PhD-FSTC-2016-7
The Faculty of Sciences, Technology and Communication

DISSERTATION

Defense hold on 24/02/2016 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG EN INFORMATIQUE

by

Sviatlana HÖHN
Born on 14 October 1978 in Vitebsk (Belarus)

Data-driven Repair Models for Text Chat with Language Learners

Dissertation defense committee

Dr. Christoph Schommer, dissertation supervisor
Professor, Université du Luxembourg

Dr. Stephan Busemann, Vice-Chairman
Professor, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Saarbrücken

Dr. Leon van der Torre, Chairman
Professor, Université du Luxembourg

Dr. Charles Max
Professor, Université du Luxembourg

Dr. Karola Pirsch
Professor, Universität Duisburg-Essen

Abstract

This research analyses participants' orientation to linguistic identities in chat and introduces data-driven computational models for communicative Intelligent Computer-Assisted Language Learning (communicative ICALL). Based on non-pedagogical chat conversations between native speakers and non-native speakers, computational models of the following types are presented: exposed and embedded corrections, explanations of unknown words following learner's request. Conversation Analysis helped to obtain patterns from a corpus of dyadic chat conversations in a longitudinal setting, bringing together German native speakers and advanced learners of German as a foreign language. More specifically, this work states a bottom-up, data-driven research design which takes "conversation" from its genuine personalised dyadic environment to a model of a conversational agent. It allows for an informal functional specification of such an agent to which a technical specification for two specific repair types is provided.

Starting with the open research objective to create a machine that behaves like a language expert in an informal conversation, this research shows that various forms of orientation to linguistic identities are on participants' disposal in chat. In addition it shows that dealing with computational complexity can be approached by a separation between local models of specific practices and a high-level regulatory mechanism to activate them. More specifically, this work shows that learners' repair initiations may be analysed as turn formats containing resources for signalling trouble and referencing trouble source. Based on this finding, this work shows how computational models for recognition of the repair initiations and trouble source extraction can be formalised and implemented in a chatbot. Further, this work makes clear which level of description of error corrections is required to satisfy computational needs, and how these descriptions may be transformed to patterns for various error correction formats and which technological requirements they imply. Finally, this research shows which factors in interaction influence the decision to correct and how the creation of a high-level decision model for error correction in a Conversation-for-Learning can be approached.

In sum, this research enriches the landscape of various communication setups between language learners and communicative ICALL systems explicitly covering Conversations-for-Learning. It strengthens multidisciplinary connections by showing how the multidisciplinary research field of ICALL benefits from including Conversation Analysis into the research paradigm. It highlights the impact of the micro-analytic understanding of actions accomplished by utterances in talk within a specific speech exchange system on computational modelling on the example of chat with language learners.

Contents

List of Abbreviations

CA Conversation Analysis

SLA Second Language Acquisition, interactionists SLA theory

L1 First Language or Native Language

L2 Second Language or Foreign Language (for me here the same)

DaF Deutsch als Fremdsprache (German as a foreign language)

ESL English as a Second Language

NS Native Speaker

NNS Non-Native Speaker

CMC Computer Mediated Communication (includes synchronous for instance, chat, instant messaging, and asynchronous, such as email)

SCMC Synchronous Computer-Mediated Communication (such as text chat, video call, instant messaging)

CALL Computer-Assisted Language Learning

ICALL Intelligent Computer-Assisted Language Learning

RI Repair Initiation

ROI Repair Other-Initiation

RCO Repair Carry-out

OISR Other-initiated Self-repair

NLP Natural Language Processing

HLT Human Language Technology

AI Artificial Intelligence

HCI Human-Computer Interaction

HRI Human-Robot Interaction

Part I.

Background

1. Setting the Scene

The ability to speak foreign languages became one of the key competences in the globalised world. People learn foreign languages for professional purposes, with the goal to study abroad or for private reasons. Distance language learning and training based on tutoring videos or online exercise books offer a convenient way to integrate language classes into a busy day. Practicing conversation with an artificial agent is seen as a good alternative if a native speaker is not available or not affordable. Computer-Assisted Language Learning (CALL) was expected to facilitate learning and teaching by providing electronic workbooks with automatic evaluation and vocabulary training. CALL technology extended with Natural Language Processing (NLP) techniques became a new research field called Intelligent Computer-Assisted Language Learning (ICALL). Language technology has been integrated into CALL application with the purpose of automatised exercise generation (Ai et al., 2015), complex error analysis and automated feedback generation (Amaral, 2011).

Education industry and language learners benefit from various deployed applications, for instance (Sagae et al., 2011; von Ahn and Hacker, 2012; Heine et al., 2007). A number of mobile applications in the AppStore and GooglePlay Store target conversation training and traditional task-based language instruction. Frequently cited real-life ICALL applications are E-Tutor for German learners (Heift, 2002, 2003), Robo-Sensei for Japanese learners (Nagata, 2009) and TAGARELLA to learn Portuguese (Amaral et al., 2011). Conversational agents, chatbots and dialogue systems for foreign language training have been developed as stand-alone conversation partners (Jia, 2009; Zakos and Capper, 2008) and as part of intelligent tutoring systems (ITS) (Petersen, 2010), serious games (Gray, 1992; Sagae et al., 2011; Wik et al., 2007; Amoia et al., 2012) and micro-worlds (DeSmedt, 1995). A new trend in technology-aided language teaching is robot-assisted language learning (RALL) (Han, 2012). Robotic language teaching assistants have been studied in a traditional language classroom (Chang et al., 2010; Kwok, 2015; Mubin et al., 2013) and for teaching autistic children to speak foreign languages (Alemi et al., 2015).

Petersen makes a distinction between a communicative ICALL and a non-communicative ICALL and sees the communicative ICALL as an extension of human-computer interaction field. His understanding of communicative ICALL is that "communicative ICALL employs methods and techniques similar to those used in HCI research, but focuses on interaction in an L2 context" (Petersen, 2010, p. 25). As a consequence, corrective feedback in communicative ICALL should incorporate corrections of L2 errors into dialogues with the user. Hence, an ability to provide corrections of linguistic errors in conversation is seen as one of the key features of such a dialogue system or a conversational agent.

Consequently, the majority of ICALL applications see the artificial agent in the role of a tutor or a teacher who have the right to provide corrections. While students and teachers in language classrooms have been taken as "role models" for user and agent models in communicative ICALL applications, there is a gap in computational simulation of types of interaction other than language classroom. Conversational agents in the role of an equal but more knowledgeable peer are under-represented in communicative ICALL research. This dissertation seeks to close this gap and to show that new data-driven user and agent models may be helpful for designing conversational agents, which act in other roles than a teacher or a tutor. To approach this objective, Conversation Analysis is included into the multidisciplinary research paradigm of ICALL.

1.1. Multidisciplinary nature of ICALL

ICALL research labs usually employ specialists bringing expertise in Software Engineering, Natural Language Processing, Pedagogy, Computer Science and CALL (Greene et al., 2004). The multidisciplinary nature of CALL research is recognised and claimed in CALL literature. (Greene et al., 2004, Sec. 3) list four research principles of the ICALL lab of the Dublin City University: "(i) reuse of existing NLP resources, (ii) reuse of existing CALL research experience, (iii) user-centred design and evaluation and (iv) interdisciplinarity". CALL works on the intersection of Computer-Mediated Communication (CMC) and Second Language Acquisition (SLA) research, which in turn are also multidisciplinary fields of work employing Computer Science, Linguistics, Psychology, Education and Sociology among others. Thus, interdisciplinary or multidisciplinary perspective and involvement of latest results in related disciplines is a must in building ICALL applications of high quality. Nonetheless, Schulze (2008) confirms (Oxford, 1993)' critique of ICALL research based on "homespun notions of language learning or notions borrowed from discourses in SLA which had long been criticised severely and/or superseded by theoretical approaches with improved explanatory power" (Schulze, 2008, p. 513). Schulze (2008) also emphasises the importance of multidisciplinary cooperation for ICALL research:

> Of course, ICALL cannot afford to ignore current discourses in SLA [...] but also needs to consider issues related to the computational implementation of SLA theories and approaches in language pedagogy. This makes it more important for ICALL researchers to foster links with researchers in SLA and NLP, and such links can only be established after fruitful discussions in which researchers from both communities participate (p. 513).

However, for realistic and individualised user models and dialogue design for communicative ICALL applications, relying barely on SLA-results is not sufficient. Markee criticises the rationalistic tradition to approach SLA studies because "an experimental, quantitatively oriented methodology inevitably loses important details of individual behavior" (Markee, 2000, p. 29). He suggest to develop an alternative, emic perspective to SLA studies and "a critical attitude toward quantified data" (Markee, 2000, p. 29).

Effects on learning with an ICALL system and learning when using text, voice and video based communication for language learning have been studied earlier from the educational perspective. Second Language Acquisition theory (SLA), mainly following the Interaction Hypothesis usually credited to Long (Long, 1996) but very similar to Krashen (1981)'s Input Hypothesis, emphasises the role of specific dialogue routines in language learning. ICALL research from its perspective is mainly focused on opportunities (affordances) provided by the technology to elicit such dialogue routines in interaction with learners. Such routines are for instance corrective feedback (Lyster et al., 2013) and meaning negotiation (Varonis and Gass, 1985). Corrective feedback and meaning negotiations are seen as important conditions for learning. Imitation as a language acquisition strategy formulated as the Imitation Hypothesis (Aguado Padilla, 2002) received less uptake in the ICALL community.

Corrective feedback is acknowledged as an important tool in language instruction to help the learner to notice the produced deviations and to give the learner a chance to improve. Classifications of corrective feedback in language classroom have been obtained from classroom data (Panova and Lyster, 2002; Lyster et al., 2013) and compared to corrective feedback in computer-mediated communication (CMC) in various language learning or practicing scenarios, see for instance (Zourou, 2012; Sauro, 2009). However, not every error is corrected even in a language classroom, specifically, if the focus is on fluency. Furthermore, in non-educational situations, corrective feedback is dispreferred, and therefore, rare. This observation has been taken up by the designers of conversational systems for SLA, see for instance (Wik and Hjalmarsson, 2009) where the authors make a difference between two embodied conversational agents (ECA). One of them plays a role of a virtual language teacher, and therefore, provides corrective feedback on pronunciation and language use. The other ECA is a role-play conversational training system with the goal to maintain an interesting conversation. Here, the agent "has the role of a native speaker, for example, a person with a service occupation, whom you need to communicate with using your new language" (Wik and Hjalmarsson, 2009, p. 1039) and therefore, the user cannot expect any corrective feedback from the ECA.

The subject of research in Conversation Analysis (CA) is, in its originals, naturally occurring interaction. However, because more and more people use a foreign language for their everyday interaction in business, educational, institutional and leisure contexts, foreign language talk came also under the lens of CA. For instance, CA analyses how participants construct together identities of language experts and language novices in interaction and whether there are differences between conversations with learners and between learners, and native-speaker-only talk. CA-driven research showed, for instance, that language learners are able to accomplish all complex social actions even with limited linguistic resources (Markee, 2000). There are also attempts to identify the process of learning by CA methods, however, not very successful till now. Kasper (2004) proposed to perform *longitudinal studies* to approach this problem. Markee (2008) introduced a learning behaviour tracking system which is not easy to implement in practice. González-Lloret (2011) proposes to change the definition of learning itself to make use of CA as a tool for research on learning:

> As for the methodological feasibility of CA to demonstrate learning, expanding the definition of learning may be necessary, [...] so that SLA is not limited only to linguistic features but also includes the social context and sequential development of interactions. In this sense, learning is understood as participation based, focusing on the improvement of the interactional resources used by learners for talk-in-interaction rather than just on their linguistic skills (González-Lloret, 2011, pp. 317-318) citing (Markee, 2008).

Studies applying methods of Conversation Analysis in Second Language Acquisition (CA-for-SLA) research show that language learners interacting with native speakers in non-educational settings engage in various types of talk focusing on linguistic matters, including error corrections provided by native speakers (Markee, 2000; Hosoda, 2006). Such types of talk are referred to as *repair* in CA. Linguistic repair is a dispreferred social action both in native speaker to native speaker (NS-NS) and native speaker to non-native speaker (NS-NNS) talk (Schegloff, 2007). However, the participants analysed in this work reported that they prefer to be corrected more often in order to learn (Danilava et al., 2013a). Hosoda (2006) points out that repair sequences distinguish interaction with language learners from NS-only talk in the following way. Repair with linguistic trouble source is found in conversations with non-native speakers more frequently and is typical even for informal conversations between learners and their native speaker peers. Multiple occurrences of linguistic repair in chat dialogues produced by NS without being instructed to correct indicate that the context of a Conversation-for-Learning seems to modify social preferences providing opportunities for repair and minimising face threatening (Tudini, 2010, p.64).

Kasper (2004) argues that there is a separate type of informal talk *Conversation-for-Learning* where conversation parties come together *because* one of them has a higher level of language proficiency than the other. The term Conversation-for-Learning reflects more precisely the application scenario where users engage in dialogues with an artificial agent in order to practice their language skills in an informal conversation. The learners would use such a system *because* they want to practice. ICALL developers need to be careful in dialogue and task modelling, because CA showed that different types of speech exchange systems work differently, and insights from research on institutional talk, such as teacher-fronted language classroom, are not necessarily valid for informal conversation or a Conversation-for-Learning (Schegloff, 1993; Markee, 2000).

Chat-Conversations-for-Learning between an artificial conversational agent and a language learner can be investigated from different perspectives and incorporates first of all technological and socio-linguistic but also learning aspects. From the technological perspective, both research results in Computer-Mediated Communication (CMC) between learners and native speakers and in Human-Computer Interaction (HCI) between learners and artificial agents may be relevant. CMC studies how communication mediated by a technology influences the interaction. In the SLA context, these studies demonstrate how technology provides opportunities for learning. SLA theory-driven studies showed that chat communication positively influences oral performance in terms of language quantity (Abrams, 2003). A small number of CA-informed studies of NS-NNS chat communication provide qualitative and quantitative analysis of different phenomena in chat interaction, such as error corrections and explanations of unknown words in Italian (Tudini, 2003, 2010) and German (Marques-Schäfer, 2013). Another study shows how Spanish learners develop the competence to interact in trouble-talk (González-Lloret, 2011). In particular, the study by González-Lloret (2011) showed that

> CA can be an appropriate tool for the study of SCMC [Synchronous Computer-Mediated Communication], depending on the focus of the study. CA is better suited for discovering patterns of how the participants carry out the interaction and how they orient to the sequences developed while they construct authentic conversation, and CMC produces large quantities of authentic materials, being one of the fastest growing communicative media in the world. (pp. 317-318)

For these reasons, expertise in Conversation Analysis, and specifically CA-for-SLA, may be an additional advantage for creating communicative ICALL applications. Looking at conversations with learners through the lens of CA-for-SLA could help to identify characteristic features of dialogues between language learners and native speakers in order to understand, to what extent the native speaker may be a model for the artificial agent and which of the key features are implementable.

1.2. Participants orientation to linguistic identities

An idealised concept of a *native speaker* was long time accepted in applied linguistics as the object for studies of communicative competence (Canale and Swain, 1980). The native speaker was put in the role of a language expert, and the non-native speakers were automatically classified as the complementary category of *language novices*. However, as Hosoda (2006) criticises and discusses in detail, these two absolute categories were "poorly supported by the sociolinguistic evidence" because "native speaker status does not necessarily correlate with a high command of the language" (Hosoda, 2006, p. 25). For a critical discussion see also (Markee, 2000; Kasper, 2006; Vickers, 2010; Wagner and Gardner, 2004).

The concept of *expertise* is seen as better reflecting the empirical research results on interaction, specifically in the context of second language communication, because the participants of an interaction may change their roles of expert or novice during the interaction. Hosoda (2006) argues that the notion of the *differential language expertise* suggested by Kasper (2004) is attractive for CA studies because

> in CA, language expertise, like any other social category or attribute, is not primarily subject to an outside observer's judgment. Instead, analysts are licensed to invoke descriptions pertaining to the participants only when the parties orient to such matters through their talk and other interactional conduct. [...] Consequently, whether or not language expertise is relevant at any point

> in the interaction is determined by the participants themselves through their observable orientation to linguistic matters" (Hosoda, 2006, p. 26).

In the same line, Brandt analyses how participants make differential language expertise relevant in NNS-NNS conversations in which no NS-NNS relationship exists (Brandt, 2011, Sec. 2.3.3). Kurhila (2006) uses the term *the knowledgeable participant* instead of *expert* and the term *non-expert* instead of *novice* allowing less extreme formulations of the differences in linguistic knowledge from both sides.

For all these reasons, I will use the terms *native speaker of language X* to refer to persons who learned a particular language X as a native language, and *language expert* to refer to participants of a conversation when they make their differential language expertise relevant in conversation. Thus, the terms *language expert* and *native speaker* are not used as synonyms in this work. As a complementary category for the "non-experts", I will use the term *non-native speaker of a language X* or *learner of a language X* to refer to persons who learn a language X as a foreign language, and the term *language novice* to refer to participants of a conversation when they orient in conversation to their linguistic identity as not-yet-fully-proficient speakers of a particular language.

Example 1.1. NS corrects an error and provides wrong metalinguistic information ("denken" is in fact not reflexive). Orig. excerpt 61 from (Marques-Schäfer, 2013, p. 185), line numbers and translation added. M learner, G tutor, native speaker of German.

1	11:39:25	M	er denkt nur an selbst *he thinks only on [*error: missing reflexive pronoun] self*
2	11:39:49	M	verstehst du mich? *do you understand me*
3	11:40:07	G	Der Chef? *The boss?*
4	11:40:39	M	du kannst mich korrigieren wenn ich fehler schreibe *you can correct me when I write [* error: lexical] mistakes*
5	11:40:53	M	ja der chef Bush *yes the boss Bush*
6	11:41:26	G	Okay, denn es heißt richtig: Er denkt nur an sich selbst. *okay, then it is called correctly: He thinks only on himself.*
7	11:41:57	G	Denken ist ein reflexives Verb. *to think is a reflexive verb*
8	11:41:57	M	seine politik macht wennig sinn *his policy makes little [* error: orthography] sense*
9	11:41:57	M	danke *thanks*

Although CA-for-SLA researchers clearly define language experts, it remains difficult to assign the status of a language expert to one of the participants in local sequences of talk where the differential language expertise is made relevant in conversation. We still need to keep in mind that the participants who position themselves as language experts in conversation may make mistakes even in the turns of talk where this positioning is performed. Example 1.1 shows a sequence from NS-NNS chat interaction that took place as part of the project *JETZT Deutsch lernen* where learners of German communicated in chat with other learners and tutors (Marques-Schäfer, 2013). The excerpt in Example 1.1 is taken form a tutored session, and G is a tutor and a native speaker of German. She is explicitly asked by M, the learner and non-native speaker to correct errors in turn 4. G corrects in turn 6 an error that M made in turn 1 providing metalinguistic information to the correction. However, the metalinguistic information contains an error: the verb *denken* (Engl. to think) is in fact not reflexive. The learner accepts the correction and the explanation in turn 9.

In Example 1.1, the native speaker G who has in addition the tutor role, is selected by the learner M as a language expert by G's request to correct errors. M accepts this role by doing correction, however, M also produces an error. There is a difference between "doing being language expert" (or "doing being a more knowledgeable participant") and "being language expert" (or "being a more knowledgeable participant"). These discrepancies may be described by the terms *epistemic stance* (expectations towards own and others' knowledge) *epistemic status* (the factual state of the knowledge) Heritage (2012). The participants who position themselves as experts when they *think* they are more knowledgeable than their partners (they do not pretend to be experts while knowing that they are wrong), however, it does not mean, that they objectively do not produce any errors. Moreover, even if they produce errors, they may still be more knowledgeable than their partners, which is already stated by the term *differential language expertise*. In Example 1.1, M appears to be more knowledgeable than G in constructing a correct German sentence, even if the theoretical explanation of the rule for this construction is wrong.

This is an extremely important difference for the remainder of the work. A conversational agent can possess as much linguistic knowledge as we can prepare for it using all available linguistic resources, but all this knowledge is useless as long as there is no need to make it relevant in conversations with the user. Even if someone positions himself or herself as a language expert in conversation, it does not mean, that he or she is no longer allowed to make mistakes. Not the actual language expertise of the conversational agent is in focus of the work, but the *practices of orientation to differential language expertise in conversation* are under the loupe.

1.3. Research objectives

This dissertation aims to contribute to the field of communicative ICALL by bringing Conversation Analysis as a research methodology into ICALL application design. Specifically, it seeks to addresss the following two research objectives:

I Find and describe interactional practices in native/non-native speaker chat-based

Conversation-for-Learning where chat participants orient to their linguistic identities of language experts and language novices.

II Create computational models of those practices and analyse technical requirements and limitations to implement the resulting models in a communicative ICALL application.

Looking ahead to the results of the first part of this research, I disclose the specific research questions which were formulated after the initial research phase of "unmotivated looking" at the data:

RQ1 Which interactional resources do language learners use in a chat-based Conversation-for-Learning with native speakers to initiate repair in order to deal with troubles in comprehension and how do native speakers deal with these repair initiations?

RQ2 How can other-initiated self-repair when the machine is the trouble-speaker be handled in a chat-based communicative ICALL system?

RQ3 Which types of other-corrections of linguistic errors exist in the dataset representing a chat-based Conversation-for-Learning?

RQ4 How can these types of other-corrections of linguistic errors be modelled in order to be implemented in a chat-based communicative ICALL system?

RQ5 Apart from the occurrence of an error, are there other factors which are relevant for the occurrence of a correction of a linguistic error in native/non-native speaker chat-based Conversation-for-Learning?

RQ6 If such factors exist, how can they be modelled in order to be implemented in a chat-based communicative ICALL system?

The restriction to chat-based communication in the research questions is necessary because of the specific speech exchange system and a specific communication medium, namely medially written dyadic Conversation-for-Learning. Based on (Hosoda, 2006)'s observation, it can be expected that there might be other interactional phenomena in native/non-native speaker chat, which distinguish them form native/native speaker chat. Specifically, it may be expected to find other types of sequences than error corrections in chat where participants orient to their linguistic identities in talk.

With regard to other-corrections of linguistic errors, I do not expect to find big numbers of occurrences of those in the data because other-corrections of linguistic errors are dispreferred, even in native/non-native speaker conversations. However, from the point of view of qualitative language research, I am interested in finding what is *typical*, but not necessarily frequent (typical sequences, typical structure of such sequences). Conversation is a cultural action, therefore, chat participants formulate their turns in a way that the action is recognisable for the partner. In the same time, they formulate their turns individually in order to be different than others. For this reason, plain prototypes of particular actions can be seen in natural data very rarely, and the variants of that prototypes can be found more often. Even for a small number of occurrences of a particular type of a sequence, it is possible to analyse, how a *typical* sequence of this kind is organised. Therefore, a generalisation based on a small number of instances is common in CA and is a valid approach to modelling.

1.4. Methodology

The methodological novelty of this research consists mainly of bringing Conversation Analysis into the ICALL research paradigm with the purpose to discover new models for the expert and learner behaviour in a communicative ICALL application. Specifically, I use methods of Conversation Analysis (Heritage and Goodwin, 1990; ten Have, 2007; Liddicoat, 2011; Schegloff, 2007) to identify typical structures in close-to-natural longitudinal chat interactions between native speakers of German and advanced learners of German as a foreign language. The data-driven research was performed in three phases. The first phase of "unmotivated looking" in order to find typical structures was followed by the phase of micro-analysis of specific structures to obtain patterns. In the final, third phase computational modelling of these structures was performed based on identified patterns. The implementability of the new models is validated by an implementation case study and an analysis of the required language technology and knowledge bases.

I provide here a short description of the dataset. A comprehensive documentation can be found in Appendix A. The participants of the data collection were 9 advanced learners of German as a foreign language and 4 German native speakers. All of the learners were Russian native speakers and students of German at a Belorussian university. Native speakers are friends and colleagues of the researcher. Each native speaker was assigned to two or three non-native speakers in pairs according to the time slots that the participants specified (when they had time to chat). Only the first appointment was arranged by the researcher. The only instruction was "just chat". The participants were expected to have a free conversation and to talk about whatever they want. The goal communicated to the participants was to produce 8 dialogues in total and to interact 4 weeks long 2 times per week, every time approximately 30 minutes.

The participants interacted using Google Talk infrastructure. A forwarding chatbot hosted on Google App Engine was used to collect the data instantly. Participants did not see each other directly, they sent the messages to the bot and the bot instantly forwarded the messages to the partner. All the chat logs were available to the researcher immediately. The participants were informed that their talks were recorded in the beginning of the data collection. The participants agreed to publish the produced chat data prior to starting the interaction.

The participants produced 73 dialogues in total (8 dialogues by each of 6 pairs, 9 dialogues by each of 2 other pairs, and 7 dialogues by the 9th pair). Besides that the participants sometimes missed each other oder forgot appointments. In these cases, the participants sent each other notifications and apologies, but in some cases several days passed between turns. The decision was made to include only full dialogue sessions in the final data set where participants met and talked. Each dialogue is between 20 and 45 minutes duration. The total size of the final dataset is 4548 messages that correspond to 236 302 text symbols. The message length ranges from 1 to 774 text symbols and of an average length of 58,5 symbols over all pairs. Table 1.1 summarises the statistics.

Metrics	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Pair 7	Pair 8	Pair 9
MaxTL	335	405	774	313	414	277	637	460	232
MinTL	1	1	2	1	2	2	2	2	1
AvTL	62,98	72,40	105,13	38,13	86,80	38,99	42,23	48,85	31,38
# Turns	365	410	346	650	218	421	730	694	714
# Symb.	22989	29683	36374	24784	18923	16413	30825	33903	22408

Table 1.1.: Corpus statistics: MaxTL - maximum turn length (# symbols), MinTL - minimum turn length (# symbols), AvTL - average turn length (# symbols) followed by the total number of turns and total number of symbols for each pair.

I will provide many examples of specific pieces of dialogues to illustrate the identified phenomena. All examples are formatted so that the reader can easily grasp the intention of the example. Each turn (message) contains a message number, a time stamp, a speaker code and a message body. Line breaks added by the speakers were kept as in the original message and marked by a line break tag. All other line breaks in the examples are caused by the page width and were not added explicitly to the formatting. Each message body contains the original wording and spelling of the message and an English translation, if needed. Messages consisting of only an emoticon or an "OK" were not translated. A word-by-word transcription is added between the original message and the translation if it was required for the understanding of the discussed phenomena. Errors are annotated in the examples only where it was necessary for the analysis. An inline error annotation containing the error desciption was performed either in the transcription line or in the translation line and is inserted in square brackets [] and marked with an asterisk *.

1.5. Contribution of the work

This research shows that Conversation Analysis as a research methodology can be effectively used for the purpose of computational modelling of dialogue. Specifically, this work shows that the multidisciplinary field of communicative ICALL may greatly benefit from including Conversation Analysis into the highly multidisciplinary field of ICALL. As a consequence, this research makes several contributions to the related research disciplines, such as Conversation Analysis, Second Language Acquisition, Computer-mediated Communication and Artificial Intelligence.

The study contributes to research on identities and membership categorisation, advances the state-of-the-art on learner and expert models, roles and personalities in dialogue research, conversational agents and communicative ICALL. The identified sub-dialogues prepare an empirically grounded basis for an informal functional specification for conversational agents in communicative ICALL in roles other than tutors or teachers. The description of what the user and the agent can do outside of these sub-dialogues remains by the intuitive concept of a free conversation.

Two types of identified sub-dialogues which are found important for SLA were selected for a deeper analysis and detailed computational modelling. These two types are:

1. Other-initiated self-repair with linguistic trouble source where the learner is the recipient of the trouble talk and
2. Other-correction of linguistic errors.

A data-driven classification of repair initiation formats has been proposed, which is distinct form existing classifications described in the academic literature in CA (Dingemanse et al., 2014) and SLA (Varonis and Gass, 1985). This model covers previously under-researched scenarios in HCI where the user may not completely comprehend machine's talk because of user's limited linguistic knowledge. Repair carry-outs were analysed from the perspective of interactional resources deployed by participants in interaction and Natural Language Technology as well as linguistic resources required to reproduce them in a communicative ICALL application. This part of the study advances the state-of-the-art in CA-for-SLA and AI and their intersections, in particular dialogue modelling, language understanding and generation by models for recognition of repair initiations, trouble source extraction and explanation generation.

In contrast to SLA-driven classifications of corrective feedback obtained from classroom data, which are usually used in communicative ICALL applications, this dissertation introduces a classification of correction formats obtained from chat Conversations-for-Learning. The advantage of the new classification is that pragmatic function of different linguistic devices for corrections are taken into account for modelling, such as highlighting, repetitions, replacements, accountings and backlinks. Moreover, sequential environments for embedded corrections were analysed. Specific data-driven models of exposed and embedded corrections provide the required basis to offer the user a variety of correction formats as part of the ongoing conversation. In addition, a first step towards formulating a feature-based decision model for corrections in conversation is made. The study of other-corrections of linguistic errors contributes to research on corrective feedback (SLA, CA-for-SLA) and learner language research (error and dialogue moves annotation), advances state-of-the-art in communicative and non-communicative ICALL by separation of error recognition from local models of correction and decisions to correct.

Overall, this dissertation is a further step toward mutually beneficial multidisciplinary collaboration between Conversation Analysis and communicative ICALL as well as Artificial Intelligence as a larger research field.

2. Learning from Other's Experience

A large number of various CALL and ICALL projects have been described in different review articles. L'Haire (2011) provides a list of 152 CALL and ICALL systems which appeared from 1970 to 2011. This list includes 3 authoring software tools, 42 educational software tools, 12 micro-worlds, 36 intelligent systems, and 22 writing assistance tools. Heift and Schulze (2007) found 119 ICALL applications which appeared between 1978 and 2004 and were documented in German and English.

Schulze (2008) criticises the situation with academic publications on ICALL, specifically mentioning that there is a small number of journals dedicated to CALL, and a large number of publications, but a very small number of authors. In addition, many CALL publications appear in a variety of other academic publishing resources, for instance in Computational Linguistics. Schulze, citing Zock (1996, p. 1002) orients to the "communication problem and a mutual lack of interest concerning the work done in the neighboring disciplines" (Schulze, 2008, p. 511) in his critical review of ICALL literature. Addressing the multidisciplinary nature of ICALL in Section 1.1 I claimed that communicative ICALL will benefit from including Conversation Analysis into the circle of related disciplines. I see differences in language (concepts and terminology) used by different communities as one of the obstacles in multidisciplinary collaboration. Therefore, this inclusion will imply conceptual work on terminology and (re-)definitions of commonly used notions borrowed from SLA, NLP, HCI, AI and CA for ICALL.

Building on Oxford's key desiderata for ICALL (Oxford, 1993), Schulze (2008) reflects on how each of them was met in the 25-30 years before he wrote his article (Schulze, 2008, p. 512). The key desiderata discussed by Schulze are:

1. Communicative competence must be the cornerstone of ICALL.
2. ICALL must provide appropriate language assistance tailored to meet student needs.
3. ICALL must offer rich, authentic language input.
4. The ICALL student model must be based in part on a variety of learning styles.
5. ICALL material is most easily learned through associations, which are facilitated by interesting and relevant themes and meaningful language tasks.
6. ICALL tasks must involve interactions of many kinds, and these interactions need not be just student-tutor interactions.

7. ICALL must provide useful, appropriate error correction suited to the student's changing needs.
8. ICALL must involve all relevant language skills and must use each skill to support all other skills.
9. ICALL must teach students to become increasingly self-directed and self-confident language learners through explicit training in the use of learning strategies. (Oxford, 1993, p. 174)

The sociolinguistic notion of communicative competence credited to Hymes (1972) converges with the concept of interactional competence in Conversation Analysis (Markee, 2000, p. 52).

Communicative ICALL is first of all focused on development of the communicative competence, interpreting the remaining 8 points of Oxford's desiderata as a condition for its development. While points 2-6 and 8-9 reflect concerns of all ICALL applications, communicative ICALL handled the requirement to provide error corrections in two ways: either simulating a language teacher in a language classroom or taking free native speaker conversations as a model where corrections of linguistic errors are absolutely dispreferred, and therefore, not corrected by the communicative ICALL system. I will discuss below in detail existing communicative ICALL systems in Section 2.1. Specifically, I will address issues in learner language understanding, user modelling, learner corpus research and error corrections.

To start the announced work on a common terminology required for collaboration, I will explain the basic important concepts and the relevant research results from CA and CA-for-SLA in Section 2.2, and discuss how they may improve the state-of-the-art in Communicative ICALL research. There are many successful attempts to integrate CA into Human-Computer Interaction (HCI) and dialogue modelling obtaining models for human-robot communication and dialogue systems from naturally occurring interaction data and experiments with robots in the wild. I will discuss related research projects in Section 2.3.

2.1. Communicative ICALL

Effects of ICALL systems on learning depend on different properties of the system itself. Is it communicative or just an electronic workbook? If it is communicative, which technology exactly is used as a communication medium? I focus here on communication modalities chosen for Communicative ICALL in earlier academic publications in Section 2.1.1, I discuss the ways chosen by researchers to approach user and activity modelling in Section 2.1.2, and to deal with issues in learner language understanding in Section 2.1.6.

Communicative ICALL research showed that learners benefit in a similar way from corrective feedback provided by a human tutor as compared to an artificial agent (Petersen, 2010). Because error correction has a prominent position in language acquisition research, it received an important role in ICALL (communicative or not). I provide a review on recent achievements in automatic error recognition and automatic feedback generation in Section 2.1.4 and 2.1.5.

2.1.1. Communication modalities

Every communication medium provides affordances and sets constraints for communication in terms of interactional resources that can be made available by participants. In addition, every way of communication with artificial agents implies technological limitations and provides opportunities for implementation. Conversational systems of different complexity and communication modalities (text, voice, video) have been employed for helping learners to practice conversation.

Even simple chatbots not explicitly designed for communication with learners have been tested as conversation practice helpers with a concluding recommendation to use them for advanced or keen learners (Fryer and Carpenter, 2006; Jia, 2004). To make chatbots more useful for a broader learner audience, additional functionality like spelling error correction (Jia, 2009) and knowledge of the learner's native language in order to facilitate the communication for beginners. For instance a chatbot presented in (Zakos and Capper, 2008) targets beginner learners of English with Greek as mother tongue. Avatars, talking heads and embodied agents as well as integration of text-to-speech engines became nice-to-have extensions for chatbots because they appeared to positively influence user's engagement in chat, though a simple chatbot was still hiding behind them (Zakos and Capper, 2008; Stewart and File, 2007; Matt Chatbot, English; Ariel Chatbot, Spanish; Diego Chatbot, Spanish). Selection of an utterance from a set of possible utterances is offered by Let's Chat conversation training system (Stewart and File, 2007), this kind of interactional resources is normally not available in human-human communication.

Speech recognition techniques became more mature in the last years, so that they can be successfully used for conversations with leaners, specifically for pronunciation and prosody training, see for instance (Wik and Hjalmarsson, 2009; Ai and Xu, 2015; Bonneau and Colotte, 2011). Multimodal interaction using humanoid robots in language classes in the role of a teaching assistant was investigated in (Chang et al., 2010; Kwok, 2015; Mubin et al., 2013; Alemi et al., 2015). Robots positively influence engagement and learning, however, modelling and evaluation of human-robot interaction increases the complexity.

Every communication modality and communication medium provides its own constraints and affordances in terms of interactional resources, and therefore, has influence on competences that can be improved when the communication is restricted to a specific modality (Darhower, 2008). In particular, pronunciation cannot be improved when the communication modality is limited to only text chat. However, text chat has been found helpful for language learners to improve language accuracy and support their vocabulary acquisition (Kost, 2004; Kim, 2003), moreover, text chat may be helpful in acquiring complex language structures, improving fluency and oral performance (Fredriksson, 2013; Abrams, 2003). This may be explained by the fact that chat interaction is medially written but conceptually oral (Koch, 1994; Beißwenger, 2002).

Chat interaction has advantages for language learners because the learners can re-read the chat history, they have more time for production and comprehension, and they even can use other tools to deal with troubles in production or comprehension, which is not possible in oral conversation. In addition, text chat helps to avoid implementation issues related to speech recognition in learner language addressed to in (Ivanov et al., 2015). Therefore, I will focus in my work on text-based chat dialogues between advanced learners of German as a foreign language (Deutsch als Fremdsprache, DaF) and native speakers of German in the data analysis phase. Based on the data analysis results, I will then focus on computational models of specific structures in text-based conversations where language learners are supposed to communicate with an artificial conversation partner.

2.1.2. User and interaction models

In the domain of dialogue systems, to which Communicative ICALL belongs, a user model is "a knowledge source [...] which contains explicit assumptions on all aspects of the user that may be relevant for the dialog behavior of the system" (Wahlster and Kobsa, 1986, p. 3). Such aspects include user's beliefs, goals and plans. The attempt to grasp individual characteristics of a particular user was approached in form of assigning specific *conversational roles* to the users and the dialogue system, like for instance a hotel guest and a hotel manager, a library visitor and a librarian or a beginner UNIX user and a UNIX consultant (see (Wahlster and Kobsa, 1986) for references to specific projects on each of them).

Many academic publications are concerned with user modelling in ICALL (Chrysafiadi and Virvou, 2013; Read and Bárcena, 2014; Martinez, 2013; Heift, 2015). Schulze (2008) and Vandewaetere and Clarebout (2014) emphasise the importance of *student* models and *expert* models in ICALL systems. This reflects the common understanding in ICALL that the role of a student has to be assigned to the user and the system has to be put in the role of a tutor (Yang and Zapata-Rivera, 2010; Gardent et al., 2013). This role dichotomy determines systems behaviour, the system is expected to mimic a teacher. This makes the interaction with an ICALL system similar to a language classroom. Other roles of ICALL systems have been rarely considered, but there are a few attempts to escape from the tutor-learner dichotomy. For instance, Greene et al. (2004) mention a artificial German co-learner. Different roles are involved in a role-play application for culture and language training of Arabic where the user is involved in a simulation of talks with local speakers in a village (Sagae et al., 2011).

Amaral and Meurers (2011) describe the student model used in a real-life ICALL application. The model includes personal information, interaction preferences and knowledge of linguistic forms. Amaral and Meurers (2007) see the need to include a set of competences into user model, which they call *strategic competences*:

> the student model needs to be extended to include the learner's abilities to use language in context for specific goals, such as scanning a text for specific information, describing situations, or using appropriate vocabulary to make requests (p. 340).

This claim is supported by findings reported in CA-for-SLA literature, which I will discuss in Section 2.2, that communicative or interactional competence goes beyond knowledge of vocabulary or grammar. Learners ability to perform specific social actions using the foreign language is as important as grammar and fluency.

Individualised instruction and adaptive ICALL have been approached by creating learner personas (Heift, 2007). The author observed how learners interacted with E-Tutor, an e-learning platform for German as a foreign language (Heift, 2002, 2003). Learners have different options, how they go through material, and Heift (2007) tracked, which links they use and which material they are interested in (more cultural or more grammar and vocabulary notes). In addition, Heift (2007) considered learner variables such as gender or level of L2 proficiency although gender did not give a significant difference. She found different interaction patterns for learners with different levels of L2. Nonetheless, all possible paths in interaction with E-Tutor are determined in advance by the system's developers. User-centered design approaches to ICALL may help to create ICALL systems tailored for learning needs of different learner groups and offering more specific options for personalisation and incremental adaptation of the system to a particular user (Petrelli et al., 1999).

Data-driven methods have been frequently used to approach user-centered design and persona-based user modelling (McGinn and Kotamraju, 2008). In light of the decision to use text-based chat as a communication medium and a free chat conversation as the only activity for the study (Section 2.1.1), the question that needs to be answered in the user modelling phase is "If there are different user types in a free conversation with a conversational agent, what are the important differences for user modelling?"

In this research, I will approach the problem of user modelling in roles other than teachers or tutors by using methods of Conversation Analysis. In particular, I will focus on informal text-based instant messaging conversations to find typical structures in conversations between advanced DaF learners and German native speakers with the purpose to obtain user models from learners' behaviour. System's behaviour will be then modelled according to patterns obtained from native speakers' talk. Because learners' and native speakers' behaviour in interaction are not independent, a special attention will be put to mutual dependencies among patterns.

2.1.3. Language understanding in conversations with learners

Learner language is mentioned as *non-canonical language* in NLP literature because "learners tend to make errors when writing in a second language and in this regard, can be seen to violate the canonical rules of a language" (Cahill, 2015). Different approaches have been used to manage the contents of the conversation with the user and to deal with learner errors. Wilske (2014) mentions constraining possible input and error diagnosis as strategies used by researchers and systems' designers to deal with the complexity of learner input.

Meurers (2009) points out that it is not easy to determine the state-of-the-art in automatic learning language analysis in terms of influence of error properties on their automatic diagnosis, kind of learner language and task type. He emphasises the importance of learner corpus research for the automatic analysis of learner language in ICALL and summarises:

> feedback and learner modelling in ICALL systems and the annotation of learner corpora for SLA and FLT research are both dependent on consistently identifiable learner language properties, their systematisation in annotation schemes, and the development of NLP tools for automating such analysis as part of ICALL systems or to make the annotation of large learner corpora feasible. (Meurers, 2009, p. 470)

Amaral and Meurers (2011) see constraining the learner input as one of the main challenges in designing an ICALL system, however a necessary step because of the need to restrict the search space for syntactic processing and meaning analysis. Constraints on input in a talk with a conversational agent can be made for instance in form of domain restrictions (Pulman et al., 2010; Gardent et al., 2013) or activity restriction(Petersen, 2010). Domain restriction in SLA context is frequently achieved through strategies like role-play in a serious game (Sagae et al., 2011; Wik and Hjalmarsson, 2009), task-based dialogues (Raux and Eskenazi, 2004b), micro-worlds (DeSmedt, 1995) and virtual worlds like Second Life (Chat, 2008) where the interaction is determined by the role of the agent or avatar.

Other approaches to deal with the limitations in learner language understanding is to offer the learner only a predefined set of possible inputs or to use pattern-based language understanding keeping the coverage as wide as possible and the responses as generic as possible, for instance by means of Artificial Intelligence Markup Language (AIML) (Wallace, 2003). A predefined set of phrases that can be used by the user are offered in (Stewart and File, 2007) where a chatbot helps the learner to acquire prototypes in specific communicative situations.

As in applications targeting mostly L1 speakers, the same general observations about the depths of the language understanding can be made for the learner domain. Namely, there is a tradeoff between deep language understanding covering only very restricted domains, and shallow language understanding with very limited understanding capabilities, like for instance pattern matching and keyword spotting, and there are also techniques combining the both (Schäfer, 2008).

Typical examples of pattern-based language understanding with a wide coverage and generic responses provide AIML-based chatbots. Free AIML sets for various languages can be easily found on the Internet, for instance German (Droßmann, 2005). AIML-chatbots are easy to use and to configure, therefore there are attempts to improve very limited conversational skills by incorporating ontologies (Hallili, 2014; Al-Zubaide et al., 2011), linguistic information and reasoning (Klüwer, 2009; Jia, 2009) and knowledge about repair initiations (Höhn, 2015)[1].

This dissertation seeks to find data-driven models of interaction patterns in conversations with language learners. Different NLP tools of different complexity may be required to make the desired models part of communicative ICALL applications or integrate them into dialogue systems and conversational agents. I will discuss the required NLP tools and knowledge bases in order to make practical applications benefit from this research.

2.1.4. Learner error recognition

Dodigovic (2005) discusses different AI and HLT techniques used by ICALL systems to evaluate the language produced by language learners. The systems discussed in her book are designed to support and evaluate production of medially and conceptually written language like essay grading. Dodigovic (2005) noticed two trends from her discussion of the automatic essay grading: the inconsistency of automatic essay graders in scoring NS writing compared to NNS writing; and the inability of parsers designed for NS language to deal with NNS language errors. She compares different grammar formalisms according to a list of criteria specified in (Matthews, 1993). The criteria are computational effectiveness, linguistic perspicuity and acquisitional perspicuity. The following formalisms have been compared:

1. Context Free Phrase Structure Grammar (CFPSG).
2. Augmented Phrase Structure Grammar (APSG) and Definite Clause Grammar (DCG).
3. Shift-Reduce Parser.
4. Principles and Parameters Theory (PPT), Principle-based parsing and Chunk parser.
5. Lexical Functional Grammar (LFG).
6. Head-Driven Phrase Structure Grammar (HPSG).

She concludes that HPSG appears to be a system with the most advantages, specifically because it combines semantic and syntactic information.

Meurers (2012) classifies methods for error diagnosis in learner language into two main categories: *pattern-matching* approaches and *language licensing* approaches. Pattern-matching based approaches work rely either on specific error patterns or on context patterns for error recognition. Licensing approaches are based on formal grammars that either provide a set of constraints that need to be satisfied or a set of rules according to which valid strings of the language can be recognised. This is usually done by a definition of explicit *mal-rules* for recognition of the deviations from the standard language. Such error recognition methods focus on errors of form (syntax, morphology, orthography) but

[1]This publication is an earlier version of Chapter 8.

cannot deal with vocabulary or pragmatics errors. Some efforts have been made towards automatic analysis of meaning in learner language, see (Meurers, 2012) and references therein.

The precision in the error diagnosis is very important for ICALL applications, this has been emphasised in multiple academic publications, for instance (Bender et al., 2004; Amaral et al., 2011). Bender et al. (2004) argue therefore that mal-rule-based error recognition techniques have advantages compared with constraint-relaxation-based techniques:

> The increased precision afforded by adding particular mal-rules rather than removing constraints on existing rules increases the utility of the grammar as a component of a dialogue-based language tutoring system (Bender et al., 2004, Sec. 4).

Amaral and Meurers (2011) criticise the research approaches to create and evaluate new techniques for parsing learner language because of the mismatch in the aims to recognise errors in student sentences and the application to "hand-constructed examples."

Petersen (2010) made use of various open source English NLP tools for automatic analysis of grammatical and semantic structures in written questions in a conversational system called Sasha. The agent communicated with learners of English as a second language (ESL). A standard NLP pipeline was applied to each user's input: spellchecking tokenisation, lemmatisation, POS-tagging (Brill-tagger) and a syntactic parsing. The Collins parser implementing a probabilistic context free grammar was re-trained on ill-formed questions which were artificially created from a set of examples. The author remarks that neither the Brill tagger nor the Collins parser were intended to work with non-native like input, therefore, a set post-parse checks were needed to analyse the integrity of the structural representation of compound nouns, prepositional phrases and embedded clauses (Petersen, 2010, p. 92). The system attempted to recognise and correct all lexical, morphological and syntactic errors in learner questions. With other words, the system had to determine the *target hypothesis*, thus reconstruct the utterance intended by the learner. Both agreement on a target hypothesis and correct parsing of learner data, is a challenging task even for human annotators (Reznicek et al., 2013; Ragheb and Dickinson, 2013).

The obstacles with parsing learner language are mastered by a prediction-driven parsing in a small-world for learning intermediate German (DeSmedt, 1995). The learner plays the role of a crime detective and has to solve a murder mystery by questioning five suspects. The parser tries first to identify the verb and then to extract all the other parts of the sentence with relation to the verb. Since questions are the most likely input type in this setting, the parser was probably not challenged by the whole palette of possible inputs occurring in a free conversation.

NLP problems with learner language are often known problems in general NLP. For instance, Bender et al. (2004) see the problem of error detection based on mal-rules as closely related to the problem of parse selection, because parse versions for an input with and without mal-rules will be concurring. As Meurers (2012) points out, recognition of errors of meaning is closely related to NLP tasks like paraphrase recognition and textual entailment.

While error detection for conceptually written language (e.g. essays) is quite advanced, error recognition for conceptually oral language (e.g. free text chat) remains very challenging. One of the reasons for that is, that oral language allows more freedom in expression so that some of deviations from written standard are no longer considered as errors in oral language, see for instance the discussions in (Schlobinski, 1997). In addition to that, there are conventions in text chat allowing even more freedom in expression. All this makes a definition of an error in oral language or text chat language to a conceptual problem. Because the standard norm for oral and text chat language are not easy to define, error recognition for these areas remains under-researched. This dissertation makes an attempt to solve this problem at least partially by description of rules for "real" errors in chat, thus errors that are severe enough to be addressed to in a Chat-Conversation-for-Learning.

2.1.5. Automatic feedback generation

Dodigovic (2005) examines some theoretical views on what the origin an L2 error is and what chances exist to correct them noting that different SLA theories have different views on the meaning of L2 errors produced by learners. She concludes that the error correction behaviour of an intelligent system supporting language learning will depend on the underlying SLA theory. The majority of academic ICALL publications ground their theoretical approach to learning and thus feedback in interactionists approaches to SLA emphasising the role of interaction, input and feedback, for instance (Heift, 2003; Petersen, 2010; Wilske, 2014). For both ICALL and communicative ICALL, two conceptually different questions exist with regard to automatic feedback generation:

1. When to provide feedback?
2. Which form of feedback should be selected?

The answer to the first question is solved for the majority of the ICALL systems by the selection of the expert and activity model where the occurrence of an error triggers generation of feedback. The number of corrected errors is maximised, therefore such feedback strategies can be called greedy. However, correcting to many unimportant mistakes may lead to learner's frustration and disengagement, therefore, even greedy feedback strategies do not correct *every* error, neither do so language teachers in a language classroom.

Amaral and Meurers (2011) see providing feedback based on linguistic, learner and activity information as one of the major challenges for ICALL. They list criteria for corrective feedback considered by human tutors:

1. Information about the learner: level, age, L1, knowledge of grammatical terminology, motivation to learn etc.

2. Information about the task: type of activity (reading, listening, composition writing, etc.), type of question item (wh-question, fill-in-the-blanks, link the columns, etc.), level of question in relation to level of student, time available, material to be consulted (dictionary, grammar book, internet), etc.
3. Information about the language: grammatical competence exhibited by the linguistic properties of the learner language (lexical, syntactic, semantic, pragmatic), the nature and type of deviations in ill-formed utterances (duplication of letters, agreement, wrong synonym, lack of anaphoric reference, etc.), level of learner language in relation to scales of language complexity and development, as well as sociolinguistic, discourse, and strategic competences. (Amaral and Meurers, 2011, p. 10)

Amaral and Meurers (2011) criticise the majority of the existing ICALL systems for their selection of only the language aspect for their correction decision and focusing exclusively on the grammatical competence. However, the authors list several exceptions where the student model plays a role in the selection of the feedback form (Heift, 2003; Amaral and Meurers, 2008). However, the occurrence of an error is still the determining event triggering an occurrence of a correction.

