checks on which channels a user is logged on, resulting in a ban as soon as something considered "inappropriate" is found, even without the user ever acting in a negative way. On one of my recordings the channel *#alnitecafe* on DALnet for example banned 69 users for being on "bad channel(s)" and 284 further users because a "banned channel" was detected. In my study I could observe that networks seem to differ regarding the frequency of kicks. DALnet seems to encounter more kicks and bans than other networks, for example two sessions from *#alnitecafe* on DALnet resulted in a total of 2039 kicks, while two equally long sessions from *#alnitecafe* on Quakenet only encountered 212 kicks. All of the DALnet kicks were also bans while only 69 bans occurred on Quakenet. Thus one can expect DALnet to be a 'rougher place' than the other networks investigated in this study.

9.4 Performance and Play on IRC

Danet et al (1998) investigated performance and play on IRC by analysing a "virtual party" taking place on an IRC channel. The playful and creative nature of communication on IRC is mentioned by many authors and is indeed a very present feature that is enriched and emphasised greatly by the substitution of nonverbal cues. Even though the study of Danet et al focused on larger scale improvisations like virtual parties and plays, their model can be applied to normal communication as well, since small scale 'performances' are relatively common. Examples include sequences with strong use of narrative action commands as discussed in some quotes of chapter 7.

The authors argue that IRC communication can take place on multiple layers, which they refer to as frames. Humans can operate in more than one frame at the same time, which means more than one can be active simultaneously even tough some might be pushed to the background. Additionally an action within one frame might also have a meaning in an outer and larger frame that incorporates it (Danet, 1998: 53). The *real life* frame is the outmost frame that is always active, despite often being pushed to the background in the course of a virtual play. While talking players might be involved in all kinds of real life activities and might make reference to what they do or what happens in their physical surroundings. People might complain that they have been dumped by their partners or that they have to leave the conversation because of some real life event. The second frame is the *let's play IRC* or *let's talk* frame, where "anything may be said and participants enjoy reduced accountability if they choose to communicate in a playful mode" (Danet, 1998: 55). People are aware that they are communicating within the artificial IRC environment and a discussion in that mode can either be playful or serious. In the "party" frame the "action and utterances are primarily in a playful mode" (Danet, 1998: 55) and an environment is established where people can have fun. Due to the further reduced accountability people can flirt, drink, fool around and make friends easily. In the inner frames, which Danet et al identify as *pretend* and *performance* frames, the imaginative behaviour of the participants is fully brought to life. In the case of Danet's study the party took place in a channel called *#weed* and its highlights included numerous references to virtually smoking marihuana. In the *pretend* frame this simulation of smoking together is pretended to be real and the

authors note that this kind of extended pretend-play would be very unusual for adults in real life. It is "usually the province of children, who play doctor, Mommy and Daddy, riding a horse and so on" (Danet et al 1998: 59) with any props they can muster. In virtual reality however adults seem to improvise and perform to bring a make-believe reality to life just like children do, with nothing but their keyboards at their disposal. In the *performance* frame these actions are played out and the message is "Let's show each other what we can do with the keyboard" (Danet et al 1998: 59). Indeed Danet et al argue that showing off creativity and getting positive feedback from the peers is very prominent in IRC performance.

It lies at hand that these two inner frames make very extensive use of substituted paralinguistic cues. The following examples from Danet et al (1998: 64-65) show an excerpt of the IRC party recorded for their study:

> <Thunder> sssssssssss *passes joint to kang*
> …
> <Kang> thanx dude *puff* *hold*
> …
> <Thunder> kang exhale.. you will die :-)
> <Kang> *exhale*
> <Kang> ;)

As one can see action strips are very common in this kind of performance. They seem to be used frequently and can be compared to stage directions. Also note the onomatopoeic expression of "sssssssssss" simulating the smoke, which is very commonly used throughout the whole party. The two emoticons are interpreted by Danet et al as feedback device to the performance of the chat users: "<Kang> winks with pleasure at his performance and at the terrific simulation game they are playing." (Danet et al 1998: 64). Later on the participants continue to creatively play with the idea of smoking by composing little emoticon stories:

<Kang> :|
<Kang> :|
<Kang> :\
<Thunder> heheheh
<Thunder> hehehehehe
<Thunder> that was great
<Kang> :\
<Kang> :)
<Thunder> hehehehehe
<Kang> *exhale*
<Kang> :()
…
<Thunder> :-Q :| :| :\ssss :)

The two participants try to imitate various stages of inhaling and exhaling smoke and Danet et al point out that the sequence :| :| :\ means something like "puff puff hold the smoke inside" while :() represents the mouth exhaling smoke (Danet et al 1998: 64). Neither of the two participants ever explains any further what the sequence :| :| :/ explicitly means, nor why *Thunder* complements *Kang* so much on his improvised performance. *Thunder* then takes up this improvised sequence to further play with it and later on also changes the channel topic to a very similar line.

This example serves to illustrate the creativity and playful use of paralinguistic content that is evident in IRC play as well as the extend to which the participants enjoy the process and value the efforts made by their peers. Danet et al argue that the boundaries of the discussed frames are playfully blurred and that within in the *performance* frame references sometimes are made to IRC or real life for example in:

<Thunder> *as smoke fills the channel again*

or

<Kang> here lucia *hands bong* *long reach (over ocean)*

In the first example a reference is made to the *IRC frame* as the channel is a substitute for the virtual room the users are in. Another similar example from Danet et al would be "*throw seeds into channel #hottub*". In the second example *Kang* makes a reference to real life and the place where *lucia* is located. He hands the bong over the ocean to her real life location even though she is at the same time also in the same 'virtual location' as *Kang* and thus only an arms reach away.

Even though the model of frames was developed with the concept of extensive virtual plays in mind, it could also be applied to general IRC discourse, because small slips into *pretend* and *performance* frames seem to be a common and relevant feature of IRC conversation and IRC users can change to these at any moment. This can be seen in *example 59* of my corpus already mentioned in chapter 7 where the expression "gun powder hugs" triggers a playful reaction by *Cann0n*, which again encourages *Bspec* to further engage in the miniature play and hand out some Kleenex, so that *Cann0n* can in turn react to *Bspecs* performance. After the short and playful interaction is over they step back from the *performance* and *pretend* frames to the *IRC frame*, and continue with whatever they were discussing before. In the case of this example the channel members were discussing a serious topic: their concerns about the harmful influence of media on our cultural values. The two performance participants (who were both not deeply involved in the discussion at that time) were trying to improve the mood by their short improvised performance. Afterwards they returned to the serious discussion and stated their own opinion on one of the most recent utterances made by their peers.

Finally every user that quits IRC inevitably steps back into the *real life* frame and all others are closed, even though references to the outmost frame might be made at any time during a conversation.

```
59.   * Bspec smothers Cann0n in gun powder hugs
      [...]
      * Cann0n sniffs and says, "I love the smell of gun powder"
      [...]
      * Bspec gives Cann0n some Kleenex
      [...]
      <Cann0n> thanks, *sniff* i uh.. i got *sniff* gun powder got in
      my eye..
```

This chapter has introduced play and performance on IRC in order to demonstrate the relevance of these concepts and their impact on IRC discussions in general. Through the example of Danet's virtual party it has been shown that paralinguistic elements play a great role in such kind of discourse and that communication on IRC can occur on different layers. Even though a full scale virtual party might be a rather rare exception, short improvised scenes seem to be a common feature. Users might jump into the *performance* layer at any time and act out a certain scene that forms in their head. These

communicative acts often use a high amount of paralinguistic cues to meet the demands of virtually improvising an interactive chain of actions. Especially the action commands shine in this area but also emoticons, action strips and other paralinguistic features such as letter reduplications and onomatopoeic expressions are used to great extent. In these creative interactions with language and keyboard the full potential of paralinguistic substitutes is utilised. When used this way they are probably not employed for their communicative role in CMC discourse but simply for the sake of playing with signs and language, as we have seen in Danet's example where emoticons were used to simulate the act of smoking.

10. Conclusion

The objective of this study was to gain new insights into paralinguistic language use in IRC as well as to provide an analysis of features already discussed by other sources, in order to show the range of application and the worth of paralinguistic cues in the IRC environment. The creation of an extensive corpus allowed meaningful generalisations about paralinguistic behaviour in IRC and provided an excellent basis for investigating various research questions. A close analysis via *Word Smith Tools* by Mike Scott made it possible to study emoticons and actionmarkers in unprecedented detail and many new findings were revealed in the course of the study.

It appeared that emoticons are by far the most common representatives of paralinguistic cues, their huge popularity being indicated by the fact that the most common smileys are among the most popular words in the whole corpus. Because IRC networks and channels differ from each other it is difficult to measure the global occurrence of all emoticons exactly. Nonetheless this study could provide a comprehensive list of the relevant items throughout IRC and it is believed that only about ten emoticons are of considerable importance. Furthermore this thesis suggests that the use of emoticons has changed over time. A comparison of data obtained from the IRC corpus and other sources indicates that the emoticon :) has lost much of its former importance as 'standard emoticon'. While it seems that ten years ago about 80% of all emoticons have been of the type :), this number has decreased to 25%, while others like :D or loans from Asian internet culture have increased considerably, with :D being used as frequently as :). This result suggests that a larger variety of different emoticons is necessary to express the most important paralinguistic cues than only the two basic smileys :) and :(and that the former lack of variability has led to the development of a limited number of further emoticons that are now widely spread and commonly used throughout IRC.

Most emoticons seem to focus on the key feelings of 'happiness' and 'sadness', even though other emotions like annoyance, irony, surprise or shock are also popular uses. The feeling of suspicion is added through an Eastern loan and did not exist before in Western emoticon culture. This might probably be caused by the fact that Asian

emoticons focus on eyes which makes it is easier to express suspicion – a feeling that is difficult to describe with the traditional Western focus on the mouth. Even though there are multiple humour markers in IRC a close investigation of these items implies that they are not true synonyms as they have individual characteristics, vary in intensity and tend to get used in different situations. Additionally emoticon use appears to be something highly personal and smileys are used differently by every person. Thus many users seem to limit the use of or avoid some emoticons altogether. The same thing can also be said about most other paralinguistic features such as actionmarkers. The individuality in employing paralinguistic cues has been indicated by the research on pause markers via dots. While some users made excessive use of this feature, others completely ignored it.

As mentioned above emoticons only seem to be used for a few of the most basic feelings while all other emotions such as boredom (*yawn*) are likely to be represented by actionmarkers. Both language features have their strengths and weaknesses and differ for instance regarding accuracy and diversity. Emoticons are characterised by their potential ambiguity and vagueness, actionmarkers on the other hand provide rather accurate semantic descriptions. The advantage of emoticons to attach basic emotions fast and with few keystrokes is counterbalanced by the necessity to memorise new combinations of symbols and thus they are limited to cover only a small number of emotions while actionmarkers can express almost everything.