Sometimes researchers in ICALL take a radical perspective towards feedback types. For instance, Delmonte (2003), building on the classification of corrective feedback proposed by Lyster and Ranta (1997), makes the following statement:

> We believe that recast, clarification request, elicitation, and repetition are totally inadequate for feedback generation on a computer. As to explicit correction, perhaps it could be done for grammar drills, but it is certainly much harder in semantically based drills. We assume that only metalinguistic feedback is fully compliant with the current state of human-computer interaction (Delmonte, 2003, p. 514).

Although this decision is not supported by socio-linguistic data nor it is supported by SLA or HCI-research, only one feedback type is then considered by the researchers, following this assumption.

Only a few forms of feedback were used in communicative ICALL systems. The majority of them do not implement the variety of correction types that can be found in classroom research, for instance (Lyster and Ranta, 1997; Lyster et al., 2013) and in learner conversation data. This is mainly because conversation practice is *not* the main focus of the system, but rather essay writing tasks, multiple-choice and template-based tests, and short-answer questions, see for instance (Nagata, 2009) and the references therein. For the cases where feedback is implemented for a dialogue interaction with the learner, it is normally done in form of a pop-up window laid over the chat window which is not part of the ongoing conversation, or in form of a retrospective summary.

A pedagogical agent Dr. Brown is a representative of the systems providing feedback in form of a retrospective summary. It was used to engage the user and the ESL learner in request games (Yang and Zapata-Rivera, 2010). The learners received a problem description and were asked to negotiate with the pedagogical agent to solve the given problem. Thus, the task of the system is providing the learner with an opportunity to practice their knowledge of pragmatics in simulated interactions in the academic context. The dialog processing is implemented as a final-state machine (FSM) able to recognise hard-coded well-formed and ill-formed learners' utterances. The FSM was created based on a dataset from a pilot study where English native speakers had to complete the same tasks. The number of potential dialogue moves for the tasks was considered as very limited. Therefore Yang and Zapata-Rivera (2010) choose to predict all possible learners' inputs for all dialogue situations. The system does not make use of any NLP tools. Keyword-based language understanding is used to evaluate the appropriateness of learners utterances. The authors defined appropriateness as the degree of directness and politeness compared to native speakers' responses in the same situations. At the end of each situation, learners received feedback based on the appropriateness score of their utterances.

The micro-world *Herr Kommisar* represents communicative ICALL systems providing feedback in form of pop-ups during the conversation. It provides corrective feedback on *every* recognised error in from of an explicit correction (DeSmedt, 1995). A small number of standard templates is used for feedback generation. User's response to correction was encoded into buttons to close the pop-up and return to the chat window. A similar approach to corrective feedback was chosen in (Lavolette et al., 2015).

Some efforts have ben taken to incorporate the research results on different types of corrective feedback in ICALL systems. Petersen's work discussed in Section 2.1.4 was only focused on recasts as a type of corrective feedback comparing recasts produces by human tutors to recasts produces by an artificial conversation partner Sasha (Petersen, 2010). Recasts are defined in SLA literature as "teacher's reformulation of all or part of a student's utterance, minus the error" (Lyster and Ranta, 1997). This definition covers a wide range of correction formats. Only one form of recasts was allowed in (Petersen, 2010)'s work: a repetition of a complete question where all morphosyntactic errors were corrected. A recast was provided for *every* recognised ill-formed user's question. Petersen sees the own study as delivering recasts in a manner comparable to recast provision in spontaneous oral interaction. However, from the perspective of CA, question elicitation cannot be seen as spontaneous naturally occurring conversation. Nevertheless, the work shows the high complexity of target hypothesis generation even when restricted to question reformulation.

Recasts in communicative ICALL have been also handled by (Morton et al., 2008; Anderson et al., 2008) and by Wilske (2014). The *SPELL* system (Spoken Electronic Language Learning) targets corrective feedback in form of recasts (Morton et al., 2008; Anderson et al., 2008). Examples of feedback provided in the academic publications are all acknowledgement-based reactions to learners' responses delivered after agent's questions. Wilske (2014) describes a text chat-based system that is able to produce recasts and metalinguistic feedback in a task-based dialogue with the learner. Only one type of grammatical errors was considered for experiments. The work focused on comparison of recasts and meta-linguistic feedback in ICALL and classroom studies.

A speech-based system *Let's Go* provides error correction within the dialogue with the user, but all of them are based on a *Did you mean* followed by a correct reformulation. The task in the communication with *Let's Go* is to plan a journey. The system produces confirmations of user's journey plans. These confirmations are preceded by corrective clarification requests in order to ensure system's correct understanding and to deal with "linguistic mismatch" (Raux and Eskenazi, 2004a).

In all these cases, the software was assumed to play a role of a teacher eligible to correct explicitly and as much as possible. This corresponds to an unequal-power speech exchange system as discussed in Section 2.2.1). However, my data show that corrections in a free conversation between peers are more diverse and in more than 50% of the cases implicit.

The opposite of the greedy correction strategies can be found in Communicative ICALL applications working with expert models other than teacher or tutor. For instance, (Hjalmarsson et al., 2007) a conversational agent in a trading role play does not provide corrective feedback. I call this correction strategy zero-correction. A combination of a pedagogical agent and a conversational agent in an ICALL system allows to handle both, accuracy and fluency as two independent tasks. For instance, Wik and Hjalmarsson (2009) describe two embodies conversational agents for L2 training. The first agent should simulate a teacher and provide feedback on phonetics, the second agent should play a role of a conversation partner and focus on fluency assuming that no feedback is given in a free conversation context. The conversation trainings system DEAL is described also in (Hjalmarsson et al., 2007) and in (Wik et al., 2007).

Because the majority of Communicative ICALL applications sees corrective feedback not necessarily as part of the ongoing conversation, they do not cover the richness of the types of corrective feedback described in the SLA classroom research. The few attempts to build upon an existing classification of corrective feedback, such as (Petersen, 2010; Wilske, 2014; Morton et al., 2008), need to restrict the variety of the feedback types in order to obtain an acceptable feedback quality. This also shows the complexity of the problem.

Conversation Analysis distinguishes between *exposed* and *embedded* corrections (Jefferson, 1983; Brouwer et al., 2004). Exposed corrections make the correcting to the interactional business while embedded corrections are accomplished implicitly, without focusing on it. Corrective feedback described in SLA literature covers only exposed corrections.

In my dissertation, I take efforts to contribute to the research on automated corrective feedback (exposed corrections) in the following way. Because there is a difference in the speech exchange system between a teacher-fronted classroom and a chat-based Conversation-for-Learning, I seek to obtain a data-driven classification of error corrections in my dataset. I will then compare the found types corrections with the classification of corrective feedback obtained from classroom data (Lyster et al., 2013). Moreover, I will suggest computational models of error corrections for communicative ICALL application resembling chat-based Conversations-for-Learning. I will document the empirical findings in Chapter 5, I will describe the model in Section 9.2 and discuss the results in Section 11.3.2.

While a lot of research efforts have been put into automatic feedback generation in form of exposed corrections, none of the academic publications in ICALL reports about studies of automated embedded corrections in interaction with language learners. This is caused on one hand by the strong connections between the ICALL scene and the interactionists SLA research. The *noticing hypothesis* emphasises the importance of the noticing of the error by the learner in order to produce modified output (Schmidt, 1990), and the evidence of noticing after embedded corrections is not easy to find. On the other hand, embedded corrections may have been out of scope in ICALL research because the empirical research did not produce any ready-to-implement model for embedded corrections. This dissertation makes an attempt to operationalise embedded corrections with the purpose to create a computational model for embedded corrections in conversations with language learners. I will report in Chapter 6 and Sections 9.3 and 11.3.3 about the success of this research endeavour.

Conversations-for-Learning combine characteristics of both, informal talk and language classroom, as explained in Section 1.1. Therefore, a new decision model for corrections is required. The desired correction decision model should act somewhere between the greedy correction and the zero-correction models allowing to come closer to the correction behaviour of the native speakers in a chat Conversations-for-Learning. This dissertation seeks to make a step towards closing this gap. I will make an attempt to create an empirically grounded decision model for corrections. Chapters 7 and 10 document the findings, Section 11.4 summarises the research results with regard to this problem.

2.1.6. Learner corpora and error annotation

As Meurers notes, the annotation of learner corpora is mainly focused on annotation of learner errors, however, annotation of linguistic categories in learner corpora is also of interest (Meurers, 2009). To create stable models of learner language for statistical NLP tools, information on occurrences of linguistic categories and their dependencies is required. This need is approached by linguistic annotation of learner corpora, similar as it has been done for native-speaker language. Examples of linguistic annotation in learner corpora are (Amaral and Meurers, 2009) who focused on tokenisation in Portuguese interlanguage, and (Díaz-Negrillo et al., 2010) addressing the problem of POS-tagging in interlanguage. Related to the annotation of conceptually oral language, the challenge of POS-annotation in chat language has been addressed by (Bartz et al., 2014).

Error-annotation of a corpus assumes a non-ambiguous description of the deviations from the norm, and therefore, the norm itself. This is quite unproblematic for errors in spelling, morphology and syntax, however, different annotators' interpretations lead to huge variation in annotation of errors in semantics, pragmatics, textual argumentation (Reznicek et al., 2013) and usage (Tetreault and Chodorow, 2008). Multiple annotation schemes and error taxonomies have been proposed for learner corpora, for instance (Díaz-Negrillo and Domínguez, 2006; Reznicek et al., 2012). Because error taxonomies are language-specific, I focus only on error annotation in German learner (L1 and L2) corpora.

The corpus of German emails posted to USENET users described in (Becker et al., 2003) consists of ca. 120 000 sentences. An error typology of orthographic, morphological, morpho-syntactic, syntactic and syntactic-semantic errors was taken as a basis for the error-annotation, however, only 16 error types from the typology were used for the corpus annotation.

Different error tagging systems for learner corpora have been described by (Díaz-Negrillo and Domínguez, 2006). The authors note that "[error] taxonomies should be grounded on the description of observable data and include well-defined linguistic categories to minimise subjectivity". Consequently, many different error taxonomies have been created to serve specific corpora. There are in total 6 error-annotated corpora of German as a foreign language. The major conceptual work on the annotation scheme and error taxonomy was done by the FALKO team (Reznicek et al., 2013, 2012) and frequently re-used or oriented to by the followers (German part of MERLIN, EAGLE, WHiG). WHiG is part of FALKO but contains texts from native speakers of British English who are intermediate learners of German (Krummes and Ensslin, 2014). LeKo is accessible though FALKO platform.

All of the error-annotated corpora consist of argumentative essays, and the developed error taxonomy is good for error-tagging in essays, but needs further elaboration to be fitted for conceptually oral language like instant messaging exchange.

The error-annotation for the mentioned corpora was approached in the following ways. The LeKo-corpus was probably a pilot project, it was created earlier than FALKO and by the same principal investigator (Lüdeling et al., 2010). The researchers elaborated an error taxonomy on a small learner corpus of 30 texts that were written manually and then re-typed to make the resources digitally available and analysable. The difficulties with error annotation that were faced by the annotators of the LeKo corpus were taken into account in the annotation definition phase for the FALKO corpus. Specifically, some of the errors can be tagged differently depending on what the learners intention was, or *target hypothesis*. Dealing with such ambiguities became an issue for learner corpus annotation.

A multilevel annotation was introduced in the FALKO corpus in order to deal with different target hypotheses (Lüdeling et al., 2005). The first target hypothesis ZH1 should only address errors in orthography, morphology and syntax, and make the sentence "understandable" for NLP tools. The second target hypothesis ZH2 should address all other types of errors, like semantics, lexical choice, pragmatics, style and so on (Reznicek et al., 2012). I will make use of the two target hypotheses in Chapter 9 to deal with errors in questions posted by the learner participants of my data collection and to come up with a model of embedded corrections in answers to such questions.

An extension of FALKO annotation schema has been suggested in the EAGLE corpus of beginning learner German where error numbering was introduced to deal with overlapping errors (Boyd, 2010). Multiple target hypotheses were approached by setting a preference for the target hypothesis which minimises the number of annotated errors.

Title	L1	GFL level	Data type	Size	Error-annotated	Available
ALeSKo	Chinese	Different	Written texts	43 texts	Partial	Yes
CLEG13	English	B-C	Written texts	731 texts	NA	Yes
FALKO	Many	Intermed. - advanced	Written texts	Under development	Yes	Yes
WHiG	English	B2	Written texts	279 texts	Yes	Yes
MERLIN		A1-C2	Written examinations	1033 texts	Yes	Yes
LeKo	Many	Different	Written texts	30 texts	Yes	Yes
LeaP	Many	Different	Speech	183 records of 2-20 min	No	Yes
EAGLE		Beginners	Online work book, essays	50 WB & 81 essays	Yes	Yes
LINCS	English	Intermed.-advanced	Written texts, longitudinal	Under development	NA	No
ADS	English	Beginner-intermed.	Threaded discussion, chat, essays, longitudinal	Under development	NA	No
Telecorp	English	Different	Email, IM, essays	1,5 mio words	No	No
deL1L2IM	Russian	Advanced	IM	52000 tokens	Partial	Yes

Table 2.1.: German learner corpora at the end of 2015

ALeSKo is a corpus of annotated essays of advanced learners of German with Chinese as L1 (Zinsmeister and Breckle, 2012), the annotation contains manual marks of topological fields (fields and error marking), referential expressions (definiteness, specificity, target hypothesis) and vorfeld use. The subject of the ALeSKo study was coherence in learner texts based on the annotation of syntactic, referential and discourse information. German-L1 part of the FALKO corpus were used for L1-L2 comparison. A specific focus of the annotation in ALeSKo lies of referential expressions (Breckle and Zinsmeister, 2010), which are also in general an important area of NLP research and relevant for this work not only from the point of view of learner language understanding, but also from the perspective of generation of embedded corrections. Models of embedded corrections will be subject of Section 9.3.

A specific characteristics of CLEG13 corpus is that it has a "truly longitudinal core" of texts produced by students from their first year to their final exams.

In contrast to the written resources described above, the LeaP corpus includes phonologically annotated speech recordings of German and English learners German (Gut, 2009). The corpus includes readings of nonsense word lists, readings of a short story, retellings of the story and free speech interviews.

German as a learner language can be also seen as standard linguistic competence acquired by German L1 speakers. The corpus KoKo is part of the project Korpus Südtirol, and focuses on L1 German learned in South Tirol by school pupils (Abel et al., 2014).

The situation with German error-annotated learner corpora is that there is a very small number of corpora, and only a small part of them are publicly available. The website "Learner Corpora around the World" maintained by Amandine Dumont and Sylviane Granger (Université Catholique de Louvain[2]) lists in Oktober 2015 only 11 German learner corpora, 10 medially written and 1 spoken. In addition, there are a few publications about German error-annotated corpora not mentioned on the web page. Table 2.1 provides an overview on German learner corpora which I was aware of at the end of 2015. The table includes only information about the German part for multilingual Corpora (LeaP, MERLIN).

The corpus requirements for the present study were defined as follows: it should be longitudinal text chat dialogue between language learners and native speakers of German. None of the previously existing corpora satisfy these requirements, therefore a new data collection was created. The deL1L2IM-corpus is a new linguistic resource which I created to serve the needs of this research. Table 2.1 shows how the new corpus fits into the German learner corpus landscape. The details of the data collection process and data quality can be found in Appendix A.

I approach the problem of error taxonomies from the perspective of the need for correction. In contrast to a writing assistance program that has to (ideally) identify and correct *every* error, only a small amount of all errors are usually corrected in a Conversation-for-Learning. Not every error is corrected even in a language classroom in a fluency context. Therefore, the artificial conversation partner needs to distinguish errors that could be potentially corrected in an instant messaging dialogue with the learner from those, which should not be addressed to. I will explain the way and describe the findings with regard to the conceptual understanding of linguistic errors in chat in Section 9.1.

2.2. What communicative ICALL can learn from Conversation Analysis

While Human-Robot Interaction benefits from CA-informed research (Section 2.3), Communicative ICALL mostly relies on Second Language Acquisition theory for user and expert modelling, activity design and interaction with the user. The aim of this section is

[2]https://www.uclouvain.be/en-cecl-lcworld.html, retrieved on 5th October 2015

to outline how Communicative ICALL research would take advantage from extending its multidisciplinary connections to CA and CA-for-SLA as related research areas.

Huge differences in organisation of different types of discourse have been emphasised in CA literature (Schegloff, 1993; Markee, 2008). In addition, there is a need to model communication roles of the system other than teacher or tutor, as argued in Section 2.1.2. For these reasons, I turn first my attention to the discussion of a chat-based Conversation-for-Learning as a specific speech exchange system in Section 2.2.1. Building on this, I will explain the problem with instant messaging research in SLA in Section 2.2.2. Earlier versions of Sections 2.2.1 and 2.2.2 were included in (Höhn et al., 2016, under review).

Several CA-driven studies of native/non-native speaker chat achieved important results in understanding of processes contributing to the construction of linguistic identities of a language expert and language novice in talk, as I outlined in Section 1.2. I will describe the most significant results in Section 2.2.3. I will continue the discussion on linguistic identities in Section 2.2.4. Because various forms of repair are employed in participants' orientation to linguistic identities, and because some types of repair are seen as important for learning in SLA reseach, I will take a closer look on relevant research on repair in Section 2.2.5 and embedded corrections in Section 2.2.6.

2.2.1. Chat-Conversation-for-Learning as a speech exchange system

Conversation Analysis (CA) research shows that it is extremely important to pay attention to the interactional rules of the specific interactional setting. Such sets of rules are referred to as *speech exchange systems* in CA literature (Schegloff, 1993; Markee, 2000). Examples of speech exchange systems are ordinary conversation, interview, business meeting, news conference, therapy, classroom, debate or ceremonial talk, and they differ in terms of turn allocation and organisation of repair (Markee, 2000, pp. 72-73, 84). As Markee citing Schegloff (1993) points out, "data collected in laboratory settings inevitably reflect a member orientation to a speech exchange system that is demonstrably different from that of ordinary conversation" (Markee, 2000, p. 33). Sawchuk (2003) analyses the "nature of informal learning as a distinct speech-exchange system" and claims that "this speech-exchange system displays features of both the formal classroom and some sort of informal conversational speech-exchange system". Putting informal learning in SLA context, Kasper (2004) uses the term *Conversation-fo-Learning* to describe interactions where the participants use a foreign language to construct talk-in-interaction. Kasper (2004) sees Conversations-for-Learning as category-bound events because the participants of the interactions come together due to their statuses as NS/language expert and NNS/language learner.

The notion of informal learning is further analysed in CA as a distinct speech exchange system which can be seen as very close or equivalent to Conversation-for-Learning in pure SLA context. Sawchuk (2003) claims that

> 'informal learning' shares qualities of 'formal classroom interaction' [...], and 'everyday conversation' [...]. More specifically, it exhibits a form of topic continuity similar to school-based (expert/novice) speech-exchange systems but which is not seen in everyday conversation. And, it exhibits shared control over turn-taking that is often associated with everyday conversation but which is not seen in school-based speech-exchange systems. (p. 295)

The analysis of technology mediated communication would not be complete if the tool itself would not be taken into account. However, only a few studies of computer-mediated communication focus on the "constraints and affordances" existing through the tool, but see for instance (Brandt, 2011). Doing more formal or more informal talk in a face-to-face interaction involves different interactional resources than doing the same in text chat or in a phone conference, as illustrated in Figure 2.1.

The fact that it is different does not make text chat interaction less powerful or participants somehow handicapped. There are just *different* sets of recourses that participants *can* make available in interaction to express themselves. From learner's perspective, text chat interaction even has many advantages because the production pace is lower than in oral synchronous communication, they have more time to thing about the wording and the grammar, and even look up in dictionaries for words that are not immediately available for them, either for the own production or for the comprehension of the partner's talk. In addition, they can re-read the chat protocol for a better comprehension or learning, the nature of text chat is that everything is recorded online and is available for the participants later. However, pronunciation cannot be heard or practiced in a text chat.

With regard to these issues, I approach the analysis of the dataset as a specific speech exchange system which is *chat-based Conversation-for-Learning*. As already several publications demonstrated (Orthmann, 2004; Tudini, 2010; Marques-Schäfer, 2013) and as I will show in Part II of this dissertation, this speech exchange system combines properties of informal oral interaction and formal oral classroom which are put into the frame of medially written interaction. However, the relation between the structures characteristic for

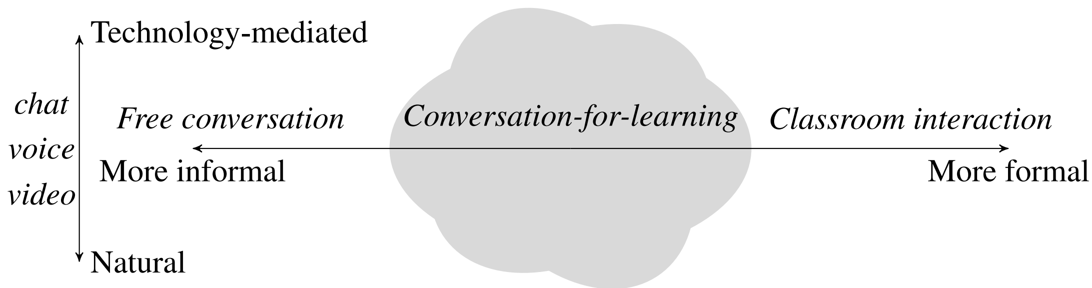


Figure 2.1.: Conversation-for-Learning combines structures of formal and informal conversation and involves interactional resources provided by the communication medium.

informal and formal interaction may vary in different chat datasets. The proportions of them are influenced by many factors. I will argue in the next section that the procedure of data collection is one of them, in support of the importance of gathering naturally occurring data emphasised in CA (Markee, 2000; Kasper, 2004; Firth and Wagner, 1997). I will discuss the quality of the corpus in Appendix A.

2.2.2. The problem with instant messaging data

Previous research on the influence of CMC on language learning occurred in experimental setups where the interaction itself was treated as a blackbox. Researchers carried out pre-tests and post-tests, and concluded on the influence of the blackbox on SLA (Abrams, 2003; Loewen and Erlam, 2006). Though this approach is valid, it does not allow for understanding the mechanism of the influences of CMC on SLA.

While group text and video chat allows for a quite easy observation and recording by the researcher in order to collect data for further analysis of different aspects of learning (Brandt, 2011; Marques-Schäfer, 2013), private dyadic text chat between learners and native speakers remains an underexplored field. This may result from the fact that data cannot be collected as easy as it is in group chat. As Marques-Schäfer (2013) says:

> The storage of the chat protocols is the most frequent data collection method in the chat linguistics and foreign language didactics research. (orig.: *Die Speicherung von Chat-Protokollen ist die häufigste Erhebungsmethode der Chat-Forschung in der Linguistik und in der Fremdsprachendidaktik.*) (p. 117).

Obtaining the stored instant messaging protocols for research purposes from the IM service provider is only possible when the study is organised in a way which allows to get the data either directly from the server or from the users. The former is normally not possible if the researcher has no relationships to the service provider. Privacy is the next issue, even when the researcher has relationships to the service provider. If instant messaging dialogues are in focus, then privacy is a bigger issue than for group chats and no provider will just give the researcher the permission to “listen” to other people’s talk. Unfortunately, researchers frequently omit to provide information about how they got the data (De Marco and Leone, 2012; Fredriksson, 2012). The data capturing methods reported in the literature base mainly on experimental settings where the participants in the dialogue have to use a specific hardware-software combination for the study, for instance PC-pools at an educational institution for task-based conversation (Kost, 2004). This makes the whole interaction to an experiment and restricts the communication. The dataset described by (Sauro, 2009) is used for several other studies, for instance (Vandergriff, 2014). Though it is a valuable resource for research on different SLA issues, however the way how the data were obtained does not allow for analysis of natural NS-NNS interaction. The native speakers in the dataset by Sauro (2009) were made familiar

> with the target form and trained in the provision of the different types of corrective feedback as well as strategies to avoid supplying the learners with positive evidence of the target form. (p. 100)

It is specifically problematic if the research focuses on the influence of natural interactions between learners and native speakers on SLA development or if researchers pose open questions like: "What happens if a native speaker of a language and a non-native speaker of the same language interact in text chat?".

Another problem is that researchers usually use a specific software as a web service where users have to create their own accounts and log in with the goal to improve the language skills (Blake and Delforge, 2004; Blake, 2008; Tudini, 2010; Marques-Schäfer, 2013). This makes the interaction closer to a classroom interaction but then, we still do not know anything about the naturally occurring chat interaction.

In order to avoid the experimental character of the CMC data and to obtain data from close-to-natural chat interactions, researchers mostly ask the participants to store the data locally and to provide them with a copy of the data log (Lee, 2007; Jin, 2013). This, in turn, puts researchers in a subordinate position and makes the whole studies dependent on the charity of the data suppliers i.e., the participants. In such situations, researchers cannot access the data immediately, they cannot control the consistency and the completeness of the data, because the participants are free to decide, what to show and what to hide. It becomes even more complicated if the study should have a longitudinal character (long-term interaction).

In addition to the described problems with the data quality, there is an issue with comparability and reusability of the datasets. Marques-Schäfer (2013) lists 40 studies related to using chat for SLA that were published between 1994 and 2010 and criticises a bad comparability of achieved results due to very individual study settings. She suggests to specify the attributes for each study, that would help to make at least the setup better comparable, even if results are still very individual. Moreover, the datasets described in the CALL and CMC literature are never available, mostly for privacy reasons. It is rarely possible to perform different studies with the same dataset, or to use the same datasets to confirm or to challenge the results presented by the researchers.

During the literature review on the description of data collection of text and video chat, 18 attributes have been identified in which the datasets may differ; namely data collection time (date and duration), participants demographic information (L1, other spoken languages, Level of L2, method of selection), target language, study description (modality of the talk, e.g.: text, voice, video, instructions to participants), number of participants, size of the dataset, data selection for analysis, software used, software availability, researcher's role, ethics, privacy protection and data availability. Variance in each of these attributes is one of the reasons for insufficient result comparability. Furthermore, technology develops rapidly and the ICALL research needs to follow all technological innovations in CMC in order to deliver up-to-date results. Hence, it is worthy of empirical investigation to pay particular attention to the impact of new technologies on interaction and learning (Brandt, 2011).

2.2.3. Conversation analysis of native/non-native speaker chat

Marques-Schäfer (2013) lists 40 studies documented between 1994 and 2010 in the major CALL journals - *CALICO*, *ReCALL* and *Language Learning and Technology* (Marques-Schäfer, 2013, Sec. 3.2.2.). While many SLA inspired studies of native/non-native speaker chat have been reported in CALL literature, relatively few CA-driven investigations of native/non-native speaker chat are documented in academic literature.

Tudini (2010) describes a long-term study of communication between Australian learners of Italian and Italian native speakers in dyadic text chat. She focused mainly on repair sequences "as this is where SLA behaviours are most evident" (Tudini, 2010, p. 6). Among other repair types, the author discusses various types of other-correction and concludes that the types of error correction build a "continuum of explicitness of exposed correction in online text chat" (Tudini, 2010, p. 101) ranging from explicit (exposed) corrections on-the-fly to teacher-like corrections accompanied by accountings and meta-linguistic information.

Vandergriff (2013) analysed dyadic task-based interactions between previously unacquainted participants in English as a second language. Specifically, she investigated how participants of a native/non-native speaker chat make their differential linguistic expertise relevant in conversation. Her findings show that these orientations primarily occur outside of other-repair sequences and therefore, do not play a role "in mitigating and/or sanctioning a face-threatening act." (Vandergriff, 2013, p. 393)

Though the following work is not CA-driven, however qualitative methods have been used for data analysis. Marques-Schäfer (2013) analysed a subset of chat protocols from *JETZT Deutsch lernen* project offered by the German GOETHE-Institute [3]. As part of this project, earners of German as a second language could chat with other learners and native speakers. The dataset consists of tutored and untutored group chat sessions (Marques-Schäfer, 2013, p. 180). She found that the number of corrections in tutored sessions was higher than in untutored sessions. Marques-Schäfer (2013) concludes, that the presence of tutors in chat emphasises the didactic character of online chat. It seems natural to expect more teacher-like behaviour from the tutors in chat due to their pre-assigned role of more knowledgeable participants.

Although CA-driven and qualitative studies of native/non-native speaker chat gained important insights in specific aspects of repair, sequential organisation and participants' identities, operationalisation of the identified structures for computational modelling still needs to be performed. This research makes a step in this direction.

[3] http://www.goethe.de/z/jetzt/

2.2.4. Constructing identities of expert and novice in conversation

Participants of an interaction may *choose* to orient to their linguistic identities in talk, but they do not have to, as I argued in Section 1.2. In addition, the membership in the categories *novice* or *expert* is not fixed. (Hosoda, 2006) shows how non-native speakers of Japanese switch their roles from language novices to language experts in informal conversations with their native speaker friends. Marques-Schäfer (2013) observed that learners corrected linguistic errors produced by other learners in tutored chat sessions of the project *JETZT Deutsch lernen*:

> die Tutorin [ist] nicht die einzige Person in einer tutorierten Stunde [...], die eine Fremdkorrektur übernehmen kann. Die anderen Chat-Teilnehmer zeigen sich [...] hilfsbereit und sprachlich kompetent, um einander zu helfen und zu korrigieren. *The tutor [is] not the only person in a tutored session [...] who can perform an other-correction. The other chat participants show themselves as willing to help linguistically competent in order to help each other and to correct* (p. 186, Engl. translation added).

Thus, non-native speakers may choose to position themselves as language experts. This observation has been confirmed for instance by a study of the use of English as Lingua Franca in non-native/non-native speaker interaction (Hynninen, 2012). The author discusses how the linguistic identity may appear in interaction due to the participant's status of a native speaker or a language professional, but also be negotiated in the course of non-native/non-native speaker interaction. Hynninen (2012) shows that "even if the courses are not language courses, language sometimes becomes the topic of discussion in the form of language correcting and commentary." (p. 13).

Hosoda (2006) focuses in his study of informal oral communication between friends who are native speakers of Japanese and learners of Japanese as a foreign language. He points out that the orientations to the roles of the language expert or novice in conversation is especially salient in "remedial" sequences like repair or correction, but the repertoire of sequences for such orientations *is not limited to repair or correction* (Hosoda, 2006, p. 26, emphasis added). The study "has demonstrated, on the occasions that participants in ordinary L1-L2 conversation orient to differences in their linguistic expertise, the structures of the conversation may become similar to those of language classrooms" (Hosoda, 2006, p. 44).

Vandergriff (2013) analysed how institutionally structured identities of student and teacher have been integrated by the study participants in their interactional roles as a language novice and language expert. She documented with multiple examples that the interactional role of a novice is not restricted to the institutional identity of a language student. In particular, face-work in the broader sense plays a major role in constructing the interactional role of a language novice, such as (re-)indexing of the L2 social self, building and maintaining social rapport with the co-participant and cultivating social presence (Vandergriff, 2013, p. 393, 404).

A one-year longitudinal study of video-recordings of native/non-native speaker interaction in Spanish are described by (Dings, 2012). Based on analysis of six conversational interactions, the author examines the interlocutors' orientation to their identities as language novice and language expert. Dings (2012) found the following forms of orientation to linguistic identities: discussion of language learning, learner error corrections initiated by the native speaker, learner error corrections initiated by the non-native speaker. In addition, (Dings, 2012) found changes in patterns for correction. While repair in the beginning was more form-focused, it tends to be more meaning-focused later.

The complexity of the notion of linguistic expertise can be understood on studies where L2 is used for work and the conversations do not have the primary purpose of language practice. Vickers (2010) examines audio-recorded team meeting data between engineering students, one native and one non-native speaker of English, jointly working on a project. The author demonstrated that language competence may be blended with other types of competence through the use of specific interactional strategies. Linguistic and non-linguistic identities of experts are locally constituted in native/non-native speaker interaction. Vickers (2010) argues that these linguistically and non-linguistically based identities are interrelated so that a linguistic identity of a novice has a bearing on the achievement of a non-linguistic identity of an expert.

In contrast to the work described earlier in this section focusing on how a participant of a conversation presents herself as member of a specific category, *interaction profiles* are concerned with *individual* shapes of interaction participants (Spranz-Fogasy, 2002). The subject of the analysis is how the conversational activities of all participants of an interaction becomes systematic and stable with the focus of one participant, and become in this way an interaction profile. Spranz-Fogasy (2002) defines *interaction profile* as

> die auf einen einzelnen Gesprächsteilnehmer bezogene Verlaufskonfiguration des interaktiven sprachlichen Handelns aller Teilnehmer in einer jeweiligen Interaktion, wie sie sich aus dem Gesprächshandeln ergibt und zugleich auf dieses Handeln zurückwirkt und es anleitet
> *a configuration in the process of conversational behaviour of all interaction participants with respect to a single interaction participants, how such a process configuration emerges from the interaction and at the same time influences and guides the interaction* (p. 47, Engl. translation added).

The author shows in a very detailed analysis of multiple interactions from various situations (public group discussions, family conflict talks and institutional arbitration talks) how specific interaction practices used by interaction participants become recognisable patterns or types of behaviour. With other words, an interaction profile is something that emerges in interaction and is influenced by all interaction participants and the interaction history when participants use specific language practices. The number and the kind of available practices is open.

Summarised, there are various possibilities in interaction to orient to linguistic identities. The identity of a language expert (or more knowledgeable participant) can only exist in a dichotomy paired with the identity of a language novice (or less knowledgeable participant). The individual shapes of these identities may vary and can be only partially influenced by the acting participant (participant who initiates such orientations). The reacting participant (who responds to this initiation) may accept or reject the role assigned to her. For the purpose of a communicative ICALL application, different local models of orientation to linguistic identities may be provided to the machine, but it is not possible and not necessary to determine in advance, which of them will be activated. A notion of interactional relevance needs to made accessible to the machine in form of a decision model for the activation of specific actions from that pool. In this way, an*individual interaction profile of a language expert* may emerge in interaction between the machine and the language learner. This idea has been developed in (Höhn et al., 2015) and will be discussed in Section 3.4.2.

2.2.5. The power of repair

(Liddicoat, 2011) describes repair as "the processes available to speakers through which they can deal with the problem which arise in talk". In Hosoda (2006)'s definition, "repair in the CA sense deals with any problems in speaking, hearing, or understanding, such ass clarification requests, understanding checks, repetitions, restatements, offers of candidate hearings, and the like, and it includes but is not limited to corrections of linguistic errors". From CA point of view, repair is not something that disturbs the interaction. It is a principal resource that participants of a conversation have at their disposal to maintain intersubjectivity, thus, to construct shared meanings (Markee, 2000; Schegloff, 1992).

Following previous work on repair (Jefferson, 1974; Schegloff, 1997; Dingemanse et al., 2014) I will use the common terminology distinguishing among practices, devices and formats. Dingemanse et al. (2014) provide the following definitions which I adopt in this dissertation:

Practices are generic, language-independent techniques like 'repetition' and 'questioning'.

Devices are particular, language-specific linguistic resources, such as 'particles', 'question prosody', or 'noun-class specific interrogatives', and rules for their application.

Formats combine generic practices and language-specific devices to deliver social actions. (pp. 10-11)

Nothing in the language is a trouble source by itself, but everything can appear to be a trouble source in a conversation if it is marked as a trouble source by the conversation participants. However, there are structures in all languages that have a greater potential to become a trouble source for the learner because they require a higher level of language proficiency to use or to understand them correctly. Such linguistic structures are for instance idioms, figurative expressions and proverbs. In addition, some social actions appear more difficult to perform in a foreign language, such as responding appropriately to compliments (Golato, 2002).

Since everything can become a trouble source, repair initiations can occur after any turn at talk, even after silence (Schegloff, 1993). However, different repair *formats* might be appropriate for different speech exchange systems. For example, a German expression *Wie bitte?* is a practice to other-initiate self-repair. The recipient of the trouble-talk asks the trouble-speaker, for instance, to repeat an utterance.

For the analysis of repair work, it is necessary to see who produced the repairable (who is the trouble-speaker), who initiated the repair, and who carried out the repair. From the perspective of the speaker who produced the trouble-source (also referred to as repairable in literature), CA researchers differentiate between self-initiated and other-initiated repair. Similarly, the repair can be carried out by the trouble-speaker or the recipient of the trouble talk. These two types of repair are referred to as self-repair and other-repair respectively.

There are specific positions dedicated to specific types of repair initiation in conversation. Due to a preference for self-repair, first-position repair initiations are dedicated to self-repair (same-turn or transition-space repair) (Schegloff et al., 1977). The first position where other-initiation can occur is the first speaker change, this is the so called second-position repair initiation. This position is dedicated to other-initiation of repair. Third position repair may be relevant when the first speaker understand from the reaction of the second speaker that the previous utterance was wrongly understood. This position is dedicated for self-initiation of repair. And the last position where repair initiations have been found till now is the forth position, which is dedicated to repair other-initiation. However, this type of repair initiation is very rare and requires the biggest amount of interaction management work of all repair types (Schegloff, 1992; Schegloff et al., 1977; Schegloff, 2000; Liddicoat, 2011).

While other-initiated self-repair and other-initiated other-repair (other-corrections) are relatively well analysed in face-to-face and chat data from CA and SLA perspectives, less is known about self-repair in chat due to difficulties in capturing the data (Smith, 2008). From the CA perspective, this is not a big problem, because "what happens before turns are posted is not relevant to the interaction unless [...] the participants themselves orient to it during the interaction" (González-Lloret, 2011, p. 318).

Benjamin (2013) contributes to the research on repair with the study of other-initiated repair in English naturally occurring social interaction. Word searches and candidate understandings are in the focus of the study of Russian and Finnish as lingua franca in (Pikkarainen, 2015). Egbert (2009) analyses the German repair mechanism in oral conversations and compares the findings to repair structure in different languages. She concludes that the basic structure of repair mechanism is language-independent, but there are language-specific and culture-specific features that can be embedded in repair activity.

> *Es gibt eindeutige Hinweise darauf, dass der Reparatur-Apparat ein sprachübergreifender Mechanismus ist, welcher in seiner Grundstruktur unabhängig von sprachlichen oder kulturellen Unterschieden ist, jedoch Möglichkeiten*

> *beherbergt, kulturelle Merkmale, Besonderheiten im linguistischen Repertoire und spezifische Handlungen im Zusammenhang mit der Reparaturaktivität zu berücksichtigen*. (Egbert, 2009, p. 166)

(Dingemanse et al., 2014) describe formats of repair other-initiations for oral data across 10 languages not including German. They show that there is a set of interactional resources in each language that correspond to open class repair initiations and restricted class repair initiations. Open class repair initiations signal *that* there is a problem with the previous utterance. Restricted class repair initiations allow to narrow the scope of the trouble diagnosis. Open class repair initiations are also analysed by (Drew, 1997; Enfield et al., 2013).

> With interjections, single question words, and formulaic or apology-based formats we have exhausted open formats. The set of restricted formats, i.e. formats that zoom in on specific items in the trouble-source, is larger and more internally varied. [...] Despite this variety, the types of linguistic resources used are relatively limited. They consist of content-question words, full and partial repetition, and various types of candidate understandings. (Dingemanse et al., 2014, p. 17)

Analysis of repair sequences in native/non-native speaker chat was performed by Tudini for a dataset of dyadic interactions between students of Italian at an Australian university and native speakers of Italian based in Italy (Tudini, 2010). This chat interaction was part of students' assessment, however, they were asked to interact with native speakers in their free time outside the classroom. Topics for the interaction were provided. The participants interacted almost every time with a different person, therefore, the long-term development of the interaction is not systematically observable.

To simulate repair sequences of different kind in dialogues between an artificial conversational partner and a language learner in a communicative ICALL application, a more detailed operationalisation of the repair components (whatever they are) is needed. Specific linguistic resources that can be used for specific types of troubles and specific types of repairs need to be described on a level that can be incorporated in a computational model of other-initiated repair. More specifically, a model distinguishing between repair initiations and all other utterances is of interest. After that, the trouble source needs to be extracted from the repair initiation and an appropriate repair needs to be generated. I accept the challenge in this work.

Specific types of repair are in focus of SLA studies because they have been considered the indicators of learning. These are meaning negotiations (Varonis and Gass, 1985) and corrective feedback (Lyster and Ranta, 1997). The trouble sources in corrective feedback and meaning negotiation sequences are normally deviations from language standard and non-native-like constructions. I call this type of repairs *repair with linguistic trouble source*.

(Varonis and Gass, 1985) proposed a sequential model of meaning negotiation which consists of a trigger (trouble source), the indicator (repair initiation), response (repair carry-out) and reaction to response. Meaning negotiations can be initiated by the trouble-speaker and by the trouble-talk recipient, thus, they include two sequentially different repair types, which makes a turn-by-turn analysis very complex. Markee (2000) criticises this model as methodologically problematic, because it is only supported by data from institutional talk like teacher-student or psychotherapist-patient. The model suggested by Varonis and Gass (1985) does not take into account the differences in speech exchange systems and is, as Markee (2000) argues, not applicable for equal-power speech exchange systems like ordinary conversation (p. 85). Building on the work by Kasper (1985), Markee (2000) differentiates between repair during the language-centered phase of classroom interaction, and repair during the content-centered phase. Kasper (1985) found, that "participants oriented to different, and indeed more complex, patterns of repair" during the content-centered phase.

Corrective feedback corresponds to other-initiated other-repair from CA perspective. Many studies on typology of corrective feedback in language classroom data and Computer-Mediated Communication have been published (Lyster and Ranta, 1997; Panova and Lyster, 2002; Zourou, 2012). One of the recent classifications of corrective feedback in language classroom was presented by Lyster et al. (2013, p. 4) and will be taken here for comparison. It includes the following nine types of corrective feedback:

1. Conversational recast is a reformulation of a non-native speaker's utterance in order to resolve a communication problem. A frequent form is a confirmation check.
2. Repetition takes the form of a 1:1 repetition of a problematic expression marked with interrogative intonation in order to highlight the error.
3. Clarification request signals a misunderstanding or a communication problem by a phrase like "I don't understand".
4. Explicit correction contains a clear identification of an error with a reformulation of the learner's expression and an explicit presentation of the correct form.
5. Explicit correction with metalinguistic information signals an error explicitly, providing a correct form and information about, for example, grammar and pragmatics.
6. Didactic recast is a reformulation of learner's utterance in the absence of a communication problem.
7. Metalinguistic clue is a a metalinguistic statement aiming at eliciting a self-correction.
8. Elicitation directly enforces a self-correction, may often take the form of a content question.
9. Paralinguistic signal is a non-verbal elicitation of the correct form.

The classification was constructed based on oral data, not everything can be expected to be found in text-based chat communication. Moreover, some of the types of corrective feedback may appear too teacher-like to be applied in a Conversation-for-Learning. In addition, the categories in the classification proposed by Lyster et al. (2013) are not disjoint. For instance, it is not easy to make a clear distinction between confirmation checks and clarification requests, conversational and didactic recasts. In addition, clarification requests may be based on a repetition which overlaps with recasts. The overlapping classes make computational modelling of various correction types problematic. Therefore, I will look at other-corrections of linguistic errors in my dataset through the lens of computational modelling and describe patterns in other-corrections.

Error typologies as a base for correction are discussed in (Vicente-Rasoamalala, 2009, Sec. 4.3.3.). However, in the previous CA research, no dependencies between error types and correction types were found (Schegloff, 1987). Moreover, Schegloff (1987) emphasised the importance of "disengaging trouble (error and nonerror) from the practices employed to deal with it" (p. 216). In addition, not every error is corrected in a language classroom, and not every error is corrected in a Conversations-for-Learning. It needs a further investigation, under which circumstances an error could and should be corrected in an informal conversation with language learners, thus, when is a correction relevant.

Corrections may be delivered immediately after the error or later. Different time points of correction delivery in a language classroom have been discussed in (Vicente-Rasoamalala, 2009, p.147). In addition to the immediate corrections, delayed and postponed corrections may take place. Delayed corrections may appear at a transition point in conversation or at a subsequent point of the same lesson. Postponed corrections may occur in another lesson. The right point of correction delivery is an important issue for correction generation in a communicative ICALL application, therefore, I will look specifically at the timing features in my analysis of other-corrections.

While very detailed descriptions of teachers' reaction to learner errors in a language classroom already exist (see for instance (Vicente-Rasoamalala, 2009) for a detailed comparison of multiple classifications), a detailed description of sequential environments, practices and turn formats for corrections in a Conversation-for-Learning need first to be created in order to prepare a basis for computational modelling.

Jefferson (1974) shows that notes that error corrections are used in native speakers' talk as an interactional resource to negotiate and reformulate the current set of identities. She argues that "error correction is a systematic feature of speech, and [...] it is achieved by the application of a specific device, the Error Correction Format" (Jefferson, 1974, p. 188). For instance, she describes Contrast Class Errors including "error correction formats involving words having the same features with opposite values" (Jefferson, 1974, p. 187). Building on this idea, I will look at different error correction types which I call *error correction formats* following the terminology suggested in the beginning of this section.

Kurhila (2001) examines the selectivity of error correction in native/non-native speaker talk based on an analysis of a corpus of naturally occurring non-pedagogical conversations. The author suggests that native speakers' decision to correct may be explained by environments in which errors can be corrected in general. Kurhila (2001) concludes that grammatical errors produced by language learners are most likely to be corrected by the native speakers in non-pedagogical talk when they occur in a 'repetition slot' or when they can be used to initiate repair. Nonetheless, "despite the rather frequent occurrence of other-correction in these environments, other-correction is still constrained in its occurrence in NS-NNS conversation, as in NS talk." (Kurhila, 2001, p. 1107).

In light of the discussion in this section, my work seeks to contribute to the analysis of repair in chat-based Conversations-for-Learning in the following way:

1. Computational modelling of other-initiated self-repair when the native speaker is the trouble speaker requires a more detailed analysis of the interactional resources used to signal and locate troubles in comprehension.
2. An empirical base for other-corrections specific to the speech exchange system in focus needs to be created. Interactional resources used for corrections of linguistic errors need to be described and operationalised for computational modelling. The timing of corrections in chat needs to be described.
3. A communicative ICALL application needs clear decision criteria for correction, for which I seek to find an empirically grounded base.

These issues are the major focus of this research. Chapters 4, 5, 6 and 7 of Part II focus on the micro-analysis of a corpus of chat Conversations-for-Learning to prepare a basis for computational modelling, which I turn to in Chapters 8, 9 and 10.

2.2.6. Embedded corrections as form of implicit feedback

The difference between exposed and embedded corrections in native speaker only talk has been first described by Jefferson (1983) where she distinguished between corrections where *correcting* becomes an interactional business and those that happen by-the-way, without focusing on the *correcting* in the interaction. Jefferson defines embedded corrections "as by-the-way occurrence in some ongoing course of talk" noticing that "the talk which constitutes embedded corrections does not permit of accountings" (Jefferson, 1983, 95). She introduces the term *accountings* to refer to all those "attendant activities" of the correctings such as instructing, complaining, forgiving, apologising and ridiculing. Kasper (2004) found that non-native speakers activate their interactional roles first of all through accountings.

The examples of embedded corrections provided by Jefferson are all recordings of native/native speaker interactions and do not contain corrections of linguistic errors. The replacements in Jefferson's examples are: police → cops, wales → threads, kilns → kils, eve → night, pretty → beautiful. I only repeat here the sequences of the initial and replaced terms, the reader is welcome to review all the examples analysed by Jefferson (1983). Similarly, the analysis of corrections (including embedded corrections) of self-references in conversation performed by Lerner and Kitzinger (2007) shows, that despite the correctness of un-repaired version (*I* instead of *we* or the other way round) "speakers can select a self-reference term so as to fit the referent to the kind of action (or personal state) formulated within the turn"(Lerner and Kitzinger, 2007, pp. 538-539). With other words, phenomena besides corrections of linguistic errors are subject to repair and are in focus of studies on embedded corrections.

Embedded corrections in native/non-native speaker conversations are in focus of only a few studies in face-to-face interactions (Brouwer et al., 2004; Kurhila, 2006; Kaur, 2011; Hynninen, 2013) and in chat (Tudini, 2010). Hynninen (2013) prefers the term *embedded repair* in her analysis of English as lingua franca because "repair (in CA) refers widely to any kind of modifications that do not necessarily focus on correcting" (Hynninen, 2013, p. 122). (Tudini, 2010) characterises embedded corrections as " ...inexplicit indirect feedback", "correction with discretion". I will interchangeably use terms *embedded correction* and *implicit correction* to refer to this phenomenon.

Kurhila (2006) discussed embedded correction in institutional and every-day conversations between native speakers of Finnish and learners of Finnish as an additional language. He found out that embedded corrections are produced in institutional setting more frequently, arguing that it is because of the need for confirmed, clear information (Kurhila, 2006).

Brouwer et al. (2004) describe the procedure of embedded corrections in oral L2 talk in the following way:

> In embedded corrections, the speaker in the ongoing turn B corrects an item in a preceding turn A while doing some possible next action to this preceding turn A. Thus, the main work of turn B is on this next action and not on the correction. It is this 'next action' which is consequential for turn C, then, and not the correction; in other words, the speaker of the trouble-source turn A does not orient to the correction in turn B. Embedded corrections therefore do not open a 'side sequence' [...] but proceed with the main business of the sequence (p. 86).

The mechanism allowing to embed corrections into ongoing business of the talk without focusing on the activity of correction is related to procedures for "consecutive reference to same objects" in conversation (Jefferson, 1983, 90). Jefferson describes three possibilities to refer to objects in conversation:

1. The next speaker uses proterms to refer to the entities named by the prior speaker;
2. A term used by a speaker is repeated by the next speaker;
3. An *alternate* term is used by the next speaker instead of the term introduced by the initial speaker. An alternate pronunciation is a variant of this for spoken interaction. An alternate writing may be an equivalent for chat interaction.

Jefferson notes, "when a next speaker produces, not a proterm or a repeat, but an alternative item, *correction* may be underway".

Academic literature documents contradicting findings with regard to occurrences of embedded corrections in native/non-native speaker interactions. Brouwer et al. (2004) identified embedded corrections of linguistic errors as typical for L2 face-to-face interaction. In contrast, embedded corrections of linguistic errors were rare in the dataset used by (Tudini,

2010) for her study of native/non-native speaker chat interaction. She concludes that "further attention is required in a separate study to determine why dyadic online intercultural chat favours exposed correction to deal with pedagogical repair".