Action commands were used in the corpus to transcribe a large variety of physical or non-verbal cues which is shown by the fact that over 150 verbs were used as collocates. While it appeared that some actions like looking, nodding, yawning or sighing are of special importance, arguably one of the most interesting observations was that some of the most frequently used actionlines represent cues of communicative closeness. Six out of the seven most frequently used verbs for action commands are represented by this type of collocation. Especially the acts of hugging and slapping are highly ritualised and widely spread throughout IRC, but also related actions like licking and poking are commonly used. This indicates a desire of the users to establish metaphorically physical contact with other chat members by ‘touching’ them and illustrates how communicative closeness cues might serve to create and maintain a friendly and light-hearted atmosphere that seems to dominate IRC discourse in most channels. The popularity of

these collocations suggests that action commands can serve to make conversations more personal and interactive and underlines the playful nature of IRC. Generally it can be observed that much creativity is employed when including actionmarkers and attempts are made to make communication more lively and interesting.

One type of action commands turned out to be of special interest: The possibility to express thoughts and opinions via indirect statements in third person form. This paper provided some suggestions for their communicative function and discursive status by analysing examples from the corpus. However these remain suggestions and further research should be done in this area as it would be very interesting to find out what factors influence the decision to paraphrase an utterance into an indirect statement. From the perspective of this research project it seems that there are numerous different reasons for expressing thoughts via a third person statement, ranging from making context explicit by sharing what is going on in someone's head to providing a certain level of neutrality or setting a certain mood. Also the desire to break monotony by using a different colour scheme and linguistic structure should be taken into consideration.

Generally paralinguistic cues are very common in IRC and one can say that many conventions to represent non-linguistic or paralinguistic information have evolved, ranging from making strategic pauses to addressing other users directly. Some of them work similar to spoken discourse and imitate features like shouting or stressing individual words. Others are an adaptation rather than an imitation of spoken discourse. Emoticons for instance cannot be seen as an equivalent to facial expressions as they have unique communicative functions that set them aside from anything encountered in face-to-face conversation.

Following this study one can come to the conclusion that in a genre which is closely associated to communicative closeness like internet chat, certain paralinguistic cues are necessary and cannot be left out completely. Substitutes had to be found for the most relevant language features and their implementation is widely used throughout IRC. Some of these conventions show that a certain level of intonation, stress and rhythm is helpful in getting meaning across in internet chat; however the most important factor seems to be able to show ones intentions and convey certain emotions. This is reflected

by the popularity of both emoticons and actionmarkers which are undoubtedly the most relevant and numerous paralinguistic cues found in IRC.

The fact that all these features are used underlines how important paralinguistic cues are to spoken discourse. Without the possibility to send information such as speaker intentions on a paralinguistic level, oral communication would not only be difficult, it would be bound to fail eventually. Words alone do not seem to be enough to avoid grave misunderstandings; otherwise CMC would not have felt the necessity to implement various substitutes – substitutes that are arguably inferior to their face-to-face counterparts in some respects. In an environment where every keystroke is an annoyance one can not stress enough the relevance of the presence of features that take time and effort to produce. Especially the addition of actionmarkers is more time consuming than producing regular plain text. Even though they might not imitate every facet of paralinguistic behaviour of face-to-face communication, internet chat cues have their own advantages and even make things possible that could not be realised in ordinary oral communication.

Whatever differences between chat networks or individual user behaviour might exist, emotions, actions and attitudes of speakers are frequently integrated into the flow of conversation. Therefore IRC users are not only capable of “typing the untypeable” – they also manage to “transcribe the untranscribable”.

Bibliography

Argyle, Michael. 1972. “Non-verbal Communication in Human Social Interaction”. In: Hinde R.A. (ed). *Non-verbal Communication.* London: Cambridge University Press, 243-268.

Baym, Nancy. 1995. “The emergence of community in computer-mediated communication”. In Jones Steven (ed). *Cybersociety: Computer-mediated communication and community.* Thousand Oaks: Sage Publications Inc.,138-163.

Beißwenger, Michel. 2001. “Das interaktive Lesespiel. Chat Kommunikation als mediale Inszenierung". In: Beißwenger, Michael (ed). *Chat-Kommunikation. Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld*. Stuttgart: Ibidem Verlag, 79-141.

Beißwenger, Michael. 2000. *Kommunikation in virtuellen Welten: Sprache, Text und Wirklichkeit.* Stuttgart: Ibidem-Verlag.

Beißwenger, Michael; Storrer, Angelika. 2008. “Corpora of Computer-Mediated Communication“. In Lüdeling, Anke; Kytö Merja (eds). *Corpus Linguistics. An International Handbook. Volume 1.* Berlin, New York: Handbücher zur Sprache und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science 29.1, 292-308.

Cameron, Deborah. 2002. *Working with Spoken Discourse*. London: Sage.

Chambers, John; Trudgill, Peter; Schilling-Estes Natalie (eds). 2002: *The Handbook of Language Variation and Change.* Maldon Mass, USA: Blackwell Publishers.

Cherny, Lynn. 1999. *Conversation and Community: Chat in a Virtual World.* Stanford, California: CSLI Publications.

Čmejrkov, Světla. 1999. “Czech on the Network: Written or Spoken Interaction?”. In Naumann, Bernd (ed.). *Dialogue analysis and the mass media*. Tübingen: Max Niemeyer Verlag, 113-127.

Cook, Guy. 1995. “Theoretical issues: transcribing the untranscribable”. In Leech Geoffrey; Myers, Thomas (eds). *Spoken English on computer : transcription, mark-up and application.* Harlow: Longman, 35-53.

Corder, Saskia; Meyerhoff, Miriam. 2007. “Communities of practice in the analyisis of intercultural communication”. In Kotthoff, Helga; Spencer-Oatey, Helen (eds). *Handbook of Intercultural Communication*. Berlin, New York: deGruyter, 441-465.

Crystal, David. 2006. *Language and the Internet*. Cambridge: Cambridge University Press.