With regard to places of embedded correction in interaction, academic literature reports different situations. Kurhila (2001) found where corrections often occur, regardless of the linguistic surface of the trouble turn: corrections *en passant*, repetition-based answers to questions and third turn by the native speaker after a question-answer pair. Corrections *en passant* are referred to as corrections *on-the-fly* and classified by other authors as exposed corrections (Tudini, 2010), but the other two types belong to embedded corrections. As opposed to this finding, Brouwer et al. (2004) report that all errors corrected implicitly in their collection were located in a first pair part, and the embedded corrections were always part of the second pair part (Brouwer et al., 2004, p. 85).

A more detailed specification of the possibilities for "the speaker in the ongoing turn B" to correct "an item in a preceding turn A while doing some possible next action to this preceding turn A" is needed to generate embedded corrections for conversational agents and dialogue systems. Building the bridge between the findings of CA and CA-for-SLA in the area of embedded corrections and the needs of communicative ICALL and AI in the broader sense is one of the contributions of this dissertation. I analyse error types and correction formats for all pairs of turns consisting of the trouble turn and the embedded correction turn in Chapter 6 preparing a base for formal modelling of these phenomena. I introduce a formal model for embedded corrections in one sequential environment in Chapter 9. The model operates on the level of grammar and turn-taking. I discuss the technological requirements and the limitations of the current state-of-the-art tools for NLP in Section 9.3.

With regard to the analysis of learning, embedded correction present a particular difficulty. As Kurhila (2006) notes:

> Embedded correction is one reason why language learning is difficult to explore in authentic conversations between native and non-native speakers. In the laboratory data, the NNSs are being tested after they have been corrected and, thus, some claims can be made about the possible improvement or learning. In real NS-NNS conversations, however, correction can be introduced in a way that specifically makes it possible for the NNS not to respond to it. Therefore, the issue of learning (or having registered the new form) remains ambiguous (Kurhila, 2006, 43).

Recordings of long-term interaction like the dataset used for this work can help to understand, how particular items are used by learners and how they develop in conversation over time. Effects of native/non-native speaker chat on learning are not the main focus of this work. However, some observations and recommendations came out during this work and will be discussed in Section 11.3.3.

2.3. Conversation Analysis and Artificial Intelligence

Application of CA methods to the field of Human-Computer Interaction (HCI) is mentioned as a field of applied CA in CA text books (ten Have, 2007, Ch. 10), and the earliest work mentioned therein comes from late eighties where the problem of plans and situated action has been approached (Suchman, 1985). Later on, qualitative research methods, such as Conversation Analysis, Interaction Analysis and Ethnography have been used to solve various design problems in Artificial Intelligence (AI), Human-Computer Interaction (Peres and Meira, 2003) and Human-Robot Interaction (HRI) (Chee et al., 2010; Plurkowski et al., 2011; Sussenbach et al., 2014). Evaluation of interaction between technical systems have been also approached by qualitative research methods (Alemi et al., 2015; Robins et al., 2004; Lee and Makatchev, 2009; Lee et al., 2010; Sabelli et al., 2011). The technical systems in focus are first of all robots and dialogue systems, however augmented reality also caught researcher's attention as a tool for linguistic research to study interactional modifications caused by the technology (Pitsch et al., 2013c).

Studies described in this section achieved valuable results by applying qualitative research methods in HCI and HRI. Some of the results can be transferred into communicative ICALL. Some of the questions, that arose from the studies discussed below, are also relevant for the communicative ICALL research. At the time of writing this dissertation, all my attempts to find academic publications about using relevant insights form CA-for-SLA for design or evaluation of a communicative ICALL system were unsuccessful.

2.3.1. Conversation Analysis for Dialogue Processing

The relationship between CA and NLP was difficult in the beginning. This is well reflected in a review article *Does Conversation Analysis have a Role in Computational Linguistics?* (Hirst, 1991) of the book *Computers and Conversation* (Luff et al., 1990).

The difficulty of the relationship between CA and HCI was caused from my perspective by several facts. First, CA always analysed naturally occurring interaction, in the beginning audio-recorded phone conversations (Sacks, 1995; Schegloff, 2007). The speech technology that time was not mature enough, the majority of dialogue systems were text-based, and CA findings from audio recordings were not directly applicable to medially written human-machine dialogues. Second, various concepts of CA were not operationalised to be directly transferrable to dialogue processing, therefore, other theories have been preferred, such as Speech Act Theory (Searle, 1969) and Belief-Desire-Intention model (Georgeff et al., 1999). This tradeoff between the knowledge about conversation gained by CA and its technical implementability was (and still is) huge, which is clearly visible in the review of *Computers and Conversation* (Hirst, 1991). Nevertheless, Hirst (1991) concludes optimistically:

> There is a sense in which it is clear that CA *must* have a role in NLU [natural language understanding], because there is a sense in which ethnomethodology

> is just a small subfield of artificial intelligence (although that might come as a surprise to the ethnomethodologists). (p. 225, original emphasis)

Despite all difficulties in the beginning, a lot of research efforts have been invested to bring various insights gained from Conversation Analysis into Dialogue Processing. Several attempts have been made to involve CA into dialogue-based human-machine interaction. Waterworth (1986) reports how CA methods were applied to design a speech-based database inquiry system. Laughters have been found in CA to occur at as topic change (Holt, 2010) and a speech recognition study was build on this insight to estimate the time within the topic change borders (Gilmartin et al., 2013). Special attention was put to turn-taking (especially for speech segmentation) and adjacency pairs in dialogue, see for instance (Thórisson, 2002; Edlund et al., 2005) and references therein. Another topic addressed in CA which received a lot of attention in HCI is repair. I will review relevant literature in Section 2.3.2.

Human-Robot Interaction makes the major benefit from CA methods applied to design and analysis of multimodal interaction. I will discuss the achievements and how they can be transferred to communicative ICALL in Section 2.3.3.

2.3.2. Types of repair for conversational agents

Types of repair for conversational agents described in state-of-the-art academic publications are closely related to the application scenarios and user models. Because language learners are usually not considered as the main user group, the assumption that human users understand everything what an artificial agent or robot may say dominates the research on repair for conversational agents. This assumption is reflected in the two main problems addressed by research on repair for conversational agents:

1. Dealing with user's self-correction which may make speech recognition difficult;
2. Managing system's lack of information in order to satisfy user's request.

These two research areas may be found under keywords *self-repairs*, sometimes *speech repairs* (Zwarts et al., 2010) or *disfluencies* (Shriberg, 1994; Martin and Jurafsky, 2009), and *clarification dialogues* in AI and NLP publications. What is referred to by the term *self-repair* in speech recognition domain corresponds to user's same-turn self-correction (self-initiated self-repair) in CA terminology. Example 2.1 illustrates this type of repair and is frequently cited in academic literature on speech recognition.

Example 2.1. Self-correction (Schegloff et al., 1977, p. 363 (5)).

L:		Is his one dollar allright or should he send more
	→	than that for the p- tuh cover the postage.

User's same-turn self-correction is subject of research in (Purver and Hough, 2014). The application in focus is incremental recognition of self-corrections performed by users of spoken dialogue systems in real time. Only same-turn (same TCU) repair is considered. The system is trained and tested on data from Switchboard corpus of English phone conversations. Skantze and Hjalmarsson (2010) differentiate between ouvert and couvert self-corrections. The authors distinguish in addition between a segment and a unit repair.

Speech recognition errors may lead to further problems in maintaining the dialogue with the user and task continuation. Modelling human clarification strategies (repair initiations) in response to speech recognition errors has been approached in (Stoyanchev et al., 2013).

The majority of academic publications on same-turn self correction for speech recognition build upon the model for such repairs proposed by (Shriberg, 1994). The model suggests that the turn containing the same-turn self-correction can be split into a sequence of units with a specific function. The units are shown in Table 2.2.

until you're	*at the le-*	‖	*I mean*	*at the right-hand*	*edge*
start	reparandum	moment of interruption	editing terms	alteration	continuation

Table 2.2.: A sequential pattern of self-repair

Shriberg (1994) uses the term *reparandum* to refer to what is calld *trouble source* in CA. The sequential model shown in Table 2.2 distinguished between pauses (moment of interruption) and lexicalised means to focus on the replacement (editing terms). Both are interactional recourses used by speakers to signal trouble in production and to make a pre-announcement of a coming replacement (alteration).

The term *clarification dialogues* or *clarification sub-dialogues* is mostly used to describe repair sequences in AI to deal with insufficient information available for the systems after speech recognition or language understanding (Allen and Small, 1982; Hayes and Reddy, 1979; Maier, 1997; Gabsdil, 2003; Kruijff et al., 2008; Jian et al., 2010; Buß and Schlangen, 2011; Kruijff et al., 2006). The term *miscommunication* was introduced to distinguish between *non-understandings* (the system could not match user's input to a representation) and *misunderstandings* (the system matched user's input to a wrong representation) (Meena et al., 2015; Dzikovska et al., 2009). These repair types correspond to other-initiated self-repair when the user is the trouble-speaker (the machine initiates and the user carries out the repair).

Clarification dialogues have been studied from the point of view of managing lack of information to satisfy user's need in task-based dialogue systems, question answering systems, information systems and robotics. Therefore, mainly the case of other-initiated self-repair where the system does not (fully) understand user's input has been covered. For instance, VERBMOBIL translation system plays only an intermediate role in communication between people, and only seldom engages in active dialogues with the user, as for instance, in clarification dialogues (Maier, 1997).

Though in AI repair initiations are frequently referred to as *clarification requests*, they should not be confused with clarification request in SLA theory. In SLA theory this term is used to refer to only a particular form of corrective feedback (Lyster et al., 2013), but also to a dialogue move in meaning negotiation sequences (Varonis and Gass, 1985).

Repair initiations are usually generated from manually created templates specific for each trouble source type (Maier, 1997). In a speech based information retrieval system SPEAK! (Grote et al., 1997) clarification sub-dialogues may be system-initiated by means of intonation to clarify user's needs. In (Kruijff et al., 2008), robots need to negotiate what people say them, they know what is unclear to them and can produce clarification requests. Clarification is seen as a continual planning process. Rodríguez and Schlangen (2004) present a multidimensional classification-scheme for form and function of clarification requests and apply it to a corpus of German task-oriented dialogues for annotation. It is an attempt to analyse the structure of repair initiations and to improve a state-of-the-art classification scheme. Clarification dialogues described in (Quintano and Rodrigues, 2008) are required to disambiguate user's request in a question answering system.

Dzikovska et al. (2009) focus on *non-understanding errors* (the cases where the system does not find any interpretation for user's utterance) in a chat-oriented tutorial dialogue system for tutoring basic electricity and electronics. The research is focused on repair initiations produced by the dialogue system. The repair initiation is generated according to a single template: excuse followed by a reference to the trouble source followed by a request to reformulate the utterance. The system tries to elicit a reformulation of the problematic utterance from the user.

The huge body of research on user's self-correction and other-initiated self-repair when the user is the trouble-speaker signals the assumption that the user always understands everything. Rarely the opposite became subject of the study, however, the importance of the types of repair for artificial agents where the user display need for clarification is acknowledged in a body of academic publications, see for instance (Purver, 2004, 2006).

> It is also very unusual for systems to be able to understand and respond when the user asks a CR [clarification requests]. Designers (very sensibly) try to avoid user CRs by making system prompts as clear and informative as possible, and some- times training users with those prompts. However, as systems start to deal with complex tasks, wider domains and wider audiences, [...] it seems inevitable that they will have to deal with users asking CRs at some point. (Purver, 2006, p. 260)

Emphasising the importance of correct recognition of user's clarification requests, Purver (2004) provides an empirical, theoretical and implementation study of various types of clarification requests. I discuss Purver (2004)'s classification here in detail because it is used in many follow-up publications, for instance (Purver, 2006; Ginzburg et al., 2007; Ginzburg, 2012). Purver (2004) uses the HPSG framework to cover the main classes of the identified classification scheme. The relationship between the form and the function of clarification requests are investigated (clarification form vs. clarification reading). The clarification forms identified by Purver (2004) from a corpus of phone conversations are:

1. *Non-reprise clarifications* including such repair initiations as *What did you mean?*, *What did you say?*
2. *Reprise sentences* including full repeats of the trouble source utterances. Some of the examples discussed in (Purver, 2004) to illustrate this type of clarification requests can be re-analysed as topicalisation (Ex. 48 p. 62) and doing surprise (Purver, 2004, Ex. 47 p. 62).
3. *WH-substituted reprise sentences* including full repeats of the trouble source utterance where the trouble source is replaced by a question word.
4. *Reprise sluices* consisting of a bare question word or question phrase such as *To where?* and *Who?*.
5. *Reprise fragments* consist of a partial repeat of the trouble source utterance, the trouble source is specified more precisely, only the relevant part of the trouble source utterance is repeated.
6. *Reprise gaps* consist of a partial repeat of the trouble source utterance which projects that the trouble source immediately follows the repeated part.
7. *Conventional* repair initiations include all open-class repair initiations based on excuses, question words and *eh?*-like expressions.
8. *Gap fillers* include repair initiations dealing with trouble in production in contrast with all other preceding types of clarification requests.

Because different functions might be expressed by a clarification request of the same form, Purver (2004) analyses the *clarification readings* to cover the correspondence between the form and the meaning of the repair initiations.

1. *Clausal* reading can be mapped to phrases like *Are you asking/asserting P?*, *Is it X about which you are asking/asserting P(X)?* or *For which X are you asking/asserting P(X)?* where *X* represents the trouble source.
2. *Constituent* reading can be mapped to one of the following template questions: *What/who is X?*, *What/who do you mean by X?* or *Is it Y that you mean by X?*.
3. *Lexical* reading targets the surface from of the trouble source utterance and requires a repetition as a response. It can be mapped to questions like *Did you utter X?* or *What did you utter?*.
4. *Correction* reading can be mapped to the question *Did you intend to/should you have uttered X (instead of Y)?*.

In light of recent academic publications on repair initiations in cross-linguistic studies in Conversation Analysis discussed in Section 2.2.5, specifically (Dingemanse et al., 2014), several points for critiques may arise regarding this classification of clarification requests. For instance, some utterances may be formatted as repair initiations but have a different interactional function, such as expressing surprise and topicalisation (not listed as possible readings). In addition, repair initiations designed to deal with troubles in understanding are put together with strategies for dealing with troubles in production (e.g. *gap fillers*). From the CA perspective, *gap fillers* correspond to self-initiated other-repair, thus is sequentially completely different from all other types.Therefore, modifications in the classification proposed by Purver (2004) needed in order to better comply with studies in CA, and therefore better reflect the state-of-the-art in CA-informed dialogue research.

The general idea of having a map between various forms of doing a particular action in conversation (e.g. repair initiation formats) and the basic meaning behind these forms (e.g. request for repetition to resolve troubles in hearing) seems reasonable, while some inconsistencies in the classification of readings and examples provided in (Purver, 2004) can be identified. For instance, The correction reading is illustrated by three examples, one of them contains an other-correction (Purver, 2004, Ex. 70, p. 69) and the other two contain self-corrections (Purver, 2004, Ex. 71 and 72, p. 69), therefore it may be confusing how to interpret this reading from the perspective of the classification of various forms of repair other-initiations.

Ginzburg et al. (2007) analyse similarities between same-turn self-correction and other types of repair. They motivate their study by Ca work, however choose to use an existing formalism described for instance in (Purver, 2004) and criticised from the CA perspective earlier in this section. Nonetheless, (Ginzburg et al., 2007) succeeded in including same-turn self-corrections into a HPSG-based formalism, which previously aimed at covering only clarification requests. In this way they showed that there is a possibility to describe the repair system by one, complete formal model. However, the critique points regarding the relationship between repair initiations and turn taking and various ways of dealing with troubles in understanding and production remain actual for this extended model.

Example 2.2. Different types of causes for clarification used in (Schlangen, 2004, Ex. (12)) to illustrate the need for an extended classification.

a.	A	I ate a Pizza with chopsticks the other day
	B	A Pizza with chopsticks on it?
b.	A	Please give me a double torx.
	B	What's a torx?
c.	A	Please give me a double torx.
	B	Which one?
d.	A	Every wire has to be connected to a power source.
	B	Each to a different one, or can it be the same for every wire?

Schlangen (2004) analyses communication problems leading to clarification requests focusing on trouble source types (what caused the communication problem). Building on the classifications by (Clark, 1996) and (Allwood, 1995), Schlangen (2004) makes clear that a more fine-grained classification of causes for requesting clarification in dialogue may be needed. Specifically, the author agues that a model distinguishing between different cases in Example 2.2 is needed.

To approach different sources of trouble from the computational perspective, Schlangen (2004) suggests to use the classification scheme presented here in Table 2.3 and compared to the classifications introduced by (Clark, 1996) and (Allwood, 1995).

From the perspective of Conversation Analysis, factors like speakers' linguistic and professional identities and preferences play a role in speaker's selection of a specific format of a repair initiation. Speaker B in Example 2.2.b. positions herself as a novice in torx matters with her repair initiation, while speakers B in Examples 2.2.c. positions herself as knowledgeable in torx matters. In addition, utterances may be designed as repair initiations, but may in fact have a different function. For instance, the repair initiation produced by B in Example 2.2.a. may be analysed as a joke not requiring an explanation.

Level	Clark, 1996	Allwood, 1995	Schlangen, 2004
1	Execution & attention	Contact	Establishing contact
2	Presentation & identification	Perception	Speech recognition
3	Meaning & understanding	Understanding	3a: parsing
			3b: resolving underspecification
			3c: contextual relevance, computing the rethorical connection
4	Proposal & consideration	Reaction to main evocative function	Recognising speaker's intentions; evaluating resulting discourse structure

Table 2.3.: Models of sources of troubles in conversation.

Other-initiated self-repair when the machine is the trouble-speaker (the user initiates and the machine carries out the repair) is explored in (Gehle et al., 2014). Based on a corpus of video-recorded human-robot-interactions in a museum, the authors analyse multimodal interactional resources used by museum visitors to signal troubles in understanding robot's talk and dealing with robot's misunderstandings. The authors conclude that "different types of trouble [...] lead to a similar way of dealing with trouble" (Gehle et al., 2014, p. 367).

The potential user of a communicative ICALL system is a language learner who may have troubles in comprehension. While user-initiated repair has been subject of research of studies in human-robot interaction and dialogue systems, not much attention has been payed to text-based human-computer interaction. A chat-based Conversation-for-Learning may provide opportunities for the language learner to engage in repair sequences with linguistic trouble source. Hence, it may be expected, that language learners will initiate repair in order to deal with troubles in comprehension. This dissertation seeks to contribute to the research on computer-mediated native/non-native speaker communication by a microanalytic study of sequences of other-initiated self-repair when the native speaker is the trouble-speaker. Specifically, repair initiation formats and practices of referencing the trouble source in the repair initiations are of interest. In addition, practices of dealing with such repair initiations deployed by the native speakers may provide patterns for repair carry-out. Based on the results of the empirical study, the problem of computational modelling of system's reaction to the learner's repair initiation will be approached. The machine will need to recognise repair initiations, to extract the trouble source and to deliver an appropriate response. The results of the study will improve language understanding for dialogue systems targeting language learners, and may be included into user and expert models for communicative ICALL applications.

2.3.3. Conversation Analysis in Human-Robot Interaction

Micro-analysis of recordings of human-human interaction have been used to create patterns for human-robot interactions. Methods of Conversation Analysis focusing on talk-in-interaction and Interaction Analysis focusing on use of multimodal resources have been employed for this purpose. Specifically, researchers put attention to moments in interaction such as dialogue openings and closings (Pitsch et al., 2009; Bono et al., 2014), interactional resource coordination (Yamazaki et al., 2008), managing participation (Pitsch et al., 2013a; Katagiri et al., 2004), referential practices (Pitsch and Wrede, 2014) and engagement (Rich et al., 2010).

Sussenbach et al. (2014) use ethnographic methods to find an empirically grounded model for motivation in a task-based interaction setting with robots where robots play a role of a fitness companion for indoor cycling. Yamazaki et al. (2008) employed methods of Conversation Analysis to develop a museum guide robot that moves its head at interactionally significant points during its explanation of an exhibit. Patterns of coordination of verbal and non-verbal actions obtained from human guide-visitor interactions positively influenced visitors reaction to robot non-verbal behaviour. These studies showed that the transfer of interaction patterns from human-to-human interaction into human-robot interaction resulted in a positive interaction experience with robot companions.

The task of creation of a functional specification for an artificial companion has been approached by ethnographic methods (Yaghoubzadeh et al., 2013; Kramer et al., 2013). The focus of the study lied on acceptability of the companion by a specific group of target users, namely, elderly people and people with mental disabilities. However, it is a quite challenging task to evaluate non-existing technology, therefore, results of studies with a similar focus may be contradicting. Nonetheless, the authors succeeded in finding out how an artificial companion can be integrated in the daily routine of potential user groups by the analysis of what the people *really do* as opposed to other projects with similar ambitions where an artificial application scenario was invented and the potential users were confronted with it.

Inspired by Conversation Analysis, conversation openings with a robot museum guide have been studied by Pitsch et al. (2009). An Aibo robot dog acting as a museum guide in a Japanese museum was provided with a simple pause and restart strategy to ensure visitor's attention. The same strategy is frequently used by people in conversations when they notice that their conversation partners do not listen. The authors showed that, similar to human interactions, openings have significant effect on user's engagement in the continuation of the interaction. However, the robot showed only in 52,9% of the cases a "contingent entry into a focused encounter" (Pitsch et al., 2009, p. 991). A similar study was set up in a German arts museum with a humanoid Nao robot focusing on the influence of robot's gaze on visitors participation in interactions with multiple visitors (Pitsch et al., 2013a,b). Specifically, the situations of inclusion of an additional participant into an ongoing dialogue and disengaging a participant from a group were in focus. Pitsch et al. (2013b) analyse video-recordings of real-world interactions of museum visitors with a robot museum guide as compared to interactions with a human tour guide. Pitsch et al. (2013b) conclude that knowledge about interactional coordination, incremental processing and strategies for proactive shaping user's conduct are required for a robot museum guide in order to manage real-world interactional situations.

Schnier et al. (2011) used methods of Conversation Analysis to investigate the influence of an Augmented Reality device (head-mounted display) on interaction. They showed that forms of interaction change as compared to face-to-face communication, but their function remains the same, and that the changed forms are the result of negotiated adaptation processes. These findings confirm results reported in studies of Computer-Mediated communication saying that the communication medium changes the set of available interactional resources, but the new set is successfully used to execute specific interactional functions, for instance, dealing with trouble in interaction (Kim, 2003; Kitade, 2000).

All studies discussed in this section show that micro-analysis of human's understanding of social actions expressed through various modalities in interaction is important for the development of the interaction process and user's engagement or disengagement. Looking at every detail in dialogues between language learners and native speakers may provide valuable insights into the role and the consequences of each single contribution of each participant at every moment of an interaction. Patterns obtained from micro-analysis of human-human dialogues belonging to a specific speech exchange system may help to make dialogue systems and conversational agents better communicators. This dissertation makes a step in this direction. Supported by the findings in HRI domain, a successful application of CA methods for dialogue design in communicative ICALL may be expected.

Concluding remark

Different terminology used in CA, SLA, AI, CHI and NLP to refer to the same phenomena in interaction signals missing multidisciplinary connections between the communities. The other way round, different terminology traditionally used by different communities to refer to the same phenomena makes academic literature research difficult. This chapter is a tiny contribution to "translation" of the terminology and and attempt to strengthen the interdisciplinary connections between CA, SLA and AI.

Part II.

Analysis of Empirical Data for Modelling

Introduction

The purpose of this part of the dissertation is to prepare an empirically grounded basis for computational modelling of sequences in conversation, where language leaners and native speakers make their linguistic identities relevant in conversation. A specific focus is on repair models in conversation with learners. Conversation Analysis is chosen as data analysis methodology, because it allows to explore the dataset without any preconceptions of what may be found. For the structures that will be found, micro-analytic analysis of their structure is expected to deliver desired prototypes which, in turn, will provide a basis for computational modelling.

This part of the research is organised as follows:

Chapter 3 documents the results of the data exploration and provides a bird's eye perspective on various ways to construct identities of language learners and language experts in conversation by analysis of different practices where linguistic identities of chat participants are made relevant in the talk.

Chapter 4 describes and discusses different linguistic resources made available by learners in chat to initiate repair with linguistic trouble source when they did not understand native speakers' utterances.

Chapter 5 presents an analysis of types and turn formats of exposed corrections produced by native speakers to correct L2 errors produces by learners in the dataset.

Chapter 6 analyses sequential environments where embedded corrections occur.

Chapter 7 makes a step towards creating a preference system for other-corrections which prepares a basis for a decision model required for computational models of corrections.

3. Patterns for Expert-Novice Chat

Due to the way in which the participants for the data collection for the present study were acquired, it seems natural that the participants of the dialogues had particular expectations regarding their potential interaction co-participants. The native speakers knew that they will communicate with young people from Belarus who study German as a foreign language. The learners had been contacted through their university, and practicing conversation was one of the reasons for them to participate. Therefore, participants on both sides interacted from the very beginning through the lens of their expectations of whom they will be interacting with. For the German native speakers, it was a person who studies their native language. For the learners, it was a person who speaks the language that they study as the mother tongue. Such expectations become visible in the interaction through the phenomenon known as *recipient design* and *stances* (Pillet-Shore, 2012; Hutchby, 1995; Sacks et al., 1974).

In addition to the stances and recipient design influencing, participants of the conversation construct their identities of a language learner and a language expert in specific sequences of talk orienting to their differential linguistic expertise. This can be done in different ways and sequences of repair with linguistic trouble source is probably the most prominent of them.

Repair sequences with linguistic trouble source have been analysed in CA and SLA literature, and the findings were discussed in Chapter 2. However, the actions and practices used by chat participants to position themselves as learners or expert are not limited to repair sequences. In my data, I saw mainly three classes of such actions:

1. Face work and evaluation: learner's level of L2 proficiency is the subject of the talk in form of learners' excuses for mistakes, fishing for compliments, native speakers' (always positive) evaluations and encouragements;

2. Meta-talk and learning together: language learning becomes the subject of the talk when, for instance, native speakers offer help in language learning tasks and exam rehearsal;

3. Repair sequences with linguistic trouble source: explanations of expressions unclear to learners, corrections of linguistic errors, word searches, vocabulary checks, definition work and similar.

In addition to the practices of talk-in-interaction where participants' linguistic identities become relevant, the form of the chat language chosen by some of the native speakers in conversations with learners differs a lot from the form of the language used, for instance, in retrospective interviews with the researcher. In conversations with learners, a variation of German close to a German standard is preferred, for instance, every turn and every noun start with a capital letter. In contrast, the same native speaker writes everything in lowercase in the retrospective interview. I will discuss examples of both in Section 7.1 more detailed with regard to the definition of linguistic error and language standard in chat. For the purpose of this section it is important to keep in mind, that there are specific points in conversations, where linguistic identities of participants are relevant. Besides of these specific points, native speakers may choose to present themselves *constantly as role models*, as persons who use the language *correctly*.

I will analyse the three types of participants' orientation to their linguistic identities in the later parts of this chapter. The unit of analysis is a turn. A turn, in my understanding, corresponds to a message in chat. This is what the participants post and receive. They cannot intervene in a message-at-production prior to receiving the complete thing, even if the message is long and could be delivered in shorter units. For instance, turn parts belonging to different threads or topics can be delivered in separate messages or put in one long message. However, the speakers can start talking (that is producing a message) at any point of time, they do not have to fight for their right to talk. They take a turn when they think it is relevant, and they make it as long as they think it is appropriate. Putting pieces belonging to different threads or topics into one turn is an interactional resource that participants make available in order to deal with, for instance, troubles with network connection.

3.1. Face work and evaluation

One of the most frequent practices used by learners to position themselves as novices in my dataset was the excuse for making mistakes which they have made and have not yet made, in combination with a self-assessment as not-yet-fully-proficient speaker. Native speakers, in turn, replied with a positive evaluation of language knowledge and encouragement, positioning themselves as language experts in talk. In fact, such excuses and self-assessments were not "real" excuses, because there were not necessarily real mistakes, and they did not do any harm to the talk. The function of these excuses and evaluations was to maintain the face, as I will demonstrate with several examples.

Example 3.1. Apologise for mistakes

8	15:30:30	L07	das macht nichts...sofort bitte um Entschulding für meine Fehlerchen...´‘=)) *it does not matter... immediately ask for apologise [* error: typo] for my errors [diminishing] ... [smile]* it does not matter... I'm immediately apologising for my small mistakes
9	15:30:57	N03	kein problem! *no problem!*

The first prototypical sequence of this kind is shown in Example 3.1. The learner L07 makes her not-yet-perfect language skills relevant in turn 8 which the native speaker N03 might notice in the future. It takes place just after the exchange of greetings in the very first conversation when the participants have just met. The learner refers to her errors as "little mistakes" using a diminishing German form which marks the potential errors as not very heavy or not to be taken seriously. The native speaker accepts the apology. The fact that the learner apologises for all the "small mistakes" that she could make in the future shows that she assumes that N03 could notice them (it is his "territory of knowledge" (Heritage, 2012)). She positions herself as a novice and assigns the role of a language expert to the native speaker. By giving the acceptance of the apology in tun 9, N03 also accepts the role of the language expert that just has been assigned to him by the learner. Thus, the adjacency pair of apology-acceptance is used to construct identities of expert and novice in conversation.

Example 3.2. Excuse for mistakes and slow typing, learner's self-assessment of comprehension competence. native speaker evaluates and encourages.

10	15:33:19	L06	ich möchte mich zuerst für meine zahlreiche Fehlern entschuldigen. und du musst geduldig sein, damit so lange Zeit auf meine Antwort warten *I would like first to apologise for my numerous errors [* error: wrong plural]. and you must be patient so that [* error: wrong conjunction] such a long time for my answer wait* First, I would like to apologise for my numerous errors. And you need to be patient because you will have to wait a long time for my responses
11	15:33:37	L06	;)
12	15:35:13	N03	also bist jetzt hast du noch (fast) keinen fehler gemacht. und das ist ueberhaupt kein problem! ich kann gar kein russisch schreiben! *well, till now you have (almost) made no errors. and this is not a problem at all! I cannot write a single bit of Russian!*
13	15:35:23	N03	ist das ok wenn ich klein schreibe? *is it ok if I write small?* is it ok if I write only lowercase?
14	15:39:18	L06	Gott sei Dank! du kannst schreiben, wie du willst, hoffentlich verstehe ich alles. wenn nicht, dann musst du dich ein bißchen bemühen, damit mir das erklären *Thank God! you can write whatever you want. hopefully, I understand everything. if not, then you must make an effort a little bit so that [* error: wrong conjunction] explain me everything* Thank God! you can write however you want. I hope, I understand everything. if not, you will have to try a little bit more to explain those things to me
15	15:39:40	N03	na klar! *of course!*

With the "warning" and the apology in turn 8, L07 gets the permission for all possible "mistakes" that may happen in the future. The permission is "granted" by the acceptance of the apology. The announcement of the possible errors allows her to save her face in communication using a language that she has not yet fully mastered.

In Example 3.2, the learner L06 excuses for all the mistakes and errors that she has produced and will produce in the future in turn 10 (at the very beginning of the conversation after the participants have just met and exchanged the greetings). In addition, she informs the native speaker that it could take longer for her to write in a foreign language using computer keyboard. The native speaker evaluates learner's German writing skills in turn 12 and positions himself as an expert in German. In addition, he says that it is okay to make mistakes in order to encourage the learner to keep on trying.

The native speaker switches his role from an expert to a non-expert in the same turn pointing out that he cannot write learner's native language at all and emphasising her area of expertise. Further in turn 13, the native speaker negotiates spelling rules suggesting that writing everything with lowercase does not count as an error. The learner agrees in turn 14 and positions herself again as not-yet-fully competent speaker. She evaluates her comprehension competence as not sufficient to understand everything. In addition, she negotiates her right to ask for explanations and negotiates the conventions that it is alright to do definition talk. The native speaker agrees in turn 15.

Example 3.3. The native speaker evaluates the level of knowledge of the learner without request in the continuation of the talk.

21	15:47:51	L06	ich habe Prüfungen in Deutsch, Englisch, Literatur und Linguistik. Also 4 Prüfungen und 4 Tests in diesen Disziplinen. Um Prüfungen abzulegen, sollen wir zuerst Tests schreiben *I have exams in German, English, Literature and Linguistics. So, 4 exams and 4 tests in these disciplines. In order to take exams, we shall first take the tests* I have exams in German, English, Literature and Linguistics. So, 4 oral exams and 4 written tests in these disciplines. In order to take the exams, we shell first write the tests
22	15:48:55	N03	das klingt nach viel arbeit. aber deutsch schaffst du auf jeden fall, dein deutsch ist sehr gut. studierst du auf lehramt? *this sounds like a lot of work. but you will make it in German any way, your German is very good. Are you studying to become a teacher?* this sounds to be much work. but you will make it in German any way, your German is very good. Do you study to become a teacher?
23	15:52:42	L06	danke. ja,aber ich werde nicht nur Lehrerin, sondern auch Dolmetscherin,wie uns versprochen wurde. Aber ich kann mich in dieser Rolle nicht vorstellen *thank you. yes, but I will not only become a teacher but rather also a translator, as it was promised to us. However, I cannot imagine myself in this role* thank you. yes, but I become not only teacher but also translator, as it was promised to us. But I cannot imagine myself in this role

Example 3.3 shows that a positive evaluation of learner's language skills is not necessary a direct response to learner's weak self-assessment. The learner tells in turn 21 about the planed exams that will happen some time after the conversation.

Example 3.4. Many errors, apology, focus on content

165	13:58:47	L03	ich meine, das hängt von unserer Mentalität. bei und wollte fast niemand in einem Dorf wohnen, weil dort zu wenig Arbeitsplätze und und überhaupt Angebote vorhanden sind. die Menschen auf dem Lande trinken zu viel, vielleicht hast du darüber gehört, deshald streben die Jugendlichen in die Städte umziehen. und die die Kinder auf dem Lande haben wenige Ansprüche und hiesige Ausbildung gilt als unqualifiziert, und die Kinder aus dem Lande lernen oft in der Oberschule in der Stadt. aber über solche Situation kann man nur pauschal genommen reden. es gebt natürlich viele Ausnamen. und überhaupt das Leben ist heutzutage imme mehr nach Wsten orientiert und viele wohlhabende Bürger ziehen in die Dürfer um, nach dem europäischen Beispiel, aber sie arbeiten doch in der stadt *I mean, it depends on our mood [* missing separable prefix]. at and [* error: probably typo] wanted almost no one in a village live, because there are too few jobs and in general opportunities there. the people in rural areas drink too much, maybe you have heard about that, therefore [* error: typo] aspire [* error: word choice] the teenager to move [* error: word choice] in the cities. and the the [* error: double det.] children in rural areas have few demands and local education counts as unqualified, and the children from rural areas learn [* error: word choice] often in secondary school in the city. but one can speak about such situation only in average. there are [* error: typo] of course many exceptions [* error: typo]. and in general the life is nowadays more and more [* error: typo] oriented to the West [* error: typo] and many wealthy citizen move [* error: typo] to the villages, according to the European model, but they still work in the city* I mean, it depends on our mood. here almost no one wanted to live in a village, because there are too few jobs and in general opportunities there. the people in rural areas drink too much, maybe you have heard about that, therefore the young people seek to move to the cities. and the children in rural areas have low standards and local education counts as unqualified, and the children from rural areas often go to a secondary school in the city. but one can speak about such situation only on average. there are of course many exceptions. and in general the life is nowadays more and more oriented to the West and many wealthy citizen move to the villages, according to the European model, but they still work in the city
166	14:03:58	L03	Entschuldigung für zahlreiche Fehler;) *Sorry for numerous mistakes [smile]*
167	14:05:26	N02	ich verstehe alles :) Das ist echt interessant [... simplified] *I understand everything [smile] It is really interesting [... simplified]*

In turn 22, the native speaker first evaluates the new information and then encourages the learner. The encouragement takes a form of an evaluation of learner's language proficiency. After the assessment of language skills was made relevant in the beginning, N03 provides an additional positive evaluation of L03's skills in the continuation of the talk in turn 22. It is approximately the middle of the first conversation, this talk consists of 55 turns. The learner thanks for the positive evaluation in turn 23 which looks more similar to a response to a compliment than to a receipt of a "grade". Here, the native speaker did not really have the time to assess learner's skills in German, especially for the assessment of the probability of passing the exams.

Learner's excuses for L2 errors can be also followed by a confirmation of comprehension, as Example 3.4 illustrates. The Learner makes many mistakes in a long narrative turn 165. The native speaker needs time to read the long message and to reply to the content of the message. The learner has therefore a chance to re-read the just-produced message and to detect errors in it, and to even produce an apology turn 166. The native speaker first replies to the apology in turn 167 with a confirmation of comprehension instead of an explicit expression accepting or rejecting the apology.

Example 3.5. The native speaker evaluates the level of the knowledge of the learner without request in the final closing sequence.

404	22:00:10	N01	Ich wünsche dir schon einmal viel Erfolg bei deinen letzten Prüfungen! *I wish you good luck on your last exams!*
405	22:00:23	N01	Und natürlich alles Gute für die Zukunft ;-) *And of course all the best for the future [smile]*
406	22:01:01	N01	Hier noch mal ein dickes Lob: Du sprichst sehr gut deutsch !!!! *Here again a thick praise: you speak very good German !!!!* Here again a special word of praise: you speak very good German !!!
407	22:02:27	L02	Vielen Dank! Das ist für mich sehr ehrenvoll das von dir zu hören!))) *Thank you very much! This is very honourable for me to hear this from you! [smile]*
408	22:02:45	L02	Ich wünsche dir auch alles Gute! *I wish you all the best, too!*
409	22:03:52	N01	Sehr gerne! (Es ist ja auch die Wahrheit!) Vielen Dank für deine Wünsche! Tschüss !!! *You are welcome! (This is just the truth!)* *Thank you very much for your wishes!* *Bye-bye !!!*
410	22:04:27	L02	Tschüss!!! *Bye-bye!!!*

In her apology in turn 166, the learner positions herself as novice because she made mistakes. At the same time, she positions herself as a person who *can* recognise the errors in the long just-produced turn. The apology is accepted indirectly with the "I understand everything" in turn 167, at the same time showing that linguistic accuracy is not that important in this chat, but the fluency and the comprehension are important. N02 returns to the topic immediately, focusing on content of the conversation.

Example 3.5 shows a different case of native spaeker'a evaluation in turn 406. The learners thanks for the evaluation in turn 407. This example presents the final closing sequence of the last conversation (farewell for ever). The learner does not make the level of language proficiency relevant in conversation (there are no excuses or self-assessments), thus, this evaluation vas not elicited by the need to maintain the face. The participants neither exchanged the contact details to continue the communication as their private business, nor they planed to meet again in some way. The evaluation is produced after 8 long dialogues, so that the native speaker had enough time to get an impression about real language proficiency of the learner. The evaluation as part of the final closing sequence discloses the view of the native speaker on the past conversations as “what we did here is practicing conversation, and you did your job very good”. The learner shares this view which is confirmed by her “thank you... very honourable...” in the response.

Example 3.6. The native speaker evaluates learner's knowledge of literature.

...	...	...	[The participants talk about literature, and come to the book "Sofies Welt.]
199	20:47:15	L02	Das Buch muss interessant sein. Es brachte mir "Den kleinen Prinz" von Exurery in Erinnerung. Ein kleiner Junge lernt mit der Welt der Erwachsenen, mit ihren Charakteren, Überzeugungen und dadurch auch mit sich selbst kennen. Früher, in der Schule, verstand ich dieses Buch nicht... *The book must be interesting. It brought "The little prince" by Exyrery [* error: typo] to my mind. A little boy learns with [* error: no prep. required] the world of adults with their characters, beliefs and in the same way with [* error: no prep. req.] himself to know. Earlier, in school, I did not understand this book...* The book must be interesting. It brought "The little prince" by Exypery to my mind. A little boy gets acquainted with the world of adults, with their moods, beliefs, and in the same way, with himself. Earlier, in school, I did not understand this book...
200	20:48:48	L02	[simplified]
201	20:49:05	N01	Ja, das ist ein sehr guter Vergleich, so ähnlich ist "Sofies Welt" vom Prinzip her. Mensch, ich muss dir mal ein dickes Lob aussprechen. Du kennst dich wirklich sehr gut aus :-) *Yes, it is a good comparison, "Sofie's World" is in principle similar. Man, I just have to express a thick praise for you. You really know a lot [smile]*
202	20:49:34	N01	Wenn du mal irgendwann Zeit hast, musst du es mal lesen. Autor ist ein Norweger. *If you have time sometime, you must read it. The author is a Norwegian.*

The evaluations can be connected to each other with backlink tokens (*noch mal*, *wieder*). This "here *again* a special word of praise" in turn 406 refers to a different conversation between these two participants where they talked about reading books and the native speaker was impressed by the learner's knowledge of German literature. A simplified excerpt is provided in Example 3.6 for a better understanding of the relationships.

In Examples 3.6 and 3.5, the native speaker does *real* teacher-like evaluation in addition to the expression of being impressed. This praise is not elicited by the learner, but is a reaction to an impressing performance, as if it was provided by a teacher or a tutor. Finally, the native speaker presents himself as someone having the right to evaluate which is accepted by the learner already earlier than the last closing sequence.

3.2. Learning in meta-talk and collaborative learning

The data collection took place when the most of the learners were at their examination period. Preparations for the exams and the results of the exams were therefore frequently topic of the talk. In such parts of the dialogues, the participants talked about learning, but this meta-talk did not necessarily contribute to the construction of the learners' identities as novices native speaker' identities as experts.

Example 3.7. Learner marks the conversation as training

94	18:05:59	N03	was hast du heut s gemacht? *what have you today [* error: typo] so [* error: typo] done?* what did you do today?
95	18:08:06	L07	heite war ich wie immer in der Uni...die letzte Vorprüfung abgegeben.... jetzt muss ich mich auf die Prüfung vorbereiten....die habe ich am montag schon-=((*today [* error: typo] was I as always at the uni... the last test submitted.... now must I refl. to the exam prepare....it have I on Monday already [sad smiley]* today I was at the university, as usuall... I had the last preliminary test.... now I have to prepare for the exam.... it is already on Monday
96	18:08:36	N03	oh
97	18:08:46	N03	na dann musst du viel lernen,oder? *ok, then you have to learn a lot don't you?*
98	18:09:47	L07	nun schaffe ich irgedwie..... erste Prüfung ist Deutsch...so habe ich gute Praxis jetzt :) *will manage it somehow... the first exam is German.... so, I have good practice now [smile]*
99	18:10:11	N03	haha
100	18:10:18	N03	muendlich oder schriftlich?!!! *oral or written?!!!*

However, meta-talk about the exam preparation can be taken up by the native speakers as an invitation to rehearse the exam. The meta-talk can be transformed by the participants in on-line exam preparation and collaborative learning.

One such sequence of meta-talk ending up with an exam rehearsal is shown in Examples 3.7 and 3.8. The meta-talk starts in turn 95 in Example 3.7 as the learner tells the native speaker about the upcoming exam and her preparation after his invitation to tell about her day. The learner explicitly declares the conversation with the native speaker as conversation practice in turn 98. The native speaker interprets it as an invitation to "practice" the exam. After the learner tells the topics of the upcoming exam, which she obviously does not like, in turn 101, the native speaker starts questioning her about these topics (I decided not to show here the entire dialogue because the content of the talk is not the focus of this section, but the reader is welcome to look at the entire sequence in the published dataset). Later in turns 145-147 the participants explicitly claim the talk-in-progress as an exam rehearsal.

Example 3.8. Continuation of the talk from Example 3.7. The participants engage in an exam role play. N03 plays the role of the examiner, L07 has the student role.

101	18:11:08	L07	leider muss ich meinem Lektor über EU, Menschenrechte oder Globalisierung erzählen... *unfortunately I have to tell my lecturer about EU, human rights or globalisation...*
102	18:11:20	L07	:'(
103	18:12:09	N03	das sind doch aber spannende themen! *but those are exiting topics!*
104	18:12:26	N03	was denkst du denn ueber die EU und ihre derzeitige situation? *what do you think about the EU and its current situation?*
105	18:13:45	L07	insbesondere EU-Gremien und so viele Daten aufeinmal viel zu viel für ein Mädchen :) *especially EU-committees and so many data [* error: lexical choice] at once [* error: missing space] to much for a girl [smile]* especially EU-committees and so much information at once to much for a girl
...	...	...	[simplified: turns 106-135 native speaker tries to elicit learner's talk on these topics]
136	18:37:44	N03	aber was ist nicht gut? *but what is not good?*
137	18:37:50	N03	in der EU? *in the EU?*
138	18:37:58	N03	an der globalisierung? *with globalisation?*
139	18:38:00	N03	...
140	18:38:02	N03	;-)

141	18:39:24	L07	hahaha........ Arbeitaplatzverlagerung, Ausbeutung im Süden, Überflutung von Informationen usw....!!! :) :) :) *[lexicalised laugh] workplace relocation [* error: typo], exploitation in the South, information flood etc....!!! [smileys]* hahaha workplace relocation, exploitation in the South, information flood etc....!!!
142	18:40:33	N03	ok
143	18:40:40	N03	usw klingt gut *etc sounds good*
144	18:40:43	N03	;-)
145	18:42:24	L07	willst du jetzt meine Prüfung mal repetieren..? :) *want you now my exam repeat [* error: lexical choice]..? [smile]* do you want to rehearse my exam now?
146	18:42:42	N03	ja *yes*
147	18:42:51	N03	also ich dachte das koennte nuetzlich sein *well I thought this could be useful*
148	18:43:02	L07	:)
149	18:43:04	N03	ich habe einen bachelor in politikwissenschaft *I have a bachelor in Political Science*
150	18:43:07	N03	:-)
151	18:43:14	N03	aber ich will dich auch nicht nerven *but I don't want to annoy you*

A variant of exam post-work that I found in my data was a sequence where the learner asks the native speaker questions that she had in her written test. She already knew the answers and just wanted to show the native speaker what the exam was about. However, in the majority of all cases when the native speakers offered help for exam preparation, it did not go beyond the meta-talk. Native speakers invited the learners to ask them questions when they prepare for the examinations. However, the learners did not make use of this opportunity to get help.

The meta-talk about language learning can also lead to discussions of grammar systems of languages spoken by participants, as illustrated in Example 3.9. The learner explains in turn 72 why English is so difficult for her to learn. The native speaker compares the English grammar to the German in turn 74 pointing to a simpler system of determiners in English. The learner makes her general difficulties with determiners relevant in turn 76. Then, the roles of expert and novice change. The native speaker positions himself as novice in Russian and Belorussian asking, how many determiners there are in those languages. The learner, in turn, positions herself as expert in Russian/Belorussian by telling, that these languages do not have any determiners (turns 79-80). She formulates it as a surprise source. The roles change back in turns 81-84 where the native speaker formulates his turn in a way that is not easily accessible for the learner. The learner positions herself again as novice by the repair initiation in turn 82 which starts a repair sequence

To sum up, the meta-talk about language learning can be used by the chat participants as a preparation to collaborative learning. However, collaborative learning does not necessarily emerge after such meta-talk. Collaborative learning can relate to practicing conversation to a particular topic of interest (e.g., to prepare for an exam or an interview) or focus on specific aspects of the language (e.g., grammar or lexicon). Specific aspects of language become topic of the talk also in sequences of repair with linguistic trouble source, which are subject of the next section.

Example 3.9. Discussion of determiners after meta-talk (simplified).

72	18:07:40	L08	)nun für mich Englisch ist sehr schwer... besonders die Zeitformen..."ein dunkler Wald"))) *[smile] well for me English is very difficult... especially the tenses... "a dark forest" [smile]* well, English is very difficult for me... especially the tenses... quite confusing
74	18:09:44	N04	dafür gibt's im Englischen nur einen Artikel, während es auf deutsch 3 sind :-) *on the other hand there is only one determiner in English, while there are 3 in German [smile]*
76	18:11:04	L08	haa)) ja Artikel ist ein Problem)) *[laugh, smile] yes, determiner is a problem*
78	18:14:05	N04	wieviele Artikel gibt es im (Weiß-)Russischen? *how many determiners are there in (White-)Russian?*
79	18:14:51	L08	für dich wird es eine Überraschung!!!=))) *for you will it a surprise!!! [smile] [* error: missing main verb]* it will be a surprise for you!!! [smile]
80	18:15:06	L08	wir haben KEINE Artikel)))) *we have NO determiners! [smile]*
81	18:16:35	N04	oh... auch gut... dann erschließt man sich das wohl aus dem Zusammenhang? *oh... good too... then one must deduce it probably from the context*
82	18:17:43	L08	verstehe ich nicht *I don't understand*
83	18:18:16	L08	aaa verstehe verstehe=)) *aaa understand understand [smile]* oh, I see, I see
84	18:19:15	N04	ah gut... hab grad überlegt, wie ich das umformulieren soll :-) *ah good... I was thinking how I should paraphrase it [smile]*
85	18:19:30	L08	wir haben Endungen des Substantivs, die Kasus zeigen *we have endings of nouns which show the case*

3.3. Repair with linguistic trouble source

As already pointed out by many authors and the discussion shown in Chapter 2, one can recognise that sequences of *repair with linguistic trouble source* also referred to as *peda-*

gogical repair, (Tudini, 2010) offer an excellent opportunity to study participants' orientations to their linguistic expertise in conversation. In this section I analyse how differently initiated and differently carried-out repair sequences help to disclose participant's orientations to their differential language expertise in chat conversation for learning.

3.3.1. Native speaker's self-initiated self-repair

Same-turn self-initiated self-repair is normally not visible to the recipients of the repaired turn, because it is completed before the turn is posted. In contrast, self-initiated self-repair which is done later than same-turn, becomes visible for the turn recipient. Both, learners and native speakers corrected some errors in their talk later than same turn if they made some errors. However, there are also multiple instances of self-initiated self-repair produced by native speakers which do not deal with errors in their own talk. They rather *anticipate* problems with comprehension from the learner's side and explain some terms in their talk to the learner without request. I referred to this type of self-initiated self-repair as *proactive explanations* in an earlier work (Danilava et al., 2013a).

Example 3.10. The native speaker explains the concept of "state exam".

59	18:26:44	L04	und hast du dein Studium in diesem Jahr absolviert? *and did you complete your studies this year?* and did you finish your studies this year?
60	18:27:46	N02	nein, ich bin 2009 fertig geworden und habe jetzt zwei Jahre mein zweites Staatsexamen gemacht, das ist eine Lehrerausbildung direkt an den Schulen *no, I graduated in 2009 and it has now been two years since my second state exam, this is a teachers' training directly at schools*

A cultural note: the concept of *Staatsexamen* (госэкзамен) exists in the Russian language, too, but it has a different meaning in the professional education system in Belarus where the learners are from.