Danet, Brenda; Ruedenberg-Wright, Lucia; Rosenbaum-Tamari, Yehudit. 1990. “Hmmm … where’s that smoke coming from? Writing, play and performance on Internet Relay Chat.” In Rafaeli, Sheizaf; Sudweeks, Fay; McLaughlin Margaret (eds). *Network and netplay: Virtual groups on the Internet*. Cambridge, MA: AAAI/MIT Press, 41-76.

Dittmann, Miguel. 2001. *Sprachverwendung im Internet, Untersuchungen zur Nutzung des Internet Relay Chat (IRC) in Deutschland und Frankreich.* Books on Demand GmbH, Norderstedt.

Döring, Nicola. 2001. “Sozialpsychologische Chat Forschung: Methoden, Theorien, Befunde”. In Beißwenger, Michael (ed). *Chat-Kommunikation. Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld*. Stuttgart: Ibidem Verlag, 141-187.

Döring, Nicola. 1995. *Isolation und Einsamkeit bei Netznutzern? Oeffentliche Diskussion und empirische Daten.* Source: http://sites.inka.de/bigred/misc/netz+einsamkeit.txt

Dortmunder Chat Korpus. Source : http://www.chatkorpus.tu-dortmund.de/

Gallery, Heike. 2000. “’bin ich-klick ich’ Variable Anonymität im Chat“ In Thimm, Caja (ed). Soziales im Netz. Sprache, Beziehungen und Kommunikationsstrukturen im Internet. Opladen, Wiesbaden: Westdt. Verlag, 71-88.

Gee, James Paul. 2003. What video games have to teach us about learning and literacy. New York: Palgrave Macmillian.

Geers, Rainer. 1999. “Der Faktor Sprache im unendlichen Daten(t)raum. Eine linguistische Betrachtung von Dialogen im Internet Relay Chat“. In Naumann, Bernd (ed.). Dialogue analysis and the mass media. Tübingen: Max Niemeyer Verlag, 83-100.

Gelléri, Péter. 1998. *The IRC Vernacular - A Linguistic Study of Internet Relay Chat*. Source: http://www.gelleri.hu/html/irc.html

Haase, Martin et al. 1997. “Internetkommunikation und Sprachwandel“. In Weingarten, Rüdiger (ed). *Sprachwandel durch Computer.* Opladen, Wiesbaden: Westdt. Verlag, 51-85.

Haralampieva, Tatiana. 2004. *The impact of internet on language: stylistic features of synchronous CMC.* MA thesis, University of Vienna.

Hentschel, Elke. (1998): "Communication on IRC". *Linguistik online 1*: 1-17.

Hess-Lüttich, Ernest; Wilde Eva. 2003. “Der *Chat* als Textsorte und/oder als Dialogsorte?“. *Linguistik online 13, 1/03.*

Herring, Susan. 2004. “Computer-Mediated Discourse Analysis: An Approach to Researching Online Behavior”. Preprint version to appear in Barab, S. A., Kling, R., & Gray, J. H. (eds). *Designing for Virtual Communities in the Service of Learning*. New York: Cambridge University Press.

Jones, Quentin. 1997. “Virtual-communities, virtual settlements & cyber-archaeology: A theoretical outline”. *Journal of Computer-Mediated Communication* [Online], 3(3). Source: http://www.ascusc.org/jcmc/vol3/issue3/jones.html.

Jones, Steven. 1995. “Understanding Community in the Information Age” In Jones, Steven (ed). *Cybersociety: Computer-mediated communication and community.* Thousand Oaks: Sage Publications Inc., 10-36.

Kelly-Holmes, Helen. 2006. “Multilingualism and commercial language practices on the internet”. *Journal of Sociolinguistics* 10/4, 507-519.

Kelly, Justin. 2009. *An Introduction to IRC commands – A quick reference of IRC commands.* Source: http://www.scribd.com/doc/12892901/A-Introduction-to-IRC-Commands

Koch, Peter; Oesterreicher, Wulf. 1994. “Schriftlichkeit und Sprache“. In Hartmut, Günther; Ludwig, Otto (eds). *Schrift und Schriftlichkeit.* Berlin, New York: deGruyter, 587-604.

Kramer, Birgit. 2008. “Laughing out loud at a newcomer. Lol @ n00b - netspeak and its impact on English outside the internet”. MA thesis, University of Vienna.

Liu, Goeffery. 1999. “Virtual community presence in Internet Relay Chatting”. Journal of Computer-Mediated Communication, 5 (1). Source: http://jcmc.indiana.edu/vol5/issue1/liu.html

Marcoccia, Michael. 2007. *Text-Centered versus Multimodal Analysis of Instant Messaging Conversation.* Source: http://www.languageatinternet.de/articles/2008/1621

Marshall, James. The canonical smiley (and 1-line symbol) list. Source: www.astro.umd.edu/~marshall/smileys.html

Ooi, Vincent. 2002. “Aspects of computer-mediated communication for research in corpus linguistics”. In Pam, Peters; Collins, Peter; Smith Adam (eds). *New Frontiers of Corpus Research*, Amsterdam, New York: Rodopi, 91-104.

Partington, Alan. 2006. *The Linguistics of Laughter – A corpus-assisted study of laughter-talk.* London, Routledge.

Reid, Elizabeth. 1991. *Electropolis: Communication and Community on Internet Relay Chat.* Senior Honours thesis, University of Melbourne, Australia. Source : http://www.scribd.com/doc/94505/Electropolis-by-Elizabeth-M-Reid

Rezabeck, Landra; Cochenour, John.1994. “Visual Cues for Computer-Mediated Communication. In *Imagery and Visual Literacy: Selected Readings from the Annual Conference of the International Visual Literacy Association*.

Rheingold, Howard. 1993. *The virtual community: Homesteading on the electronic frontier*. Reading, MA: Addison-Wesley.