Example 3.10 contains in turn 60 an explanation of the word *Staatsexamen*, the learner did not mark this word as somehow unclear, but the native speaker anticipated potential difficulties in comprehension and explained the concept. In this way, the native speaker positions herself as an expert in German language (explanation of a word) and in German culture (knowledge about detailed of teachers' education) and in her profession (teacher). When N02 provides additional information on teacher's training, which was not specifically inquired, with the assumption that the concept *Staatsexamen* might be not (fully) clear for the learner, the learner is put by the native speaker to the category of potential novices in German language and in German teacher's education systems. This example also contains an embedded correction in turn 60.

The assumptions about the chances for the learner to comprehend the just-produced talk can also be based on the previous interaction history. For instance, if a particular expression was not clear to the learner and was marked in a repair sequence as trouble source, and then the same expression is used with a different sense, an explanation (self-repair) might be produced without request. I illustrate in Example 3.11 a proactive explanation of an expression that caused a repair initiation in the preceding talk of the participants.

The native speaker uses an expression in turn 386 in which the meaning is not completely clear. The learner initiates a repair in turn 388 providing a candidate understanding. The learner positions herself as a novice in this way. N04 clarifies the meaning of the expression in turn 390 confirming the candidate understanding suggested by the learner. Later in the same conversation, the native speaker uses the same expression again in a different sense. He provides an explanation of the expression and positions himself as a language expert who also understands and anticipates potential difficulties that the learner might have with a different meaning of the same expression. The native speaker contributes to the construction of his own identity as a language expert and also as a helpful and collaborative conversation partner.

Example 3.11. Self-initiated self-repair as a proactive explanation.

386	20:39:52	N04	alles klar? *everything clear?* everything alright?
387	20:40:05	L08	ist es nicht zu spät? *is it not too late?*
388	20:40:53	L08	alles klar- alles in Ordnung? *all clear- everything fine?*
389	20:41:38	N04	nein, es ist nicht zu spät, hier ist es ja 1 Stunde früher :-) *no, it is not too late, here it is 1 hour earlier [smile]*
390	20:41:55	N04	und ja, alles klar = alles in Ordnung. *and yes, all clear = everything fine.*
391	20:42:33	N04	gerade spielt Schweden gegen England, sehr spannend! *Sweden is playing against England right now, very exiting!* Sweden is playing against England, very exiting!
392	20:42:41	L08	au fsolche Weise, alles ist klar!) *in [* error: typo] this way everything is clear! [smile]* in this way everything is clear!
			`later in the same session`
473	22:04:47	L08	bestimmt)ok, NATIVE04_FN ich muss jetzt meine Spickzettel zur nächsten Prüfung ausschneiden) *sure [smile] ok, NATIVE04_FN I must now my scribs for the next exam cut [smile]* sure, ok, NATIVE04_FN I need to cut out my cheat sheet for the next exam now
474	22:06:36	N04	alles klar. (das heißt diesmal: "verstanden"). Dann wünsche ich Dir viel Erfolg bei der Prüfung und eine gute Nacht! *all clear. (this means this time: “understood”). Then I wish you good luck for the exam and a good night!*

Proactive explanations in the form of self-initiated self-repair are not limited to the repeated use of a previously utilised expression. The inferences about the need to repair can be done on a higher level of abstraction, as the dataset suggests. For instance, if a learner has difficulties in comprehension of abbreviations in the past, and the native speaker uses a new abbreviation in the just-produced turn, the native speaker can choose to explain the new abbreviation without request from the learner's side.

To sum up, the need for self-initiated self-repair is not only determined by actual "errors" in the just-produced turn, but also by participants' stances and the interaction process. The native speaker assumes that there might be a need for an explanation because their co-participants somehow oriented to their language expertise as not-fully-professional. The decisions for proactive explanations can be made based on the past interactions or assumptions about learners' language knowledge.

3.3.2. Other-initiated self-repair with linguistic trouble source

If we ask someone a question, we normally expect, that this person can answer this question, thus, this person has the necessary expertise or necessary knowledge to answer ("territories of knowledge" or "epistemic stances" (Heritage, 2012)). If the learner asks the native speaker for explanation of a just-produced turn or part of it, then the leaner demonstrates that (a) his/her own knowledge is not sufficient and (b) the native speaker's expertise is expected to be high enough to clarify the problem. Example 3.12 illustrates this distribution of expertise. I will analyse this excerpt later from the point of view of turn design for repair sequences. Here I look at this excerpt from the perspective of how expert-novice relationship is constructed in talk-in-interaction.

Example 3.12. Other-initiated-self-repair when the native speaker is the trouble-speaker.

122	13:08:39	N04	zugegeben, ich war dieses Jahr auch noch in keinem See, aber so langsam könnte man das mal ins Auge fassen :-) *admitted, I was this year not in a lake, either, but slowly one could reach this into the eye [smile]* I admit I was this year not at a lake, either, but I could slowly consider it [smile]
123	13:10:47	L09	ins auge fassen? *reach into the eye?*
124	13:11:56	N04	das heißt _hier_ etwa soviel wie "planen" oder "bald mal machen" *it means here something like "to plan" or "to do soon"*
125	13:12:16	L09	:) :)
126	13:12:22	L09	klar *clear* okay

In turn 122, the native speaker uses an idiomatic expression which is marked as problematic by the learner in turn 123. The native speaker has to determine what kind of trouble does the learner have with this expression. In turn 124, the native speaker explains the meaning of the expression. This makes visible (for the learner and for the researcher) that the native speaker interpreted the repair initiation as the need for filling a gap in learner's language knowledge. Turns 125-126 show that it was a correct interpretation.

Example 3.12 presented an instance of other-initiated self-repair (OISR) when the native speaker is the trouble-speaker. However, it is also possible for the learner to be the trouble-speaker and for the native speaker to other-initiate self-repair. The practice of questioning may be chosen by native speakers to initiate repair (clarification requests, repetition-based repair initiations). Such repair initiations produced by native speakers sometimes deal with real communication problems, and sometimes function as didactic corrective feedback. The goal of the questions is either to clarify a communication problem (caused by non-native-like expressions) pushing the learner at the same time to correct, or the goal is to elicit the production of a correct expression in absence of a communication problem (not equivalent to classroom elicitations "How do we say that in German?"). Since the native speaker is the participant who has the right to judge about what is the correct and what is not correct, it is the native speaker who positions himself as an expert, assigning the role of a language novice to the learner.

Example 3.13. Other-initiated self-repair when the learner is the trouble-speaker

416	20:59:09	L08	nun...Moskau wie Stadt(im großen und ganzen) hat mir nicht beonders gefallen: sehr viel Lärm,die Leute sind "ohne Gefühl" *well... I did not like Moscow as a city [* missing blanc] (in general) much [*error: typo]: very much noise, [*missing blanc] the people are "without feelings"*
417	21:00:13	L08	aber die Architektur ist wunderbar!!man kann eine Gebäude eine Stunde lang beobachten. *but the achitecture is wonderful!! [* missing blanc] one can observe one building [*error: wrong gender] one hour long.*
418	21:01:36	N04	ich war leider noch nie dort. Von Fotos kenne ich den Kreml. *unfortunatelly, I have never been there. I know the Kremlin from pictures*
420	21:02:32	L08	ja, ich war dort. es ist sehr maßstäblich *I was there. it is very full-scale [* error: lexical choice]*
421	21:03:30	N04	maßstäblich? *full-scale?*
422	21:04:02	L08	hmm...sehr sehr groß, umfangreich *[discourse marker]... very very big, wide* hmm...very very big, wide
423	21:07:34	N04	ah, ok!

A detailed look at the Example 3.13 will help to understand how the profiles of expert-novice are constructed in OISR-sequences when the learner is the trouble-speaker. The learner choses to use in turn 420 the word *maßstäblich*, which is a legal German word but has a different sense. It does not make much sense in the context of the talk. The native speaker initiates a repair by copying the trouble-source and marking it as problematic (question mark after the word) in turn 421. The learner is forced to describe the sense of what she was going to say with other words (and not to explain the meaning of *maßstäblich*). The native speaker accepts the explanation in turn 423 signalling that the communication problem is resolved.

Why do the repair initiations look similar but the repairs are different? Participants' expectations and attitudes towards their chat partners (stances) play a role in their interpretation of the trouble, and influence their choice of interactional resources for repair. First, both native and non-native speakers are expected to know, what they wanted to say. In addition, native speakers are expected to understand what learners say in native speakers' native language. Therefore, a different word with a *different meaning* is delivered by the native speakers to carry out repair initiated by the learners. In contrast, learners deliver the same meaning expressed by different words after repair other-initiated by the native speakers.

3.3.3. Self-initiated other-repair: word searches and definitions

In this section I discuss interactional practices used by learners to deal with difficulties with the production of a particular meaning. As reported by the participants in the retrospective interviews, they usually use the most common strategy of looking up words from dictionary. However, in some cases the learners preferred to ask the more knowledgeable peer for help. Example 3.14 shows how a word search sequence may be organised.

Example 3.14. Word search in two stages

430	21:14:17	L08	wie heißt es auf D., wenn man selbt mit dem Auto fährt? *how is it called in G., if oneself drives a car?* how do you say it in German, if you drive a car?
431	21:15:22	N04	ich fahre mit dem Auto *I drive a car*
432	21:15:50	N04	hab ich Deine Frage richtig verstanden? *did I understand your question correctly?*
433	21:16:01	L08	nein, und wenn man KANN mit dem Auto fahren *no, and if one CAN drive a car*
434	21:16:19	L08	führen? *to lead?*
435	21:18:19	N04	ich fahre, du fährst, er/sie/es fährt, wir fahren, sie fahren, ihr fahrt *I drive, you drive, he/she/it drives, we drive, they drive, you [pl.] drive*
436	21:18:45	N04	Verzeihung, stehe ich gerade auf dem Schlauch? *Sorry, am I not getting it?*
437	21:20:19	L08	=)auf Engl- drive *[smile] in Engl- drive* in English - drive

438	21:23:15	N04	ah, ob es heißt "mit dem Auto fahren" oder ob anstelle von "fahren" etwas anderes gesagt wird? *ah, if it is called "to drive a car" or if we use something else instead of "drive"?*
439	21:23:21	N04	nein, es heißt "Auto fahren" *no, it is called "to drive a car"*
440	21:23:40	N04	war das die Frage? *was this your question?*
441	21:24:34	L08	ach, ja)so, ist das richtig:du kannst mit dem Auto fahren, d.h. du hast ein Fahrerschein. ich habe richtig verstanden? *ah yes [smile] so, is is correct: you can drive a car, this means you have a driver [* error: overgeneralization] license. did I get it right?* ah yes [smile] so, is is correct: you can drive a car, this means you have a driver license. did I get it right?
442	21:25:50	N04	ja, ich habe einen Führerschein und ich fahre selbst :-) *yes, I have a Führer license and I drive myself [smile]* yes, I have a driver license and I drive myself
443	21:26:08	N04	(kein blöder Witz: es heißt wirklich "Führerschein") *(no stupid joke: it is really called Führer license)* (this is not just a stupid joke, we really say Führer license)
444	21:29:31	L08	klar. ich will im August in die Schule gehen um einen Führerschein bekommen. HAst du schon lange F. bekommen? *clear. I want in August in the school go in order to a driver license obtain. Have you already long D. get?* I see. I want to go to a driver's school in August and get my driver license. Did you get your D. long ago?
445	21:30:31	N04	ja, vor mehr als 10 Jahren *yes, more than 10 years ago*

The word search sequence is divided into two stages. First, the learner wants to know if the word *fahren* (En.: *to drive*) is used in German to say that a person drives a car by herself. As it becomes clear later on, she wants to tell the native speaker that she is going to get her driver license and to learn how to drive. It is not completely clear at the time point of turn 430 where this request is posted, that she might have difficulties with *production*, and that the request in turn 430 is in fact a self-initiation of other-repair. After multiple attempts to clarify the meaning of this request (turns 432-437), the native speaker finally understands the question and presents the word which the learner has been searching for.

In the second stage, the learner uses the German for drive - *fahren* - to formulate drivers license in turn 441. However, a different base word in German is used for that. The native speaker corrects this created word by presenting the correct German word for driver license in turn 442, emphasising that it was not just a joke in turn 443 because the word has a semantic relation to the word *Führer*. In turn 444, the learner finally uses the word that she was searching for and say what she probably wanted to say closing the repair sequence in this way.

Word searches can be performed in a less prominent manner. For instance Kurhila (2006) describes "repair as response to uncertainty" in oral data. Language learners may display hesitations and uncertainty during production, and the native speakers may respond to it with help to find the target (Kurhila, 2006, Sec. 3.3.1). I did not find such close-to-oral displays of hesitation in my chat data, only explicit searches (like "I don't know what is the German for X"). It is possible that my participants preferred to search in the web or look up in dictionaries for words that they want to say. Specific characteristics of responsiveness in chat allow the participants to take their time for production, as opposed to oral conversations. However, Marques-Schäfer presents instances of self-initiated other-repair when the learner has troubles with production and displays hesitations by symbolic means in chat (Marques-Schäfer, 2013, Excerpt 51, p. 171).

3.3.4. Other-corrections of linguistic errors

Exposed corrections of linguistic errors produced by the learners are obviously *the* places in conversation where native speakers position themselves as language experts. Embedded corrections, in contrast, offer a lot of space for interpretation. The differential linguistic expertise is just applied in talk, but it is by itself not the business of the talk. This is a kind of hidden, indirect way of being expert. However, the learners took up the corrected units after embedded corrections. Therefore, it can be also seen as implicit positioning. Doing the things right is also a kind of being expert. I look at these two types of corrections from the point of view of making differential language expertise relevant in conversation in the remainder of this section.

Exposed corrections

Exposed corrections are probably the most "teacher-like" type of positioning as language expert in conversation. The participants reported that they preferred to be corrected by the native speakers more frequently so as to improve their language learning. However, exposed corrections disrupt the conversation, therefore, the native speakers corrected learners' error very sparingly, and some of them almost did not correct errors at all.

Example 3.15 shows a prototypical sequence of an exposed correction. The learner L09 produces an expression in turn 49 which is marked as error by the native speaker N04 in turn 51. N04 rejects the word used by the learner (sein) and replaces it with a correction (haben). This replacement is framed by accountings (left-hand side: *falls ich korrigieren darf*; right-hand side: *smile*). He labels what he was doing in turn 51 as correction. The learner thanks for the correction and takes up the corrected expression (turns 52-53).

Besides this prototypical structure of exposed correction (trouble source – correction – uptake), correction variants come with more or without any accountings, without uptake or with delayed uptake. However, the exposed corrections always come with a component which is marked explicitly as wrong for some reasons, and another component which would be a better alternative from the corrector's perspective. This change is described by Jefferson (1983) as a sequence of replacements in the corresponding sequence of turns: Speaker A uses term X, next speaker uses a different term Y, speaker A, in turn, may keep on using the new term Y or go back to X. I will discuss various turn formats of exposed corrections in Chapter 5 in detail.

Example 3.15. An exposed correction of a linguistic error.

49	19:40:59	L09	Zur Zeit bin ich frei um Diplomarbeit zu schreiben *at this time I am [* error] free in order to write [* error: 0-determiner] diploma thesis* at the moment I have free time to write my diploma thesis
50	19:41:15	L09	danke schön! :) *thank you very much [closing of a previous sequence]*
51	19:41:58	N04	Falls ich das korrigieren darf: Du "hast" frei, nicht "bist". :-) *If I may correct you: you “have” free, not “are”. [smile]*
52	19:42:48	L09	Danke *Thank you*
53	19:43:16	L09	Ich habe frei :) *I have free [smile]* I have free time

Embedded corrections

If embedded corrections are opportunities for the native speakers to position themselves as experts implicitly, then they may offer chances for learners to position themselves as novices implicitly on the same token. By comparing the structure of exposed correction (trouble source – correction – uptake), one can find that embedded corrections match this prototypical structure, too.

Example 3.16. An embedded correction of a linguistic error. Chat on 11.06.12.

228	20:15:33	L02	Dann brauchst du irgendetwas Beruhigungsmittel einzunehmen))) *Then you need to ingest something tranquilliser [* error] [smile]* Then you need to ingest something tranquillising
229	20:16:02	N01	Genau, zur Beruhigung meiner Nerven :-) *Exactly, to tranquillise my nerves*
...	...	...	14.06.2012 uptake
291	20:03:44	L02	Ja! Ist dein Herz in Ordnung? *Yes! Is your heard all right?*
292	20:04:57	L02	Hast du keine Artzneien zur Beruhigung deiner Nerven eingenommen ?) *Did you ingest no remedy [* error: spelling] to tranquillise your nerves ? [smile]* You did not ingest any remedy to tranquillise your nerves, did you?

In both exposed and embedded corrections, learners may take the correction up immediately or later. The repaired version of the trouble source was used by learners in my dataset later than immediately after the embedded correction. Delayed (later in the same conversation) or postponed (in a different chat session) uptake make the analysis of learning difficult.

Example 3.16 shows an instance of embedded correction with learner's postponed uptake. This sequence is spread over two conversations, one on 11th June and the other on 14th June 2012. In the conversation from 11th June, L02 uses a non-native-like expression in turn 228 which has been implicitly corrected by N01 in form of an acknowledgment in turn 229. In the conversation the other day, L02 uses the correct expression in turn 292, as it was presented to her in native speaker's embedded correction. This example also provides evidence that *noticing* (Schmidt, 1990), which is considered important for learning (Mackey, 2006), is possible even without focusing on errors explicitly. Consequently, embedded corrections provide opportunities for learning.

Example 3.17. The learner encourages the native speaker to correct linguistic errors

130	19:01:18	L08	danke für Besonderheitn der Sprache=)für mich ist es sehr gut , dass du meine Fehler korrigierst) *thank you for the specialities [*error: typo] of the language [smile] for me it is very good that you correct my errors [smile]*
131	19:05:01	N04	wenn es nicht nervt, gerne :-) *if it is not annoying, with pleasure [smile]*

Because other-corrections are dispreferred and face-threatening, face work becomes an important part of exposed corrections in form of accountings. However, learners perceive error corrections as helpful, and encourage native speakers to correct, as Example 3.17 illustrates.

After a preceding lengthly sequence of several repairs initiated by the learner and the native speaker (see Example 4.5), learner L08 encourages her partner N04 to correct her errors in form of an expression of thanks and explicitly addressing her preference for correction. N04 agrees to correct "if it is not annoying" making visible that the dispreferred character of corrections is clear to him.

What else can be done by corrections

My dataset contains only positive examples of the chat users as helpful and supporting nice persons. This might be explained by the instruction to interact over many weeks, therefore chat participants payed attention to their relationships with their partners and handled the "face" of the others very carefully. The interaction setup in the study of Marques-Schäfer was a different one, the participants could come and go when they wanted and the chat was a group chat with many participants. The author reports cases of "mis-use" of exposed corrections for aggression in an untutored session (non-native speaker-non-native speaker group chat) (Marques-Schäfer, 2013, p. 192).

In Example 3.18, the learner L makes an orthographic error (probably just a typo) in turn 1 which L self-corrects in turn 2. Typos are considered unimportant and are normally not corrected in chat. However, the result of the typo is a legal German word, and it provides the other learner N room to interpret it as a morphosyntactic error (missing verb) and to produce a correction. The other-correction in turn 4 contains a correct pair of pronoun + verb "sein" in Präsens Indikativ (present simple indicative), which is the very basic level of German and is taught in the very beginning when people start learning German. From turns 1 and 2 it is clear that L has already mastered this level and is more advanced. When N teaches L how to conjugate "sein" in turn 4, it cannot be interpreted by L as "help to learn the language" but rather as face-threatening and aggressive. It is not surprising that L reacts with aggression in turn 5. N continues to teach L how to conjugate the verb "sein" in turn 6 referring to this activity as "teaching for free" in turn 8, which is pure provocation.

Example 3.18. Aggression in form of an exposed correction. Original excerpt (Marques-Schäfer, 2013, p. 192, Excerpt 67), turn numbers added.

1	19:14:53	L	S: ich ja nicht so schlimm oder?) *S:I [*error: typo] surely not so bad, or? [smile]*
2	19:15:00	L	*ist **is*
3	19:15:34	S	nee. ich wusste es eigentlich schon - und es gibt auch schlechtere maniere :) *nope. Actually, I new it already - and there are also worse manners [*error: wrong plural]*
4	19:15:36	N	ich bin* *I am**
5	19:15:54	L	N: halt deine klappe ok? *N: shut your trap, ok?*
6	19:16:00	N	er ist* *he is**
7	19:16:07	N	ich schreibe ich spreche nicht *I am writing, I'm not speaking* (...)
8	19:16:26	N	dazu lehre ich dich kostenlos *in addition, I am teaching you for free*

With the correction in turn 4, N positions himself as a language expert and makes the differential language expertise relevant in conversation. With this positioning, he presents himself as more advanced and more knowledgeable as L, which does not correspond the actual situation and is not accepted by L. The correction in turn 4 was not meant as "willing to help to learn" as it was the case in all the corrections in my dataset. There might be different reasons for this behaviour of N, which are not analysable for me from this excerpt. Important for this work here is, that there are also other things which can be done by corrections, besides a correction.

3.4. Preliminary findings and discussion

The goal of this chapter was to describe and to analyse how linguistic identities of a language novice and a language expert emerge in a chat-based Conversation-for-Learning between native speakers of German ad advanced DaF-learners. Three parts were found important for the analysis:

- Participants' expectations and attitudes towards their partner's language knowledge,
- Participants' own style of language use, specifically in chat allowing more freedom in orthography,
- Participants' orientations to their linguistic identities in conversation.

Although sequences of repair with linguistic trouble source are the most obvious way to construct linguistic identities of language novices and language experts, other forms of positioning as a novice or an expert can be found in instant messaging dialogues between language learners and native speakers. These forms include face-work, evaluation, meta-talk about learning and collaborative learning.

The purpose of this analysis is to prepare an empirical basis for data-driven modelling of conversations with language learners with an artificial conversation partner. I summarise and discuss the findings in Section 3.4.1 and make a first attempt to transfer them into context of computational modelling of an artificial conversation partner that helps a language learner to practice conversation in Section 3.4.2. I will use the concept of *interaction profiles* introduced by Spranz-Fogasy (2002) in order to build a bridge between the empirical reality and the world of computing.

3.4.1. Participants' orientation to their linguistic identities

I described and analysed three major types of practices where linguistic identities of a language expert and a language novice are co-constructed in text-based instant messaging dialogues between advanced learners of German and German native speakers.

1. Face work and evaluation: learners' language proficiency is the subject of the talk.
2. Meta-talk about learning and collaborative learning: language learning becomes the subject of the talk.
3. Repair with linguistic trouble source: linguistic issues become subject of the talk or are handled implicitly.

Independently of sequences where participants orient to their linguistic identities in conversation, participants can choose to present themselves as representatives of a specific category. For instance, native speakers can mark their own membership in the category of fully-proficient language users and become a role model for the learner.

Face work and evaluation

Learner-initiated face-work sequences frequently exploit practices of apology: learners apologise for their L2 errors. The purposes of these excuses were fishing for compliments, warnings in order not to be laughed at (face preservation), justifications of errors and correction post-work. Native speakers handle the apologies not only as apologies doing nothing else but apology, for which an accepting or a dis-accepting response may be the preferred next. Instead, native speakers position themselves as more knowledgeable language users and helpful and encouraging interaction co-participants. They do it in form of a positive evaluation and assessment of learners' language proficiency, which frequently have the function of encouragement and empowerment. A ritualised "no problem" was rarely a response to such an excuse.

In addition, learners put their limited linguistic competences in focus in form of a negative evaluation of their production and comprehension abilities. The social actions wrapped in these practices are mostly related to face work. Learners check if they are accepted as competent interactants despite their limited linguistic competence. They present themselves as not-yet-fully-proficient but willing to learn the native language of their partners. Table 3.1 presents examples of adjacency pairs for apology-based learner-initiated face work sequences.

First pair part: learner	Second pair part: native speaker
Excuse for mistakes	Positive evaluation, encouragement Positioning as language novice in learners mother tongue Compliments Ritualised "no problem"
Negative self-assessment in production	Encouragement, empowerment Compliments Positive evaluation
Negative self-assessment in comprehension	Promise to explain Encouragement, empowerment

Table 3.1.: Learner-initiated adjacency pair prototypes for apology-based and self-assessment-based face work sequences

I did not find any negative evaluation provided by native speakers to learners in the dataset. In addition to the positive evaluations in a second pair part in face-work sequences initiated by the learner as summarised in Table 3.1, there are teacher-like performance evaluations initiated by the native speaker after special "achievements" and in the end of the communication. It needs further analysis in a separate study, which metrics for the evaluation can be used in a free chat, and where in conversation such evaluations are appropriate. For the models of long-term interaction, sequences of sequences as shown in Examples 3.6 and 3.5 raise specific questions in terms of turn design with back-links to events in the past (like *noch mal* in Example 3.5 turn 406).

In addition, participants may negotiate what the conversation should be. Is it expected to be more focused on form and accuracy or do only the content and the fluency matter? Are errors and error corrections acceptable? Every answer to these questions leads to a specific variation of the speech exchange system and to a different "incarnation" of linguistic identities of a language novice and a language expert, and a different *interaction profile* (Spranz-Fogasy, 2002).

Meta-talk about learning and collaborative learning

In Section 3.2 I discussed various sub-dialogues in which language learning becomes subject of the talk. Such sub-dialogues include talk about comparison of languages spoken by participants of the conversation, talk about "doing conversation practice" and study and doing exam preparation in form of a role play and question discussion. Three distinct corresponding roles emerge for the native speakers: a peer, a more knowledgeable peer and examiner in a role play, respectively. The learners may take one of the three complement roles: an equally knowledgeable peer, a less knowledgeable peer and a student in a role play, however, the participants might not always accept the roles assigned to them by the partner, as Example 3.8 shows.

Role play differs from the other two types of talk in the following way. The practices deployed in examiner-student role play are restricted to the interview-like talk. Then the "examiner" has the right to ask questions and to evaluate, and the "student" has the obligation to answer. In other words, participants simulate a non-equal power speech exchange system of an exam within their equal-power speech exchange system. They embed a classroom-like talk in a free conversation, which does not oblige them to such form of talk by itself.

When learners and native speakers talk about differences in their mother tongues or about particular difficulties in languages that they speak as foreign languages, they become sensitive to specific issues in grammar or lexicon. However, they do it in chat in a very "light" form, they do not discuss rules of grammar for hours, like it sometimes happens in a language classroom. In contrast to short face-work sequences described in the preceding paragraphs, sub-dialogues where learning becomes subject of the talk can become very long. Models of such sub-dialogues can be created similarly to models of task-based dialogues for dialogue systems. However, an artificial conversation partner will need to determine, when one of such sub-dialogues or topics becomes relevant. The dataset used for this work may provide the necessary information for the model.

The participants engaged in lengthly discussions of properties of shared languages when:

- They collected information about the partner in the beginning of communication and wanted to know, which other languages their partner speak.
- They mentioned events related to language learning, like exams at the university or travel plans.

Such lengthly continuations of the topic sequentially were described in CA literature as *non-minimal post-expansions* like topicalisation and reworking of the first pair part or the second pair part (Schegloff, 2007, pp. 155-168).

Practicing specific conversational situations becomes relevant when events related to performance and evaluation of learners' language performance outside of the chat become relevant. For instance, when the learners have to prepare for language tests or exams, it might be appropriate to offer help in preparation and to practice in form of a role play. Results in the area of task-based dialogue systems (Hjalmarsson et al., 2007), serious games (Wik and Hjalmarsson, 2009; Sagae et al., 2011) and micro-worlds (DeSmedt, 1995) for SLA can be applied for designing such sub-dialogues. Topic identification, topic tracking and generation of a topical talk are also related research areas (Makkonen et al., 2004; Breuing et al., 2011; Breuing and Wachsmuth, 2012).

Repair with linguistic trouble source

Examples in Section 3.3 present various types of repair where the linguistic identities of a language novice and a language expert have been made relevant. Repair with linguistic trouble source was found in all types of repair (self-initiated, other-initiated, self- and other-repair). Different types of repair discussed in Section 3.3 have different purposes in dialogues with learners. I found the following systematics, which are summarised in Tables 3.2 and 3.3:

1. Native speakers' self-correction pro-actively repairs potential troubles, anticipating learners' difficulties in comprehension.
2. The learner self-corrects errors of form.
3. Other-initiated self-repair. Depending on the trouble speaker, such sequences have different purpose.
    a) The learner is the trouble-speaker, the native speaker corrects errors.
    b) The native speaker is the trouble-speaker, the learner seeks explanation of unclear expressions.
4. Self-initiated other-repair (the learner is the trouble-speaker). The learner addresses own difficulties in production initiating various kinds of definition work.
5. Other-correction (the learner is the trouble-speaker). The native speaker provides implicit or explicit corrective feedback on learners linguistic errors.

Tables 3.2 and 3.3 show that while native speakers are eligible to initiate and carry out repair caused by own and learners' linguistic choices, learners are only responsible for their own linguistic troubles.

RI\ RCO	Self	Other
Self	Correct errors of form and typos	Word searches and definition work
Other	*OISR-based corrections of learner errors*	*Corrections of learner errors*

Table 3.2.: Types of repair with linguistic trouble source when the learner is the trouble-speaker. RI: repair initiation, RCO: repair carry-out

RI\ RCO	Self	Other
Self	Pro-active explanations of potential troubles	
Other	*Explanations of unknown words and not completely understood expressions*	

Table 3.3.: Types of repair with linguistic trouble source when the native speaker is the trouble-speaker. RI: repair initiation, RCO: repair carry-out

Different strategies can be used by language learners to deal with trouble with understanding in text chat: they can search for translations and explanations using various online and offline resources, and they can ask their partners for help. Section 3.3.2 illustrates how learners may initiate repair to elicit an explanation of an unclear idiomatic expression in form of a repair proper (Ex. 3.12). On the other hand, the environment of other-initiation of repair may be used by native speakers for the purpose of correction of linguistic errors (Ex. 3.13). In both cases, the environment of other-initiated self-repair is used to construct linguistic identities. However, depending on who initiates and who carries out the repair, the emerging linguistic identities are different. The learner initiates the repair in Example 3.12 and positions herself as a language novice. The native speaker initiates the repair in Example 3.13 and positions himself as a language expert. The corresponding repair carry-outs are different with respect to their function and speaker's linguistic identity.

Language learners may use different strategies to deal with problems in production. Similarly to dealing with troubles in understanding, learners may use various online and offline information resources, machine translation tools and dictionaries to find the word that would in their opinion express their thoughts and emotions as close as possible. In addition, language learners may choose to ask their more knowledgeable conversation partners to help them. Multiple instances of word searches and definition work document that the sequential environment of learners self-initiated other-repair may be used by the learners to find the required vocabulary. Linguistic identities of a language novice and a language expert emerge in this way in informal conversations.

Native speakers appear as more knowledgeable language users in sequences of other-corrections of different types. If a deviation from language standard is worth a correction from the perspective of a native speaker, the native speaker may initiate and carry out a repair of such an error. They may do it in a more explicit (e.g. exposed corrections) or a less explicit form (e.g. embedded corrections). Although other-corrections of linguistic errors are generally dispreferred in informal conversations, learners perceive them as helpful. Nonetheless, the participants share the understanding of the face-threatening potential of such corrections. This becomes visible in form of accountings and explicit negotiations where participants make clear that corrections are acceptable, and even desirable.

In addition to sequences where participants' orient to their linguistic identities in talk, other possibilities of showing the linguistic identities exist. Specifically, native speakers may position themselves as role models and proficient language users throughout all conversations constantly by paying specific attention to their own language production.

Native speaker as a role model

Text-based instant messaging conversations with language learners may raise a question "What is the most appropriate language form to be used across conversations?" The answers that the native speakers who participated in the data collection chose were different, for instance:

1. We use instant messenger, chat is the same as "talk in written". Then more oral style in syntax and orthography was chosen, and a faster typing pace had a higher priority than correct orthography. This strategy of language use was chosen by the majority of the native speakers in the dataset and is frequently reported in the literature.
2. I communicate with a language learner and want to use the language as correct as possible to provide positive examples for the learner. Then the correct form of the language had a higher priority than a faster typing.

In both cases, native speakers become role models: the former shows the learner how to do text chat in German, the latter emphasises the correctness. However, the decision, what is better for the learner, is not just a simple selection between these two variants of language form. As I will show in Chapter 7, orthography is an interactional resource which can be used in chat to regulate social closeness.

3.4.2. Doing being a language expert: towards interaction profiles for an artificial conversation partner

Correction of L2 errors is not the only way for an artificial agent to construct the identity of a language expert. Other types of sub-dialogues described in this chapter need to be taken into account, too. However, the identity construction is a joint product of the machine and the user. If the user does not accept the machine as an expert, the machine cannot do much.

The identity of a language expert is constructed by two things simultaneously: 1) doing being expert - as discussed in this chapter, and 2) applying the own expertise in overall doing (e.g. overall correct "native-like" language use - whatever it is). The grade of acceptance of the artificial conversation partner as a language expert by the user can be seen in user's reaction to corrections, explanations, but also in presence or absence of other forms of orientation to linguistic identities in talk described in this chapter. Technical limitations in language understanding and generation which are available to the current dialogue technology such as conversation agents, dialogue systems and chatbots, will play a role in interaction. Learners accepted the native speakers as "role models", and therefore they took up corrections, asked for explanations and engaged in definition work sequences. It needs to be investigated in a separate study whether users of an artificial conversation partner would accept it as a more knowledgeable user of the language that they learn. However, further analysis of long-term interaction between language learners and native speakers is required prior to setting up such a study.

All of the native speakers became a language expert in conversations with learners, but the specific language expert "shape" is different for each of the native speakers. The concept of the *interaction profiles* introduced by Spranz-Fogasy (2002) and discussed in Section 2.2.4 can help to build a bridge between the empiric reality and the technology. The work on interaction profiles was an attempt to grasp and to operationalise the individual factors of each interaction participants and the adaptive behaviour of the interaction participants towards other participants taking into account the interaction history. One of his findings has a special relevance for this work. He showed that each participant of the interaction has only a *limited* influence on his own interaction profile and image. As Spranz-Fogasy (2002) mentions, if someone becomes an outsider, it is less the result of her their own interactional activities but rather a product of interactional activities of *others*. Spranz-Fogasy (2002) illustrates with his analysis of a TV debate, how a participant of the debate becomes a troublemaker in interaction through the contributions of the other participants, and not through his own talk (Spranz-Fogasy, 2002, Sec. 2.1). The consequences for computational models of a conversational partner that simulates a language expert are that the conversational system will have to decide on one hand which "incarnation" of a language expert from all the available options to chose, and on the other hand to adapt to the possibilities provided by the user in the interaction and the opportunities for such behaviour that emerged trough the interaction history.

Every native speaker of the dataset can be characterised in terms of practices chosen to orient to the linguistic identities. All of them positioned themselves as language experts in conversation or have been put in this position by learners. However each of them used different resources for that and combined the resources in different ways. This, in turn, lead to different pictures of the persons behind. The arsenal of the resources that can be used to position oneself as a language expert in conversation was described in the previous chapters of this part. Now I summarise how different interaction profiles can emerge in conversation by using parts of the arsenal.

The Teacher

1. Writes full sentences all the time, does not change this even if the learner makes it more "oral". Is a role model with his correctly formed German, very polite and politically correct.
2. Evaluates learner's performance from time to time. Evaluations are connected by backlink tokens.
3. Creates environments in conversation for language learning tasks, e.g. uses difficult idiomatic expressions consciously with the goal to give the learner a chance to learn them. Checks their understanding with comprehension checks. Tells jokes with double word meaning (= elicitations of repair initiation).
4. Selects topics for talks related to German language and German literature.
5. Recommends books for reading.
6. Corrects errors sometimes.

The Adaptive

1. Very reactive to learner's style. Imitates learner's self-repair formats, turn length, orthography.
2. Very reactive to learners emotions. Uses more emoticons if the learner does use a lot, and less emoticons, if the learner does not use them much.
3. Prefers to let-it-pass. Mutual understanding is more important than form and accuracy.
4. Corrects either in embedded form or by repair-initiation formats when learner's errors make the understanding difficult.
5. Reacts to learner's fishing for compliments by compliments, positive evaluation of language proficiency and encouragement.
6. Reacts to learners excuses for errors by confirmation of comprehension.

The Chaotic

1. Is always in hurry, no time to talk and establish a kind of closeness, making appointments is the most frequent activity.

2. Is not consistent in spelling and does not commit to conventions (writes sometimes ue instead of ü, sometimes ü, sometimes nouns with an initial capital, sometimes everything lowercase).
3. Consciously smuggles grammar errors into his messages, but does not react to learners repair initiations.
4. Error corrections are rather random and formatted without any visible system.
5. Very creative in self-repair carry-outs after learners' other-initiations.
6. Embeds language learning tasks into the conversation in form of exam rehearsal.

The More Knowledgeable Friendly and Helpful Peer

1. Writes very "oral" utterances, writes lowercase mixed up with correct spelling and grammar, uses stretches and emoticons to express emotions.
2. Offers help in learning tasks and exam preparation regularly.
3. Corrects many errors in both, exposed and embedded way.
4. Encourages the learner in form of positive evaluation, diminishing the errors.
5. Provides detailed explanations for requests for definitions and word searches.
6. Helps to remember spelling of difficult cases by memos.
7. Reacts to excuses for errors with positive evaluation and encouragement.
8. Uses rich palette of means for repair carry-outs: metalinguistic information, examples, paraphrasing, hyperlinks to examples, code switching.

The construction of linguistic identities in an interactional achievement and product of joint activities of both participants of each dialogue. As (Hosoda, 2006) reports, it is normally the learner who starts making the differential language expertise relevant, and sometimes the learner needs several re-initiations of such "doing being novice" before the native accepts the role of the expert. My dataset does not confirm this finding. The learners, L03, L04 and L05 did not ask much their partner N02. They preferred to search for words in online dictionaries instead of repair initiation. These three learners did not position themselves as novices. This is why N02 did not have a chance to position herself as a language expert.

However, N01 and N04 are deviant cases in my dataset. N01 *proactively* introduced more complex linguistic material to provoke repair initiations, he initiated assessments of learners' linguistic performance and praised the learners for their high level of proficiency. This may be explained for N01 by his profession of teacher of German. N04 corrected a lot of errors without being asked, which contradicts to (Hosoda, 2006) observation and the correcting behaviour of the other three native speakers in my dataset. It might be explained for N04 by his previous experiences with learners of German in chat. In addition, N04 and L08 explicitly negotiate that corrections of linguistic errors are allowed (Example 3.17).

This information may be transferred to conversations with his other partners. As a result, N04 engaged in other-corrections with L09 without being invited to correct. In the retrospective interview, N04 told that he sometimes confused L08 with L09 "because they come from the same town" (orig,: *glaube, ich hab die beiden am Anfang sogar miteinander verwechselt, weil sie beide aus der gleichen Stadt kommen).* All these factors may explain, why N04 started to correct L09's errors without her invitation or prior negotiations as opposed to Hosoda's findings (Hosoda, 2006) and other native speaker's behaviour in my dataset.

Summarised, there is no such thing like *the* language expert and *the* language novice. There are various practices where participants of native/non-native speaker conversations make their linguistic identities relevant in talk. Linguistic identities emerge in conversation in form of face work, meta-talk about learning, role-play and repair with linguistic trouble source. Different practices can be deployed to realise them, as examples in this chapter illustrate. Each combination of these practices will lead to different interaction profiles of conversation participants, however, all these different profiles will fall into bigger categories of language novices and language experts.

4. Other-initiated Self-repair with Linguistic trouble source

Participants of a conversation sometimes need clarification in order to assure that they understand their partners correctly. As it was shown in Chapter 3, learners of a language interacting in a foreign language may have the additional need to clarify meaning of particular linguistic constructions in order to complete their learning and to fill the gaps in language knowledge. In L1-only interaction, similar situations may occur if, for instance a participant is not familiar with specific terminology and asks for explanation. A prototypical sequential organisation of the interaction in such situations in oral and face-to-face settings described in academic literature (Varonis and Gass, 1985; Schegloff, 2007) and confirmed by the dataset is:

1. Participant A produces a turn with a unit which participant B cannot understand;
2. Participant B initiates a repair sequence asking for clarification or explanation;
3. Participant A carries out the repair providing a clarification or an explanation;
4. Participant B accepts the explained meaning explicitly or implicitly.

However, differences in resources made available by participants for trouble identification may occur due to virtual adjacency in text chat (Tudini, 2010). From all possible constellations, I focus on the case where the native speaker of German produces a turn containing a linguistic trouble source, the learner of German initiates repair, and the native speaker carries out the repair. Thus, I speak about other-initiated self-repair with linguistic trouble source when the native speaker is the trouble-speaker (OISR).

I will analyse in this chapter, which interactional resources the learners make available in chat in order to other-initiate repair with linguistic trouble source, that is to signal trouble and to reference the trouble source. I will specifically focus on turn formats for repair initiation, because this is important for the future recognition of such sequences by conversational agents. Section 4.1 illustrates the use of interactional resources for repair other-initiation on multiple examples from the dataset. I will examine the interactional resources used by the native speakers in order to carry out the self-repair, because this is required to model the responses of the conversational agent. Section 4.2 is dedicated to various ways to carry-out self-repair after an other-initiation in chat. I discuss the insights and the findings as illustrated by examples in Section 4.3.

4.1. Repair initiation: signalling trouble in chat

The majority of the repair other-initiations in my dataset belong to one of two abstract types of repair other-initiations: statements of non-understanding where a part of partner's utterance is marked as unclear, and meaning checks where the own version of understanding of the problematic unit (candidate understanding) is provided in the repair initiation. The former requires an explanation of the trouble source in the repair, the latter requires a yes/no answer (a confirmation of the candidate understanding or its rejection and then probably explanation of the original version). I analyse the interactional resources that the learners use to implement these two types of repair initiation below in detail.

I found two distinct types of position for repair other-initiations in my dataset. The first type comes immediately after the trouble source turn (adjacent position). The second type comes later than the adjacent turn. In this case, one or many turns produced by the trouble-speaker or trouble-talk recipient may be between the trouble-turn and the repair initiation. I termed the former as *immediate* and the latter as *delayed* repair initiation. Sequentially, both correspond to the next-turn repair initiation or second position repair described in CA literature as the first structurally specified place for other-initiated repair (Schegloff, 2000; Liddicoat, 2011). Delayed repair initiations may occur because participants can produce turns simultaneously and follow multiple distinct conversation threads so that the threads are interleaved.

I will argue in this section that there is a dependency between the position of the repair initiation and the interactional recourses that need to be involved to produce a repair initiation. I will show that some sorts of resources are dedicated to initiate repair in the immediate adjacent position.

4.1.1. Interactional resources available in chat for repair initiation

In my dataset, I observe similarities between the resources for repair initiation described in CA literature for face-to-face interaction (see. Section "2.2.5, specifically open class repair initiations and specific repair initiations), but there are also specific interactional resources in chat influenced by the sequential organisation of the chat (virtual adjacency) and medially written mode of communication. It is observable that the repair initiations are conform to the conversation quantity rule suggested by Grice: they are as short as possible but as informative as needed to understand the problem (see (Auer, 1999) for a short summary).

In Example 4.1, the learner initiates a repair by posting three question marks (and nothing else) directly after the trouble source turn. The native speaker N04 is able to locate the trouble source, which is the abbreviation. Why does it work? The abbreviation "gn8" is a second pair part of the farewell. Three question marks do not contain any reference to anything that has been said before, therefore, they are interpreted by the native speaker as a signal of non-understanding of the whole previous turn, which contains only the abbreviation and a smiley. The symbolic resource - question mark - is normally used to mark something as a question, which is something that requires a clarification or an explanation or more information. The native speaker makes his interpretation of the reading of "???" by providing an explanation of the abbreviation in turns 619 and 620. The learner accepts the explanation with a news receipt token in turn 622.

Example 4.1. Open class repair initiation symbolic means

615	20:41:24	L08	danke. good night) *thank you. good night [code switching] [smile]*
617	20:41:38	N04	gn8 :-)
618	20:41:48	L08	??? *??? [repair initiation]*
619	20:42:58	N04	gn8 ist ein zusammengeschrumpftes "gute Nacht" (lies: "g" = "gut" und "n8" = "N-Acht") *gn8 is an abbreviation of "good night" (read: "g"="good" and "n8" = "n-ight")*
620	20:43:27	N04	oder englisch, g=good, n-eight *or Englisch, g=good, n-eight*
621	20:43:29	L08	aach sooo)) *stretched discourse marker [smiley]* I see [smiley]
622	20:43:42	N04	:-)

Example 4.2 shows an alternative possibility of a repair initiation referring to the whole previous turn, this time an explicit statement of non-understanding accompanied by a sad emoticon "(". The learner L08 makes visible in turn 505 that something in turn 504 is problematic for comprehension, which is, in fact, the whole turn containing colloquial tokens like "halt" and "nix". N04 paraphrases the whole message because he interprets the repair initiation as referring to his whole previous turn.

Example 4.2. Open class repair initiation: lexical means

504	18:13:50	N04	ich war halt arbeiten, da gibt's nix zu berichten *I was just at work, there is nothing to report*
505	18:15:01	L08	ich verstehe nicht(*I do not understand [sad smile]*
506	18:15:46	N04	damit wollte ich sagen: ich war bei der Arbeit und habe nichts erlebt, was man erzählen könnte *with this I wanted to say: I was at work and did not experience anything that one could tell* I just wanted to say: I was at work and nothing happened that would be worth telling

Repair initiations of this type correspond to open class repair or unspecific initiations described in CA literature and discussed in Section 2.2.5. In my data, they occur always in the turn adjacent to the trouble source turn and refer to the whole trouble source turn, which may contain only one token or a longer utterance. This shows that the research results described in Section 2.2.5 obtained for oral conversation are also valid for chat conversation, although they do not cover delayed repair initiations.

A bit more specific repair initiations can be done with demonstrative determiners and demonstrative pronouns as Example 4.3 illustrates. In turn 168, the learner asks for clarification of the expression that the native speaker posts in turn 167. As in the examples discussed above in this section, the reference *dieser Ausdrück* (Engl.: *this expression*, * contains a spelling error) relates to the whole turn and is adjacent to the trouble source turn. In addition, L02 provides her own version of the understanding in turn 169 which related to the topic "dance" discussed in the previous conversations.

Example 4.3. Use of a demonstrative determiner to reference the TS

167	20:01:24	N01	Was macht die Kunst? *What makes the art?* How are you?
168	20:03:29	L02	Was bedeutet dieser Ausdrück? Ich verstehe nicht (((*What does this expression [* error] mean? I do not understand [sad smiley]*
169	20:04:09	L02	Meinst du Tanzen? *Do you mean dance?*
170	20:04:44	N01	Das habe ich mir schon gedacht :-) Wir verwenden diesen Fragesatz in Deutschland als Synonym für "Wie geht es dir?". Man könnte auch sagen "Was macht die Kunst (des Lebens)?" *This have I me already thought [smiley]* *We use this question sentence in Germany as a synonym for "How are you?".* *One could also say "What makes the art (of life)?"* That's what I was already thinking about [smiley] We use this question in Germany as a synonym to "How are you?". One could also say "How is the art (of life) going?".
171	20:05:32	L02	Ach so! Ich habe das nicht gewußt)) *I see! I did not know this [smile]*

However, it is not always necessary to mark the whole turn produced by the native speaker as a problem. If the learners are able to identify a smaller segment of a trouble-turn as problematic, they use more specific references to the trouble source in their repair initiations. Such specific references contain all required information about the nature of the trouble. In Example 4.4, the native speaker uses in turn 640 a word to describe his kinship relations that the learner does not understand. This turn is followed by a return question *Und du?* in turn 641. Even without the insertion sequence in turns 642-646, an open class repair initiation after the counter question in turn 641 would not address the problem with the token *halbBruder* because of the delayed position. Therefore, something more specific is needed to mark the particular word as unclear. The learner recycles the term used by the native speaker adding a question mark to it in turn 647. The question mark signalises the need for clarification of this particular term.

Example 4.4. Using repetitions of the trouble source to initiate a repair (simplified)

634	19:45:33	L07	und hast du geschwister? *and have you brothers and sisters?*
639	19:47:05	N03	Einen älteren Bruder *An older brother*
640	19:47:34	N03	Und einen jüngeren halbBruder *And a younger halfBrother*
641	19:49:29	N03	Und du? *And you?*
			`[turns 642-646 build an insertion sequence]`
647	19:53:35	L07	halbbruder?? *halfbrother?*
648	19:54:56	N03	Sohn meines leiblichen Vaters und der Lebensgefährtin nach meiner Mutter *Son of my biological father and his partner for life after my mother*
649	19:55:53	L07	aaa... dann ich habe auch einen... und wie ich schon geschriben habe eine schwester..)) *discourse marker ... then I have also one... and as I already have written one sister [smile]*

Recycling of one or many tokens combined with symbolic means (mostly question marks, but also combined with dashes and quotes) to signalise trouble with one word or a part of a message is frequently used by chat participants in my dataset. However, lexical means to signalise trouble are also present in my data (explicit statements of non-understanding) as shown in Example 4.5.

Example 4.5. Reuse part of a message (simplified)

118	18:50:30	N04	Ja, Kommissar Rex hab ich früher auch geschaut... Das muß 15 Jahre her sein, da kannst Du dich dran erinnern? *Yes, Kommissar Rex have I earlier also watched... This must 15 years back be, there can you reflexive particle on recall?* Yes, I used to watch Kommissar Rex, too... This must be already 15 years ago, you can still remember that?
120	18:52:53	L08	her sein, da kannst Du dich dran erinnern? verstehe nicht) *back be, there can you reflexive particle on recall?* *understand not* back be, there can you reflexive particle on recall? don't understand

121	18:53:33	L08	her=früher *back=earlier*
122	18:54:26	N04	etwas ist schon lange her = "etwas ist vor langer Zeit passiert" *something is already long back = "something is long time ago happened"*
123	18:55:57	N04	und "dran" (umgangssprachlich) = "daran" (hochdeutsch) *and "dran" (colloquial) = "thereon" (hochdeutsch)*
124	18:56:29	L08	aha...nun in Belarus wurde es später passiert. *discourse marker...well in Belarus was it later happened [error: use of passive not grammatical with the verb passieren].* I see... well, it was happened later in Belarus.
125	18:56:59	L08	oh danke für solche Erklärungen)) *oh thank you for such explanations [smile]*

The learner L08 copies a part of the partner's message and adds an explicit statement of non-understanding in a new line of her message followed by a smile (the closing parenthesis) in turn 120.

To sum up, Examples 4.1 to 4.5 illustrate a common prototypical structure of repair initiations where a unit of the trouble source turn is marked as unclear by lexical or symbolic means (signalling non-understanding):

reference to the trouble source + signalling non-understanding

Candidate understanding or meaning check is another possibility to mark a unit of an utterance as not (completely) clear. In this case the repair other-initiation is designed as a polar question requiring a confirming or a disconfirming answer and checking the similarity of two different descriptions of an object. Example 4.6 shows a fragment of a chat where the native speaker N04 uses the word *überfülltes* to describe an event in Munich (turn 221). The learner L08 checks her understanding of this term in turn 223 by copying the trouble source and providing her own candidate understanding of the word. A confirmation follows in turn 224.

Example 4.6. Many many people

221	18:45:26	L08	ja ich habe über Oktoberfest gehört, etwas lustiges und buntes)) *yes I have heard about Oktoberfest, someting funny and*
222	18:46:25	N04	ja, und teures und überfülltes ;-) *yes, and expensive and overfilled*
223	18:47:48	L08	))überfülltes bedeutet "viele viele Leute"? *overfilled means "many many people"?*
224	18:48:14	N04	genau *exactly*

Example 4.6 illustrates a prototypical repair initiation format for a meaning check:

left-hand part + EQUALS-*token* + *right-hand part* + *turn-final element*

The EQUALS-tokens that I found in my dataset can be lexical like *bedeutet, heisst* or symbolic like *"="* and *"-"*. The borders of the left-hand part and / or right-hand part may be also marked in addition with quotes (*framing*) or uppercase (*highlighting*). The turn-final element may be symbolic, such as one or many questions marks, or lexicalised, like for instance the question *habe ich das richtig verstanden?*. The turn-final element has the function to mark the turn as a question seeking confirmation.

The repair initiations produced by the learners in my dataset always try to resolve problems with the meaning of something in native speaker's utterance. None of the repair initiations was concerned with the form by itself. Given that repair other-initiations can occur everywhere in conversation because everything can potentially cause problems with understanding of the talk by the recipient, there are some specific units in every language that become trouble sources for non-native speakers more likely. These are in particular idiomatic expressions and abbreviations. These two categories of lexical units frequently cause repair other-initiations in my dataset. Sometimes the learners marked the whole messages of the native speakers as unclear because they were not able to identify the problem more specifically. Examples 4.2 and 4.13 contain instances of such repair initiations. However, as soon as the learners understand, what exactly is the problem in the native speaker's turn, they make their repair initiations very specific, see for instance Examples 4.4 (halbbruder), 4.10 (in sachen essen) and 4.9 (eckball). I found the following types of trouble source in my dataset:

1. The whole one-word message.
2. The whole longer message.
3. A part of a longer message.
4. An idiomatic expression.
5. A single word which is part of a longer message.
6. A single word which is the entire message.

These types play a role for the production of the repair as I will show in Section 8.1.2.

4.1.2. When the recognition of the repair other-initiation fails

As Schegloff pointed out, "there appears to be a determinate set of turns formats used to initiate such [other-initiated] repair" (Schegloff, 1993). In addition, it has ben shown that other-initiated repair is present in all languages (no language is known till present that does not have other-initiated repair), and that the majority of repair formats exists in all analysed languages (Dingemanse et al., 2014). Learners of a language, however, sometimes have difficulties to formulate a repair initiation properly in the target language, which increases the variance in referencing and signalling and makes the intended natural language understanding task more challenging because of the learner language. I discuss one of such instances in Example 4.7.

Example 4.7. Learner failed to reference the trouble source properly

303	17:10:14	N02	Danke, bis Dienstag ist noch Schule und dann sind endlich Ferien :) aber die zwei Tage krieg ich jetzt auch noch rum! *Thanks, till Tuesday is still school and then are finally holidays [smile] but the two days get I as well around!* Thanks, we have classes till Tuesday, but then we have holidays, finally [smile] but the two days will pass as well!
304	17:12:11	L03	ach so, ja, ab dem 1. August, ich erinnere mich jetzt... dieses Phraseologismus habe ich noch nie gehört :) *interjection so, yes, from the 1st of August, I recall now... this* [neutrum] phraseologism [maskulinum] have I yet never heard* I see, yes, beginning on the 1st of August, I recall now... I have never heard this phraseologism [smile]
305	17:12:26	N02	welchen? ;) *which [maskulinum] one? [smile]* which one?
306	17:12:32	L03	dei Phraseologismus :) *[typo: intended "den" - determiner maskulinum] phraseologism* the phraseologism
307	17:12:51	L03	über zwei Tage krieg *about two days war*
308	17:13:10	N02	oh ach so ;) haha, dann lernst du hier ja sogar noch etwas :) *newsmarker interjection so [smiley + lexicalised laugh], then you learn here even something [smiley]* oh I see [smiley + lexicalised laugh], then you will even learn here something [smiley]
309	17:14:38	L03	hahaha, aber ich zweifele jetzt, ob es wirklich eine Idiomatische Wendung ist ist *[lexicalised laugh], but I doubt now if it really an idiomatic phrase is is* [lexicalised laugh], but I doubt now it it is really an idiomatic expression
310	17:16:52	N02	nein, das ist wohl eher Dialekt... *no, this is rather dialect...*
311	17:16:53	N02	;)
312	17:18:27	L03	ach so... in jedem Fall interesant. *interjection so ... in any case interesting.* I see... interesting anyway.

There are several problems with the repair initiation in turn 304 of this example. First, L03 wrongly identifies the expression *zwei Tage krieg* as an idiomatic expression. However, she refers to this trouble source as *dieses* Phraseologismus* (En.: *this [* wrong determiner-noun congruence] phraseologism*). The reference cannot be resolved by the native speaker, therefore, she starts a nested repair initiation sequence in turn 305, which also contains an exposed correction (see Chapter 5 for a detailed discussion of exposed corrections and this example in particular). Turn 306 contains an uptake of the corrected congruence error. Due to a small difference in time stamps of turns 305 and 306 it is hard to say whether it was a self-correction or an acceptance of the explicit correction provided by the native speaker. Second, even the explicit repair initiation in turn 307 does not properly locate the trouble source because the learner segmented the utterance wrongly, *zwei Tage krieg* (En.: *two days of war*) instead of *zwei Tage rumkriegen* (En.: *to get around two days*).

In addition, by using the term *Phraseologismus* to refer to the trouble source in turns 304 and 306, L03 positions herself as an expert in linguistics, and, at the same time, she produces a grammatical error combining a neutral gender determiner with the male gender noun. This is in my eyes one of the reasons why the native speaker refuses to provide a proper clarification until the end of this sequence. L03 provides an explicit reference to the trouble source repeating the problematic unit in turn 307. However, the native speaker still does not provide any repair of the linguistic trouble source, commenting on learning in turn 308 instead. In order to explain the part of the message that really caused the problem, native speaker would have to show the learner, that she is NOT an expert (she wrongly identified the expression as idiomatic and she even wrongly segmented the utterance), which might be too much face threatening.

4.2. Repair carry-out strategies

This section describes the strategies that the native speakers deployed to carry out the repair with linguistic trouble source in response to repair initiations. Repair carry-out strategies depend on the type of the trouble source and the repair initiation format. As explained earlier in Section 4.1, some of the repair initiations require a yes/no answer, and some of them require definition work or paraphrasing of the trouble source. The definition work can be carried out in different ways, as I will show in the remainder of this section.

If the trouble source is an abbreviation, the definition work carried out by the native speaker normally contained a full spelling of the abbreviated words and their explanation. For chat abbreviations, a full reading of the abbreviation was normally provided and enough for explanation, as Example 4.1 demonstrates. The problematic abbreviation is always repeated in my dataset, then a full spelling or (as for chat jargon) an intended reading is provided.

If the trouble source is one semantic unit (one word or an idiomatic expression), a dictionary-like definition (synonyms + examples) is often selected to provide a repair. For longer messages or longer parts of longer messages, a strategy of splitting the message into smaller semantic units and separate explanation of each unit can be chosen, as illustrated in Example 4.5 turns 121-123. Paraphrasing is also one of the strategies used by the native speakers to explain longer messages, see Example 4.2 turn 506.

4.2.1. Providing definitions, paraphrases and synonyms

Examples 4.3, 4.4 and 4.8 show instances of such repairs. Turn 170 of Example 4.3 contains a synonym of the problematic expression. Turns 172 and 173 provide in addition an explanation about the use of the expression.

Example 4.8 illustrates another sequence where the native speaker uses synonyms to construct an explanation for an idiomatic expression emphasizing in turn 124 ("_here_") that this expression was used in this meaning in this particular context.

Example 4.8. Ins Auge fassen (prev. Ex. 3.12)

122	13:08:39	N04	zugegeben, ich war dieses Jahr auch noch in keinem See, aber so langsam könnte man das mal ins Auge fassen :-) *admitted, I was this year not in a lake, either, but slowly one could reach this into the eye [smile]* I admit I was this year not at a lake, either, but I could slowly consider it [smile]
123	13:10:47	L09	ins auge fassen? *reach into the eye?*
124	13:11:56	N04	das heißt _hier_ etwa soviel wie "planen" oder "bald mal machen" *it means here something like "to plan" or "to do soon"*
125	13:12:16	L09	:) :)
126	13:12:22	L09	klar *clear*

4.2.2. Translations and demonstrations

Other strategies implemented by the native speakers for definition work that I found in my dataset are translation into a different language and hyperlinks to examples of the objects referenced by the trouble-sourse expressions. The translations were done manually into a shared foreign language or automatically into the native language of the learner using a machine translation service. Hyperlinks as explanation strategy were used in combination with textual explanation of the trouble source expressions, or stand-alone. I analyse these four strategies for repair carry-out below on examples from the dataset.

Example 4.9 illustrates how a translation into English, which was a shared foreign language for N04 and L08, can be used in addition to a textual explanation of the problematic unit in native speaker's talk. Turn 528 contains a trouble source "Eckball" as it is made relevant in turn 530 by the learner in form of a repair initiation. Turn 531 contains the English translation of the German word "Eckball". The follow-up turn 532 provides in addition an explanation in German, how the ball must be played in this situation.

Example 4.9. Use English translation to repair.

528	18:47:09	N04	und schon ein Eckball für Deutschland! *and already a corner kick for Germany*
529	18:47:28	N04	doch nicht :-(*not yet [sad emotion]*
530	18:47:29	L08	eckball? *corner kick?*
531	18:47:42	N04	corner kick auf englisch *corner kick in English*
532	18:48:05	N04	der Ball wird von einer Ecke des Spielfelds vor's Tor gespielt *the ball is being played from a corner of the play field to the goal*
533	18:48:22	N04	aber diesmal noch nicht, denn der Schiedsrichter hat sich noch mal umentschieden *but this time not yet, because the referee changed his decision again*
534	18:48:25	L08	aga, etzt klar) *[acknowledgement token], now [typo] clear [smile]* I see, now it's clear [smile]

Example 4.10 shows how a machine translation service can be used for definition work. It might be important to know for the analysis of this sequence, that N03 learns Russian as a foreign language, but he has just started and cannot do conversation in Russian.

Example 4.10. In Sachen Essen: repair is carried out by an automated translation.

376	07:40:24	N03	gibt es irgendwas moskau typisches in sachen essen? *is there something typical for moscow in things food?* is there something of food which is typical for moscow?
377	07:41:03	L07	in sachen essen??? *in things food???*
378	07:41:31	L07	übersetze bitte))) *translate please [smile]*
379	07:44:36	N03	какая пища является типичным Москве? *which food is typical [* wrong noun-adjective congruence] [* missing preposition] Moscow?* which food is typical for Moscow?

Turn 376 contains an expression that the learner does not (fully) understand: "in sachen essen" which is being officially made to a trouble source in the repair initiation in turns 377 and 378. Turn 377 locates the trouble source and marks the expression as unclear. Turn 378 contains an instruction what kind of explanation is desired.

The learner was aware of the fact that N03's level of knowledge in Russian is not enough to translate this, but she challenges him in her repair initiation, however, the word "translate" can also mean "say it in a language that I understand". The native speaker chooses the playful version of translation and translated his question using an online machine translation service having no idea about the quality of the translation.

Example 4.11. Demonstration in addition to the textual definition work.

541	18:52:01	L08	du bist heute ein internationaler Fussballkommentator) *you are today an international football commentator*
542	18:52:43	N04	ja, ich bin grad sowas wie der Live-Ticker *yes, I'm now something like a life-ticker*
543	18:54:49	L08	Ticker–nicht besond. verstehe Bedeutung *Ticker – not special. [specially] understand meaning [* multiple errors]* Ticker - I don't exactly understand the meaning
544	18:57:19	N04	der Begriff kommt vom "newsticker", den man früher in Zeitungsredaktionen stehen hatte *The term comes from "newsticker", which they had in the past in newspaper editorial offices*
545	18:57:30	N04	und heute meint man damit sowas: HYPERLINK *and today we mean with it something like: HYPERLINK*
546	18:57:37	N04	oder sowas: HYPERLINK-11freunde *or something like: HYPERLINK-11freunde*
547	18:58:13	N04	(der 11freunde-Ticker ist allerdings nicht ganz so ernst gemeint) *the 11freunde-ticker is however not meant very seriously*
548	18:59:23	L08	nun im großen und ganzen verstehe ich) *well, in general, I understand [smiley]*

In addition, participants of the chat conversation may use hyperlinks to further information resources on the Web for repair cary-out. In Example 4.11, the learner compares the native speaker to an international football commentator (turn 541) because he commented in chat everything that happened on the TV in a football game (before turn 541). However, the term *Life-Ticker* is from N04's point of view conceptually closer to what he was doing (textual comments instead of oral comments as usually done by commentators on the TV), therefore, an embedded correction is performed in turn 542. Unfortunately, the learner does not understand the term *Life-Ticker* and initiates the repair in turn 543 marking only a part of the term as unclear *Ticker* and producing L2 errors in her repair initiation in addition. The native speaker explains the origin of the word in turn 544 and provides two hyperlinks to websites with life-ticker examples to demonstrate what they look like in turns 545 and 546. Tern 547 explained how the learner should interpret the demonstration from the turn 546.

Example 4.12 illustrates how the explicit definition work can be avoided (indeed, it is not easy to explain what a life stream is) still providing a preferred response to a repair initiation, which is in this case just a demonstration of a instance of a life stream. The term *Lifestream* was used by the native speaker in turn 653. The learner marks this term as unclear in turn 656 (delayed repair initiation). The native speaker provides a link to the life stream that he is watching. This action is implicitly accepted as an appropriate clarification by the learner in turn 659.

Example 4.12. Use of a hyperlink to an example

653	19:22:14	N04	[simplified] ich schau mir gerade den Livestream von einer Preisverleihung an *I am watching the live stream of an award at the moment*
654	19:22:43	N04	es geht da um einen Preis für Unternehmensgründer *it is about a prize for the company founder*
655	19:22:59	N04	meine Chefs haben auch etwas bekommen :-) *my bosses got something, too [smile]*
656	19:23:22	L08	Livestream? *Live stream?*
657	19:23:28	N04	der interessante Teil ist vorbei, aber hier ist der Link: *the interesting part is over, but here is the link:*
658	19:23:34	N04	HYPERLINK
658	19:24:23	N04	was gerade läuft ist ein Portrait des Preisträgers in der Kategorie "Lebenswerk" *what is running now is a portrait of the winner in the category of "lifetime achievement"*
659	19:28:18	L08	und wo sind deine Chefs?)oder sie fehlen) *and where are your bosses? [smile] or they are missing [smile]*

4.2.3. Non-present repair carry-out after other-initiation

Repair carry-out is a preferred response to a repair initiation, but not the only possible. In Example 4.7 I showed that not every repair initiation contains a proper reference to the trouble source and therefore has little chance to be satisfied in the repair response.

Example 4.13. Non-present self-repair carry-out after other-initiation

395	20:43:29	L08	oh, Spiel ist am Ende od. am Anfang? *news marker, the match is at the end or in the beginning?*
396	20:43:43	N04	ist in etwa 5 Min. aus *is in about 5 min. over*
397	20:44:04	L08	unklar *unclear*
398	20:44:07	N04	Schweden 2:3 England *Sweden 2:3 England*

Example 4.13 shows another case where a dispreferred response to a repair initiation is provided, namely, a comment on the topic of the talk, but not a specific reaction to a repair initiation. The problem with the trouble source - the whole turn 396 - is never clarified in the chat between L08 and N04. The non-present explanation of linguistic trouble is omitted by the native speaker. Instead, he provides information about the state of the game, addressing a potential problem with understanding the situation in the game.

The explanation *why* the repair has not been carried out may be very simple: there are only 3 seconds between the repair initiation in turn 397 and the next turn of native speaker. It is very likely that there was an overlap in production of the turns 397 and 398. The native speaker probably did not want to ignore the repair initiation. However, the turn containing the comment on topic "Sweden 2:3 England" came just *after* the repair initiation and could be interpreted by the learner as a response to it, even if not explaining the problem. The fact *that* a repair carry out is non-present (or maybe even missing) and the problem is not clarified seems not to have any impact to the continuation of the interaction. The learner does not re-initiate repair, and with the continuation of the talk, the problem becomes irrelevant. Thus, the non-present repair carry-out does not have any sequential consequences in this case.

4.3. Preliminary findings and discussion

In this chapter, I analysed a collection of other-initiated self-repair sequences where native speakers produced something in their turns that language learners could not (completely) understand and asked for explanation or clarification. Specifically, I looked at interactional resources that language learners made available in text-based instant messaging dialogues in German to produce repair other-initiations, at positions relatively to the trouble-turns where the repair other-initiations were placed, and at types of troubles addressed. This analysis was performed with the purpose of computational modelling of such sequences, therefore, turn formats for repair other-initiations in chat are of special interest. In particular, the intended model needs to distinguish between repair other-initiations and all other turns, and to identify the trouble given a repair initiation.

Repair initiation formats combining interactional practices and devices were analysed. Specific devices are used to signal trouble and other devices are used to reference the trouble source. More specifically, it was found that:

- Questioning is *the* practice to initiate repair in chat, confirming the results in academic literature for oral interaction(Dingemanse et al., 2014). Other practices are declarations of lack of understanding such as *unklar* and *ich verstehe nicht* (labelling).
- Devices for signalling are question marks, dashes, explicit statements of non-understanding and presenting candidate understandings.
- References to trouble source may be realised through the adjacent position, demonstrative proterms and phrases and repetitions.
- Though all repair initiations were second-position initiations, they were not all immediate. Delayed repair initiation require more specific referencing to trouble source, open-class repair initiations cannot be used in a delayed second position.

- Repetition-based repair initiations may contain repetitions of one specific unit from the previous turn and contain a copy of the preceding turn regardless the unit boundaries. The latter may be placed between open class and restricted class repair initiations. Such types of repetitions have not been previously described in academic literature and must be typical for non-native speakers.
- Communication medium influences repair initiation types and formats. In particular, repair initiations eliciting a repetition of the trouble source are uncommon in chat. Misreadings are possible, but they are made visible through mis-productions in repetition-based repair initiations.
- Non-native speaker identity influences the format of candidate understandings which differ from those in L1 talk.
- Repair initiation is one option to deal with trouble in comprehension. Learners may make use of other options, too (dictionary look-up, let-it-pass...). Leaner's interactional and linguistic competence influence the selection of a repair initiation format and its successful recognition.

Repair carry-out is the preferred and the most frequent response to a repair initiation but other forms of responses are also possible, for instance a new repair initiation to deal with difficulties in identification of the trouble and responses which do not address the trouble. The following was found with regard to repair carry-out formats:

- Explanations of the meaning, translations and demonstrations are the most frequent forms of repair carry-outs.
- Repair design is linked to expectation of what is known to the repair recipient. Consequently, repairs are design for the language learners targeting difficulties in linguistic matters.
- Repair carry-outs may be immediate and delayed.
- References to trouble source may be realised by the same resources as for repair initiations.
- There are dependencies between types of trouble source and participants' selection of resources for referencing the trouble source. For instance, abbreviations are usually repeated.

I found, that each repair initiation contains some symbolic or lexical signs which are dedicated to *signal* the trouble, and other resources which are used to point to the trouble source, which I call *referencing*. Depending on the position of the repair other-initiation relatively to the trouble-turn, different sets of resources may be used to reference the trouble source. Normally, the next turn position is dedicated for repair other-initiation after which a self-repair follows (Schegloff, 2000). Because text-based chat allows for virtual adjacency (Tudini, 2010), next-turn repair-initiations may appear immediately after the trouble-turn, as in oral talk, or a few turns later, where turns belonging to a different thread in conversation may intervene.

Open-class repair initiations do not contain an explicit reference to the trouble source. The referencing takes place by the adjacent position immediately after the trouble turn. Use of demonstrative pronouns and determiners makes sense only in immediate repair initiations, too. In both cases, the complete trouble-turn is marked as problematic. More precise references use a copy (reuse) or a modified copy (recycle) of a part of the trouble turn to reference the trouble source. Such repeat-based references of the trouble source can be used in immediate and in delayed repair other-initiations. Various means to frame (e.g. by quotation marks) or to highlight (e.g. by uppercase) the trouble source in repeat-based repair initiations may be used in addition to make the turn format more clear and to emphasise the its social action, namely, repair other-initiation.

For signalling the trouble, a variety of symbolic and linguistic means were used by language learners. Symbolic means include question marks, dashes, quotation marks and equal signs. Lexical means include explicit statements of non-understanding such as *Ich verstehe nicht* or *unklar*. Lexicalised equivalents of paralinguistic cues were not used by language learners to initiate repair, as opposed to native speakers' *hä?* used for this purpose. This shows that models based only of native speaker data may not reflect the reality in interaction with learners.

The dataset contains examples where native speakers had difficulties with recognition of learners' turns as repair initiations because learners had problems with referencing the trouble source. Such cases will make automatic recognition of repair initiations even more difficult. Sometimes, native speakers failed to produce a repair proper, however, it is very likely, that they did not notice learner's repair initiation.

Furthermore, I analysed interactional resources which were made available in chat by native speakers to carry out self-repair after other-initiations. Repair carry-outs were tailored to the type of the trouble source. If it was an abbreviation, the trouble was resolved by a complete writing of the abbreviated words. If a word or an idiom caused trouble in understanding, an explanation of the meaning of these items was provided. If a longer part of a turn or a longer turn caused a problem with comprehension, it was paraphrased or only a few words, which were supposed to be the trouble source, were explained. Sometimes translations in other shared languages was used to explain unknown terms.

Native speakers were very creative in providing explanations of unknown terms. Although using synonyms, examples and paraphrasing is a frequently used form of self-repair carry-out, various external resources from the internet were involved in explanations. Native speakers used hyperlinks to show examples of entities referred to by words which were unknown to the learner, like *Lifeticker*. In addition, machine translation tools were used to present a translation of unknown expressions to the learner. Summarised, various information resources from the Internet were used by native speakers in instant messaging dialogues *as an interactional resource*. This is impossible in a face-to-face interaction without technology.

Focus on form is found to be very rare in a Conversation-for-Learning (Hosoda, 2006; Hauser, 2010), however correction of language form is something that participants may chose to do. In example 4.5 the native speaker choses to correct the erroneous use of a passive verb form. This correction follows immediately a repair-sequence with focus on meaning initiated by the learner. Because the participants were already engaged in discussion of linguistic matters, it may have made the shift from the focus on meaning to the focus on form easier for them.

In order to "serve computational interests" (Schegloff, 1996), the following needs to be taken into account for the purpose of modelling. Because repair initiations may occur *everywhere*, each user's utterance may be a repair initiations. Therefore, a repair initiation recognition routine needs to be activated after *every* user's turn. Two essential problems must be solved by a computer program in order to react to a repair initiation properly:

1. Recognise that user's utterance is a repair initiation,
2. Recognise the source of troubles and the kind of the problem.

Repair initiation formats need to be "translated" into patterns and then into computational models of repair initiations to make the findings applicable for computational purposes.

5. Exposed Corrections

Exposed corrections of linguistic errors are dispreferred in both L2 and L1 conversation, however, they occur in L2 conversation more frequently and are perceived by language learners as helpful for learning. This has been reported by the participants of the data collection in the retrospective interviews. The position of the native speaker as more knowledgeable in linguistic matters and the asymmetric relationship with the learner in those matters make corrections possible, but the asymmetry does not explain, why or when they occur (Kurhila, 2006, p. 35). In addition, no dependencies were found in the previous research on exposed corrections between the error and the correction (Schegloff, 1987). Error taxonomies and automatic error recognition are considered to be a preliminary step for error correction. However, as pointed out by Schegloff, "the occurrence of repair is not prompted only by the occurrence of error, an important step was taken in disengaging trouble (error and nonerror) from the practices employed to deal with it" (Schegloff, 1987, p. 216). For these reasons, I focus in this chapter on correction practices, decoupled from the errors and other reasons for correction aside from the occurrence of an error. Nonetheless, I will come back to the analysis of the factors influencing the occurrence of a correction other than just the occurrence of an error in Chapter 7.

Jefferson (1983, p. 88) made the observations, that first, "whatever has been doing on prior to the correcting is discontinued" and "correcting is now the interactional business", and second, there are such "attendant activities as e.g. instructing, complaining, forgiving, apologising, ridiculing". She introduced the term *accountings* for this class of activities.

Kurhila (2006) describes a type of correction on face-to-face communication in which a correction is provided without accountings and the talk in progress is continued immediately so that there is no space left for the non-native speakers to start any kind of negotiations. Kurhila (2006) calls this type of repair the *en passant* correction. Tudini (Tudini, 2010) observed a similar type of correction in her study of intercultural chat. She mentions two types of exposed corrections: correction *on-the-fly* where the social talk immediately follows the correction so that there is no space left for the learner to contribute to the correction sequence, and explicit exposed corrections with accountings where the learners have a chance to contribute to the correction side sequence. I analysed these two types of exposed corrections in Section 3.3 from the perspective of construction of identities of experts and novices in chat and the ways to make the differential language expertise relevant in conversation. I use the term *on-the-fly* correction suggested by Tudini.

A conceptual definition of an error in chat will be included in the discussion and will be subject of Section 7.1. Section 5.1 shows that there are some dependencies between the type of error and the type of trouble caused by it, and correction format chosen by native speakers to deal with it. I will discuss the variations in turn design and the influence of the turn taking on the turn design in exposed corrections in Section 5.2. Finally, in Section 5.4 I will summarise and discuss the findings.

5.1. Types of corrected errors

The participants of conversations in the analysed dataset engage in sequences of exposed corrections of orthographic, grammatical and lexical errors, however, the majority of all corrections in the dataset deal with lexical errors. I found 22 instances of exposed corrections in the data set, which are not equally distributed throughout all conversations. As already mentioned in Chapter 3, different native speakers selected different strategies in interaction sometimes correcting a lot of different errors, and sometimes choosing to let it pass although some of the learners made a lot of mistakes.

Schegloff examined problematic references and problematic sequential implicativeness as classes of trouble based on a collection of third position repairs (Schegloff, 1987). He described his study as "a brief account of several such trouble sources" and "neither an exhausting nor a systematically representative display" (Schegloff, 1987, p. 204). He observed that "there do not seem to be systematic relationships between the types of trouble source and the form taken by repairs addressed to them"(Schegloff, 1987, p. 216). However, some dependencies between the errors and their corrections can be seen, as I will argue in the remainder of this section.

5.1.1. Lexical errors

Lexical errors have been corrected by native speakers in both situations where misunderstandings occurred and where intersubjectivity could still be maintained despite suboptimal lexical choice. Example 5.1 shows a case where intersubjectivity is not in danger however the native speaker chooses to correct the wrong lexical choice.

Example 5.1. Pedagogical repair: the native speaker corrects wrong lexical choice in the absence of a communication problem.

150	13:32:57	L09	jeztz ist dein gesichtskreis mehr) *now is your horizon more [* error: lexical choice]*
151	13:33:15	N04	größer :-) *wider [smile]* [simplified]
153	13:33:35	L09	ok))grösser) *ok [smile] wider [smile]*

In contrast, Example 5.2 shows one of the instances of a correction where learner's turn is ambiguous due to a collocation error. L09 produces a lexical error "* leicht gefallen". This collocation may have two correct targets: "gefallen" (to like) or "leicht fallen" (be easy for someone). This trouble has been resolved by N04 by means of contrasting two possible interpretations of the erroneous collocation. Thus, the type of trouble (ambiguity of the learner's expression) required a repair strategy which allows a disambiguation.

Example 5.2. A correction of a lexical error: a collocation creation leads to ambiguities and has sequential consequences.

135	13:21:09	L09	gefiel dir das studium leicht? *did you like [* error] your study easy [* error]* did you like your study? / was your study easy for you?
136	13:21:45	N04	es gefiel mir, aber es fiel mir nicht immer leicht ;-) *I liked it but it was not easy for me [smile]*
137	13:22:14	N04	("gefallen" = "etwas schön finden", *("to like" - "to find something pretty",*
138	13:22:41	N04	etwas fällt jemandem leicht = man hat keine Mühe damit) *something is easy for someone = one has no effort with it)*

Example 5.3 illustrates the case where a real communication problem exists and a lexical error (a creation, the word does not exist in German) needs to be corrected. However, the communication problem is not resolved in this repair sequence from the point of view of the researcher.

Example 5.3. native speaker produces a correction formulated as other-initiated self-repair. The communication problem remains unresolved.

119	17:31:47	L04	es kommt manchmal vor, dass einige Studenten viele Male ihre Diplomarbeit wiederdrücken müssen!! *it happens sometimes that some students many times their diploma thesis againpress [* error: creation] must!!* it happens sometimes that some students must re-print their diploma thesis many times
120	17:32:25	N02	meinst du wiederholen? also nochmal schreiben? *you mean repeat? well, write again?*
121	17:32:43	L04	jja! mit Korrigierungen *yes [* error: typo]! with correctings [* error]* yes! with corrections

L04 creates a word 'wiederdrücken' (wieder - again, drücken - press). Probably, the word re-print (print again, wieder/neu drucken) was intended. However, from the repair initiation by the native speaker in turn 120 it is analysable that native speaker's interpretation of the creation is "to write again", so to start the complete work from the beginning and to write the whole thesis once again. The learner accepts the correction of the word in turn 121 adding the clarification "with correctings" (re-write the thesis with corrections). It is not clear from the data whether the communication problem was resolved and whether the native speaker understood the intended target meaning first expressed with the creation.

Example 5.4. Correction of a lexical error in form of repair other-initiation.

286	08:47:07	N03	alle pruefungen ueberstanden? *got through all examinations?*
287	08:49:32	L07	am montag hab ich englisch abgegeben.... für eine 10.... es sind jetzt 2 letzte geblieben..und endlich sommerferien...!! die sibd auch die letzten— nachstes jahr absolviere ich DIE UNI-))) *on Monday I delivered [* error???] English.... for a 10.... there are the last 2 left.. and finally summer vacations...!! but they are [* error: typo] the last ones— next [* error: typo] year I complete [* error: lexical choince] THE UNI [smile]* I passed English on Monday... I got a 10.... there are 2 more left.. and finally summer vacation...!! but they are the last ones— I'm graduating from university next year
288	08:58:43	N03	wow.
289	08:58:56	N03	was heisst das, absolviere ich die uni? *what does it mean, I complete the uni?*
290	08:59:04	N03	bist du fertig, schliesst du ab? *are you done, are you finishing?*
291	08:59:52	L07	ja... du bist ein guter übersetzer..)))))) *yes... you are a good translator [smile]*
292	09:00:23	N03	hahaha

In contrast, a lexical error is corrected in form of repair other-initiation in Example 5.4. The learner uses the word "absolvieren" in a wrong context in turn 287, however, the native speaker guesses the intended meaning. He "wraps" the correction of this error into a repair initiation in turn 289 suggesting the correction in turn 290.

5.1.2. Morpho-syntactic errors

Errors in syntax and/or morphology build the second largest category of all explicitly corrected errors. The errors in word order, verb tenses and conjugation, noun-determiner congruence and government are examples of troubles on the morpho-syntactic level that have been made relevant for corrections by the native speakers in the dataset. All instances of corrections of morpho-syntactic erros in my data set deal with pure pedagogical repair in the absence of any communication problems.

Example 5.5 illustrates an exposed correction of a wrong use of an infinite verb form. The learner produces the error in turn 323. The error does not have any sequential implications for the next relevant action. The native speaker produces the response turn which also incorporates the correction. The response is not formulated as a projected next after a good-wishes-expression (*relax well, too*, for instance), but as an explicit correction with the form of the verb that should be used from the perspective of the native speaker highlighted with quotes. The emoticon belongs to the class of accountings.

Example 5.5. A correction of an error in grammar, use of the infinite verb form is replaced by the imperative.

323	17:55:04	L08	alles Gute)und sehr gut sich erholen!!))
			all the best [smile] and to[error] relax very good!! [smile]*
324	17:55:51	N04	Danke, und "erhol Dich gut" ;-)
			Thank you and "relax very good" [smile]
325	17:56:02	L08	=))

However, morpho-syntactic errors can lead to problems with maintaining intersubjectivity and therefore have sequential consequences. A situation of this kind but with the native speaker as trouble-speaker and the learner who initiates repair is discussed in Example 5.6. I analyse it from the perspective of the relationship between the error and the chances to produce the next relevant turn for the trouble-talk recipient.

Example 5.6. A morpho-syntactic error produced by the native speaker leads to a communication problem but remains unresolved.

652	19:57:33	N03	Der hast auch einen halb Bruder?
			Determiner [maskulinum III p. sg.] aux. II p. sg. too one half brother?
			He has a half-brother, too?
653	19:58:17	L07	r hat oder du hast??
			r [probably the end of der = he] has or you have?
			he has or you have?
654	19:58:28	L07	wer von uns?
			who of us?

The native speaker produces an error in turn 652 in a question. The wrong subject-verb congruence does not allow for the understanding of the intended meaning of the question. The learner cannot produce the next relevant action (an answer to a question) and has to initiate a repair sequence in turn 653. The repair initiation contains both variants of possible correct readings of the erroneous pronoun-verb pair. Turn 654 is a repair re-initiation which explains the kind of the trouble more precisely, and this appears to be an attempt to resolve an ambiguous reference to a person.

5.1.3. Spelling errors and typos

Typos are normally not considered to be errors in instant messaging. Chat communication requires fast typing. As a consequence, mis-typings happen quite often to both learners and native speakers. Normally spelling errors have been ignored, too, or corrected in form of embedded corrections, but there are two cases when they have been corrected in form of exposed corrections. I will discuss them in detail here with the goal to explain in which situation an exposed correction in an informal conversation might be relevant for the class of "unimportant" errors.

One of the cases is when the spelling of a word is difficult to learn even for the native speaker, and a sort of memo expression exists for native speakers to remember the spelling better. Example 5.7 illustrates one such case.

Example 5.7. Ziemlich ohne h

167	18:54:38	L08	ja, bin ich ziehmlich einverstanden *yes, am quite [* error] agreed* yes, I'm pretty much in agreement
168	18:55:11	N04	ziemlich ohne h ;-) *ziemlich without h [smile]* [simplified]
170	18:57:06	L08	ich zweifelte daran)mit oder ohne=) *I questioned [smile] whether it was with or without [smile]*
171	18:57:51	N04	es gibt da so einen Merkspruch, aber der ist fast ein bißchen unverschämt... *there is a mnemonic, but it is almost a little bit rude...*
172	18:58:02	N04	ich schreib's Dir trotzdem, nicht böse sein: *I write it nonetheless, don't be angry:*
173	18:58:27	N04	wer ziemlich oder nämlich mit h schreibt ist dämlich *He who writes ziemlich or nämlich with h is stupid*

The second case of an exposed correction of a spelling error concerns a repeated deviation in spelling of the same word twice in one turn. Because the word *Theory* was wrongly spelled *twice* in one turn, the native speaker did not consider this error as a simple typo any more. The error was corrected in turn 348.

Example 5.8. A repeated occurrence of a typo is considered to be a lack of knowledge in spelling and corrected.

346	18:37:10	L08	Teorie der Literatur und Teorie der Sprachwissenschaft *teory [* error: orthograpy] of literature and teory [* error: orthograpy] of linguistics* theoryof literature and linguistic theory
348	18:38:57	N04	Sehr gut. Und "Theorie" mit "h" :-) *Very good. And "theory" with "h" [smile]*
349	18:39:30	L08	of course)ich bin nicht aufmerksam wie immer) *of course [smile] I am not attentive as usuall*
351	18:40:13	N04	doch, du machst das gut *No, you are doing a good job*

Of course there are cases in many languages where the edit distance between two different words is very small and mis-typing could lead to funny sequential implications. The two errors analysed above are not of this kind, there is not mis-understanding or problem in communication. Both cases of exposed correction of spelling errors are instances of pedagogical repair. Marques-Schäfer (2013, Sec. 7.3) discusses the difference between errors and mistakes. In her understanding, errors are gaps in learner's interlanguage, mistakes are lapses. One of the possibilities to test if a deviation is an error or a mistake is to see continuously if the learner produces the same deviation constantly or only occasionally. In Example 5.8, the learner repeats the deviation in spelling twice within one short turn. Therefore, it seems natural to assume that the native speaker interprets this deviation as an error, and not as a mistake.

Having more examples would be an advantage or a reasonable generalised model of exposed corrections of spelling errors. However, a simple rule-based heuristic based on these two examples would be a good start. Since spelling errors and typos normally were not addressed in corrections, having less exposed corrections of spelling errors would be better than having too many.

5.2. Turn design and turn-taking in exposed corrections

This section analyses if there are any structural relationships or dependencies in the format of the trouble-source turns and the corresponding correction turns. The properties of the turns containing corrected errors may be relevant for the intended modelling of exposed corrections for an artificial conversation partner. Correction turn formats are of interest, specifically possibilities to reference the trouble source, to present the correct version and accountings. Finally, the relationship between the trouble-turn and the correction turn(s) in terms of turn design is also relevant for the intended modelling.

I found the following types of exposed corrections:

1. Simple explicit corrections only presenting the correction. Corrections *on-the-fly*, simple explicit corrections with minimal accountings and with rich accountings are the variations that I found in this class.
2. Explicit correction by means of contrast comparing the wrong and the correct item.
3. Exposed corrections integrated into the next relevant turn.
4. Exposed corrections formatted as other-initiations of repair.

I discuss the identified correction types below and illustrate my findings on examples.

5.2.1. Simple explicit corrections

I order the examples in this section according to my understanding of the explicitness of the correction. The types are corrections on-the-fly, explicit corrections with minimal accountings, and explicit corrections with rich accountings. Corrections *on-the-fly* are the least explicit of all exposed correction that have been described in literature and that I found in my data. Example 5.9 shows how a correction on-the-fly is organised.

Example 5.9. Simple correction on-the-fly with minimal accountings.

444	08:06:25	L07	es wäre prima... aber wir haben hier nur seen... die gibt es mehrere in unserer umgebung...) man kann zu fuß gehen oder mit dem auto fahren ein bisschen weiter...weil sich so viele leute jetzt zum strand begeben)) *it would be great... but we have here only lakes... they [* error] there are many in our area... [smile] one can walk or go by car a little bit farther... because so many people betake themselves now to the beach [smile]* it would be great... but we only have lakes here... there are many of them in our area... one can walk or go by car a little bit farther... because so many people go to the beach now
445	08:07:16	N03	[[von denen gibt es mehrere]] *[[of them there are many]]* [[there are many of them]]
446	08:07:19	N03	:-)
447	08:07:25	N03	ach so *I see*
448	08:07:37	N03	wenn du meer haettest waere ich mal vorbei gekommen *if you had a sea, I would come over*
449	08:07:40	N03	:-)

The design of the correction turn 445 can be characterised as minimal: only the correct version is presented, symbolic means are use to mark it as a correction. The correction turn is followed by a turn containing minimal accountings - a smiley (turn 446). The turn design of the correction sequence does not leave any space for the learner to react to the correction turn. The correction and the accountings turns are immediately followed by turns relevant for the interpersonal trajectory therefore, it is the expert who makes the reaction to the correction to a non-relevant action in turns 446-449.

Examples of corrections on-the-fly provided by Tudini (2010, Chapter 5) show similar sequences, however, the correction turns in her examples do not contain any accountings and the interpersonal trajectory is continued in the same turn as correction. It seems important, that the participant who positions himself as language expert in form of an exposed correction on-the-fly, immediately continues the interpersonal trajectory, with or without accountings, and the accountings should be minimal.

A prototypic sequential structure of corrections on-the-fly is:

T1 Learner produces an error.

T2 Native speaker presents the correct version, possibly with minimal accountings.

T3 Native speaker continues the interpersonal trajectory.

In Example 5.10, the learner tells the native speaker about her hometown and the city where she studies, and the capital of Belarus. Turn 150 contains a trouble source which is corrected in turn 151. The native speaker produces a correction with minimal accountings.

The turn consists of the correct version not specifically marked and the smiley. Here, in contrast, the learner reacts to the correction. Turn 152 is produced by the learner and continues the interpersonal trajectory started earlier (Belorussian geography). In turn 153, the learner reacts to the correction of the error (with one turn intervening between the correction and the reaction).

Example 5.10. Simple explicit correction with minimal accountings.

148	13:30:39	L09	schaust du google map?)) *are you looking up google map? [smile]*
149	13:31:29	N04	ja, um mal einen Überblick zu kriegen :-) bisher kannte ich nur Minsk, weil es halt die Hauptstadt ist. *yes, to get an overview [smile] till now I new only Minsk because it is the capital.*
150	13:32:57	L09	jeztz ist dein gesichtskreis mehr) *now [* error: orthography] is your horizon more [* error: lexical choice]*
151	13:33:15	N04	größer :-) *wider [smile]*
152	13:33:25	L09	Vitebsk ist gebietstadt)) Vitebsk is a province city [* error]
153	13:33:35	L09	ok))grösser) *ok [smile] wider [smile]*
154	13:34:27	N04	ja, wenn man "Vitebsk" in googlemaps eintippt, leuchtet der ganze Norden Weißrusslands rot, aber die Stadt hab ich auch gefunden :-) *yes, if you type "Vitebsk" in google maps, the whole North of Belarus shines red, but I found the city, too*

Accountings allow to highlight the correction (in addition to the action, that they are dedicated to, like excuse...). This kind of highlight can be of different intensity, for instance, minimal like in Examples 5.10 and 5.9, or "rich" like in Example 5.11 presenting a sequence where the participants are making an appointment. Here, the learner produces an error in turn 148 (incorrect future form) in her inquiry about native speaker's availability. The native speaker choses first to produce the next relevant turn continuing the interpersonal trajectory in turn 149. This is a repair initiation, because clarification is needed with regard to the intended time slot. After the appointment has been made in turn 152, the native speaker renews the context of the error turn by marking his turn as something additional, coming with a delay ("PS:"). He recycles a part of the learner's utterance and uses symbolic and lexical means to frame the correct version (*es sollte heißen "..."*).

Since the correction is made not immediately after the error and the context of the error had to be renewed, more interactional management had to be made in order to make the correction recognisable as correction. In addition to making the correction recognisable, the native speaker highlights the action of correcting by symbolic means such as quotations and uppercase. The accountings ("Verzeih mir, aber") make the correction even more prominent.

Example 5.11. Explicit correction with rich accountings.

148	18:31:40	L08	NATIVE04_FN, wirst du hier um 23:00 d.Z? *NATIVE04_FN, will you [* error: missing main verb] here at 23:00 G.t?* NATIVE04_FN, will you here at 23:00 German time?
149	18:31:59	N04	heute? *today?*
150	18:32:04	L08	ja *yes*
151	18:32:19	N04	keine Ahnung, mal sehen, schreib dann einfach mal, falls ich da bin, antworte ich *no idea, let's see, just write me, if I'm there, I will reply*
152	18:32:34	L08	ok)abgemacht!) *ok [smile] agree!*
153	18:33:08	N04	PS: Verzeih mir, aber es sollte heißen: "... wirst Du hier SEIN um ..." *PS: I'm sorry, but you should say: "... will you BE here at ..."*
154	18:35:47	L08	klar)das große Problem der d-en Sprache ist: man vergißt am ende des Satzes, was man am Anfang sagen wollte) *clear [smile] the big problem of the G-n language is: one forgets in the end of a sentence, what one wanted to say in the beginning*
155	18:36:10	N04	verstehe, diese zweigeteilten Verben sind sicher verwirrend *I see, these split verbs are surely confusing*

A prototypic structure of simple explicit corrections with minimal accountings is:

T1 Learner produces an error.

T2 Native speaker presents the correct version accompanied by minimal accountings.

T3 Learner continues either on the interpersonal trajectory or by an uptake.

I differentiate between immediate and delayed correction in the prototypical turn structure for corrections with rich accountings. For an immediate correction:

T1 Learner produces an error.

T2 Native speaker presents the correct version accompanied by rich accountings.

T3 Learner continues either on the interpersonal trajectory or by an uptake.

A delayed correction will appear a few turns later. The exact number of turns is not given in advance, therefore I denote it by i. A prototypical structure for a delayed correction will have the following form:

T1 Learner produces an error.

T+i Native speaker presents the correct version with rich accountings and backlinks.

T+i+1 Learner continues either on the interpersonal trajectory or by an uptake.

These prototypes will be taken into account in the modelling phase in Part III of this work.

5.2.2. Contrasting explicit corrections

The examples discussed earlier in this section deal with corrections that only present a correct version of the trouble-source. However, a direct comparison between the trouble source and the correct version in the correction turn is also a possible way to present the correction.

Example 5.12 illustrates this type of turn design for exposed corrections. The learner produces an error in turn 49. The native speaker corrects the error in turn 51 (turn 50 is a sequence closing of a different thread in talk). The correction turn contains accountings, a turn-initial request for permission to correct and a turn-final smile. Both the original trouble source unit and the replacement are part of the turn. They both are marked with quotes and the erroneous unit rejected by the negation (you X not Y).

Example 5.12. Contrasting the error and the correction (prev. Ex. 3.15).

49	19:40:59	L09	Zur Zeit bin ich frei um Diplomarbeit zu schreiben *at this time I am [* error] free in order to write [* error: 0-determiner] diploma thesis* at the moment I have free time to write my diploma thesis
50	19:41:15	L09	danke schön! :) *thank you very much*
51	19:41:58	N04	Falls ich das korrigieren darf: Du "hast" frei, nicht "bist". :-) *If I may correct you: you “have” free, not “are”. [smile]*
52	19:42:48	L09	Danke *Thank you*
53	19:43:16	L09	Ich habe frei :) *I have free [smile]* I have free time

This turn format allows highlighting of the correct version and making the action of correction more prominent. The learner thanks for the correction in turn 52 and repeats a part of her original utterance taking up the correct word.

The prototypical structure of corrections based on contrasting is the same as for simple corrections with rich accountings with a difference in the native speakers' turn T2 in an immediate correction (or T+i+1 in a delayed correction):

T2 Native speaker presents the correct version as opposed to the original version with deviations accompanied by rich accountings.

5.2.3. Explicit corrections integrated into the next relevant action

Embedded corrections as discussed in Chapter 6 do not contain accountings. Examples discussed in this section do. However, they integrate (or embed) the corrections into the next relevant action, similar to embedded corrections. The corrections are not explicitly made to the new interactional business (compare Sections 5.2.1 and 5.2.2), but they contain accountings and explicit markers for the corrected unit, which is not characteristic for embedded corrections as defined by Jefferson (1983). The majority of examples of corrections integrated into next relevant action are located in the second pair part (SPP) of an adjacency pair (question-answer, farewell-farewell).

Example 5.13 illustrates this type of correction in a location different that an SPP. Turn 304 is a repair initiation, and it contains a congruence error: neutral determiner *dieses* combined with male noun *Phraseologismus*. However, the native speaker does not understand what is meant by this referring expression, because she did not use any phraseological expressions in her previous turn. Therefore, the next relevant action for the native speaker is not a repair proper, but another repair initiation, which is produced in turn 305.

Example 5.13. The native speaker other-initiates learner's self-repair correcting a grammar error (prev. Ex. 4.7).

303	17:10:14	N02	Danke, bis Dienstag ist noch Schule und dann sind endlich Ferien :) aber die zwei Tage krieg ich jetzt auch noch rum! *Thanks, till Tuesday is still school and then are finally holidays [smile] but the two days get I as well around!* Thanks, we have classes till Tuesday, but then we have holidays, finally [smile] but the two days will pass as well!
304	17:12:11	L03	ach so, ja, ab dem 1. August, ich erinnere mich jetzt... dieses Phraseologismus habe ich noch nie gehört :) *interjection so, yes, from the 1st of August, I recall now... this* [neutrum] phraseologism [maskulinum] have I yet never heard* I see, yes, beginning on the 1st of August, I recall now... I have never heard this phraseologism [smile]
305	17:12:26	N02	welchen? ;) *which [maskulinum] one? [smile]* which one?
306	17:12:32	L03	dei Phraseologismus :) *[typo: intended "den" - determiner maskulinum] phraseologism* the phraseologism

The question word in the repair initiation is put into the correct congruence form with the word *Phraseologism* (correction) and accompanied by an emoticon (accountings). Without accountings it would become an embedded correction. The correction was noticed by the learner, she starts an attempt to self-correct in turn 306 and produces a typo, so that the error could not be self-corrected in this case. It is not fully clear from the data, if it was a self-initiated self-repair by the learner in turn 306 or if it was an uptake after the correction turn 305, because the difference in the time stamps for these two turns is very short. Parallel production of these two turns could be possible.

The correction turn is designed in order to perform the next relevant action. The correction is integrated here into a repair other-initiation, but the trouble source addressed by the repair initiation (the idiomatic expression) is not equal to the trouble source addressed by the correction (determiner-noun congruence). In contrast, in the following section I discuss variants of exposed corrections by means of other-initiations of repair where the trouble-source addressed by the repair initiation is the same as that addressed by the correction.

Minimal accountings are characteristic for integrated corrections, however, metalinguistic information may be provided after the correction. I will analyse possibilities to integrate a correction into the next relevant turn without accountings in Chapter 6 dedicated to embedded corrections. Integrated corrections can be seen as a variation of embedded corrections with accountings in terms of their integrity into the relevant next action. However, due to the present accountings, they are explicit, and can be classified as a type of exposed corrections.