Schönfeldt Juliane. 2001. “Die Gesprächsorganisation in der Chat-Kommunikation“. In Beißwenger, Michael (ed). *Chat-Kommunikation. Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld*. Stuttgart: Ibidem Verlag, 25-55.

Schulze, Markus. 1999. “Substitution of Paraverbal and Nonverbal Cues in the Written Medium of IRC”. In Naumann, Bernd (ed). *Dialogue analysis and the mass media*. Tübingen: Max Niemeyer Verlag, 65-82.

Scott, Mike. 2004-2006. *Oxford WordSmith Tools.* Version 4.0. Oxford: Oxford University Press.

Storrer, Angelika. 2001 a. “Getippte Gespräche oder dialogische Texte? Zur komunikationstheoretischen Einordnung der Chat-Kommunikation“. In: Lehr, Andrea; Wiegand, Herbert-Ernst (eds). *Sprache im Alltag. Beiträge zu neuen Perspektiven in der Linguistik.* Berlin, New York: deGruyter, 439-465.

Storrer, Angelika. 2001 b. “Sprachliche Besonderheiten getippter Gespräche: Sprecherwechsel und sprachliches Zeigen in der Chat-Kommunikation“. In Beißwenger, Michael (ed). *Chat-Kommunikation. Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld*. Stuttgart: Ibidem Verlag, 3-24.

Storrer, Angelika. 2000. *Schriftverkehr auf der Datenautobahn: Besonderheiten der schriftlichen Kommunikation im Internet*. Source: www.studiger.tu-dortmund.de/images/Storrer_2000_Schriftverkehr_auf_der_Datenautobahn.pdf

Troest, Mads Orbesen. 1998. *Computer mediated communication – lingua ex machine.* Source: http://earth.subetha.dk/~eek/museum/auc/marvin/www/library/uni/papers/cmc/cmc.htm

Yuki, Masaki; Maddux, William; Masuda, Takahiko. 2007. “Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States”. In *Journal of Experimental Social Psychology* 43(2), 303-311. Source: http://hdl.handle.net/2115/22527

Wenger, Etienne. 2004. *Communities of Practice.* Cambridge: Cambridge Univ. Press.

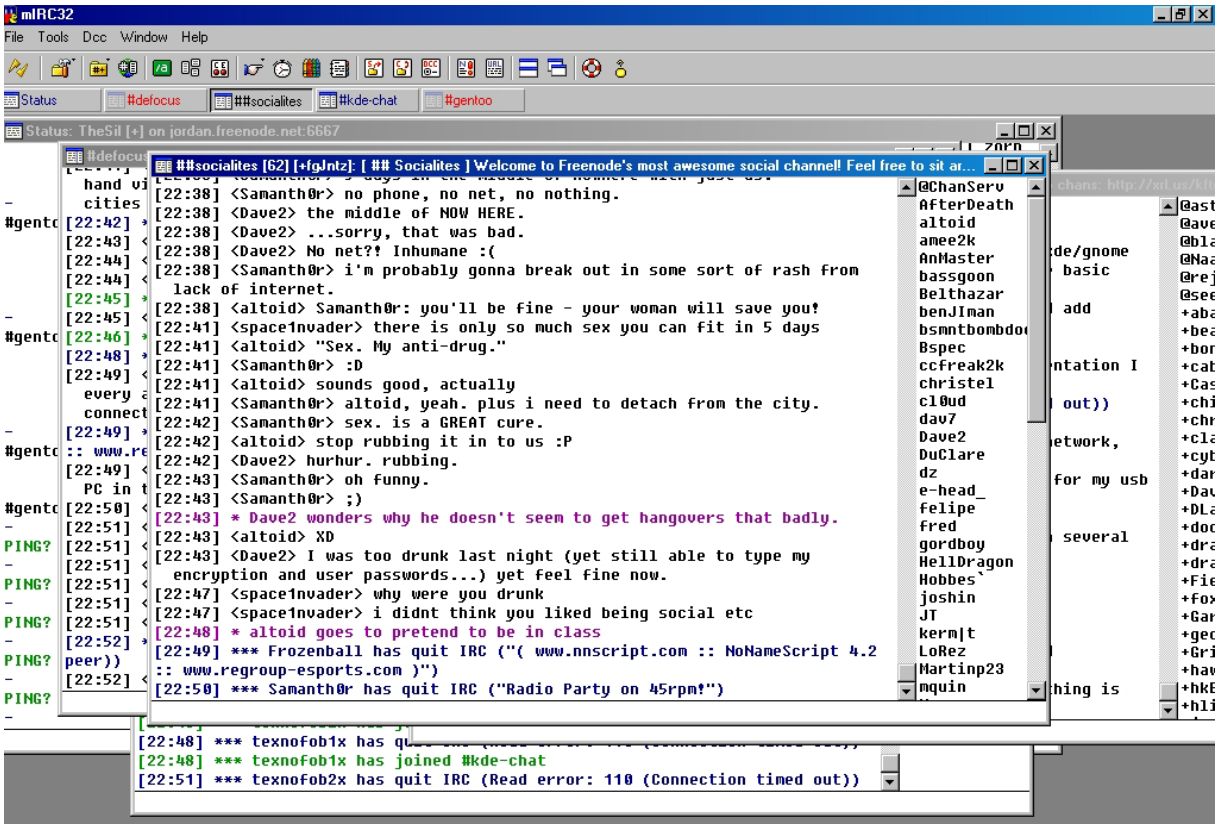
Werry, Christopher. 1996. “Linguistic and interactional featuers of Internet Relay Chat”. In Herring, Susan (ed). *Computer-Mediated Communication: Linguistic, social and cross-cultural perspectives*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 47-65.