5.2.4. Other-corrections by tools of OISR

Some of the corrections are delivered by the native speaker in form of repair other-initiations which is responded by the learners' self-repair. In Example 5.14 the learner produces a problematic lexical unit in turn 420. This unit is marked as trouble source by the native speaker in turn 421. There is no explicit correction (no explicit correct version of the trouble source) in this turn, however, we can assume that the native speaker understands the meaning of the word itself. This type of correction was described in language classroom research as repetition (Lyster et al., 2013) where the teacher repeats the problematic part of learner's utterance adjusting intonation to draw learner's attention to the error. The rising intonation can be marked in chat by adding a turn-final question mark (as in turn 421). It is not analysable from the data if the target meaning was clear to the native speaker or not.

Example 5.14. Repair other-initiation as a way to correct language errors (prev. Ex. 3.13).

420	21:02:32	L08	ja, ich war dort. es ist sehr maßstäblich *I was there. it is very full-scale [* error: lexical choice]* I was there. it is very full-scale
421	21:03:30	N04	maßstäblich? *full-scale?*
422	21:04:02	L08	hmm...sehr sehr groß, umfangreich *[discourse marker]... very very big, wide* hmm...very very big, wide
423	21:07:34	N04	ah, ok!

The format of this repair sequence equals to one of the formats for repair other-initiation when the native speaker is the trouble-speaker as discussed in Section 4.1 (e.g. Example 4.4), but with the difference, that the learner is here the trouble-speaker. The main difference here is in the design of the repair proper. The learner does not explain the meaning of the original reference to an object or action which caused the trouble, but tries to express the intended meaning in a different way by a different reference. Thus, recipient design plays a role in how repair carry-outs are designed after other-initiations. The identified prototypical sequential structure for repair other-initiation-based error corrections is the following:

T1 Learner produces a turn x_1 ... TS ... x_n.

T2 Native speaker marks a part of the trouble turn as unclear: $unclear(TS)$.

T3 Learner selects a different reference to the object or action instead of the trouble-source.

Candidate understandings can be made part of repair other-initiation-based correction formats. For instance, Example 5.4 discussed earlier in this chapter from the point of view of error types contains a correction of a lexical error produced by the learner in turn 287. The native speaker produces first the next projected action - a news marker in turn 288 confirming a receipt of new information and a kind of astonishment (so, the message is clear, there is no problem with the intended meaning of turn 287). The contribution to the interpersonal trajectory is immediately followed by a repair initiation marking a part of the learner's utterance as problematic (recycling) which, in turn, is followed by candidate understandings. The sequential prototypical structure can be described as follows:

T1 Learner produces a turn containing an error TS.

T2 Native speaker corrects the error by providing a candidate understanding.

T3 Learner confirms or disconfirms.

The learner only needs to produce a confirming or a disconfirming answer to a correction formatted as a polar question. There is no difference compared to repair sequences where the learner produces a repair initiation containing a candidate understanding and the native speakers just needs to confirm or to disconfirm.

5.3. Types of accountings and emphasis on correction

Native speakers used a variety of interactional resources in chat to emphasise corrections. All these resources can be categorised as those for framing and those for highlighting. Framing resources were typically quotes, dashes, brackets and parentheses, which were put on the left-hand side and on the right-hand side of the copy of the trouble source. In contrasting corrections, framing was also found to mark the corrected trouble source. Highlighting was typically realised by uppercase typing of the correction. Various examples discussed in the preceding sections contain instances of each of the two classes.

Non-symbolic highlighting was found only once in the dataset after an integrated correction. The trouble source and the correction were spelled in a similar way and differed in only one symbol. To highlight the correct version, the native speaker posted an increment to the correction where the correct version was highlighted. Example 5.15 shows the sequence from the dataset where a native speaker produces a correction in turn 442. However, he adds accountings to emphasise the correction in the subsequent turn 443, which makes the correction to an exposed correction.

Example 5.15. Accountings highlight the correction

441	21:24:34	L08	ach, ja)so, ist das richtig:du kannst mit dem Auto fahren, d.h. du hast ein Fahrerschein. ich habe richtig verstanden? *oh yes well, is it correct: you can drive a car, this means, you have a driver license [*errors: gender, lexical]. I understood correctly?* oh yes well, is it correct: you can drive a car, this means, you have a driver license. Did I understand correctly?
442	21:25:50	N04	ja, ich habe einen Führerschein und ich fahre selbst :-) *yes, I have a driver license [correction] and I drive by myself*
443	21:26:08	N04	(kein blöder Witz: es heißt wirklich "Führerschein") *(this is not a stupid joke: it is called really "Führerschein")*

Regarding a classification of accountings, I found correction formats accompanied only by emoticons, which I call minimal accountings. I found corrections where several other social actions were performed in form of accountings, such as request for permission to correct, excuses for corrections and justifications of corrections. I call them rich accountings. Some of the types of rich accountings are typically correction-initial, such as requests for permission, correction announcements and instructing. Other types are found to be typically correction-final, such as encouragements, justifications of correction and declarations of intentions.

5.4. Preliminary findings and discussion

Conversation-for-Learning as a speech exchange system, with which I have to deal during this analysis, turned out to be a supporting environment for exposed corrections. They probably would occur less frequently and be less preferred in other native/non-native speaker communication contexts, such as online dating. The participants of the chat conversations engaged in exposed corrections of various types of errors on the levels of orthography, morpho-syntax and vocabulary. A detailed analysis of practices for error corrections built the core part of this chapter.

Native speakers corrected learners' lexical errors if they caused a communication problem, but also in the absence of communication problems. Different correction formats have been used by native speakers to deal with different levels of trouble caused by wrong lexical choice. Correction formats designed as information requests (repair initiations) are preferred if the intended meaning is not (completely) clear. Pedagogical corrections presenting multiple target hypotheses in one correction are suitable to deal with ambiguity errors. Simple corrections are chosen by native speakers to deal with lexical errors in absence of a communication problem, thus, for barely pedagogical purposes.

Compared to the turn sequence in Example 5.2, the strategies applied by the native speaker and the learner to resolve ambiguities can be different. This is because some parts of turns become trouble sources for different reasons. For instance lack of knowledge causes the production of an error in Example 5.2 as opposed to to accidental deviations in spelling in Example 5.6. In addition, participants' stances and identities are relevant for the formulation of the repair.

The following correction formats were used by native speakers for corrections of linguistic errors produced by the learners:

- Simple explicit corrections presenting only the correct target to the recipient of the correction. Such corrections can be accompanied with *minimal or rich accountings*. The turn-taking can be organised in a way that the learner does not have a space to react to the correction, so called correction *on-the-fly*.
- Contrasting explicit corrections presenting to the recipient of the correction both the initial deviation and the correct target. This type of correction can be also accompanied by minimal or rich accountings.
- Exposed corrections integrated into the next projected turn are those corrections which syntactically and semantically incorporate the correct target but highlight it by symbolic means and expressions of contrast.
- Exposed corrections by means of other-initiated self-repair. These corrections are formulated as clarification requests providing the trouble-speaker the opportunity for self-correction. Candidate understandings can be also presented to the trouble-speaker replacing the reference with deviations by a correct one.

Accountings help to deal with dispreferred social actions performed by corrections of linguistic errors. As in face-to-face interaction, various types of accountings occur in text-based chat. I found corrections where accountings were minimal and were expressed first of all by symbolic means of highlighting the correct target like uppercase spelling, quotes or blinking smileys. In contrast, corrections with rich accountings include expressions of excuse, instructing, justification of correction and similar.

I found explicit corrections of lexical, morpho-syntactic and orthographical errors, however, the majority of all exposed corrections were produced to correct lexical errors. The lexical errors which became trouble sources in conversations from my dataset are of two types:

1. Problematic references to objects, their attributes or actions so that the native speaker is still able to guess what object or attribute was intended (more-wider, astrology-astronomy), see Example 5.1. This type makes no restriction for correction types.

2. Problematic references to objects, attributes or actions that make it difficult to guess what was intended and have sequential consequences: the recipient of the trouble-talk needs clarification of what is the next relevant action for the response to the trouble-turn. These errors, in turn, can be classified further in those which allow to project a small number of alternatives (e.g. Example 5.2), and those which leave the intended meaning completely open.

Since the collection of the errors is - similarly to that by Schegloff (Schegloff, 1987) - neither systematically representative nor exhausting, general conclusions would not be valuable, however, I can see the following tendencies. First, the projection only restricts the corrections to those involving the projected alternatives or are formatted as repair other-initiation. Second, for the open type, corrections formatted as repair other-initiations (with or without the candidate understanding) are the preferred type of correction. Summarised, the correction type involves some sort of a target hypothesis (what the learner is supposed to mean). The following variants are observed in the dataset:

1. A unique target hypothesis,
2. Multiple target hypotheses from a small set of possibilities,
3. Unclear targets.

Strategies to deal with multiple target hypotheses include disambiguation attempts and clarification requests. I did not find any cases in the dataset where an unclear target had sequential consequences and had to be resolved "at any cost". Let-it-pass was the preferred strategy for the continuation of the conversation and was chosen by the native speakers. The problem with determining a correct target hypothesis will be discussed from the perspective of learner language annotation in Section 9.1 and will play a role in the modelling phase in Section 9.4.3 .

Sequential positions of the trouble-turn are relevant for the selection of the correction formats. In the analysed examples, corrections on-the-fly are preferred after questions containing errors. Such corrections allow to deal with the error and immediately deliver a projected turn type, thus an answer. Delayed corrections are an alternative form to corrections on-the-fly, because they allow to close the open sequence first and then to come back to the linguistic matters, such as Example 5.11. Participants' choice of a correction format is influenced by the sequential position of the trouble-turn and the projected next action. This is specifically important for the correction formats which are integrated into next relevant turn. The correction in this case is repetition-based, but the repeated unit contains modifications which correct the error.

Tudini suggests to analyse corrections as different options "in the continuum of explicitness of exposed correction in online text chat" (Tudini, 2010, p. 101). The grade of explicitness of the corrections is determined by type and number of accountings and by the grade of integrity into ongoing talk. An operationalisation of the explicitness and possibly a corpus-based machine learning model for explicitness-based correction classification may be approached in a separate study.

6. Embedded Corrections

As already discussed in Chapter 2, embedded corrections are performed in a way that allows to avoid the loss of face and replaces the erroneous item by a new one in a discrete way. In this case, "... correction occurs, but is not what is being done interactionally" (Jefferson, 1983).

To understand the phenomenon of embedded correction we need to look again at possibilities to reference objects in conversation (see Chapter 2 for literature review). Jefferson (1983, pp. 90-93) mentions 3 possibilities for referencing: use proterms, repeat the term, and replace the term by a different one. Jefferson's hypothesis is that "when a next speaker produces, not a proterm or a repeat, but an alternative item, *correction* may be underway.

Examples 6.1 and 6.2 illustrate how an incorrect spelling of a word may corrected without making the correction interactional business (Example 6.1), and how the error can be ignored (so called let-it-pass strategy) and replaced by a preposition pointing to the object (Example 6.2).

Example 6.1. Embedded correction: the learner L04 uses a non-German version of the name of Rome (probably Italian). The native speaker N02 replaces it by the German version.

163	18:06:05	L04	ich habe 2 Wochen in Roma verbracht)) *I spent 2 weeks in Roma [* error: spelling]*
164	18:06:12	L04	das war wunderschön *it was wonderful*
165	18:06:32	N02	oooooooooooooooooooh, Rom ist die schönste Stadt der Welt!!! *[stretched discourse marker]* Rome is the most beautiful city in the world!!!
166	18:06:53	L04	jaja!! *yes yes!!*

Thus, a replacement can be used as a way to correct in an embedded way, however, not every replacement is an embedded correction of a linguistic error. Jefferson (1983) provides examples of embedded corrections where the term "nigger" is replaced by the term "Negro", the same for "police" and "cops" (Jefferson, 1983, p. 93)). These replacements are rather of stylistic or conceptual nature, and do not correct any errors in orthography, grammar or L2 lexical errors. However, it is apparently the lexical choice that is corrected in Jefferson (1983)'s examples. I found replacements similar to Jefferson (1983)'s examples in my dataset. However, the majority of all embedded corrections in the dataset is composed of embedded corrections of linguistic errors (ca. 70%).

Example 6.2. Example of avoiding an embedded correction: the learner L04 uses a non-German version of the name of Rome (this time the English version). The native speaker N02 uses a pronoun "das" to reference it in the response.

417	16:49:33	N02	in welcher Stadt in Italien bist du denn? *in which city in Italy are you now?*
418	17:04:09	L04	in Rome *in Rome [* error: spelling]*
419	17:05:36	N02	ach super, das ist sooooooooooooooooo eine schöne Stadt! *[discourse marker] super, this is such [stretched] a beautiful city!*
420	17:05:43	N02	Da wünsch ich dir ganz viel Spaß! *I wish you to enjoy it there!*

In this Chapter, I describe the types of embedded corrections found in the dataset. I discuss types of errors corrected in this way. I analyse the sequential environments where embedded corrections occur. In addition, I argue that opportunities to produce embedded corrections depend on the sequential position of the trouble turn and the unit of the trouble turn where the error occurs, and the error type. The next Section 6.1 provides a detailed analysis of embedded corrections of linguistic errors. Section 6.2 describes other types of embedded corrections that are designed to handle other types of trouble sources. The working mechanism of embedded corrections is analysed in Section 6.3. The findings are discussed in Section 6.4.

6.1. Embedded corrections of linguistic errors

An embedded correction of a linguistic error is a pair of utterances of the form:

1. Speaker A produces an utterance a containing one or more linguistic errors;
2. Speaker B produces a response r reusing the unit from the preceding turn where the error occurred, and modifying this unit in such a way that the error is corrected.

Thus, embedded corrections are recognisable as *corrections* only in a direct comparison with the corresponding trouble turn.

To prepare the empirical basis for a computational model of embedded corrections, I examine different types of embedded corrections and errors corrected by them. I describe the types of errors that were corrected implicitly in Section 6.1.1. I analyse the turn design of sequences where embedded corrections were made in Section 6.1.2.

6.1.1. Errors addressed by embedded corrections

The majority of all embedded corrections were produced for lexical errors. The collection of embedded corrections of lexical errors consists of 18 instances. In addition, there are 11 embedded corrections of morpho-syntactic errors and 8 embedded corrections of spelling errors and typos. Table 6.1 summarises the numbers of corrections per error type.

	Embedded Corrections				
Error Type	Total	N01	N02	N03	N04
Lexical	18	4	4	3	6
Morpho-syntactic	11	3	3	2	3
Spelling & typos	8	3	1	2	2
Total all error types	37	10	8	7	11

Table 6.1.: Types of errors addressed by embedded corrections in the dataset.

Example 6.3 illustrates a possible way to correct a lexical error implicitly. The term *Lieblingsgemeinshaft* (En.: favourite community) produced by the learner in turn 366 is replaced by the native speaker by the term *Lieblingsmannschaft* (En.: favourite team) in turn 367. Despite the spelling error in the last part of the composite *Lieblingsgemeinschaft*, both the trouble source and the correction are legal German words. However, the latter is a common term in the context of sports which is the topic of the talk[1].

Example 6.3. Embedded correction of a lexical error.

366	18:49:22	L08	und welche deine Lieblingsgemeinshaft für heute?) *and which [error: missing verb] your favourite community [errors: wrong lexem and wrong spelling] for today?)* and which is your favourite team today?
367	18:59:14	N04	hab heute keine Lieblingsmannschaft. ich freue mich auf ein entspanntes Zuschauen und warte auf irgendetwas Spektakuläres oder Schönes oder Furchtbares oder sonst irgend etwas, wovon man noch in Jahren spricht *don't have today any favourite team [embedded correction]* *I'm looking forward to a relaxed viewing of the game and waiting for something spectacular or beautiful or ugly or what ever that people will talk about for years to come*

The first turn-constructional unit in the native speaker's response *hab heute keine Lieblingsmannschaft* documents more precisely the information requested in learners' turn while filling the verb gap and correcting the lexical error. This part of the native speaker's turn is clearly doing more than simply producing the projected second pair part to the question.

[1]The interaction took place in the time of the European Soccer Cup in the summer 2012.

Example 6.4 shows an embedded correction of a morpho-syntactic error. The native speaker replaces the erroneous plural form of the word Test - *Teste* - by the correct one *Tests* in turn 142. The repeat-based response to a news re-formats the trouble source to a surprise source.

Example 6.4. Embedded correction of the wrong plural form.

141	18:29:10	L08	leider auch nicht(((morgen schreibe ich 2 Teste in Deutsch und Englisch. und wie du verstehst, habe ich noch nicht sie gelernt=)) *unfortunately not, either [sad smiley] tomorrow write I 2 tests [* error: wrong plural] in German and English. and as you understand, have I not yet they [*error] learned [smiley]* unfortunately not, either. I'm writing tomorrow 2 tests in German and English. and as you understand, I haven't prepared for them yet.
142	18:30:02	N04	2 Tests? ok, klar, dann erstmal viel Erfolg dabei! *2 tests? ok, I see, then good luck with them for now!*

Example 6.5 shows that spelling errors can be corrected implicitly. In this case, a farewell token *Tschüß* is replaced in the farewell reply by *Tschüss*.

Example 6.5. Embedded correction of a spelling error (only regionally used form).

82	20:53:17	L02	Gegenseitig und ja, Montag passt gut. Tschüß! *Mutually [* error: lexical choice] and yes, Monday suits good. Bye-bye [* error: spelling]!* For me too and yes, Monday suits good. Bye-bye!
83	20:53:34	N01	Tschüss und gute Nacht! *Bye-bye [correction] and good night!*

Embedded corrections of spelling errors, typos and some morpho-syntactic errors have a form of repetition with some editing, while embedded corrections of lexical errors are realised through replacements. In the Examples 6.4 and 6.5, we see replacements *Teste*→*Tests* and *Tschüß*→*Tschüss*, which are repetitions of the same term with replacements in it. In contrast, embedded correction of lexical errors replace the entire tokens or expressions, as Example 6.3 demonstrates: *Lieblingsgemeinschaft*→*Lieblingsmannschaft*.

Sometimes, a learner's utterance contains many errors. A complete correction of each of the errors could have made the interaction too much classroom-like. In addition, some of the errors are no easy to correct. The native speakers are fluent users of the language, but are not always experts in linguistic matters. If the meaning of the learner's utterance is clear, the strategy of partial correction can be selected. Then only the error should be corrected which can be corrected *easily*. As opposed to this, some erroneous units may have sequential consequences. Then a correction might be necessary. I illustrate this strategy in Example 6.6.

Line 145 of Example 6.6 contains multiple errors, a correct formulation of the question could be *Wie war dein Wochenende?*. The native speaker N02 decides not to replace as much as possible in for instance a response like *Mein Wochenende war schön aber anstrengend.* or even *Wie mein Wochenende war? Schön aber anstrengend.*, but just to make sure that only the last weekend was meant, not many of them, and to rephrase the rest. On the other hand, not-correcting the erroneous use of plural verb and possessive pronoun aligned with the singular form of the noun, and referring to the object using a proterm would mean - at least for the native speaker - that multiple weekends are mentioned, and not only the last one (there was only one weekend between the previous chat and the current one). Therefore, a replacement was necessary here.

Example 6.6. Partial embedded correction.

			L05 tells about problems with internet connection
145	17:18:41	L05	Wir waren deine Wochenende? *We [*error: typo] were [*pl. aligned with weekends] your [*pl. aligned with weekends] weekend?* How was your weekend?
146	17:21:59	N02	alles kein Problem! Ich hatte ein schönes, aber anstrengendes Wochenende. War in Darmstadt auf Wohnungssuche und hatte einen Termin nach dem anderen, das war nervig... aber abgesehen davon war es ganz schön, war abends mit Freunden weg (was trinken und - wie immer - Fußball schauen). Was hast du gemacht? *no problem at all! I had a nice but exhausting weekend. [continues] What did you do?*

To sum up, I found that embedded corrections are used to correct only those cases of errors where the recipient has access to at least one target hypothesis, as opposed to exposed corrections which may deal also with unclear targets. Examples of such error classes are wrong plural form, use of plural instead of singular noun, spelling errors and typos, missing auxiliary verb, missing determiner and use of a non-native-like expression where the intended meaning is projectable.

6.1.2. Embedded corrections of linguistic errors and turn-taking

Examples 6.3 and 6.5 of the previous section have in common the high-level principle of embedding a correction in the second pair part (SPP) of an adjacency pair as a response to the first pair part (FPP). Examples of such pairs are greeting-greeting, question-answer, farewell-farewell. This is the most frequent sequential architecture of embedded corrections in my dataset. In addition, I found instances of replacements in return questions, acknowledgements, expressions of surprise, questions in positions later than an FPP and sequence closings such as assessments and evaluations. Besides replacements, I found several instances of insertions, where a syntactic position empty in learners' utterance was filled in by the native speaker. Insertions of missing determiners and verbs are examples of such corrections. I will discuss the identified sequential environments supporting embedded corrections in the reminder of this section.

Embedded correction in a second pair part

Embedded correction in the second pair part of an adjacency pair is the most frequent type of embedded correction in my dataset. One of the instances is shown in Example 6.7. The learner L01 uses a non-standard German[2] spelling version of the farewell token *Tschüß* in turn 35. The native speaker N01 replaces the spelling version by a standard version *Tschüss* in turn 37. The correction is taken up by the learner in turn 38. I refer to this type of embedded corrections as FPP-SPP correction.

Example 6.7. Replacement of the form FPP-SPP : embedded correction of a spelling error.

35	20:53:29	L01	Ja, das passt mir) Dann Tschüß! *Yes, this suits me [smiley] Then bye [spelling error]!*
36	20:54:01	L01	Ja, das klappt bei mir *Yes, this works out for me*
37	20:54:22	N01	Okay, super! *Okay, super* Dann bis dahin. Tschüss. *See you then. Bye [embedded correction].*
38	20:54:44	L01	Tschüss. *Bye [accepting the correction]*

Non-present auxiliary verbs and determiners are frequent learner errors. Embedded corrections of such omission errors are insertions. In Example 6.8 the learner L01 drops the verb *war* in her question (FPP) in turn 277. The native speaker N01 inserts the missing verb in his answer in turn 280.

Example 6.8. Insertion of the form FPP-SPP: embedded correction of a morpho-syntactic error.

277	21:08:00	L01	Wie deine Arbeit? *How [error: missing auxiliary] your work?* How was your work?
278	21:08:04	N01	Wann hast du denn die Prüfung? *When do you have the exam?*
279	21:08:27	L01	um 15.00 *at 3 p.m.*
280	21:09:57	N01	Meine Arbeit war ganz ok. Wie immer eigentlich. Die Schüler waren brav. Die Kollegen waren nett. Und es waren keine Eltern da, um sich zu beschweren ;-) *My work was [insertion] quite ok. As usually actually. The pupils were good. The colleagues were nice. And there were no parents there to complain [smile]*

A prototypical structure of this type of the FPP-SPP correction consists of two steps:

[2]Nothern-German according to Duden online dictionary. The dictionary lists two correct versions of this farewell token, *tschüs* and *tschüss*.

1. Speaker A produces an FPP containing a problematic unit.
2. Speaker B produces an SPP containing a replacement or an insertion.

Deletions of some tokens which may be present but not required by the utterance structure, such as zero-article errors, were not found in the dataset. Nonetheless, such deletions may be also produced by means of replacement in an SPP.

Embedded corrections in counter questions

Environments suitable for counter questions such as *and you?* provide an opportunity to embed a correction of a linguistic error. The sequential place of the embedded correction depends on who is the initiator of the initial question and who produces the counter question, which turn contains the trouble source, and which turn(s) can contain the correction. I explain these sequential variants in the following two examples.

The first sequential variant can be mapped to the following prototypical structure:

1. The native speaker produces the initial question: Q_1.
2. The non-native speaker produces a response containing an error followed optionally by a counter question: $R(Q_1)^{error}$, $Q_{counter}$.
3. The native speaker responds to the counter question with an embedded correction: $R(Q_{counter})^{correction}$. Even if Q_{return} in the ancestor turn is omitted, the response still has a form of $R(Q_{return})$ and contains a replacement.

Example 6.9 illustrates this schema: L05 uses a word by word translation from Russian *Ausgehtage* (выходные дни, going-out days) instead of a standard German *Wochenende* (weekend). N02 replaces this token in her answer *Wochenende war toll* and reuses the same token in her return question *Wie war dein Wochenende?*

Example 6.9. Embedded correction in a counter question.

89	16:58:34	L05	Hallo! Wie geht es? Wie waren deine Ausgehtage? *Hello! How are you? How were your going-out days [error]?*
90	16:59:55	N02	Hallo! Danke, mir geht es gut, und dir? Wochenende war toll, war auf dem Geburtstag einer Freundin, gestern hab ich Wohnung aufgeräumt usw, also alles erledigt, was so angefallen ist, das musste auch mal sein. Wie war dein Wochenende? *Hello! Thank you, I am fine, and you? The weekend [replacement] was excellent, was at a birthday party of a friend, [continuing] How was your weekend?*

In some contexts, *Ausgehtage* could be a legal expression, for example if a speaker makes it relevant that an action of going-out took place on a particular day. This expression is not restricted to weekends (*Der beliebteste Ausgehtag ist Donnerstag*, En.: The most popular going-out day is Thursday). However, the chat participants L05 and N02 did not make the action of going-out relevant in the previous dialogue. In addition, they met on 13th of June, and the subsequent time (the chat where Example 6.9 was taken from) took place on the 18th of June. Thus, both participants knew that only one weekend was between their chats, none of the participants could talk in this context about multiple weekends. In Russian, the phrase выходной день (vykhodnoy den') can denote every day where one does not have to work (weekend, legal holiday, a day off), but normally it is used to refer to weekend. Therefore, this replacement is analysable as a correction of an error in lexical choice, and not as just a disambiguation.

The second prototypical structure of a sequence involving embedded corrections in counter questions has the following form:

1. The non-native speaker produces a question containing an error: Q_1^{error}.
2. The non-native speaker responds to the question. The response is followed by a counter question. Both, the response and the counter question may contain an embedded correction: $R(Q_1)^{correction}$, $Q_{counter}^{correction}$.
3. The non-native speaker responds to the counter question. The response may contain an uptake $R(Q_{counter})$.

An instance of this type is shown in Example 6.10. In turn 248 a wrong plural form *Abenteuers* is produced and replaced by a correct plural form *Abenteuer* in turn 252 taking the form of an answer to an imaginary counter question.

Example 6.10. Replacement of the form "response to an (imaginary) return question".

248	08:12:52	L06	manchmal sind die Träume erschrecklich. besonders mag ich nicht, wenn ich meine toten Verwandten sehe. aber meistens habe ich irgendwelche Abenteuers *sometimes are the dreams erschrecklich [*error: creation]. especially I don't like, when I see my dead relatives. but in the most cases I have some adventures [* error: wrong plural]*
249	08:14:26	N03	krass! *Gosh!*
250	08:14:41	N03	ich habe noch nie tote verwandte in traeumen gesehen! *I have never seen dead relatives in my dreams*
251	08:14:46	N03	das klingt schrecklich! *it sounds terrible [correction of the lexical error]*
252	08:14:53	N03	ich habe meist abenteuer *I have adventures [corrected plural form] in the most cases*

To sum up, the counter question environment provides opportunities for embedded corrections regardless of who produced the first question. A successful production of an embedded correction in a counter question depends on the appropriateness of the counter question itself.

Embedded corrections in expressions of surprise

In the examples of embedded corrections in expressions of surprise, the trouble source is reused to become a surprise source. In Example 6.4 discussed earlier in this chapter, the non-native speaker produces a message containing a problematic unit *2 Teste*. The native speaker's response starts with a repetition of the trouble source formatted as surprise source "2 Tests?" followed by a wish of good luck.

Example 6.11 shows another case of embedded corrections formatted as an expression of surprise. Turn 186 contains the trouble source *in dem Kindergarten*. In turn 188, N03 recycles the trouble source to display surprise and corrects the error. The same turn contains an evaluation. The native speaker keeps the floor and provides information about himself repeating the corrected version again in turn 189.

The prototypical sequential structure of using the surprise environment to embed corrections of linguistic errors can be described as follows:

- T1 The non-native speaker tells a news (e.g. a new fact about herself) that contains a trouble source which is a potential surprise source.
- T2 The native speaker produces an expression of surprise by reusing the trouble source from the preceding turn as a surprise source. An evaluation, a ratification, an assessment, an information request, a "me too" statement may continue the surprise turn.

Example 6.11. Embedded correction instead of a 1:1-repetition in expressions of surprise.

186	19:37:51	L06	ich arbeite als Erzieherin in dem Kindergarten *I work as a nursery nurse in the [*error] kindergarden*
187	19:38:12	L06	wie war dein Urlaub? *how were your holidays?*
188	19:39:15	N03	Im Kindergarten? !Super. *In a [correction] kindergarden? !Super.*
189	19:39:22	N03	Ich habe auch mal ein Jahr im Kindergarten gearbeitet *I used to work a year long in a [repetition of a correction] kindergarden, too*
190	19:39:34	N03	Der Urlaub war toll *The holidays were awesome*

An expression of surprise, as everything else, cannot be placed randomly in conversation. There are turns designed to elicit a surprise response, for instance "negative observations, and extreme case formulations are common components of turns treated as surprise sources" (Wilkinson and Kitzinger, 2006, p. 157). If a surprise reaction follows a turn not designed for a surprise reaction, an elaboration is needed, see (Wilkinson and Kitzinger, 2006, p. 159). An elaboration locates the surprise source (often a repetition or a paraphrase of the surprise source) and explains, why it is surprising.

Corrections embedded in an information request

Repair initiations and counter questions provide opportunities for embedded corrections. Another type of question where embedded corrections were found is information request, as shown in Example 6.12. The non-native speaker L02 produces a problematic unit *Halbezeit* in turn 283 (En.: *half of the time*). The native speaker N01 requests more information about something related to the current topic of the talk. This information request contains a replacement *Halbzeit* (half-time) in turn 288.

A prototypical sequential structure for this type of embedded correction may be the following:

T1 The non-native speaker produces a second pair part containing an error.

T2 The native speaker initiates a post-expansion by an information request, which contains a correction.

Example 6.12. Replacement of the form "request more information".

283	20:00:15	L02	Hallo, NATIVE01_FN! Ich statte dir meinen Glückwunsch zum Sieg deiner Mannschaft ab! Ich habe gestern eine Halbezeit angesehen. ;) *Hello, NATIVE01_FN! I am paying you my congratulations [* error: lexical] to the win of your team! I watched a half-of-time [*error: orthography] yesterday. [smile]* Hello, NATIVE01_FN! Congratulations on the win of your team! I watched until halftime yesterday.
284	20:00:32	N01	Vielen Dank! Vielen Dank! *Thanks a lot! Thanks a lot!*
285	20:00:41	N01	Ich bin begeistert :-) *I am excited [smile]*
286	20:01:30	N01	Jetzt sind wir fast schon im Viertelfinale. *We are almost in the quarterfinal.*
287	20:01:36	L02	Obwohl ich kein Profi bin, denke ich, dass die deutschen Fußballsieler gestern der Situation gewachsen waren. *Although I am not a professional, I think that the German football players [* error: spelling] were up to the task yesterday.*
288	20:01:54	N01	Hast du die erste oder die zweite Halbzeit angesehen? *Did you watch the first or the second half-time?*
289	20:02:04	L02	Erste *First*
290	20:02:38	N01	Gut, das war die bessere Hälfte. Mit zwei deutschen Toren :-) *Good, this was the better half. With two German goals [smile]*

Embedded corrections in information requests after an SPP represent embedded corrections in initiations of non-minimal post-expansions. However, embedded corrections may also occur at other places in post-expansions as the next section shows.

Embedded correction accompanied by acknowledgement tokens

The example discussed in this section represents embedded corrections in non-minimal post-expansions later than its initiation. Such non-minimal post-expansions are referred to as *topicalisation* in CA literature (Schegloff, 2007). Example 6.13 shows how a turn with an embedded correction can be formatted. The non-native speaker produces a message containing a problematic unit. The native speaker replies with an acknowledgement / confirmation token followed by a repetition of the trouble source with correction. I found 3 sequences of this type in the dataset, however, more examples of this kind can be found in the CA literature, see for instance (Jefferson, 1983, p.94 (17), (18)).

Example 6.13. Embedded correction after an acknowledgement token (prev. Ex. 3.16).

228	20:15:33	L02	Dann brauchst du irgendetwas Beruhigungsmittel einzunehmen))) *Then you need to ingest something tranquilliser [* error] [smile]* Then you need to ingest something tranquillising
229	20:16:02	N01	Genau, zur Beruhigung meiner Nerven :-) *Exactly, to tranquillise my nerves*

Although the error and the correction turns are part of a non-minimal post-expansion, the correction turn is analysably a projected second pair part to the learner's suggestion, which is a first pair part. Such pair parts may be placed also elsewhere, not only in post-expansions. A prototypical structure of an embedded correction in combination with acknowledgement tokens can be specified as follows:

1. The non-native speaker produces a turn projecting agreement or disagreement in the response.
2. The expert produces a confirmation / agreement containing an acknowledgement token and a corrected version of the trouble source.

Similar to an expression of agreement, variants with expressions of disagreement or rejection of an offer are potential candidates for embedding a correction. However, I did not find any instance of this variant, neither in my dataset nor in the related literature.

Embedded correction in an assessment

Examples 6.1 and 6.10 provided earlier in this section contain embedded corrections in assessments. Example 6.1 is reproduced below to make its discussion more convenient. In current Example 6.14, a non-standard German spelling of the word *Rom* (En.: Rome) in turn 163 is replaced by the standard German spelling in turn 165. The native speaker shares her opinion about the place in form of an assessment. The place is referred to by a full name in both turns. The correction of the spelling error is realised through the repetition of the full reference to the place which the participants were talking about.

Example 6.14. Embedded correction in an assessment (prev. Ex. 6.1)

162	18:05:50	L04	Vor 2 Wochen kam ich aus Italien) *2 weeks ago I came from Italy*
163	18:06:05	L04	ich habe 2 Wochen in Roma verbracht)) *I have 2 weeks in Roma [spelling error] spent* I spent 2 weeks in Rome
164	18:06:12	L04	das war wunderschön *it was wonderfull*
165	18:06:32	N02	oooooooooooooooooh, Rom ist die schönste Stadt der Welt!!! *[discourse marker] Rom [correction] is the most beautiful city of the world* oooooooooooooooooh, Rome is the most beautiful city in the world!!!

Assessments and evaluations are frequently found in the position of a minimal post-expansion or "sequence-closing thirds" (Schegloff, 2007). The turn sequence analysed in Example 6.14 is part of a non-minimal post-expansion (topicalisation). The correction is, in turn, embedded in the projected next action after the action of telling in turns 162-164.

To sum up, there are two levels of analysis required for embedded correction. First, on the level of turn pairs to find the mechanism of the "correcting", and second, on the level of the place of this turn pair within the larger turn sequence, such as post-expansion.

6.2. Other types of embedded corrections

The purpose of this section is to support a conceptual understanding of the error-correction dichotomy from the radically emic perspective on data analysis postulated by Conversation Analysis. From this perspective, particular units in talk are qualified as errors if they are *handled* as errors by interaction participants. More specifically, not only language errors can be handled by conversation participants as something that needs a correction. The recipient of the talk may find some features in style, political correctness, preciseness or emotional marking of a reference to an object, attribute or action as requiring a replacement by another term which may be more precise, more polite or more neutral. Such other classes of corrections may be conceptual replacements and stylistic replacements. In this way, the types of embedded corrections discussed in this section are closer to types discussed in (Jefferson, 1983)'s analysis of native speaker data.

6.2.1. Conceptual embedded corrections

In conceptual embedded corrections, a unit is replaced by a different unit, although the first unit does not contain any formal linguistic error. To better understand the difference between embedded corrections of linguistic errors and conceptual replacements, the examples below are provided.

In Example 6.15, the native speaker N01 replaces the neutral word *Anhänger* (supporter, follower) by a stronger concept *Fan* (very enthusiastic supporter, fan). This is even emphasised by the adjective phrase *sehr sehr großer* (very very big). It was not wrong to use the word *Anhänger* in this case. However, the native speaker chose to use a different concept, maybe reflecting more precisely his passion for soccer.

Example 6.15. Replacement of a concept by another concept with no linguistic error.

98	20:11:38	L02	Bist du Fußball-Anhänger, spielst du selbst? *Are you football-supporter, do you play yourself?*
99	20:13:12	N01	Ja, absolut. Ich bin ein sehr sehr großer Fußball-Fan :-) Vor allem wenn eine Welt- oder Europameisterschaft stattfindet. *Yes, absolutely. I am a very very big football fan [smile] Most of all when a world or Europe cup takes place.*

I found three different kinds of conceptual replacements in my data:

1. Replacement of a concept by a more specific concept, for instance *read* $\rightarrow$ *read books*. Such replacements may be figurally described as zooming in or out.
2. Replacement of a concept by a stronger / weaker concept. Example 6.15 contains such a replacement. Other examples of this kind are *interesting* $\rightarrow$ *very interesting, exciting* and *you are today a ...* $\rightarrow$ *I am something like a ...* .

Further research on typology of replacements would probably discover more different kinds of what I call conceptual replacements. Although this type of correction is closer to embedded corrections in native speaker data described by (Jefferson, 1983), differences in comparison to native/non-native speaker data may exist and can be detected in a future study.

6.2.2. Embedded correction of style

Instant messaging communication is a conceptually oral type of discourse where participants tend to use short phrases instead of full grammatically correct sentences. This is also a way to adjust the grade of social proximity. The regulation of the social proximity may be also performed by the lexical choice where participants select among more formal or marked and more colloquial or unmarked expressions.

I found three different kinds of conceptual replacements in my data:

1. Replacement of a marked expression by an unmarked expression, for instance *to take pleasure in something* $\rightarrow$ *to enjoy something*. Example 6.16 contains an instance of a replacement of this kind.
2. Replacement of more official, formal terms by more colloquial, for instance *White Russia* $\rightarrow$ *Republic of Belarus*

In Example 6.16 the non-native speaker uses a marked expression an etw. Gefallen finden (En.: *to take pleasure in something*). The native speaker replaces it by a more neutral, unmarked *hat mich auch gefreut* (En.: *I enjoyed it, too*). The use of the adverb *auch* makes the replacement analysably to a correction saying "I did the same thing as you did, but I use a different name for it".

Example 6.16. Replacement of a marked expression by an unmarked.

35	20:35:57	L05	OK, es passt mir ganz gut)))) So, bis morgen, ich fand Gefallen an unserem Chatten, gute Nacht!)) *OK, it fits me quite well [smile] Well, see you tomorrow, I took pleasure in our chat, good night! [smile]* translation
36	20:37:12	N02	ok schön, dann bis morgen, gute Nacht! :-) *ok great, then see you tomorrow, good night! [smile]*
37	20:37:16	N02	hat mich auch gefreut *I enjoyed it, too*

Non-native speakers tend to use additional language resources to deal with troubles in production. When using dictionaries, they may face the problem of selection of an appropriate expression to serve a particular pragmatic need. This may lead to an incorrect use of correct expressions in a second language, as illustrated in Example 6.16. Therefore, such cases may be also classified as linguistic errors.

6.3. The mechanics of embedded corrections

Repetitions and replacements do not happen incidentally, but perform a particular kind of work in conversation. Repetitions can be used, for instance to formulate a topical focus, to confirm receipt of a prior turn (Tannen, 1987), to claim epistemic authority over the matter (Stivers, 2005), to initiate repair (Schegloff et al., 1977; Rieger, 2003) or to resist presuppositions generated by questions, to mark questions as problematic and display a problem in question comprehension (Bolden, 2009). To perform these types of interactional work, some particular parts of a preceding turn may need to be repeated, and some other may not. In addition, some units in a turn may exist, that cannot be repeated. This, in turn, may influence the opportunities to embed a correction in the repetition. With this motivation, the questions that this section seeks to answer is, if there is any connection between the repeatable units and embedded corrections.

As discussed earlier in this chapter, there are various sequential environments in chat where an embedded correction can occur. However, the majority of embedded corrections was found in question-answer pairs. Therefore, the following specific questions are addressed in this section:

1. Which errors produced in a question can be addressed in an embedded correction in a response to the question?

2. Do any properties of the error in a question exist that make an embedded correction of this error in the answer impossible?
3. Are there cases where embedded corrections, are unavoidable?

I use a collection of question-answer pairs to get an intuition from the data, if there are dependencies between error types and embedded corrections. The collection was composed of all questions produced by learners and responses to these questions produced by native speakers. Questions that did non receive any response were not included into collection, because no correction can be embedded in a non-existing response.

6.3.1. A modified coding scheme for question-answer pairs

For the decision what is a question and what is not a question but looks like a question I use the question coding scheme introduced by Stivers and Enfield (2010). According to their criteria for questions, an utterance qualifies to be coded as a question in the following cases:

1. It is a formal question if it employs lexico-morpho-syntactic interrogative marking.
2. It is a functional question if it effectively seeks to elicit information, confirmation or agreement regardless of the sentence type.
3. It is a news marker like “Really?”. News markers qualify as formal questions because they are treated as seeking confirmation.

Other types of utterances formatted as questions (questions seeking acknowledgment, continuers, requests for immediate physical action or questions offered in reported speech) were not classified as questions. In addition, I found special cases of questions that I did not considered as questions, for instance, repetitions of the same question because of connectivity problems. Then, only the first occurrence of a question was included in the collection of questions.

The unit of analysis is here a question, and not a message, because a question in chat can consist of multiple turns or be only a part of a turn. Due to the medially written communication mode, the following had to be added to the definition of questions:

1. Turns may contain phrases related to different threads. In this case only the phrase formatted as a question was considered as one question for the analysis, and not the complete message.
2. The opposite case when many phrases from one or many messages form one question. In such cases, the question is frequently formed by a declarative utterance followed by a confirmation request or an element corresponding to a turn-final element (Stivers and Enfield, 2010). The response is then related to the whole question stretching over multiple turns, and not only to the part of it formally designed as a question. Such sequences of messages are handled as one question.

I found 481 question-answer pairs in the dataset. All questions were independently annotated by two human annotators. The annotators were one German native speaker and one non-native speaker with a native-like fluency in German, both experienced IM users. The annotation was performed in the following way:

- Each of the annotators marked first all questions that contained some deviations compared to German language standard or non-native like constructions. The annotators used Duden online dictionary as a reference.
- Because not all deviations from German language standard can be addressed in chat as linguistic errors, the annotators intuitively marked questions containing "real" errors with a "real" error flag.
- All questions containing "real" errors were coded as polar, content or multiquestion according to the extended question coding scheme based on the scheme proposed by Stivers and Enfield (2010).

The intuitive "real" error annotation is sufficient for the purpose of this section, however, a proper definition of the concept of linguistic errors in chat is required for an automatic error recognition and will be subject of Section 7.1.

For each "real" error, I annotated the error types (what exactly is the problem), the trouble source in the original question and the references to the trouble source in the original NS's responses. Only 17 questions containing real errors received an embedded correction in native speakers' responses. Therefore, in order to make the collection of examples, a methodological change was introduced, which is normally not applied in CA-driven research. In order to "simulate" an artificial dialogue system which should generate embedded corrections, I modified native speakers' responses to questions that did not receive any embedded correction in the original corresponding response in the following way. I added embedded corrections to the responses trying to preserve the social action of the original response.

This methodological modification was motivated as follows. First, embedded corrections were interspersed into answers to questions, which are *the* place in which embedded corrections have been found in naturally-occurring data. Second, only repeatable units of the question could be reused in the answer, otherwise the modified responses were no longer valid answers. In addition, the analysis is restricted to the local question-answer pair, it does not address the reasons for a present or non-present embedded correction in a larger interactional context. Consequently, responses modified by the researcher could not be modified in a way, which could lead to a falsification of the result.

After the error annotation and response modification I had three parts in the collection of the questions with "real" errors:

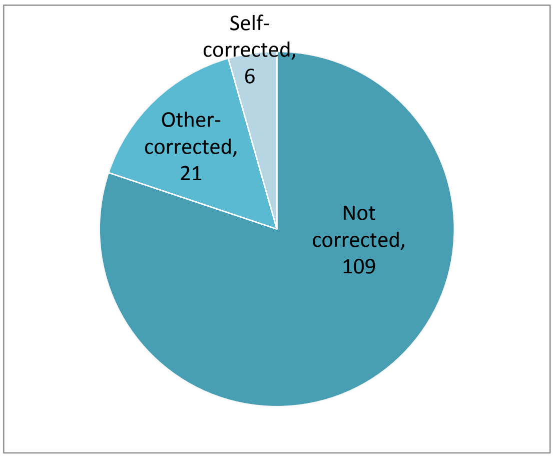
1. EC question-answer pairs where a question is produced by the learner, and an answer is produced by the native speaker. These are answers which did not contain any corrections and which were extended by an embedded correction. This collection includes 74 examples.

Metrics	L01	L02	L03	L04	L05	L06	L07	L08	L09	Sum
Total questions	22	38	51	75	41	37	52	90	75	481
Errors Duden	11	13	46	73	35	38	50	87	75	428
Real errors	2	12	19	14	26	8	11	28	16	136
EC	0	9	15	6	15	3	5	11	10	74
$NOEC$	1	1	2	6	8	5	5	11	6	45
EMB	1	2	2	2	3	0	1	6	0	17

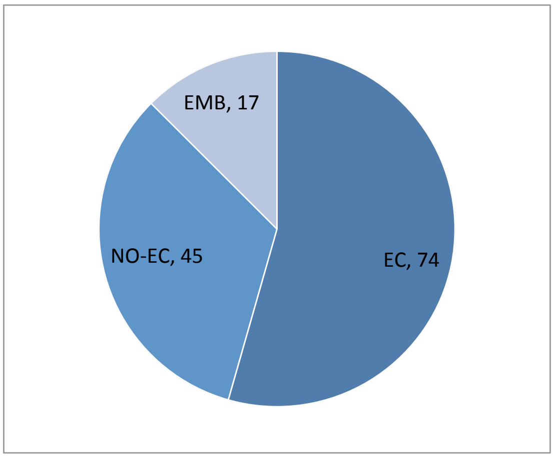
Table 6.2.: Questions, errors and embedded corrections in responses

2. $NOEC$ question-answer pairs where a question is produced by the learner, and an answer is produced by the native speaker without any embedded correction, but the answers could not be extended by an embedded correction. This collection includes 45 examples.
3. EMB question-answer pairs where a question is produced by the learner, and the answer produced by the native speaker contains an embedded correction of a linguistic error. This collection includes 17 examples.

Table 6.2 summarises numbers of questions, questions with deviations from the German standard as specified by Duden and "real" errors. Figure 6.1 illustrates the composition of the collection of questions with real errors. From 136 questions in total, 6 received a self-correction by the learner, and 21 received a response containing either an exposed or an embedded correction (Fig. 6.1a). The composition of the same set after the interspersal with embedded corrections is shown in Fig. 6.1b.



(a) Composition of the original collection



(b) Composition of the enriched collection

Figure 6.1.: Questions with real errors: partitions

Responses to questions which received a self-correction were modified because self-corrections do not change the basic correctability of the *error type*. However, none of self-corrected errors was other-corrected in the subsequent continuation of the interaction. This must be taken into account in the final decision model which should determine if a correction should be produced or not. I will come back to this problem in Section 10.2.4.

To understand what distinguishes errors that can be corrected implicitly from those that cannot, I analysed error types in questions. I describe below the properties of the errors for which an embedded correction can be produced as opposed to errors for which embedded corrections are not possible. In addition, there are examples of embedded corrections for errors which do not count as errors in chat. Such corrections appear because there is a need for a repetition of the reference to an object and the first mention of the object contained an error, for instance a typo. I discuss possible, impossible and unavoidable embedded corrections in the remainder of this section. Because polar questions were the best operationalised type of questions covered in the coding scheme by Stivers and Enfield (2010), I focus only on *polar questions* and *answers* to them.

According to the question coding scheme by Stivers and Enfield (2010), a polar question can receive a response which handles the question "as put" by a *confirming* or a *disconfirming* reaction regardless of form. Such responses are classified as *answers*. All other responses to polar questions, including I-dont-knows, indirect responses and responses requiring inferences from the response recipient are classified as *non-answers*. Answers can be further divided into interjection-based answers (e.g. *yes*, *no*), marked interjection-based answers (e.g. *absolutely*) and repetition-based answers (e.g. *Are you going to the cinema tonight? - I'm going...*). The relevant part of the scheme with modifications is provided in Table B.1 of Appendix B. This annotation step specifically targeted responses to questions based on repetitions and replacements, because opportunities for repetitions and replacements provide conditions for embedded corrections.

6.3.2. Errors for which embedded corrections are produceable

I will argue in this section that an embedded correction is produceable in a question-answer pair if the error-constituent from the question can be repeated in the answer, and if there is a replacement in it that corrects the error. I presented and discussed many examples of successful embedded corrections in Chapter 6. I consider the joint set $EC \cup EMB$ to find commonalities in embedded corrections and to describe them on a more abstract level.

Error categories not excluded from embedded correction in Section 6.3.4 can be potentially implicitly corrected. I cannot provide a general description of *the* error types in question which are potentially implicitly correctable, because there are different types of questions (content, polar, alternative and through-produced multi-questions) and each of these types can be further classified by, for instance, form, social action and possible response form.

The question coding scheme suggested by Stivers&Enfield and modified in this dissertation (Appendix B) provides a good basis for further investigations of these dependencies in a separate study.

Repetition-based answers to polar questions

The set of all polar questions from the set $EC \cup EMB$ which received an answer as a response will be referred to as $QAPA$. The size of the set is $|QAPA| = 30$ and it is composed of $|EC \cap QAPA| = 25$ and $|EMB \cap QAPA| = 5$. Examples of polar questions with different types of answers are provided in Table 6.3. I made the following observations after the analysis of the $QAPA$ set (number of examples in parentheses):

1. Confirming answers (22):
   a) based on repetitions can correct errors (4);
   b) based only on interjections (marked or unmarked) do not correct errors (14);
   c) based on replacement do not correct errors (4).
2. Disconfirming answers (7):
   a) based on repetitions do not correct errors (1);
   b) based only on interjections (marked or unmarked) do not correct errors (3);
   c) based on replacement can correct errors (3).
3. Jein (1): this sort of response was found doing exposed correction of learners errors by a disambiguation (sequence presented in Example 5.2). However, an answer version with an embedded correction is possible. Both answer variants (exposed and embedded) to this question are based on repetitions.