Witmer, Diane; Katzman, Sandra. 1990. “Smile When You Say That: Graphic Accents as Gender Markers in Computer-Mediated Communication”. In Rafaeli, Sheizaf; Sudweeks Fay; McLaughlin Margaret (eds). *Network and netplay: Virtual groups on the Internet* Cambridge, MA: AAAI/MIT Press, 3-12.

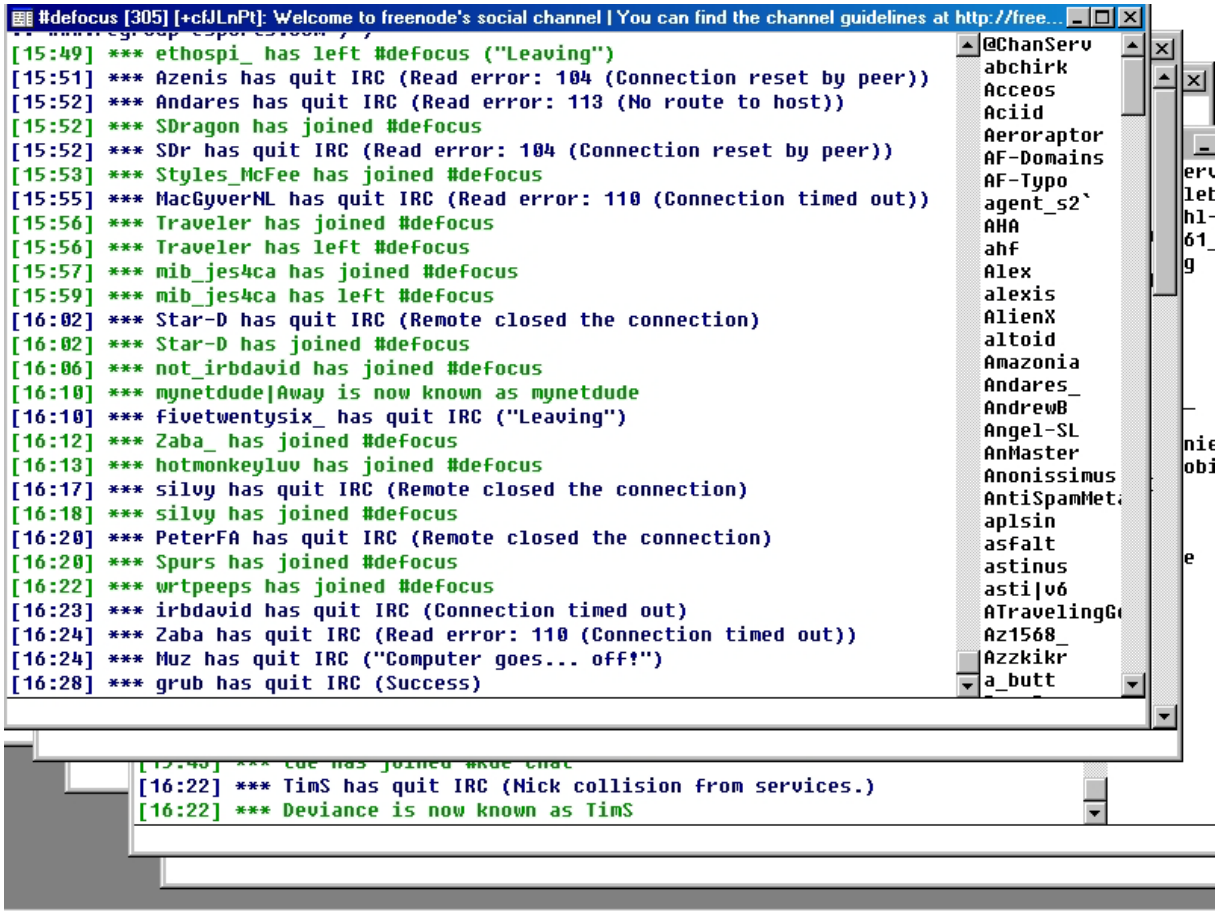
Yates, Simeon. “Oral and written linguistic aspects of computer converencing”. In Herring, Susan (ed). *Computer-Mediated Communication: Linguistic, social and cross-cultural perspectives*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 29-47.

Appendix

Appendix 1: IRC Screenshots



The IRC client *mIRC*. Several channel windows are open, ##socialites is the currently active window. Note the purple action command lines.



Another IRC screenshot. No one is talking in this channel and the only messages are the green and blue system messages of people joining and leaving. On the right hand side of each channel window all users currently in this channel are listed.

Appendix 2: IRC Dialogue – Example text from the corpus

<jhonijim> FUCK!!!
<jhonijim> FUCK!!!
<troubled> grrr
<jhonijim> FUCK!!!
<jhonijim> FUCK!!!
<jhonijim> FUCK!!!
<jhonijim> why are you looking at me
<Cann0n> why are you spamming?
*** GaintSura has quit IRC (Connection timed out_)
<Cann0n> !admin
<ErrantEgo> please quit flooding the channel
<troubled> jhonijim: please, just let it go
<ErrantEgo> please quit swearing
<jhonijim> whats so wrong about swearinh
<ErrantEgo> jhonijim, #defocus is a family friendly channel (meaning young eyes could be watching), please watch your language while you are in #defocus...thank you. :-)
<ErrantEgo> in addition, you were flooding
<jhonijim> i may have been flooding
<DrJ> you think?
<ErrantEgo> which is unnecessary
<jhonijim> but aleast i got someone to talk to me
<troubled> jhonijim: was uncalled for, and staff have better things to do :)
<DrJ> *sigh*
*** hd_ has joined #defocus
<cal> f word flooding? o.O
<jhonijim> its just a word
<ErrantEgo> its an INAPPROPRIATE word.
<cal> OMFG SOMEONE IS SWEARING WE'RE GOING TO DIEEEEEEEEEEEEEEE!!!!!!!!!
<mynetdude> why do you guys bother feeding the troll anyway?
<jhonijim> who defins it as inapropriat
<Pritchard> Well, I don't think swearing's all that bad. But that's irrelevant here.
<mynetdude> cal shhhh really
* cal dies
<mynetdude> jhonijim great... yeah troll
* Nicholas snuggles mynetdude
<troubled> cal: please dont mock ErrantEgo. he's right , that was uncalled for
<Nicholas> :3
<Pritchard> This channel defines it as inapprorpiate. By joining this channel, you agree to abide by our rules.
<Pritchard> It's as simple as that.
<Pritchard> If you want to debate freenode channel policy, you can take it up with the admins.
<neyo> exactly
<cal> troubled: sorry, but the internet _is_ serious business :)
<Pritchard> Now if you're going to argue that, you're a troll who's about to be ignored by everyone in this channel.
<troubled> cal: which is why im smilin ;)
<mynetdude> cal or unserious depends on where you go
<Cann0n> well, i mean, i've been yelled at for some stupid stuff in this every chat..
<Cann0n> spamming and using bad language is a no no, and if i get in trouble for it, i think the rules should apply to everyone. :P
<DrJ> I've been yelled at before for talking excessivally about bacon
<jhonijim> about baccon
<Kitsune> bacon!?
<cal> bacon ftw :D

Appendix 3: Examples for Intensity Scale of Humour Markers

Intensity 3
<+WhiteDove1> have to find a mic <+WhiteDove1> somewhere in hubby's mess <+WhiteDove1> he took mine <JRz> well tell him to leave your shit alone <+WhiteDove1> oh, that would go over soooo well. <@Jayde> they're married <@Jayde> its communal property <@Jayde> especially gadgets lol <+WhiteDove1> yeah, true <+BryanXXVIII> hopefully your toys dont start smelling funny [...] <+WhiteDove1> BryanXXVIII!!! Lol
<+triviabot> Question 9 (biology): Every human has one of these on their tummies. <+Julies85> bellybutton <@HHC> beer <+Julies85> belly button <+Julies85> :) <+triviabot> Hint: n**** <@^^^fIzI^^^> lol
Intensity 2
<@istok> easy to google them, but if you add them there is always a chance you did it wrong or the script conflicts or doesn't do what you want it to do. <easy> easy again?! <@DAHwoman> sorry all i gotta go need to eat finish laundry and get to work <Ghassan> take it easy <Ghassan> lol [...] <+undone> u're quite an easy target, easy. <easy> i think i have to change my nick lol

<snooken> I was talking to a customer (who is a girl: boobs, vagoo, etc.) where I work (has nothing to do with computers) but we got into the discussion of computers and I tell her I built mine.
[...]
<snooken> Girl: I built mine too.
[...]
<snooken> Me: Oh? Intel or AMD?
<snooken> Girl: IBM
<snooken> Me: Okay.
<snooken> Then I stopped talking to her.
<snooken> drinks out nose
<snooken> omg
[...]
<fincognito> snooken i thought that you were gonna say like "and then the girl said instead of a floppy i mounted a 12" dildo instead"
<fincognito> or atleast something almost funny
<fincognito> that was just shit
<Wilitus> :D
<fincognito> i give you an F for that one
<Wilitus> sounds like copy pasta fincognito
<Wilitus> i meannnnnn
<Wilitus> like, stfu
<fincognito> Wilitus you get an A+ for doin' nothing
<fincognito> :D

Intensity 1

<+kungpo> i got two packages coming i wonder which one will arrive firrst
[...]
<dasirene> I bet that the package from the farer place will arrive first :P
<DyslexicGhost> probably, lol

<BonesolTeraDyne> Reminds me of a certain image on FurAffinity, though I can't post it. NSFW
[...]
<Minataku> Post it anyway
<Minataku> :D
<dsmith_> lol

Intensity 0.5

<Porta> do you know what country they're from
<Depressed> rosja
<Porta> wat
<Depressed> Motherland Russia
<Porta> nice ^^

<@Black_Dahlia> am gonna renew mine in a few days...
<@Black_Dahlia> even if it's my last money, ever lol i will still have my flickr account haha
<+flipflop> Black_Dahlia: It's the way you show your photos, it makes sense for you.
<@Black_Dahlia> :D

Appendix 4: Examples for variety of emoticon :X

<@JackSchitt> the workshop looks like a slaughter house
<@JackSchitt> blood all over the touch screen and job documents :)
<@Nisha> ouchie :x

<`jules> so if I put strawberries and whipped cream on utah_chic, you'd eat her?
<utah_chic> hey now
<+Kameleon`> she's so tiny it'd be like a snack
<scouty> mmm
<+Kameleon`> :x

<@Black_Dahlia> Nisha, I want those earrings I showed ya last time hahaha
<+Nisha> sorry :x

<ooonY> damn onions, im farting like badass :x
<keit> eww

<+myHB|PhilMe> verdammt xD
<+myTB|Gina1987> DU NOOB :D
<moeni> das lied <3
<+myHB|PhilMe> NOBODY SAID IT WAS EASY!
<+myTB|Gina1987> nobody said it was easyyyyyyyyyy :X

* +romanu` I'm a gangster for life :x

<exordium|wizard> PUNBUSTER FTW.
<exordium|wizard> -.-
<exordium|wizard> PUNKBUSTER*
<Leissi> PUNBUSTER
<Leissi> ba-dum-tish
<nikomo> Why's quakelive.linux invite-only atm?
<Moggy> lol
<EvoldicA> nikomo, elite only?
<exordium|wizard> can help me punkbuster error...
<Leissi> go read bugreport forum first
<dnorm> rott was ugly compared to doom
<exordium|wizard> many pages :x