As explained in the beginning of this section, it was possible to find modified answers performing the same social action as the original answers, but in addition providing embedded corrections for at least one error identified in the corresponding question. All modified answers with embedded corrections and original embedded corrections are based on partial, full or modified repeats of the question. In addition, the following regularities were found in the modified answers:

1. Embedded corrections addressing spelling errors are based on repeats of the erroneous word with replacements of letters.
2. Embedded corrections addressing morpho-syntactic errors are based on repeats of the erroneous constituent with replacements of morphemes or changes in the word order.
3. Embedded corrections addressing errors in vocabulary and pragmatics are based on repeats of parts of the question with replacements of words or phrases.

To sum up, partial modified repeats and full modified repeats can implicitly correct errors.

I illustrate three types of errors in learners questions with original answers produced by native speakers and modified answers with embedded corrections on examples. The errors and the replacements are underlined. The modified response is marked by the word *modified* and *EC* (embedded correction) on the left.

Answer form		Question-answer pair
Repetition	Q	Hast du keine Artzneien zur Beruhigung deiner Nerven eingenommen ?) *Did you take no medicine [*error: spelling] to tranquilise your nerves?*
	A	Ich habe 1000 Tabletten eingenommen ;-) *I took 1000 pills*
Interjection + repetition	Q	ach, ja)so, ist das richtig:du kannst mit dem Auto fahren, d.h. du hast ein Fahrerschein. ich habe richtig verstanden? *oh yes well, is it correct: you can drive a car, this means, you have a driver license [*errors: gender, lexical]. I understood correctly?*
	A	ja, ich habe einen Führerschein und ich fahre selbst :-) *yes, I have a driver license and I drive by myself*
Interjection	Q	hast du etwas über deutsche Gruppe PUR gehören? *Have you heard [*error: wrong form] something about German band [*error: zero-article] PUR?*
	A	ja, kenn ich *yes, I know [them]*
Marked interjection	Q	studiertest du physik dort? *did you study [*error: wrong tense] physics there?*
	A	genau *exactly*
Replacement	Q	kennst du vielleicht irgendwelche studiumsprogramme für ausländer_ *do you maybe know some study programs [*error: composite form] for foreigners?*
	A	hm, da muss ich mal überlegen, aber so spontan fällt mir nichts ein *hm, I need to think about it, but for now nothing comes into my mind*
Interjection + Replacement	Q	ja, vielleicht ist diese Periode in Deutschland ein echtes Volksfest? *yes, maybe this period in Germany is a real folksfest?*
	A	ja, das kann man so sagen!!! Die Stimmung ist echt gut *yes, you can say like this!!! The mood is really good*

Table 6.3.: Forms of answer-responses to polar questions.

Example 6.17 shows an instance of a morpho-syntactic error which is wrong subject-verb congruence in person. The original answer in turn 126 is interjection-based and does not address the error. The modified response is based on a full modified repeat of the question. The error is verb-subject congruence. To correct this error, the verb-subject constituent needs to be repeated with correction in order to present the correct congruence relationship.

Example 6.17. A modified answer correcting a morpho-syntactic error: embedded correction based on a full modified repetition.

125	20:57:49	L02	Hast dir dieser Film gefallen? *Have [*error: wrong person] you liked this movie?* Did you like this movie?
126	20:58:29	N01	Ich muss ehrlich sein: Nein, nicht besonders ;-) *Frankly speaking: no, not really [smile]*
	modified	EC	Ich muss ehrlich sein: Nein, dieser Film hat mir nicht besonders gefallen ;-) *Frankly speaking: no, I have [correction] not really liked this movie [smile]* Frankly speaking: no, I did not really like this movie

Because it is just one additional letter, inserted into the correct form of *haben - hat -* it is also likely, that the error was just caused by quick typing, and not by lack of knowledge of German verb conjugation. However, *if* this error is a morpho-syntactic error and should be corrected implicitly, the verb-subject phrase needs to be repeated, and the wrong verb form needs to be replaced by the correct one.

Example 6.18 illustrates how errors in orthography and lexical choice can be corrected implicitly. L03 uses a non-native-like expression to refer to the school vacation in turn 79. This non-native-like expression is used in a wrong number which leads to a verb-subject congruence error. In addition, she spells the German name of Bavaria in a wrong way, which can be interpreted as lack of knowledge and not as a typo, because it is a special case of spelling.

Example 6.18. A full-repeat-based response correcting errors in vocabulary an spelling (simplified).

79	13:22:49	L03	)))) ja, wahrscheinlich! Sind die Grenze des Schuljahres von Urlaubssaison in Beiern abhängig? *[smile] yes, maybe! Are the border [*error: wrong number] of the school year [*error: lexical] depending from the vacation time in Bavaria [*error: spelling]?* Does school vacation time depend on the vacation time in Bavaria?
81	13:23:41	N02	ja genau! ist das bei euch auch so? *yes, exactly! is it the same in your place?*
	modified	EC	Ja genau! Die Ferienzeiten sind von der Urlaubssaison in Bayern abhängig. Ist das bei euch auch so? *Yes, exactly! The school vacation time is depending on the vacation time in Bavaria. Is it the same in your place?* Yes, exactly! The school vacation time depends on the vacation time in Bavaria. Is it the same in your place?

To correct the error in the subject of the question in turn 79, two variants of correction are possible:

1. Repeat the non-native-like expression using the correct form of the noun *Grenzen* or
2. Replace the phrase by a different phrase focusing on the native-like expression and not on the form of a non-native-like expression.

I focus on the native-like expression, because the majority of all corrected errors in the dataset were lexical errors, thus, errors in meaning is more important than errors in form. The congruence error is not specifically addressed to in the chosen correction. To make it visible, that it is a *replacement* of the subject constituent, the verb-subject constituent needs to be repeated. Therefore, a full repeat with modifications is used to produce an answer with an embedded correction. To correct the spelling error in the German name of Bavaria, it would be enough to use the correctly spelled word in the response.

The minimal correcting replacement

I make here first attempt to formulate abstract rules for embedded corrections in answers to polar questions. For this purpose, I introduce the concept of the *minimal correcting replacement*. Let us look again at Examples 6.17 and 6.18. *Hast* is a valid German verb form, *dieser Film* is a valid German noun, only if they appear together as a verb-subject pair in a sentence, it comes to an error. On the other hand, *die Grenze des Schuljahres* by itself may sound strange, but the congruence error cannot be detected. However, put together with the verb in the verb in plural *sind*, the congruence error becomes visible. And both errors can only be corrected, if the relationship between these words (dieser Film hat, die Grenze des Schuljahres / die Ferienzeiten sind) is corrected.

The minimal correcting replacement is a repetition of a constituent containing an error with the following properties:

1. It is the smallest constituent so that the error scope is within the constituent.
2. The replacement corrects the error.

Then, if there is a question containing an error, this error is potentially implicitly correctable, if there is a way to use the minimal correcting replacement in the response to this question. Specifically, an answer to a polar question will contain an embedded correction, if the answer is based on a modified repeat and this repeat is a minimal correcting replacement.

6.3.3. Embedded corrections dedicated to counter questions

Some morpho-syntactic constructions used in questions cannot be repeated in the direct responses to questions. However, they can be repeated in a counter question. Such morpho-syntactic constructions are pronouns and verbs in the second person (pl. and sg.), specific word order in questions, question words (wh-words). The set $NOEC$ contains 6 questions where errors could be corrected in counter questions. However, the "correctability" is not determined just by the error types, but by the information which the users already exchanged, the information type requested in the question and "territories of knowledge" (Heritage, 2012). Therefore, these questions are included into $NOEC$ set of questions.

Example 6.19. An error that can be corrected only in a counter question.

162	18:47:26	L07	im allgemeinen beenden sie Uni später als wir...wie lange denn studieren sie-? *in general they [*error] finish the university later than we... how long do they [*error] study-?* in general you finish the university later than we... how long do you study-?
163	18:48:16	N03	mein bachelor war eigentlich 3 jahre *well, my bachelor was 3 years*
164	18:48:32	N03	fuer mich aber 4 jahre weil ich ein jahr in bruessel studiert habe *but for me 4 years because I studied one year long in Brussels*
165	18:48:37	N03	und du kannst ruih *and you can feel free to [misspelling, cut off]*
166	18:48:43	N03	ruhig du sagen *[corrects and continues] feel free to say du*
167	18:48:44	N03	;-)
168	18:49:50	L07	ja ich hab verallgemeinert-... kannst du andere sprachen–? *yes, I generalised-... du you speak other languages–?*

In Example 6.19 the learner uses a wrong form of person reference: *sie* (they, III p. pl.) instead of *ihr* (you, II p. pl). However, a polite form *Sie* is also grammatically and semantically possible in this place. This error is corrected b the NS in turns 165-167 explicitly showing that he interprets the original person reference as a misspelled *Sie*. *Du* (the informal form of address) is the opposite of *Sie* (the formal or polite form of address). The native speakers's correction shows that his interpretation of learners intention is *Sie*. She corrects his interpretation in form of a self-correction in her turn 168: her intention was not the polite form but a generalisation. Her intended person reference should refer to people in Germany, and not only N03. To sum up, a first-person reference in the subject is required in the answer. There is no way to embed a second-person reference in the subject into answers to such questions. However, a counter question may help to correct such errors implicitly. A possible counter question after turn 164 would be *und wie lange studiert ihr normalerweise?* (En.: *and how long do you [II p. sg.] normally study?*).

Example 6.19 is interesting also from the point of view of a correction: the correct version of the intended learners expression (ihr beendet, ihr studiert) is presented neither in the correction in turns 165-167, nor in the subsequent self-correction in turn 168. N03 tested in his correction one of the possible target hypotheses, and it became clear later, that this target hypothesis was wrong. However, L07 started a new topic in turn 168 just after the self-correction rejecting N03's target hypothesis. N03 chose not to continue the discussion of the linguistic matters, however, it might be a possible continuation of the talk to teach L07, how a correct person reference in a generalisation can be formed. This example provides a strategy, how an artificial agent could act in situations where multiple target hypotheses are possible.

6.3.4. Errors for which embedded corrections are impossible

A common characteristic of all such error types is that the constituent in the learners question where the error is located cannot be repeated in the response. In contrast to the examples of possible embedded corrections discussed just above, certain changes of some linguistic features a required in the answers, for instance person and negation. Therefore, if the errors are located in these parts of the question, embedded corrections are not possible. In addition, there are elements connecting two actions in conversation, which cannot be repeated in the responses, as illustrated by Example 6.20. Such elements are normally not repeated even in the repetitions of the same question in repair sequences.

Example 6.20. Elements connecting two actions cannot be repeated.

26	18:12:09	L04	aber sowieso wie heißt du?
			but anyhow [error: lexical] what is your name?*
			but anyway, what is your name?
			ZH2: Aber egal, wie heißt du?
27	18:12:13	L04	))
28	18:12:26	L04	ORGANIZER hat mir nicht gesagt))
			ORGANIZER did not tell me [smile]
30	18:14:06	N02	oh, nicht? ich heiße N02_FirstName
			oh, no? my name is N02_FirstName
			oh, she didn't? my name is N02_FirstName

Turn 26 of Example 6.20 starts with an element which marks unit boundaries: one action is closed and another action is initiated just after *aber sowieso**. However, the adverb *sowieso* is not used as a boundary marker in German, and is probably an incorrect translation from Russian. The target hypothesis ZH2 for this utterance contains a more native-like expression *aber egal* which can be used in German to mark unit boundaries. However, the unit boundary has been already marked in the beginning of turn 26, and even with the error in it, the job of this expression is done in the beginning of turn 26. A repetition of such an expression with a correct token in the response would create a *new* unit boundary in a place where it is not relevant.

Counter questions cannot be produced for every learners question due to restrictions in pragmatics, sequential organisation or the information state (in other words, counter questions sometimes make no sense). This turns out to be a further restriction for the errors that can be implicitly corrected only in form of a counter question, as discussed in the previous part of this Section. If a counter question makes no sense for such questions, then the errors of this class cannot be addressed in an embedded correction, but only in an exposed correction.

Other types of errors in questions that cannot be corrected implicitly in the responses fond in the dataset:

1. Errors in negations if the answer is negative, for instance *magst du nicht spiele? - normalerweise schon*; *bleibt alles unveränderlich? - doch, es gibt änderungen*).
2. Errors in demonstrative expressions referring some unshared activities or other unshared entities. For example, the demonstrative *solche* in the question *macht dir spaß solche arbeit?* points to native speaker's work.
3. Word order errors in questions that cannot be replied with a counter question.
4. Errors in question words.
5. Person, case and gender errors in reflexive particles and personal pronouns that require person change in the response when a counter question does not make sense.
6. Errors in counter questions which only can be corrected in a counter question.
7. Errors which only can be corrected in a counter question if a counter question does not make sense.

In addition, turn-taking sequential organisation of the talk may influence the presence or absence of corrections. For instance:

1. If the question contains a request for action and an unmarked response is the required action, and not just information, for instance *Kannst du was was über sich erzählen?** (En.: *Can you tell something about self?*).
2. If a different action becomes more urgent than a response to question. For instance, if the learner initiates making appointments, and the native speaker has to apologise first because of being away from keyboard.

6.3.5. Unavoidable embedded corrections

Sometimes a reference needs to be repeated in order to focus only on a part of the previous turn. Sometimes it needs to be repeated because the turn is delayed and the context needs to be renewed. Embedded correction may be forced in such cases. Example 6.21 illustrates one such case. The learner produces a typo in turn 116 *Romana* and a counter question in turn 117. The native speaker replaces the trouble source by the correct spelling in turn 118 performing an embedded correction. Because typos are normally not corrected in chat, something else must have been more important than the need to omit the correction.

What alternatives did the native speaker have? The learner references to two categories of literature in her utterance, the native speaker references, however, to only one of them. Therefore, using *it* in the answer to the counter question posted in turn 117 would not reflect the reality because *it* would refer to *both* categories. What are the possibilities to refer to only one of the categories in the answer to such a counter question without repeating the term used by the learner with the error? Obviously, it could be a synonym or a referring expression like *the former* or *the latter*. Using such references in conceptually oral talk is uncommon. Another alternative would be omitting a part of information. The NS could just say what he is reading now, but not say, what he used to read when he was young. However, finding similarities helps to create rapport. All of the alternatives would be either impossible or less preferred than this embedded correction of a typo, which are normally unimportant and not corrected in chat as long as they do not change the meaning.

Example 6.21. Historische Romana

115	20:36:06	N01	Welche Bücher liest du gerne? *Which books do you like?*
116	20:39:12	L01	ich lese historische Romana und manchmal Science-fiction, aber auch kann Krimis oder Liebesromane lesen. Alles hängt von dem Inhalt ab *I read historical novels [*error: typo] and sometimes science-fiction, but I can also read crime novels or romantic novels [* error: word order]. Everything depends on the content*
117	20:39:32	L01	Und du??? *And you???*
118	20:41:59	N01	Klingt gut, historische Romane habe ich als Jugendlicher auch viele gelesen. In letzter Zeit lese ich aber eher keine fiktionalen Bücher, sondern eher wissenschaftliche Geschichtsbücher. Obwohl...auf meinem Nachttisch liegt aber im Moment ein Sherlock-Holmes-Buch :-) *Sounds good, I used to read historical novels [correction] a lot as I was teenager. In the last time, I don't read fiction books, but rather scientific historical books.* *Though... A Sherlock-Holmes-book lies on my bed table at the moment [smile]*

Alternative questions and delayed responses can make embedded corrections unavoidable. This is the case if the minimal correcting replacement has to be part of the response.

6.4. Discussion and preliminary findings

This chapter presents the results of the analysis of a collection of embedded corrections. The collection of embedded corrections which I found in the dataset contains two in principle different types of trouble sources:

1. L2 errors which are clear deviations from the linguistic standard;
2. Imprecisions of expression, which do not not deviate from a linguistic standard but rather do not completely correspond to the repair-speakers attitude towards particular references or actions.

I summarised my observations related to such imprecisions under categories of conceptual and stylistic embedded corrections. Stylistic embedded corrections can potentially address non-native like expressions, too, as demonstrated by Example 6.16. Compared to replacements in Jefferson (1983)' examples, *police→cops*, *nigger→Negro*, embedded corrections of lexical errors are closer to embedded corrections in native speakers' talk. In contrast, embedded corrections focused on form (morpho-syntactic and spelling errors) appear to be typical for native/non-native speaker interaction, while the concept of spelling errors is only relevant in medially written communication.

I found embedded corrections of linguistic errors in second pair parts of adjacency pairs and in post-expansions. Second pair parts are answers to questions and greetings in response to greetings. Such corrections address errors in the first pair parts. Post-expansions are "sequence-closing thirds" - evaluations, assessments, news markers, surprise and beginnings of topicalisation such as requests for more information. Such embedded corrections address errors in the second pair part. Summarised, I found the following types of turn pairs where embedded corrections occurred in the dataset:

1. A correction in a second pair part correcting an errorfrom the first pair part.
2. An embedded correction in a counter question environment.
    a) An error in the initial question and the correction in the answer repeated in the counter question;
    b) an error in the answer to the initial question followed by a counter question, and the correction in the answer to the counter question.
3. Embedded corrections in expressions of surprise as a response to new information provided for instance in a second pair part.
4. Embedded correction in information requests after a second pair part.
5. Embedded correction in an assessment may be placed in the second pair part and after the second pair part.
6. Embedded corrections in combination with acknowledgment tokens in a second pair part and after a second pair part.

Embedded corrections of linguistic errors were found in corrections of lexical, morpho-syntactic and orthographic errors. To understand if there are some restrictions by error type for embedded corrections, a collection of question-answer pairs was analysed. I found that there are error categories which cannot be corrected implicitly. These are errors in constituents which cannot be repeated in the relevant next action. Thus, the *correctability* is determined not by the error type, but by the error location, sequential environment, information state and availability of a target hypothesis.

In order to make visible that something is implicitly corrected, one needs to keep something unchanged, and change something else. Therefore, embedded corrections are noticeable only in a direct comparison with the trouble turn. Exposed corrections, in contrast, are recognisable as corrections due to the specific correction turn formats.

The working mechanism of embedded corrections is described on an abstract level by the concept of the minimal correcting replacement introduced in Section 6.3. The minimal correcting replacement is a repetition of the smallest constituent in the trouble turn which completely contains the error by the corrected version of this constituent in the response to the trouble turn. Every constituent that contains the minimal correcting constituent will be a correction too, but it will not be minimal. Every constituent which is smaller than the minimal correcting constituent will not correct the error.

Because embedded corrections repeat the trouble source with modifications, they potentially can correct errors in all constituents that can be repeated in the next relevant turn. Specifically for question-answer pairs I found that repeat-based answers to polar questions support embedded corrections. Partial or full repeats with modifications in answers to polar questions contain embedded corrections if they contain the minimal correcting replacement. In contrast, errors in word order in questions, for instance, can be only corrected, if the complete question can be repeated in the next relevant turn. Counter questions (variants of *and you?*) can do this job. However, counter questions are sometimes not relevant due to the information state.

Repetitions of parts of the previous turn may be required in conversation without the purpose to correct an error implicitly. Such repetitions may be necessary in order to update the subject of the talk in a delayed response caused by the virtual adjacency in chat. The need to restrict the scope of the question in an answer may also be a reason for repetition. If linguistic errors occur in references to such objects or actions in the questions, such repetitions in the answers will either repeat the original reference "as is", with all errors, or they will repeat them with modifications, and thus, with embedded corrections. I call such embedded corrections unavoidable.

As in exposed corrections of linguistic errors, the problem with multiple target hypotheses exists for embedded corrections, too. Multiple errors may exist in one constituent which can become the minimal correcting constituent for all the errors. However, as illustrated by Example 6.18, it is possible that only one of the errors can be corrected at once, and a decision has to be taken, which of the errors should be corrected and which of the errors should be ignored. Computational models of embedded corrections need to be prepared to face this issue.

The analysis of assessments and acknowledgements shows that two levels of analysis are required for embedded correction. The analysis of turn pairs allows to understand the working mechanism of the "correcting", as it was done for question-answer pairs. The analysis of the place of this turn pair within the larger turn sequence allows to discover dependencies between sequential organisation and references to objects in conversation, including repetitions.

7. To Correct or not to Correct

In Chapters 6 and 5, I discussed practices deployed by native speakers in German instant messaging chat to correct learner errors. Error correction, however, is not the only possibility to deal with an error. Possible alternatives to a correction are for instance:

1. Talk off-topic: the recipient of the trouble talk changes the subject of the talk.
2. Expressions of emotions (emoticons or lexicalised equivalents such as *hahaha*).
3. Silence.
4. Non-correcting talk on-topic:
    a) Use only discourse markers and evaluations (wow, cool, okay...);
    b) Let-it-pass strategy: continue the talk as if nothing had happened;
    c) Paraphrases;
    d) Repeat the error (behaviour rarely observed here but reported also in (Marques-Schäfer, 2013).

Silence provides opportunity to self-correct (Schegloff et al., 1977). However, due to a special responsiveness characteristics of an instant messaging chat, silence does not provide additional opportunities to self-correct in chat. Trouble-speakers can self-repair later than in the next turn in chat, and longer breaks between turns have mostly reasons other than giving the trouble-speaker a chance to self-correct (Danilava et al., 2013b).

If a linguistic error is produced by a participant of a conversation, the recipient of the trouble-talk can choose to address the error in interaction in one of the available forms, or not to address it in the interaction at all. The opposite to a correction is not an absence of a correction, but "normal talk". Many alternative ways to do "normal talk" are potentially available for conversation participants, for instance, to stay on-topic but ignore the error (let-it-pass strategy), repeat the TS "as is" not focusing on the error, paraphrase or reply by discourse markers. Other possible ways are off-topic talk, expressions of emotions such as laugh (Vöge, 2008) and addressing to troubles with the technology (Rintel, 2015). Figure 7.1 visualises different ways of handling L2 errors in talk-in-interaction.

Not all errors are being corrected in a language classroom, and not all errors can and should be corrected in a chat Conversation-for-Learning, see e.g. discussion in (Marques-Schäfer, 2013, pp. 154, 163). To correct a learner error in chat, the native speaker (with the assumption the correction is native speakers job) needs first of all to recognise an

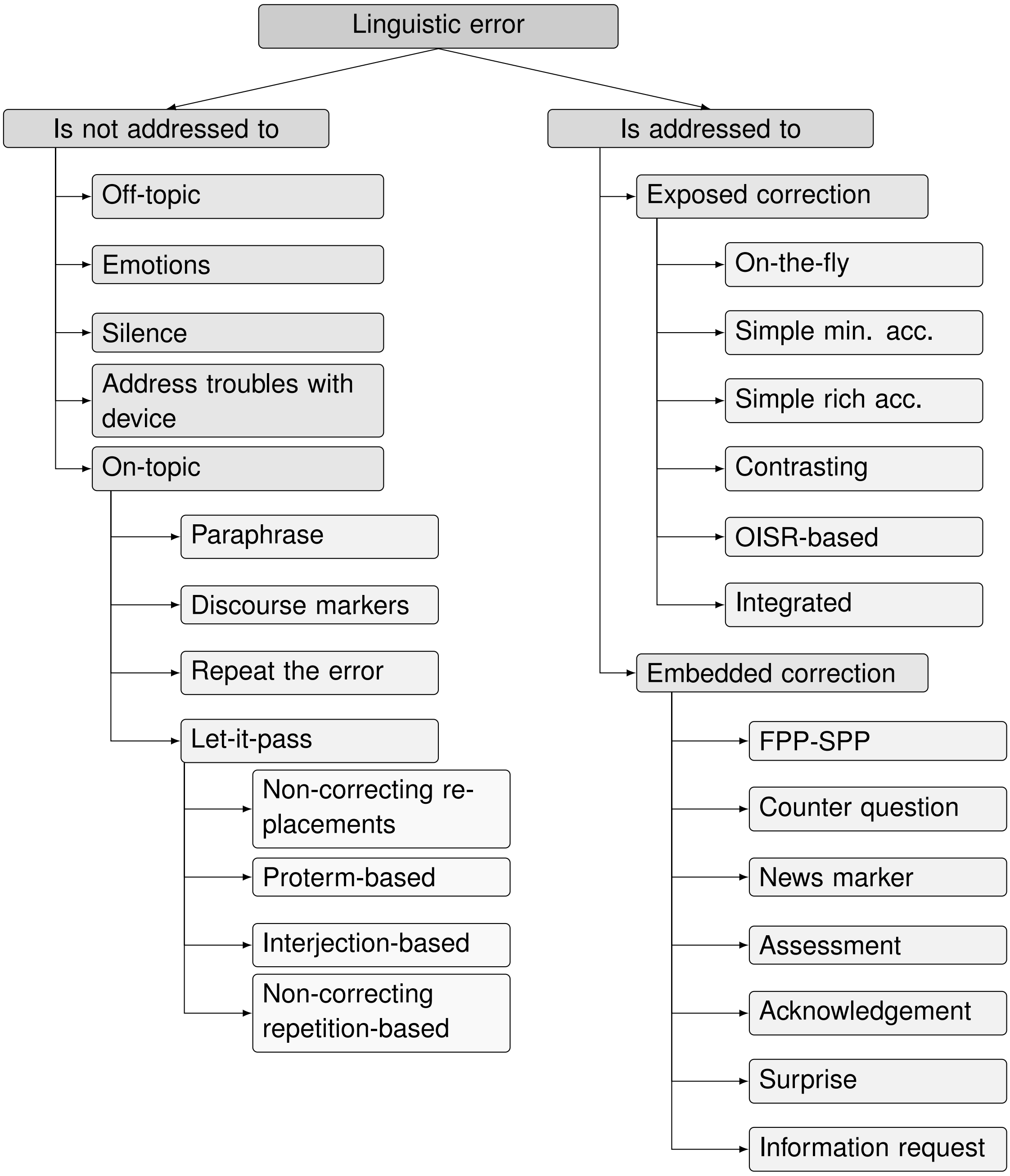


Figure 7.1.: Some options for dealing with linguistic errors in conversation

error. Thus, native speakers need to have at least a basic understanding of what is correct language and what is a deviation from the correct language. Then, even if a native speaker identified some deviations in learners utterance, there might be reasons to omit a correction and to continue the talk using one of the alternatives listed just above.

An understanding of what is correct, what is acceptable and what is an error is a preliminary step for every correction. Unfortunately, there are a few problems with a conceptual definition of a linguistic error in chat. First of all, chat language is conceptually oral language. Therefore, one should not expect to find in chat only full sentences like in an essay. Second, chat language includes symbolic means of expressivity like smileys, word stretches and uppercase typing, which are clear deviations from the standard orthography. Moreover, quick typing often leads to misspellings and typos for both parties, learner and native speaker, but it does not mean, that there is a gap in the linguistic knowledge. In addition, chat participants try to increase their typing pace and omit message-initial capital letters, as well as initial capital letters for German nouns. All these deviations do not count as errors in chat, although they do not satisfy norms of grammar and orthography.

Because the actual language expertise and attention of the native speakers may very, it is possible that they do not notice the error or do not know the corresponding grammar rule, or do not know the correct version, thus, they may have similar difficulties with the language as the learners have. The following reasons may influence the presence or absence of corrections in chat (see also (Marques-Schäfer, 2013, pp. 163-145)):

1. The recipient of the trouble-talk does not notice the error because he/she focused on the continuation of the talk or cannot recognise the error as error due to lack of language competence or is busy with other tasks not related to chat;
2. The recipient of the trouble-talk noticed the error but does not know the rule to explain the correction;
3. The recipient of the trouble-talk noticed the error and knows the rules but prefers to focus on the interpersonal trajectory and keep the interaction going (fluency context).

My dataset does not contain explicit requests for corrections in conversation, in contrast to the dataset analysed by Marques-Schäfer (2013). However, in the retrospective interviews, learners expressed their expectations to be corrected more frequently during the conversations, as opposed to the native speakers' preference to focus on the content rather than on linguistic matters. The dataset used by Marques-Schäfer (2013) for her study of learning German as a foreign language in chat contained tutored and untutored sessions. There was a predefined role of the tutor who is allowed to correct in the tutored sessions, which makes the speech exchange system more similar to a teacher-fronted classroom and represents an unequal power speech exchange system. The tutors had to reflect on their correction behaviour in a questionnaire. The decision to correct errors in chat was taken by the tutors based on the following self-defined criteria:

1. Error frequency. A frequently occurring error is corrected.

2. Relevance for comprehension (compare (Schegloff, 1987) sequential consequences for the next relevant action).
3. Learner's wish to be corrected made relevant in conversation.

If a more knowledgeable chat participant decides to correct, he or she still has to make a choice among various correction formats. For each pair of participants in my dataset, for all exposed corrections of the pair, each error addressed by an exposed correction was corrected only once, even if this error is frequently produced. In contrast, embedded corrections of the same error were produced more than once per participant pair.

Embedded correction seem to be a preferred correction format in a Conversation-for-Learning, because none of the exposed corrections repeatedly corrected the same type of error in the talk of the same learner, but there are several instances of embedded correction repeatedly correcting the same error in learners talk. As Chapter 6 shows, embedded corrections are not always possible as opposed to exposed corrections. There are dependencies between error types and turn-taking and the possibilities for embedded corrections. Incidental replacements/repetitions are not desired if the already produced embedded corrections should be taken into account for the model of future corrections. Sometimes, embedded corrections are the only way to respond if the participant does not want to repeat the error (however, it is a possible strategy, too). These dependencies will be taken into account in the decision model for corrections.

With this motivation in mind, the purpose of the present chapter is to identify a set of parameters relevant for a decision model determining whether an error in chat should be corrected or not. In addition to the error properties influencing opportunities for corrections discussed in Chapter 6 I seek to answer the following questions in the present chapter:

1. What is an error in chat?
2. Which factors may influence a presence or a non-presence of a correction?

I discuss the issues with the definition of linguistic error in chat and propose a solution in Section 7.1.1. I will show that orthography is used by chat participants as an interactional resource in Section 7.1.2. I will look at learners' self-corrections and their role for follow-up corrections in Section 7.1.3.

7.1. Learner errors in chat and opportunities to correct

Because chat interaction is conceptually oral, and because there might be a pressure to produce turns in chat as quickly as possible, the error tolerance in chat is very high. On the other hand, there is still a language standard, and there are errors which are "too heavy" even for chat, because corrections were found in the dataset. Therefore, there must be at least two levels of description for chat learner language:

1. Objective linguistic errors produced by learners compared to the language standard in terms of rules of orthography and grammar.
2. Chat conventions that make some of the identified deviations from the standard acceptable in chat as communication medium.

As argued earlier, the dataset in focus represents an equal-power speech exchange system, Conversation-for-Learning. There is no pre-assigned roles of tutors or a teachers. Both native and non-native speakers already had experiences in instant messaging. These experiences probably influenced their behaviour in the data collection phase. The applied conventions must be something on the intersection of the conversation training, chat communication expertise and stances. Different conventions are applied by different pairs of participants in my dataset. Some of the pairs produced chat messages where all rules of German standard were taken into account. Some of them ignored spelling rules for German nouns and used only small letters, started the messages with a small letter and did not follow the punctuation rules. However, none of the corrections of learner errors addressed such fussy deviations as errors. Native and non-native speakers have a similar understanding of the concept of a linguistic error in chat, no matter which standard they use for their own language production.

I discuss chat conventions in Section 7.1.1 from the following perspectives:

1. Analysis of examples of explicit negotiations of language conventions which I found in the dataset.
2. Retrospective interviews: native speakers' position towards learner errors and native speakers' own role in chat.
3. A case study of mutual dependencies within pairs of participants: how N03 and her partners make use of the chat conventions and language standard, and how N03s language is influenced by her partners' language.

From the computational perspective, for an automated error identification an error taxonomy and error models are required. There are already many projects on error-annotation for learner corpora, some of them for German. I test here the applicability of the annotation scheme created for conceptually written language (Reznicek et al., 2012). Because embedded corrections of learner errors are of specific interest, and because the majority of identified embedded corrections occur in question-answer pairs, a collection of question-answer pairs was built from the dataset. In all question-answer pairs in this collection, a learner produces a question and a native speaker answers it. In this section I will use this collection for the discussion of "real" errors in chat. This collection and the insights from the analysis will be also used later for the computational modelling of embedded corrections in Part III of this dissertation.

7.1.1. Language standard, chat conventions and L2 errors

In chat data, I see that some deviations from the standard German do not count as an error. Sometimes it is even explicitly negotiated by the participants that, for instance, writing everything small will be declared as correct. Therefore, in addition to the objective linguistic error (difference between the produced language and the language standard) must be then seen through the lens of conventions that is valid for the specific communication medium (chat in this case) and accepted by the interaction participants. This means that it cannot be completely defined in advance for chat, what will be an accepted deviation covered by conventions and what will "count" as an error that could be corrected. The following may be included:

1. Quick typing: everything that speeds up the typing pace does not count as errors: ignore capital letters in sentence and noun beginning, sentence punctuation.
2. Expressivity: word stretches, uppercase, special symbols, punctuation symbols, quotes and parentheses, as well as various combinations of all of them are used as means of expressivity.
3. Minor misspellings: typos are not important.
4. Oral style: not every utterance is a full sentence, word order is similar to oral (a question has a question mark but the word order as in an assertion).

Explicit negotiations

In Example 7.1, N03 engages in negotiation of spelling conventions in turn 10 (very beginning of the first talk). N03 argues that his typing pace is otherwise too slow. Such way of addressing spelling matters in chat is uncommon in L1 chat or in chat where linguistic identity of non-native speaker as language novice is not important. In turn 10, the native speaker shows that (1) it is clear to him that writing everything small violates the language standard and (2) to satisfy the language standard is somehow relevant for this interaction.

Example 7.1. Alles klein (everything small)

10	15:31:05	N03	ist es ok wenn ich klein schreibe? *is it ok if I write small?*
11	15:31:08	N03	geht schneller. *is quicker*
12	15:31:16	N03	ich bin sonst sehr langsam... *I'm too slow otherwise*
13	15:33:12	L07	=))) ich auch dann... nun erzähl etwas von sich.. wie steht es ? *[smile] then I will do the same... well tell something about yourself.. how is it going?*

In Example 7.2, N04 shows that he is familiar with German language standard, and that his language in chat does not satisfy the standard, but that it is acceptable in chat due to its conceptual orality (turn 201).

Example 7.2. Chatten ist wie schriftlich reden (to chat is like to talk in writing)

192	13:58:39	N04	streng genommen mache ich auch fehler: bei den meisten Verben in der "ich"-Form lasse ich oft das letzte "e" weg: "hab", "mach" usw. *stricktly speaking, I'm making mistakes, too: in most of the verbs in the I-form I often omit the final "e": "hab", "mach" and so on*
193	13:58:53	N04	oh, und "Fehler" hab ich gerade klein geschrieben... *oh, and I've just written "Fehler" small*
194	13:59:00	L09	umgangsform *way of behaving [* error: lexical choice]* colloquial form
195	13:59:04	L09	nicht? *not?* isn't it?
196	13:59:15	N04	ja, das kommt dem gesprochenen deutsch näher *yes, it is closer to oral German*
197	13:59:22	N04	süddeutsch *South-German*
198	13:59:26	N04	glaub ich *I think*
199	13:59:27	L09	angewöhnt *a habit*
200	13:59:32	L09	:)
201	13:59:54	N04	ja, chatten ist ja so ähnlich wie "schriftlich reden" *yes, to chat is similar to "talking in writing"*

I did not find too many sequences of this kind, but in all sequences that I found, I observed the following:

1. If participants engage in negotiations of spelling conventions, such negotiations are always initiated by the native speaker.
2. Production pace and conceptual orality of the interaction are the reasons for deviations addressed in chat, but not a lack of knowledge.
3. Deviations from language standard for the purpose of expressivity are not perceived as deviations by chat participants.

Insights from retrospective interviews

It is not directly observable from the data why native speakers choose to stay on-topic and to contribute to the interpersonal trajectory not addressing the errors, and why they do not decide for one of the alternatives (off-topic, silence...). However, it can be assumed that the participants did not want to be perceived as impolite, they wanted the talk to go on and chose to produce preferred responses (a preferred response to a question is an answer, and not silence or laugh or a news marker). N01 writes in the retrospective IM-based interview about his motivation not to correct (original orthography preserved):

> eine korrektur habe ich unterlassen, weil ich die fehler eher für unbedeutend gehalten habe (manchmal wurde das falsche tempus oder das falsche geschlecht eines wortes gewählt. es kam nie vor, dass ein satz vor fehlern nur so gestrotzt hätte) und - was noch wichtiger ist - ich wollte den inhaltlichen gesprächsfluss nicht unterbrechen. verstanden haben wir uns ja. das war mir wichtiger, als kleinliche fehler auszubessern.
> *I omitted a correction, because I considered the errors to be unimportant (sometimes a wrong verb tense of a wrong gender of a word was selected. It was never the case that a sentence teemed with errors) and - what is even more important - I did not want to interrupt the substantial flow of the talk. We understood each other. It was more important for me than correcting of petty errors.*

In contrast to N03, N01 saw himself in chat with learners as a role model with respect to orthography, like for instance capital letters at the beginning of the nouns and utterances. He explains in the retrospective interview that he tried to write in according to German standard

> weil ich gegenüber nicht-deutschen-muttersprachlern versuche, die deutsche sprache so gut wie möglich in wort und schrift zu verwenden.
> *because I am trying to use written and oral German language as good as I can in communication with Non-German native speakers.*

As these two quotations show, N01 uses lowercase-only spelling during the interviews as opposed to the standard-compliant spelling that he chose to use during the data collection. A sample of his spelling in the dataset can be found for instance in Example 6.21. The orthography in chat which N01 uses with different partners is *recipient-designed*.

7.1.2. Orthography and social closeness

The presence of a high number of deviations from the language standard in text chat has been explained by a pressure to type quickly and demand for a high production pace in CALL studies (Loewen and Reissner, 2009). However, language learners report that they had (or took) their time to use additional resources (such as dictionaries) for dealing with trouble in comprehension and production. Hence, the production overhead necessary for a standard-conform language in chat might have a particular interactional import and may have impact on participant's understanding of their social roles and used for the regulation of social closeness.

Example 7.3 presents the very beginning of the talk between L03 and N02. Because the participants have never met before, L03 does not know, who is on the other side of the connection. She comes too late to her first appointment and formulated her first message (turn 1) to her chat partner in a very polite way using a polite German form of address *Sie* (III p, pl., no English equivalent). In addition, she produces an email-style turn - conceptually closer to written than oral language - according to German spelling standard and closes if with a "best regards + signature" non-typical for instant messaging.

Example 7.3. Mutual dependencies between orthography and social closeness.

1	19:57:31	L03	Hallo! Entschuldigung, Ich weiß nicht, wie heißen Sie. Ich bitte um Verzeihung, ich habe total über heutige Unterhaltung vergessen. Ich schäme mich, wirklich, aber ich war beschäftigt, und musste dringend einige Probleme lösen, deshalb habe ich total über den Chat vergessen- ich bitte noch ein Mal um Entschuldigung, und verspreche, dass es nie wiederholen wird. Ich hoffe, dass unser Chat wird uns Spaß machen. mit freundlichem Gruß, L03_FirstName L03_LastName! *Hello! I am sorry, I don't know your [III p. pl.] name. Please forgive me, I totally forgot about [* error: wrong preposition] today's conversation. I feel ashamed, really, but I was busy, and had to solve several problems urgently, this is why I totally forgot about [* error: wrong preposition] the chat - please forgive me again, I promise that it will never happen again. I hope that our chat will be pleasant. best regards, L03_FirstName L03_LastName!*
2	19:59:57	N02	Hallo L03_FirstName, das ist überhaupt kein Problem! Ich hoffe, alle Probleme sind gelöst und wir können ein bisschen chatten. *Hello L03_FirstName, it is absolutely no problem! I hope, all the problems are solved and we can chat a little bit.*
3	20:01:58	L03	Ja, natürlich! wie heißt du? *Yes, of course! what is your [II p. sg.] name?*
4	20:02:21	N02	oh Entschuldigung, ich heiße N02_FirstName, bin 27 Jahre alt und wohne in München. *oh, I'm sorry, my name is N02_FirstName, I am 27 and live in Munich.*
5	20:03:45	L03	sehr angenehm! und ich bin 21 und wohne in Vitebsk, Belarus! *very pleasant! and I am 21 and live in Vitebsk, Belarus!* nice to meet you! and I am 21 and live in Vitebsk, Belarus
6	20:04:37	N02	oh, ich bin schon alt ;) *oh, I am already old [smile]*
7	20:04:54	N02	warst du schon mal in Deutschland? Ich war noch nie in Belarus *have you already been to Germany! I have never been to Belarus*
8	20:05:11	L03	ja, aber ich bin schon verheiretet))) *yes, but I am already married [smile]*
9	20:05:22	N02	oh echt?? wow! seit wann denn, wenn ich fragen darf? *oh really?? wow! may I ask you, how long?*

L03 produces multiple morpho-syntactic and semantic errors, however, her phrases start with a capital letter (except of the closing expression), and she is doing her best in positioning herself as a competent German speaker. N02 answers with a "no problem", and her message satisfies the German language standard, too. L03 switches from *Sie* to *du* (you, II p. sg.) in turn 3. In addition, she changes the spelling in the second phrase starting with a small letter instead of a capital. N02 responds with changed applied standard in turn 4 writing only nouns with an initial capital letter.

The participants continue with the rule "write only nouns with a capital letter". Shorter time intervals between turns 5-9 in Example 7.3 show how higher engagement leads to higher talk pace and therefore higher production pace. Deviations from language standard are the price for the typing pace, but in addition, they express a higher grade of engagement and social closeness.

There are mutual dependencies between participants. A closer look at N02 and her partners L03, L04 and L05 helps to understand how participants deal with spelling and punctuation conventions, and how they influence each other. N02 behaves differently with her different partners:

L03 Both participants start with the standard-compliant spelling and shift then to a version where they move between standard-compliant spelling and "write-only-nouns-with-a-capital". L03 starts with "Sie" but switch to "du" in turn 3.

L04 starts with a "relaxed" version of spelling: only nouns are written with a capital, a very oral style. N02 starts with a norm-compliant version but adapts to non-native speaker's spelling version after 10 turns. Later on, both participants even use lowercase for all words. L04 starts with "du". Overall chat of this pair can be characterised as very oral: short phrases, quick, many short turns.

L05 starts with a norm-compliant orthography and "Sie". L05 makes lexical errors in her first turn. N03 replies with "Sie" but she decides to write the first word in each sentence small. Later on, L03 changes between a norm-compliant spelling and the relaxed "first-letter-small" version. L05 adopts this way of spelling from time to time. In the second chat, L03 start with "du" (first turn in this meeting) using proper spelling, but switches later to the relaxed "first-letter-small" version. It remains an open question if N03 noticed that L05 is not that much an independent language user (compared to the others) and shows her, how to do "chat-in-German".

The other native speakers in the dataset prefer to keep the same orthography style with all their partners: N01 presents himself as a role model, N03 prefers to optimise the spelling to increase the typing pace and types everything with lowercase, and N04 normally types all nouns with an initial capital, but starts all new sentences with a small initial letter.

7.1.3. The role of learners' self-corrections

Even if an error is potentially correctable in chat (is a "real" error and corrections are welcome etc), learners may notice the deviations in productions (no matter if the deviations are just mistypings or "real" errors) and correct themselves. Normally, no second correction is produced after a self-correction. Example 7.4 illustrates a typical case of a non-present other-correction after a self-correction, even if there are other, non-self-corrected errors which might be corrected by the native speaker. Learner L03 produces several errors in turn 79 and self-corrects one of them in turn 80. The native speaker N02 does not correct remaining uncorrected errors and provides an interjection-based response.

Example 7.4. No correction after a self-correction (prev. Ex. 6.18).

79	13:22:49	L03	)))) ja, wahrscheinlich! Sind die Grenze des Schuljahres von Urlaubssaison in Beiern abhängig? *[smile] yes, maybe! Are the border [*error: wrong number] of the school year [*error: lexical] depending from the vacation time in Bavaria [*error: spelling]?* Does school vacation time depend on the vacation time in Bavaria?
80	13:23:16	L03	* Bayern ** Bayern [self-correction]* Bavaria
81	13:23:41	N02	ja genau! ist das bei euch auch so? *Yes, exactly! The school vacation time is depending on the vacation time in Bavaria. Is it the same in your place?* Yes, exactly! The school vacation time depends on the vacation time in Bavaria. Is it the same in your place?

Example 7.5. Delayed embedded correction after multiple self-correction attempts. Replaced units are underlined.

310	17:48:08	L08	ok)siehst du "Eurovision"? *see [present tense] you "Eurovision"?* do you see Eurovision
311	17:48:36	L08	richtiger "sah" *more correct "saw" [*error: wrong tense and person - self-corrected tense]* more correct "saw"
312	17:49:05	L08	o nein)"sahst" *o no [smile] "saw" [*error: wrong tense - self-corrected person]* oh no, "saw"
...	...	...	`[Simplified: three turns belonging to a different sequence are hidden]`
316	17:50:39	N04	ja, den Eurovision hab ich gesehen, die russischen Omas mit den Plätzchen waren toll :-) *yes, I have seen [correction: change to unmarked verb form] the Eurovision, the Russian grannies with cookies were great [smile]* yes, I watched the Eurovision, the Russian grannies with cookies were great

Example 7.5, as opposed to all other cases of self-correction, illustrates how an embedded correction may appear after a self-correction. L08 produces a grammatically correct utterance in turn 310. However, she finds the verb tense that she uses in this utterance wrong and makes a first attempt to correct in turn 311: she replaces present by a past form, this time with a grammatical error. In addition, German imperfect (past simple) is a marked verb tense (not neutral).

Learner's second self-correction attempt in turn 321 is successful in producing a grammatically correct verb form, however, an unmarked verb tense is more appropriate. Finally, in turn 316, two replacements are performed by N04: *Eurovision* $\rightarrow$ *den Eurovision* (determiner added) and *sahst* $\rightarrow$ *habe gesehen* (verb tense changed from marked to unmarked).

Because three turns belonging to a different dialogue thread (a different topic) are placed between the question in turn 310 and the answer in turn 316, a repetition-based answer to the polar question is needed. The repetition has first of all the purpose to update the subject of the talk, and the correction is a secondary business. However, because of the made replacements, an embedded correction is produced in an answer to a question despite the presence of the self-correction. Subject update may have a preference over a preference not to other-correct after a self-correction.

7.2. Discussion and preliminary findings

I started this chapter with a bird's eye view on various possibilities to do non-correcting talk and the place of corrections on their side. Some reasons for non-present corrections were already discussed in Chapters 5 and 6. I expected to go deeper in understanding of all possible reasons for a preference to correct at a specific point in talk. I found that the following factors play a role in the decision to correct or not to correct:

1. Speech exchange system:
   a) Conversation-for-Learning combines properties of classroom talk and free conversation, however, there is a variance within this space.
   b) Focus on fluency has normally a higher preference than focus on form and accuracy.
2. Communication medium:
   a) Both, deviations from language standard and meeting all language standards can be used as interactional resource.
   b) Communication pace may cause more deviations during productions, conscious and unconscious.
3. Participant's linguistic expertise:
   a) Is the recipient of the trouble talk able to recognise the errors?

   b) Does the recipient of the trouble talk see the errors as severe enough to be corrected?
   c) Is the recipient of the trouble talk able to produce a correction? Is there a clear target hypothesis?
4. Participants' negotiations and agreements:
   a) Learner's may explicitly ask their more knowledgeable conversation partners to correct.
   b) Participants may agree on not-counting some deviations as errors.
5. Participants' engagement:
   a) Other, more interesting things may prevent the recipient of the trouble talk from noticing the errors.
   b) Moving the conversation forward has a higher priority than focusing on errors.
6. Sequential organisation of chat:
   a) Delayed responses may lead to corrections where they are normally dispreferred, e.g. after self-corrections or after typos.
   b) Relevance for comprehension: a correction (e.g. a clarification request) may be the only possible next relevant action.
7. Error properties:
   a) Is there a clear deviation from the standard language?
   b) Is this deviation covered by chat conventions?
   c) How frequent does this error occur?
   d) Which types of correction formats can be produced to correct this specific type of error?
8. Correction history:
   a) Has this error already been explicitly corrected?
   b) Has this error already been implicitly corrected?
   c) Has this error already been self-corrected?

As demonstrated in Section 7.1.1, there is no such thing like *the* chat conventions. It is rather a range, or a step-wise relaxation of the standard. This shows, that many acceptable spelling versions exist at the same time. In addition it shows, that it is not important, because there is no consistency in use of the convention rules. The conventions can be negotiated or assumed, they can change, be accepted or rejected by the participants. However, mostly deviations in orthography related to capital letters (including upper and camel case) and incidental misspellings due to typing pace, as well as deviations in punctuation are covered by the category "chat conventions". Other oral features such as omission of the grammatical subject in turns are also allowed. In general, it is expected by chat participants, that the chat language is morpho-syntactically and semantically and pragmatically "correct".

I rarely observed that participants write everything with lowercase, however, it is a common practice in L1 German chat or in native/non-native speaker German chat where language learning is not relevant for the interaction. The most common range of acceptable orthography in the dataset is between the standard and the version where at least nouns are written with an initial capital letter.

As demonstrated in Section 7.1.2, participants of an instant messaging chat use deviations from language standard as a interactional resource to regulate social closeness and to present themselves as members of specific categories, such as a native speaker who positions herself as a role model, as well as a competent non-native speaker who is a competent instant messaging user.

All factors influencing the presence or non-presence of corrections in chat ideally need to be taken into account in decision models for corrections. I will come back to this problem from the point of view of computational decision models for corrections in Chapter 10.

Part III.

Computational Models of Explanations and Corrections

Introduction

Various forms of making linguistic identities relevant in conversation were described in PartII of this research. Such forms include excuse-based face-work sequences, assessment of linguistic proficiency, collaborative learning in role-play sub-dialogues and repair sequences with linguistic trouble source. Part II of this work describes how such repair sequences may be structured locally. In addition, I analysed under which circumstances such repair sequences may appear, and what may be a trigger for each of them.

In this part of the dissertation, I will make a step from the empirical findings towards computational models of specific sequences where linguistic identities of IM chat participants can be made relevant in talk. The purpose of this modelling is to make some first steps towards providing a conversational agent with a mechanism that allows the agent to behave like a language expert, as inspired by the dataset. I continue here the work started in Part II towards computational modelling of the following two repair types:

1. Other-initiated self-repair when the agent is the trouble-speaker. These are situations when the learner who communicates with an artificial agent does not understand agent's talk and initiates a repair. The agent is expected to produce an appropriate explanation.
2. Other-corrections of linguistic errors produced by the learner. I distinguish here between exposed (explicit) and embedded (implicit) corrections, as motivated in Part II of this work.