<Anon> "IdrA recalls a time when he had a heart. He now lives to ruin other people's dreams"
<ooonY> the only dream he ruin is his own of beeing a pro in korea :x
<Clowe> (ooonY): :X
<ooonY> there is just one word to say: Broodsports :x

<+deekay> what the fuck am i going to do with my life
<@Chill[GGBaby]> what the fuck am i doing with my life
<@Chill[GGBaby]> :X

Appendix 5: Example of a spam attack in an IRC channel

<+undone>
<+Ghassan> hmmmmmm :)
* @Toi out sleepin ..money money.. be rich or die like a fich lool
* +babyboo sighs
<+Ghassan> Toi ull die like a fish then :P
<classicl> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says ki
-classicl:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says ki
<usernamec> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte

-usernamec:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte
<leftwingz> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte say
-leftwingz:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says
<cashp> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says k
-cashp:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says ki
<enterprisep> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon

:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte sa
-enterprisep:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte say
<bravom> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says ki
-bravom:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says ki
<marijuanaz> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte sa
-marijuanaz:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon

:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte say
<founderu> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says
-founderu:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says
<lucusl> @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says ki
-lucusl:#facebook- @@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon :$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says kis khwetkon
:$@@@@@@@@@@@@@@@@@@@@ :D sorte says ki
-Ageless:#facebook- Channel has been locked due to flood, sorry for any inconvenience this may have caused.
<+Ghassan> we r already lag
<+Ghassan> :@
<+undone> :)
<+Ghassan> :)

Appendix 6: Analysed IRC Dialogue

This is the same dialoge as in appendix 1, but every occurrence of a paralinguistic feature discussed in this paper is marked with a corresponding colour:

```
<jhonijim>
FUCK!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
<jhonijim>
FUCK!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
<troubled> grrr
<jhonijim>
FUCK!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
<jhonijim>
FUCK!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
<jhonijim>
FUCK!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
<jhonijim> why are you looking at me
<Cann0n> why are you spamming?
*** GaintSura has quit IRC (Connection timed out_)
<Cann0n> !admin
<ErrantEgo> please quit flooding the channel
<troubled> jhonijim: please, just let it go
<ErrantEgo> please quit swearing
<jhonijim> whats so wrong about swearinh
<ErrantEgo> jhonijim, #defocus is a family friendly channel (meaning young
eyes could be watching), please watch your language while you are in
#defocus...thank you. :-)
<ErrantEgo> in addition, you were flooding
<jhonijim> i may have been flooding
<DrJ> you think?
<ErrantEgo> which is unnecessary
<jhonijim> but aleast i got someone to talk to me
<troubled> jhonijim: was uncalled for, and staff have better things to do
:)
<DrJ> *sigh*
*** hd_ has joined #defocus
<cal> f word flooding? o.O
<jhonijim> its just a word
<ErrantEgo> its an INAPPROPRIATE word.
<cal> OMFG SOMEONE IS SWEARING WE'RE GOING TO DIEEEEEEEEEEEEEEE!!!!!!!!!
<mynetdude> why do you guys bother feeding the troll anyway?
<jhonijim> who defins it as inapropriat
<Pritchard> Well, I don't think swearing's all that bad.  But that's
irrelevant here.
<mynetdude> cal shhhh really
* cal dies
<mynetdude> jhonijim great... yeah troll
* Nicholas snuggles mynetdude
<troubled> cal: please dont mock ErrantEgo. he's right , that was uncalled
for
<Nicholas> :3
<Pritchard> This channel defines it as inapprorpiate.  By joining this
channel, you agree to abide by our rules.
<Pritchard> It's as simple as that.
<Pritchard> If you want to debate freenode channel policy, you can take it
up with the admins.
<neyo> exactly
<cal> troubled: sorry, but the internet _is_ serious business :)
```

<Pritchard> Now if you're going to argue that, you're a troll who's about to be ignored by everyone in this channel.
<troubled> cal: which is why im smilin ;)
<mynetdude> cal or unserious depends on where you go
<Cann0n> well, i mean, i've been yelled at for some stupid stuff in this every chat..
<Cann0n> spamming and using bad language is a no no, and if i get in trouble for it, i think the rules should apply to everyone. :P
<DrJ> I've been yelled at before for talking excessivally about bacon
<jhonijim> about baccon
<Kitsune> bacon!?
<cal> bacon ftw :D

Paralinguistic features:

- emoticons (7)
- action command (2)
- action strip (1)
- reduplicating letters (4)
- shouting (3)
- emphasis (1)
- pauses (3)
- addressing directly (9)

(repeated messages (spam) were not counted)

Appendix 7: List of emoticons available in the message board www.justfindout.de


WAYNE
INTERESSIERTS
GUILTY!
give me all your cookies!!

Lebenslauf

Persönliche Daten

Name: Peter Waldner
Geburtsdatum: 11. Mai 1984
Geburtsort: Rum bei Innsbruck
Nationalität: Österreich
Familienstatus: verheiratet, keine Kinder
Adresse: Mariahilferstraße 202, Stiege 1, Top 9
1150 Wien
Österreich
Telefonnummer: +43 (0)680 2115284
e-mail: p.waldner@gmx.at

Ausbildung und Studium

Volksschule in Innsbruck, 1990 – 1994
Gymnasium in Innsbruck, 1994 – 2002,
Matura am 10. Juni 2002 bestanden.

Universität:
- 2002: Beginn des Bakkalaureatsstudiums Informatik an der Universität Innsbruck
- 2004: Zusätzlich Unterrichtsfach Englisch an der Universität Innsbruck (in Kombination mit UF Mathematik)

Seit 2006: Unterrichtsfach Englisch / Informatik an der Universität Wien.