As noted by (Schegloff, 1993) and discussed in Chapter 6, everything may potentially become a problem in conversation. This is why other-initiations of repair can occur everywhere, after each turn, they were even observed after silence. For this reason, each learner's turn can *potentially* contain a repair initiation. Other-corrections, in contrast, should not be produced every time they could be potentially produced. Many reasons for that were discussed in Chapter 7. Hence, a decision model for corrections is required. Therefore, I decompose the modelling into models addressing the practices of dealing with trouble, which I call local models, and high-level decision models, which help to determine, if one of the local models should be activated.

This part of the work is structured as follows. Chapter 8 describes a local computational model of other-initiated self-repair when the machine is the trouble-speaker (OISR_M). Chapter 9 describes local computational models of exposed corrections and embedded corrections in answers to polar questions. Chapter 10 describes a rule-based decision model for corrections of L2 errors in conversation.

8. A Local Model of Other-initiated Self-repair

Chapter 4 described the types of trouble that the learners may have in understanding native speakers' utterances. I analysed the ways of signalling trouble used by the learners (called repair other-initiations), and the ways of dealing with such requests chosen by the native speakers (called self-repairs or repair carry-outs). This chapter presents a computational model of other-initiated self-repair when the artificial agent is a trouble speaker. The model is based on the findings of Chapter 4. The artificial agent will need to do the same as the native speakers, namely, recognise repair other-initiations and carry out self-repair after that in order to resolve the trouble in learner's comprehension.

Although the main focus is on German, possibilities to use the same model for other languages are of interest. It might be possible because repair other-initiations are structured very similarly in all languages for which other-initiated repair was studied till present (Dingemanse et al., 2014). Therefore, one of the challenges in the modelling phase is to decouple language-specific devices (e.g. specific tokens used for repair initiation in German) from sequential models and turn formats that could be used across languages (e.g. repeats and specific symbolic resources available in chat).

Building on the prototypical structures in sequences of other-initiated self-repair (OISR) described in Part II, I introduce a sequential computational model of other-initiated self-repair when the machine is the trouble-speaker (OISR$_M$) which I describe and discuss in the remainder of this chapter. Section 8.1 describes a local model of OISR$_M$for German text-based chat dialogues with language learners. The model is divided into two steps:

1. Recognition of repair initiation and extraction of the trouble source described in Section 8.1.1.
2. Generation of a repair proper (repair carry-out) explained in Section 8.1.2.

In Section 8.2, I describe an implementation case-study where the proposed model of OISR$_M$ was realised as a repair manager in a simple conversation program - a chatbot. The purpose of this section is to understand, what is required from the computational perspective to simulate sequences of OISR$_M$. Specifically, it is important to know for an implementation, what kind of meta-linguistic information might be required, what kind of NLP tools might be helpful, and where the limitations are. I discuss preliminary results and discuss the issues in Section 8.3.

8.1. Two steps for dealing with trouble

Repair initiations normally contain all the necessary information for human participants to understand that there is a problem with comprehension, to locate the trouble source (TS) and to identify the type of trouble. However, sometimes additional work is required for the participants to clarify what exactly is unclear. This process can become very complex and the trouble source may or may not be identified at the end of the clarification sequence. However, such nested repair sequences where repair initiation becomes itself a new trouble source still have the same prototypical structure. There is something in a preceding turn which is marked as a trouble source in the ongoing turn and is resolved in the subsequent turn. The basic structure including a trouble source, a repair initiation and a repair carry-out is kept in nested repair sequences as well. Therefore, for simplicity, I focus here only on repair sequences where no additional work for the identification of the trouble source is needed.

Because every repair other-initiation also contains information about the kind of trouble which the trouble-talk recipient has, it is possible for the recipient of the repair other-initiation (who is also the trouble-speaker) to recognise, *that* there is a problem with understanding and *what kind* of problem occurred. A repair proper is delivered after that. Because nothing in conversation is dedicated to be a trouble source by itself, but everything in conversation may appear to be a source of trouble for the recipient of a piece of talk, repair initiations may occur after each turn. This means for chat-based communication with an artificial agent that every message of the user may be a repair initiation. Therefore, each user's turn needs to be checked, if it is a repair initiation.

With this motivation, I divide the handling of OISR$_M$ in two steps:

1. Recognition of repair other-initiation and extraction of the trouble source;
2. Self-repair carry-out.

I explain the two steps in the remainder of this section.

In Chapter 2 I provided an overview of the most important results in research on repair in naturally occurring and technology-mediated interaction. In contrast to oral conversations, medially written chat conversations do not contain problems with hearing (and need for a simple repeat), participants can re-read the history. Therefore, the range of potential trouble sources is modified in chat as opposed to face-to-face or oral technology-mediated conversation. In addition, repair sequences with linguistic trouble source (when the trouble in understanding is caused by insufficient L2 knowledge) need to be distinguished from all other troubles in understanding (caused by other factors in talk). The model described in this chapter concerns specifically OISR$_M$-sequences with linguistic trouble source, thus, when troubles in learner's comprehension occur due to not-yet-fully-proficient level of knowledge of the foreign language.

8.1.1. Recognition of repair initiation

Sequential organisation of repair in chat is closely related to the chat-specific turn-taking system and adjacency pair structure in chat. The participants of a chat interaction have the possibility to re-read what has been previously written and to connect messages which build interactional adjacency pair but are not adjacent on the screen - so called virtual adjacency or split-adjacency pairs (Tudini, 2010). As opposed to oral data, next-turn repair other-initiations in chat may be immediate or delayed. Immediate repair initiations take place in the next, adjacent turn in the timeline. Delayed repair initiations occur with one or more turns between the trouble-turn and the repair initiation (virtual adjacency). This has implications for how the trouble source can be referenced. Therefore I distinguish between immediate and delayed repair initiations.

The most frequently used device for marking the trouble source and signalling a problem with understanding in the dataset was reusing the problematic token or a phrase and appending one or more question marks to it. In such repair initiations, the question marks are the way to signal that something is wrong with the copied part of the trouble-turn. The copied part is used to point to the trouble source. Each class of repair initiations discussed in Chapter 2 also contains a specific form of referencing the trouble source. To extract a trouble source from the repair initiation and the corresponding trouble-turn, I look at possibilities to reference the trouble source that I found in the dataset. The following types of referencing appear in the OISR$_M$-sequences with linguistic trouble source:

1. Repeat-based initiations: *reuse* (a 1:1-copy of the trouble source), *recycle* the trouble source (rewriting it in a slightly different way),
2. Demonstratives-based initiations: using demonstrative adverbs, determiners and pronouns, such as *dieser Ausdruck* or simply *das*.
3. Open-class initiations: referencing by placing a statement of non-understanding in a turn adjacent to the trouble-source turn. The adjacent position of the repair initiation references the whole preceding turn as a trouble turn. Therefore referred to this type of referencing as *reference by position*.

To extract the trouble source from a repair initiation, it is good to know, how big the trouble source is, and what is the scope of the search. Each class of repair initiations listed just above is dedicated to reference trouble of a particular size: either is is the whole preceding message (open-class and demonstratives-based repair initiations) or it is only a part of it (repeat-based and recycle-based initiations). Then, there are three cases of trouble sources: single word (part of a longer message or a one-word message), part of a message (PoM) of two or more words and a whole message consisting of two or more words.

Various resources are used by learners in chat for signalling trouble. These include symbolic means such as question marks, dashes, quotes, uppercase writing, but also lexicalised means such as explicit statements of non-understanding and requests for help. Repair initiations may also contain candidate understandings. Thus, repair initiations can be formatted to simply mark something as unclear or to mark something as unclear and to provide a candidate understanding by the trouble-talk recipient. Summarised, signalling trouble requires involvement symbolic and/or lexicalised means of signalling and a specific format designed either to mark something as unclear or to compare the trouble source with the own version of understanding. I call the combination of these resources *signalling format*.

Table 8.1 summarises the different possibilities to reference the trouble source and to signalise troubles with comprehension that I found in the data set. Demonstrative determiners (DD) or demonstrative pronouns (DP) may be used in combination with the adjacent position to point to the trouble source.

Time	Ref.	Signalling	Used for
Immediate	Reuse, recycle	One or more "?" also combined with one or more"-", explicit statements of non-understanding, lack of knowledge, requests for help, candidate understandings	Single word, POM
	Position	multiple "?", non-lexical tokens of non-understanding, explicit statement of non-understanding, requests for help	whole message
	Position, DD/DP	DD/DP in combination with explicit statements of non-understanding, candidate understandings	whole message
Delayed	Reuse, recycle	one or more "?" also combined with "-", explicit statement of non-understanding, lack of knowledge, ask for help, candidate understandings	whole messages, single words, POM

Table 8.1.: Referencing trouble source and signalling trouble.

The architecture of the repair initiation (RI) for this type of OISR can be formalised as follows. Depending on the time, different formats for the repair initiation may be used:

$$RI = TIME \times RIFormat \tag{8.1}$$

Time may be immediate or delayed: $TIME = \{immediate,\ delayed\}$. A repair initiation format is a combination of a reference to the trouble source and a selected signalling format:

$$RIFormat = REF \times SignalFormat \tag{8.2}$$

As mentioned earlier, repeat-based references such as reusing $reuse(x)$ and recycling $recycle(x)$ the trouble source. They can also be based on demonstrative determiners or demonstrative pronouns $DD,\ DP$. In addition, adjacent position-based AP references may be used to refer to the trouble source in the repair initiation. Signalling format may mark something in the trouble-turn as unclear $unclear(x)$ or present a candidate understanding in comparison to the trouble source $equals(x, y)$. The trouble source x and the candidate understanding y may be a single word, an idiomatic expression, part of a message or a complete turn (utterance).

$$REF = \{reuse(x),\ recycle(x),\ AP,\ DD,\ DP\} \tag{8.3}$$

$$SignalFormat = \{unclear(x),\ equals(x, y)\} \tag{8.4}$$

$$x, y \in \{word,\ idiom,\ PoM,\ utterance\} \tag{8.5}$$

For each repair initiation, if only adjacent position or only a demonstratives-based reference was used in the repair initiation to point to the trouble source, then the scope for the trouble source extraction is limited to the preceding turn.

```
if REF =∈ {AP, DD, DP} then
| TIME = {immediate} ;
else
| REF ∈ {reuse(x), recycle(x)};
| TIME = {immediate, delayed};
end
```

This repair recognition procedure is also expected to differentiate between ordinary questions related to the subject of the ongoing talk and repair initiations. It works because ordinary questions are not formatted as $unclear(x)$ or $equals(x, y)$.

If a complete turn is recognised as a trouble source and this turn is a longer message, further filters may be applied to identify more precisely, which of the parts of the longer message may cause a problem with comprehension. Such more precise trouble source identifications may be influences by the learner model (which words is the learner supposed to understand), but also by systems capabilities to generate a repair proper (are paraphrases available or only synonyms for single words). I will address this problem in Section 8.2 providing examples of possible further filters.

After a successful identification of a repair other-initiation and a trouble source extraction, a generation of a repair proper is required. I describe the model for repair carry-out in the next section.

8.1.2. Self-repair carry-out

Repair carry out in conversations with learners differs depending on the trouble-speaker and repair-initiator. As illustrated on multiple examples in Part II of this dissertation, different types of repair proper are designed to solve learner's problems with comprehension of the native speaker's talk, and the other way round. While explanations of the meaning of the trouble source are delivered to the learner (the trouble source is not replaced), *different* descriptions of the objects or actions are delivered to the native speakers (the trouble source is replaced).

Similarly to repair initiations and its relative position to the trouble source, the repair carry out part of a repair sequence can occur immediately after the repair initiation or a few turns later. Additionally, it can contain a lexical reference to the trouble source, such as repeat-based and demonstratives-based references, or point to it just by occurring in the adjacent position just after the subsequent repair initiation. The format of the repair proper depends on the format of the repair initiation and on the type of the trouble source.

<table>
<tr><th>TS Type</th><th>RI Signaling</th><th>RCO Time</th><th>RCO Reference</th><th>RCO Type</th></tr>
<tr><td rowspan="2">Abbreviation</td><td rowspan="2">$unclear(x)$</td><td>immediate</td><td rowspan="2">reuse</td><td rowspan="2">spell out, synonym, example</td></tr>
<tr><td>delayed</td></tr>
<tr><td rowspan="4">Noun, verb, adverb, adjectiv, idiom</td><td rowspan="2">$equals(x, y)$</td><td>immediate</td><td>position</td><td rowspan="2">confirm, disconfirm</td></tr>
<tr><td>delayed</td><td>reuse</td></tr>
<tr><td rowspan="2">$unclear(x)$</td><td>immediate</td><td>position</td><td rowspan="2">explanations by definition, synonym, example etc.</td></tr>
<tr><td>delayed</td><td>reuse</td></tr>
<tr><td rowspan="2">Long turn, long part of a turn</td><td rowspan="2">$unclear(x)$</td><td rowspan="2">immediate</td><td>split-reuse</td><td>$explain(x)$ for all x</td></tr>
<tr><td>position +/- DP</td><td>paraphrase</td></tr>
</table>

Table 8.2.: Dependencies between repair other-initiations and self-repair carry-out.

Table 8.2 contains dependencies between repair other-initiations and the subsequent self-repairs that I identified in the data set. I use the abbreviation RI to refer to repair initiation, and the abbreviation RCO to refer to repair carry-out in the table and in the remainder of this chapter. *Split-reuse* is a type of a reference to the trouble source which did not appear in repair other-initiations but was found in corresponding self-repair carry-outs. This way of referencing corresponds to self-repairs where native speakers only explained a few words from a longer turn or longer part of a turn marked as a trouble source. The trouble source was split in tokens, and only tokens which were supposed to cause the trouble were explained. This is shown in the table in form of the $explain(x)$ function where x is the extracted trouble source. The corresponding trouble signalling format is denoted by $unclear(x)$ and $equals(x, y)$, as in the preceding section.

All repair other-initiations in repair sequences with linguistic trouble source that I found in the dataset can be interpreted as either a content question "What does x mean?" or a polar question "Does x mean y?". Table 8.2 reflects this observation. A confirmation or a disconfirmation is an appropriate type of self-repair carry-out after a repair other-initiation presenting candidate understandings $equals(x, y)$. All other self-repair carry-outs are expected to provide an explanation of the unit that is marked as problematic, a simple yes or now would be not sufficient.

The following decision routine for self-repair carry-out results from the dependencies summarised in Table 8.2. Because different options are available for referencing trouble source in immediate and delayed self-repairs, time needs to be taken into account in the abstract description of the self-repair carry-outs:

$$RCO = TIME \times RCOFormat \tag{8.6}$$

As for the repair other-initiations, time may be immediate or delayed:

$$TIME = \{immediate,\ delayed\}. \tag{8.7}$$

A self-repair carry-out is a product of a reference to the trouble source and the function, which it is expected to perform: confirming/disconfirming answer or an explanation.

$$RCOFormat = REF \times RCOFunction \tag{8.8}$$

$$REF = \{reuse(x),\ recycle(x),\ AP,\ DD, DP, \tag{8.9}$$

$$splitReuse(x),\ splitRecycle(x)\} \tag{8.10}$$

$$RCOFunction = \{explain(x),\ confirm(equals(x, y))\} \tag{8.11}$$

Depending on the timing of the self-repair production, different types of referencing may be appropriate: delayed self-repairs need to update the focus of the talk, and therefore, a repeat-based reference makes more sense than other types of referencing.

if $TIME = \{immediate\}$ **then**
  $REF = \in \{reuse(x),\ recycle(x),\ AP,\ DD,\ DP\}$;
else
  $REF \in \{reuse(x),\ recycle(x)\}$;
end

In practice, the function $explain(x)$ needs to be implemented differently for different types of trouble source (see column RCO Type in Table 8.2). For instance, abbreviations from chat jargon are typically explained by spelling out the intended reading of the abbreviation. For all other abbreviations a full version of the word(s) is presented, sometimes in combination with examples, synonyms and comments on the topic. Practices used to explain whole messages consisting of two or more words include splitting the message into single words and explanation of a couple words of the message and paraphrasing. The quality of the response is highly dependent on the linguistic resources available for the chatbot. I discuss various practical issues in the next section.

8.2. An implementation case study: a chatbot extended with a repair manager

The purpose of this section is to prove a practical applicability of the abstract model described in the preceding section. Because language understanding and generation capabilities of each artificial system supposed to have a conversation with the user determines the possibilities for implementation of the OISR_M model, I decided to take the simplest form of such a system, namely an AIML-based chatbot (Bush, 2006). In this kind of systems, natural language understanding and generation are covered by the Artificial Intelligence Markup Language (AIML) (Droßmann, 2005) and is provided by pattern-template pairs as shown below. If the chatbot finds an input that matches to WIE GEHTS, the utterance stored in the template tag will be delivered to the user as a response.

```
<category>
   <pattern>WIE GEHTS</pattern>
   <template>Gut, und selbst? Alles paletti?</template>
</category>
```

Example 8.1 illustrates how a chatbot can benefit from patterns extracted from the dataset and modelled in Section 8.1 to come closer to the behaviour of a language expert.

Example 8.1. A sub-dialogue with the chatbot: other-initiated self-repair where the chatbot is the trouble-speaker.

1	`User`	wie gehts? *how are you?*
2	`Bot`	Gut, und selbst? Alles paletti? *I'm fine, and you? Everything okay?*
3	`User`	paletti?
4	`Bot`	umgangssprachlich alles gut, alles in Ordnung, alles okay. *colloquial everything good, everything fine, everything okay.*

The bot uses a colloquial expression in turn 2 which is not clear for the user. The user initiates the repair in turn 3 using a specific repair initiation format: repetition of a part of the trouble source combined with a question mark. The bot recognises turn 3 as a repair initiation and extracts the trouble source: the repeated word *paletti* and the corresponding idiomatic expression *alles paletti*. Bot's response in turn 4 is a repair carry-out generated from a linguistic database, which I call ExplanationDB.

The baseline chatbot (Bush, 2006) was extended with a repair manager. A free German AIML set (Droßmann, 2005), which builds the "brain" of the chatbot, was extended by specific AIML categories responsible for matching repair other-initiations to corresponding self-repair formats and generation of self-repair based on prepared templates.

The work of the repair manager is organised in two steps determined by the model:

1. Recognition of repair other-initiation and trouble source extraction,
2. Self-repair carry-out.

Based on the rules for repair initiation, as described in the previous section, patterns for the recognition of the repair initiations have been developed. From these patterns, specific AIML categories for the two main classes of signalling have been created:

1. $unclear(x)$. Every user's input that requires an explanation of a single entity (word, idiom) is redirected to the category that implements this function. A new AIML tag has been introduced for the purpose of this work : `<explain>`. An additional processor named `explanation processor` has been implemented to generate a response.
2. $equals(x, y)$. Every user's input that corresponds to an inquiry "does x mean y?" is redirected to the AIML category implementing meaning checks. An additional tag `<meaningcheck>` and a `meaning check processor` have been added to carry out the repair of this type.

I describe language-specific resources and knowledge bases used by the repair manager below.

8.2.1. Recognition of $unclear(x)$ and repair generation

The repair manager searches in every user's input for symptoms of a non-understanding that could be used for a RI (multiple question marks, lexical and non-lexical expressions of non-understanding, demonstrative determiners and pronouns). If an immediate repeat-based repair other-initiation could be identified, the trouble source is extracted from the repair initiation immediately. However, users may repeat only a part of an idiomatic expression in the repair initiation, and the complete expression needs to be extracted for a proper explanation, as it was illustrated in Example 8.1. Therefore, other turns are taken into account for the trouble source extraction as follows:

1. If the type of the signal found is typical for an immediate repair other-initiation, and the reference is not repeat-based, the program tries to locate the trouble source in the last utterance of the chatbot.
2. If the type of the signal found is typical for a delayed repair other-initiation, the program searches the last three utterances of the chatbot for a trouble source.

The heuristic starts the identification of a repair initiation with symbolic, medium-specific resources for signalling trouble like for instance ---, --, -, *?, ??, ???*, and lexical, language-specific recourses, such as *Ich verstehe nicht*, *unklar* or *was bedeutet?*. Then, if it is likely that the utterance in focus is a repair initiation, the program searchers for the trouble source. Uppercase writing or quotation marks may be used to highlight the trouble source, but also emoticons, to display own dissatisfaction with the trouble. To deal with learner language in repair other-initiations produced by language learners, Levenshtein distance of

20% of the average length of the word pair is set to recognise recycled words (e.g. retyped by the learner with errors).

Difficulties can occur in locating a TS that is longer than one word. This case is given when the user does not understand an idiom (collocation, proverb) or the meaning of an entire utterance or a longer part of an utterance. The assumption is that idiomatic expressions are more difficult to learn and require a higher level of language proficiency than single words. All idiomatic expressions used by the chatbot have been linked to their explanations in the Explanation DB and to a keyword list. The program searches for keywords in user's message and checks if they could form an idiom that the bot possibly used in one of the last three utterances. If so, the program generates an explanation from the corresponding entry in the ExplanationDB.

The ExplanationDB has been created from Wiktionary and contains information about German nouns, verbs, adverbs, adjectives, abbreviations, idioms and proverbs. To make the search faster and more efficient, a Wiktionary dump was preprocessed and only information in German and about German words was extracted, which was related to linguistic features such as synonyms, meanings, examples and notes about use. Every explanation was automatically generated from these fields according to AIML templates.

Whole sentences of longer parts of sentences can be sources of non-understanding if they contain unknown words or all words are familiar to the learner but the meaning of the sentence is not clear. A satisfying explanation can be a paraphrase of the sentence or its part, or a word-by-word explanation of potentially problematic tokens. Paraphrase generation / recognition (a.k.a. textual entailment) is a research direction, in which many authors invested lots of effort and which was discussed in Chapter 2.

Our data set contains only a small number of instances of OISR_M where paraphrasing or split-reuse is used for explanation. Most repair-initiations produced by learners address single words. Only simple versions for each of these self-repair carry-out strategies were implemented, namely paraphrasing and word-by-word explanation.

Paraphrasing as an RCO strategy is useful for sentences that contain collocations. For those sentences I created a set of corresponding paraphrases. Paraphrasing is also another way to explain idiomatic expressions instead of definitions from ExplanationDB.

A word-by-word explanation only makes sense for words that could be difficult for the learner. I use a list of 100 and 1000 most frequently used German words[1] to filter those words that are supposed to be well known to everybody. If the remaining set of words contains two or three words, they will be explained separately using the Explanation-DB.

[1]http://wortschatz.uni-leipzig.de/html/wliste.html

8.2.2. Recognition of $equals(x, y)$ and response generation

The repair manager searches in every user's input for sequences that match the pattern $X = Y?$. This pattern is a generalisation of multiple practices used by NNS in chat to perform a meaning check. A set of AIML categories helps to identify meaning checks and to extract x (the expression used by the system) and y (corresponding user's understanding of this expression). Once extracted, x and y will be redirected to the main AIML category responsible for the response generation.

The heuristic starts with the search for the middle part, which may be realised by symbolic, medium-specific means such as --, -, = or lexicalised, language specific means, for instance *heißt, bedeutet*. Left-hand part and right-hand part are checked in addition, if they contain any framing or highlighting, like for instance quotation marks or uppercase writing. Moreover, the left-hand side must be a repeat-based reference to the trouble source. Only if all these conditions are satisfied, the utterance is classified as an other-initiation of repair requiring a confirmation/disconfirmation as a response (otherwise, utterances such as *Wie heißt du?* could be recognised as repair initiations.

To generate a response, the chatbot needs to answer the question if x means the same as y? This is again an instance of the textual entailment problem. If x is a single word, an idiom, a collocation or a proverb, the system can check the list of the synonyms of the corresponding entry in the ExplanationDB (and do the same for y). If x and y are listed as synonyms in the ExplanationDB, a positive answer will be generated (yes, x=y). Otherwise, the system will explain the meaning of x using $explain(x)$ method described in Section 8.2.1.

8.3. Preliminary findings and discussion

In this chapter, I introduced an abstract computational model of other-initiated self-repair when the machine is the trouble speaker tailored for conversations with language learners. The conversation-analytic concept of *recipient design* is reflected in this model in a way that the self-repair carry-outs are tailored for language learners: the delivered explanations clarify the meaning of the used words and expressions.

Because repair is a building block of conversation, it is absolutely eligible to expect a conversational software to be able to perform proper handling of repair initiations and the ability to distinguish between questions with the function of repair initiation and all other questions. However, the great majority of the conversational agents and dialogue systems were designed with the assumption, that the user understands everything, and only the machine might have difficulties in language understanding due to technological limitations. I discussed a number of studies targeting these issues in Chapter 2. However, if a language learner is involved in a dialogue with a conversational agent or a dialogue system, the assumption that the user understands everything may no longer be valid. This work made a step towards closing this gap.

In order to determine what is required from the computational perspective to simulate sequences of OISR_M, an implementation case study was set-up. It was tested in the course of the case-study what kind of meta-linguistic information might be required, what kind of NLP tools might be necessary, and where the limitations are. With a very simple conversational technology such an AIML-based chatbot, it is already possible to make recognition of repair other-initiations part of the system using the model presented in this chapter.

To set-up a life chatbot test, many adjustments need to be performed in the chatbot besides of those related to OISR_M: learner language understanding, dealing with errors in repair initiations, what to talk about, how to deal with user's frustration because of the limited conversational capabilities. The local models of OISR and corrections can be evaluated by standard quantitative methods using precision/recall metrics with the accuracy of the repair initiation recognition and trouble source extraction as evaluation criteria. However, even human conversation participants may have difficulties in recognition of an other-initiation of repair due to learner's limited linguistic knowledge in the target language. I discussed this issue in Section 4.1.2.

Moreover, a non-present repair or correction does not mean that it is missing. Other things may be done instead of a repair initiation, and other strategies may be chosen by conversation participants to deal with trouble in understanding. The learners ask native speakers to explain something unclear because they assume, that the native speakers are able to do it, because it is native speakers' "territory of knowledge" (Heritage, 2012). However, in the tests with the machine, the users may choose to look up the dictionary instead of a repair initiations, if they do not trust machine's language understanding and language generation capabilities. Therefore, the most reasonable way to evaluate these dialogue models is qualitative analysis of chat protocols combined with other methods of qualitative social research, like interviews and questionnaires.

With regard to the areas of applicability of the proposed model, similar types of repair carry-outs can be found in professional talk, where novices in other areas than foreign language acquire professional terminology. It can be checked in a separate study, if the model presented here for SLA application can be transferred (maybe with modifications) into other domains where conversational agents are used, for instance as tutors.

9. Modelling Corrections of L2 Errors

In light of the discussion in Chapter 7 about the issues with the definition of an error regarding language standards, language dynamics, conceptually oral language and chat conventions, I will assume for the purpose of this chapter that something like *the language standard* exists and can be provided to a computer system as a reference for corrections.

Following the need to disengage "trouble (error and nonerror) from the practices employed to deal with it" pointed out by Schegloff (1987, p. 216), I looked at practices used by chat participants for dealing with learners' linguistic errors in Chapters 6 and 5. Because "the occurrence of repair is not prompted only by the occurrence of error," (Schegloff, 1987, p. 216), I separately analysed possible other reasons for occurrences of corrections, in Chapter 7. For the purpose of computational modelling, it appears reasonable to separate between models for dealing with errors from models for error recognition and models responsible for the decision, whether a correction should be produced. Therefore, I distinguish between models for practices of correction which I call *local models of correction* and factors influencing the decision if a correction should be produced and which form should it take. This chapter focuses specifically on local models of corrections of linguistic errors.

Correction type \ Attribute	Explicitness	Integration	In a projected turn
On-the-fly	x	-	-
Simple	x	-	-
Contrasting	x	-	-
OISR-based	x	x	-
Integrated	x	x	x
Embedded	-	x	x

Table 9.1.: Correction type classification and partition for computational modelling.

I presented and discussed various correction formats in Part II of this work. For the purpose of computational modelling, all identified correction types have been partitioned in two sub-classes according to their grade of integration into the next relevant projected turn. Table 9.1 shows the differences in classes with regard to the identified dimensions as well as the result of the partition.

For the remainder of this Chapter, I assume that an error recognition function (ERF) detects errors in learners' utterances. ERF may incorporate techniques for error recognition discussed in Section 2.1.4. However, error recognition in learner chat language needs to deal with a huge variance in orthography in syntax, which do not count as error in chat. I will address the issue of the error annotation in learner language in Section 9.1 in order to show the necessary modifications in correction modelling.

Computational models of exposed and embedded corrections will be subject of Sections 9.2 and 9.3, respectively. I will first explain how specific exposed correction formats can be generated from a set of language-specific and medium-specific resources. Error correction formats obtained from the data build a basis for the correction types. Because of the complexity of embedded corrections, which I discussed in Chapter 6, I will restrict modelling of embedded corrections to pairs of polar questions and answers. As in the corresponding empirical part, I consider polar questions produced by a learner and containing at least one "real" error. The corrections are then incorporated into answers to these questions. The model evaluation will take the form of a technical specification in which I analyse how the model can be implemented using available NLP-tools. The specification will be described in Section 9.4. The findings will be discussed in Section 9.5.

9.1. Learner error annotation

The error-annotation of the questions was performed according to the annotation guidelines for FALKO Corpus of German learner language (Reznicek et al., 2012, 2013). To prepare a basis for an error annotation, two types of target hypotheses were introduced in FALKO. The minimal target hypothesis ZH1 aims at sentence normalisation and is limited to only orthography and morpho-syntax. ZH1 was constructed according to the rules of standard German grammar and orthography with Duden dictionary as a reference. Semantics, lexical constructions and pragmatics are the subject of the extended second target hypothesis ZH2. Example 9.1 shows the two target hypotheses for a sample question.

Example 9.1. Creating target hypotheses for error correction in questions.

402	20:47:31	L08	und um wieviel Uhr gehst gewöhnlich zum Bett?
			and at what time do you normally go to bed?
		ZH1	Und um wie viel Uhr gehst du gewöhnlich zum Bett?
		ZH2	Und um wie viel Uhr gehst du normal ins Bett?

The following issues were faced when annotating errors in chat according to FALKO guidelines. First, special symbolic and orthographic means of expressivity used in chat must be classified as errors according to Duden and FALKO error annotation guidelines. Second, FALKO annotation guidelines do not provide any specific instruction for the cases where the errors in the verb make more than one target for the verb possible. Example 9.2 illustrates one of the cases. This error has been corrected by the interaction partner of L09 in the dialogue and both possible targets for the erroneous question were addressed in the correction. Therefore, having in mind the application where corrections should be automatically generated in a conversation, we add both target hypotheses to the annotation.

Example 9.2. Ambiguous target hypotheses.

135	13:21:09	L09	gefiel dir das studium leicht?
			Unclear target: Was the study easy for you? or Did you like your study?
		ZH1a	Gefiel dir das Studium?
		ZH1b	Fiel dir das Studium leicht?

The differences between the original learner's utterance and the two target hypotheses help to classify the errors and to generate corrections. In addition, it allows to analyse empirically what normalisation steps are really required for automated language understanding. However, chat conventions allow writing everything with small letters only and do not consider typos as errors that need a correction. This is why information about potential correctability of the error in chat need to be encoded in the error annotation. Additional rules for exceptions need to be specified when deviations in orthography and punctuation are used as a means of expressivity. Therefore, I introduced the "real" error flag with the purposer to identify all errors that are *potentially addressable* in chat. The conventions that I take into account for the "real" error flag are discussed earlier in Section 7.1. They all are restricted to orthography and allow to

1. start an utterance, a new sentence and nouns with a small letter,
2. write lowercase or uppercase or camel-case,
3. use punctuation and special symbols for the purpose of expressivity (emoticons),
4. omit punctuation and to use emoticons to separate turn-constructional units,
5. produce word stretches.

These rules are consciously applied by chat participants while typing. In addition, there are misspellings which are the result of a high typing pace and not lack of knowledge. They also do not qualify as errors in chat and are not considered as "real" errors. There are two exceptions that I take into account:

1. If a speaker repeats the same misspelling several times and the misspelled word sounds exactly as the correctly spelled word (comp. Example 5.8).
2. If it is a special, difficult case where even native speakers often make mistakes, for instance *ziemlich*.

In Section 6.3.1 I introduced a method for creation of a larger set of embedded corrections by data enrichment. Every answer to polar question containing a real error was modified by including an embedded correction in it. During this correction authoring process, it was not always clear, which ZH should be used for the embedded correction in the modified response. Example 9.3 illustrates an error in plural in a non-native-like expression. Two response versions with embedded corrections are suggested in EC1 and EC2. Each of the suggested embedded corrections corrects only one of the errors at the same time. It will be also hardly possible address these two errors with a single exposed correction.

Example 9.3. Different target hypotheses correct different errors. Trouble sources are underlined. Target hypotheses and respective embedded corrections are added in the bottom.

79	13:22:49	L03	)))) ja, wahrscheinlich! Sind die <u>Grenze des Schuljahres</u> von Urlaubssaison in <u>Beiern</u> abhängig? *yes, probably! Do <u>the border of the school year</u> [* errors: verb-subject number congruence, uncommon expression] depend on the holiday season in Beiern [* error: spelling]* yes, probably! Do the borders of the school year depend on the holiday season in Bavaria
80	13:23:16	L03	* Bayern * Bayern [self-correction] *Bavaria
81	13:23:41	N02	ja genau! ist das bei euch auch so? *yes, exactly! is it like this in your place, too?*
		ZH1	Sind die Grenzen des Schuljahres von der Urlaubssaison in Bayern abhängig?
		EC1	*Ja, die Grenzen des Schuljahres..* [corrects congruence error]
		ZH2	Sind die Ferienzeiten von der Urlaubssaison in Bayern abhängig?
		EC2	*Ja, die Ferienzeiten...* [corrects error in lexical choice]

Dealing with multiple target hypotheses will also be an issue for a computer program that should produce a correction. I will come back to this problem in Section 9.4.3.

9.2. Generation of exposed corrections

This section seeks show how exposed corrections for linguistic errors can be generated. I assume that for each error in learner utterance, the error recognition function ERF is able to recognise the error and to determine the target hypothesis. Chapters 5 and 6 showed that many correction formats are at participants' disposal. However, it is neither feasible nor desirable to specify each correction format for each error "hard-coded" in a conversational system. The aim of this section is therefore to describe a generalised model for each type of exposed corrections from a given set of language specific and medium specific resources. It needs a further investigation in a separate cross-linguistic study, which adjustments in the model are required to adapt the model to a language different than German.

9.2.1. Turn format and turn taking

Since only one error was addressed in each correction produced by the native speakers, I assume for the purpose of this section that user's message contains at most one error. Let user's message containing an error $M = m_1...m_n$ and $X^* = [m_i : m_k]$, $1 \leq i \leq k \leq n$ the deviant part of the message (the trouble source may be one or more tokens). Let X the correct expression in the target language, or simply target, for X^*. I use A to denote the set of accountings. I will distinguish between minimal accountings A_{min} and rich accountings A_{rich}. F_l and F_r denote the sorted sets of framing means to highlight the repetitions of the trouble source and the targets, so that $f_{l,i}$ $f_{r,i}$ build a pair of framing symbols for each $f_{l,i} \in F_l$, $f_{r,i} \in F_r$ and $1 \leq i \leq |F_l|$, $|F_l| = |F_r|$. I will use the $\centerdot$ operator to denote the concatenation of the parts of the correction. Some of the parts of turn formats are optional. This is taken into account in this model by the empty element ϵ included in each set where the presence of the elements of the set is optional. If the presence of an element is mandatory at a particular place, ϵ is not included. I use the letter $\mathscr{C}$ for correction format. The specific models for each correction format are described below.

Exposed corrections on-the-fly

Corrections on-the-fly are delivered in a chunk with a subsequent turn continuing the interpersonal trajectory. Specifically, the trouble source turn is followed by the target optionally framed with highlighting symbols, immediately followed by an emoticon $e \in A_{min}$ and a response to the interpersonal trajectory R_{topic}. Then the correction on-the-fly $\mathscr{C}_{onf}$ will have the following format:

$$\mathscr{C}_{onf} = f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e \centerdot R_{topic} \tag{9.1}$$

The order of two last elements e and R_{topic} can be reverse, and either the framing or the position of emoticons e or even both may be empty. Either the generated correction can be broken down to three messages posted to the user with very short breaks one after the other, or they all can be posted as one message. This leads to the following two models for the turn taking for this correction format:

1. Three separate turns, one per part:

   T1 $f_{l,i} \centerdot X \centerdot f_{r,i}$
   break

   T2 e
   break

   T3 R_{topic}

2. The correction and the response to the interpersonal trajectory are delivered in one turn: T1 $\mathscr{C}_{onf}$ as specified in Equation 9.1.

The length of the breaks can vary. Empirically grounded values can be obtained from the chat data in an additional study focused on responsiveness in instant messaging based communication. One attempt to get insights from the dataset into the issues of responsiveness in chat was discussed in an earlier work (Danilava et al., 2013b).

Simple corrections with minimal or rich accountings

In simple corrections with minimal accountings, only a correction is presented, optionally followed by a member of the set of minimal accountings A_{min}. The correction may be optionally framed or highlighted. The correction turn $\mathscr{C}_{sma}$ has then the following form:

$$\mathscr{C}_{sma} = f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e \tag{9.2}$$

The recipient of the correction has then the choice to react to the correction or to continue on the interpersonal trajectory.

Rich accountings of different kind are involved in corrections to express different social actions and intentions. As shown in Section 5.2, more than one element from the set of rich accountings may be placed in a chunk as part of a correction. I use a_{rich} to denote such a chunk of rich accountings composed from a non-empty subset of all possible types of rich accountings A_{rich}. If the chunk contains more than one element, they are sequentially ordered by their interactional function. A generalised format of a simple correction with rich accountings $\mathscr{C}_{sra}$ may then look as follows:

$$\mathscr{C}_{sra} = a_{rich} \centerdot e \centerdot f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e \centerdot a_{rich} \centerdot e \tag{9.3}$$

Depending of the type of rich accountings (apology, correction justification etc.) they may be correction-initial or correction-final. The correction may be presented to the user in multiple turns:

T1 $a_{rich} \centerdot e$
break

T2 $f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e$
break

T3 $a_{rich} \centerdot e$

Correction-initial accountings, such as requests for permission to correct, may be followed by a pro-forma break offering the user the chance to react. However, they are not real requests for permission, and therefore, there is no need to wait for permission. Nonetheless, such requests can be used for the purpose of negotiation of conventions. In this way it can be clarified if it is okay for the user to receive error corrections.

Specific social actions performed by each type of accountings need to be taken into account when each correction turn is generated. Therefore, the simplified model specified in Eq. 9.3 must become more specific for an implementation. I will come back to this question in

Section 9.2.2 where I address language-specific devices and social actions performed by them.

Highlight by register change (e.g. uppercase) can be used instead of or in addition to framing. In practice, however, the language resources in chat were used rather economically. Because the typing pace is an issue, the participants used either framing or uppercase to highlight the important parts in the correction. The model reflects this observation. I denote the register change to uppercase by $upper(X)$ for the trouble source X. Then the correction format will be as follows:

$$\mathscr{C}_{sra} = a_{rich} \centerdot e \centerdot upper(X) \centerdot e \centerdot a_{rich} \centerdot e \tag{9.4}$$

A presentation of the correction in multiple turns will be similar to the variant without uppercase. This way of highlighting was never used by NS participants for corrections on-the-fly or corrections with minimal accountings, probably because writing everything in uppercase is interpreted as shouting in chat, and for these two correction formats the correction equals to the entire message. As opposed to these short correction formats, uppercase writing of a part of the turn helps to highlight the target in corrections with rich accountings.

Since delayed corrections require a higher effort in terms of interaction management, only corrections with rich accountings (both simple and contrasting described below) allow for delayed corrections. In addition, it is necessary to renew the context, which was done by the native speakers in the dataset by specific back-linking tokens. I use BL for the notation of the set of all back-linking tokens and bl for its elements, thus $bl \in BL$. Then a delayed correction with backlinks will have the following format:

$$\mathscr{C}_{sra} = bl \centerdot a_{rich} \centerdot e \centerdot f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e \centerdot a_{rich} \centerdot e \tag{9.5}$$

The variant of presentation of the correction to the learner in multiple turns will have a back link in the first turn.

Contrasting corrections with minimal or rich accountings

Corrections formatted in a way that both, the trouble source and the target expression are presented to the learner in a direct comparison (contrast) may require minimal or rich accountings. The trouble source produced by the learner is explicitly rejected and replaced by a different unit within this correction format. The basic elements used for the contrasting correction formats are the same as for simple corrections with accountings with an additional element: negatively marked repetition of the trouble source which I denote by $\neg X^*$. A contrasting correction with rich accountings will then be formed as follows:

$$\mathscr{C} = a_{rich} \centerdot e \centerdot \neg(f_{l,i} \centerdot X^* \centerdot f_{r,i}) \centerdot f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e \centerdot a_{rich} \centerdot e \tag{9.6}$$

A contrasting correction with minimal accountings will differ from the correction model specified in Eq. 9.6 only in the type of accountings used:

$$\mathscr{C} = e \centerdot \neg(f_{l,i} \centerdot X^* \centerdot f_{r,i}) \centerdot f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e \tag{9.7}$$

Similar to the simple corrections with rich accountings, the highlighting of the trouble source and the target can be done with register change (uppercase). A contrasting correction with rich accountings will then have the following changes:

$$\mathscr{C} = a_{rich} \centerdot e \centerdot \neg(upper(X^*)) \centerdot upper(X) \centerdot e \centerdot a_{rich} \centerdot e \tag{9.8}$$

The contrasting correction can be presented to the user in one turn or in a turn chunk, for instance the variant with rich accountings corresponding to 9.8:

T1 $a_{rich} \centerdot e$
break

T2 $\neg upper(X^*)$

T3 $upper(X) \centerdot e$
break

T4 $a_{rich} \centerdot e$

A similar chunk-wise output may be produced for 9.6. Again, the duration of the time breaks between the turns in the chunk are motivated by the responsiveness values between short turn in users' chunks. However, responsiveness values in chat needs a further investigation to provide a more stable, empirically-grounded model for breaks between turns. Possible combinations and sequential order of rich accountings will be discussed in Section 9.2.2.

Corrections formatted as repair other-initiations

To generate repair initiations by a computer system, the reverse process of those described in Chapter 8 needs to be specified here. A recognised trouble source needs to be wrapped into a repair initiation. For this purpose, resources for signalling trouble and referencing trouble source need to be involved in this type of correction. In the examples found in the dataset and discussed in Part II of the work, two distinct formats of repair initiation as a way to correct errors were found:

1. Repetition or recycling of the trouble source combined with marking it as unclear,
2. Presenting of one or several candidate understandings.

However, additional repair other-initiation formats cannot be excluded in general from using them for the purpose of correction. Therefore, the grammar for error corrections in form of repair other-initiation must take different possible types of trouble source into account. In contrast to the repair other-initiations produced by the learner, explicit statements of non-understanding will not fit into corrections. For instance, a repair initiation-based correction "X? Nie gehört..." is either too teacher-like or learner-like. Therefore, a different set of language-specific resources needs to be specified here. I use $CR(X)$ for all trouble signals formatted as clarification requests, and $candidate(X, Y)$ for all signals involving candidate understandings. Then a correction $\mathscr{C}$ can be defined as an element-wise product of referencing trouble source and signalling trouble.

$$\mathscr{C} = REF \times SignalFormat \tag{9.9}$$

Below I explain the generation of REF and $SignalFormat$ for each type of trouble source.

A part of a longer message: specific formats based on reuse or recycling or demonstrative determiners and proterms are suitable for this type of the trouble source. A part of a longer message may be, for instance, one word, a phrase consisting of several words, or a copy of a longer part of the message.

if *TIME=immediate* **then**
  $REF \in \{reuse(x),\ recycle(x),\ DD,\ DP\}$;
else
  $REF \in \{reuse(X*),\ recycle(X*)\}$;
  $SignalFormat \in \{CR(X*),\ candidate(X*, X)\}$;
end

The whole message: open class repair initiation formats are appropriate for this case (reference by adjacent position), however, they only allow to unambiguously reference the trouble source if used in the adjacent turn. Otherwise, formats based on recycling are preferable. In Chapter 8 I distinguished between one-word messages that became trouble sources, and longer messages that became trouble sources.

if *TIME=immediate* **then**
  $REF \in \{AP,\ DP\}$;
else
  $REF \in \{reuse(X*),\ recycle(X*)\}$;
  $SignalFormat \in \{CR(X*),\ candidate(X*, X)\}$;
end

Signalling devices can be placed in the turn on the left of the reference to the TS or on the right, or both. Open class repair initiators do not contain an explicit reference to the TS, so it is only relevant for more specific types of references (reuse, recycle, DD, DP).

A right-hand side of a signal format set may contain all symbolic and lexical means that can be placed after the reference to the trouble source. Normally, they can be found in a turn-final position, on the right-hand side of the trouble source, which I denote by $CR(X*)_{right}$. For instance:

$$CR(X*)_{right} = \{\text{"?","???", "--???", "? Wie meinst du das?", "-- unklar"}\}.$$

Repair other-initiation-based formats are functionally questions, and therefore, they are normally formatted as questions, which is done by a question mark in chat. Question marks are a very important mean of expressivity in chat, they are rarely omitted by chat users, and should be used by conversational agents to ensure a better understanding of agents' intentions. Therefore, $CR(X*)_{right}$ does not contain ϵ.

In contrast, signalling means which can be placed on the left-hand side of the trouble source (normally, turn-initial) contain ϵ, because it is optional. An example of a left-hand side of a signal format set may be the following:

$$CR(X*)_{left} = \{\epsilon\text{"wie,", "was heisst", "was bedeutet"}\}.$$

Repair other-initiation-based correction formats found in the dataset were all repeat-based: their purpose was to replace a reference that caused troubles in understanding by another reference that resolved these troubles. It was always possible for the native speakers to locate the trouble source very specifically. In the cases where the meaning of an entire learner's utterance was not understandable, native speakers preferred to let-it-pass. However, sometimes native speakers initiated immediate repair with a lexicalised paralinguistic cue such as *hä?* (open class).

Repetition of the trouble turn is the usual repair carry-out type after such repair initiations in oral interaction. Learners, in contrast, delivered a *different* expression to replace the trouble source, or additional explanation to clarify their intention. It was just not a repeat in any case. Dialogue models need to include a decision what happens with the repaired talk and how to handle it sequentially, if the machine cannot understand learner's utterance even after the replacement. This problem has already been addressed to in academic literature on so called *clarification requests* which I discussed in Section 2.3.2.

9.2.2. Language-specific and medium-specific devices

The machine needs to guess, which resources for framing and highlighting the user might have used, or the recognition of repair other-initiations by learners. As opposed to this, a set of resources typical for German chat needs to be given to the machine for correction generation. Scientific studies focusing on language resources in German chat may provide a more comprehensive description of medium-specific resources, for instance (Orthmann, 2004). However, technological development and the growth of the variety of device-specific and application-specific means of expressivity can hardly be completely mirrored in scientific articles which have been published a decade ago and tell about data which are even a few years older. Nevertheless, there are "good old smileys" which are device-independent and application independent, and are comprehensible for all chat users. I will take into account only this sort of emoticons in this section when I talk about accountings.

Resources to mark the boundaries of corrections

As discussed earlier in this section, corrections and repetitions of trouble sources are frequently highlighted by symbolic and lexical means. Symbolic resources include specific symbols to mark the boundaries of the correction (framing) and register change, such as uppercase spelling. Lexical means include special words and expressions. Resources for emphasis have been taken into account in correction patterns. An example framing set may be provided to a computer system as a medium-specific resource and include the following elements:

$F_l = \{\epsilon$, [", ', (, -, --$\}$, $F_r = \{\epsilon$,], ", ',), -, --$\}$.

Uppercase typing was also used to highlight the error and the correct form in exposed corrections. I denoted uppercase spelling as $upper(X)$ in preceding sections.

For repair other-initiation-based correction formats, a similar set of signalling may be used for correction generation as provided to the machine for repair other-initiation recognition. Only native-like expressions should be used by the machine for repair other-initiation-based correction formats.

Accountings as resources to manage corrections as social actions

Minimal accountings include smileys :-) and more frequently ;-) (or their variants). These are medium-specific devices used across languages for chat communication. However, many different variants of them exist and some of them are preferred in particular cultures. For instance, Belorussian participants used only one or more parenthesis ")" for smile or laugh ")))))))" while German participants preferred the version with the eyes :) or with the eyes and a nose :-). Orthmann (2004, p. 145) describes other variants of emoticons used in German youth chat, which I did not observe in my dataset. This might me explained by factors like participants' age, intercultural character of the communication, changes in chat language over time and dyadic interaction instead of group chat, which was put in focus of research by Orthmann (2004).

The purpose of the minimal accountings is to focus learner's attention on correction while managing its face-threatening. Therefore, minimal accountings are restricted to framing possibly combined with a single smiley. Angry or somehow negatively polarised smileys were never found in the role of accountings in my dataset. Therefore, an example set of minimal accountings may include the following elements:

$A_{min} = \{\epsilon$, :-), :), ;-), ;)$\}$

Rich accountings may express apologising, ridiculing, accusing, instructing, forgiving, admitting, complaining. I found in the data only requests for permission to correct (e.g. "kurze Anmerkung zum Deutsch (wenn ich darf)" turn 237 chat N03-L07), apologising, instructing (e.g. a combination of the last two "Verzeih mir, aber es sollte heißen:" Example 5.11) and justifications of corrections (e.g. "ich sage das nur weil du sonst sehr gut Deutsch sprichst/ schreibst" turn 247 chat N03-L07). A set of phrases for each type of rich accountings may be the easiest solution. In the following example sets A_i, the index i expresses the social action performed by the set members:

A_{excuse} = {"Sorry, aber", "Verzeih mir, aber", "Entschuldigung, wenn ich dich damit nerve, aber", ϵ}

$A_{request-for-permission}$ = {"Falls ich das korrigieren darf...", "Eine kurze Anmerkung zum Deutsch (wenn ich darf)...", ϵ}

$A_{instructing}$ = {"Es sollte heißen", Es heißt", ϵ}

$A_{instructing-neg}$ = {"So was gibt es nicht in der Form", ϵ}

$A_{intention}$ = {"Ich will dich aber nicht nerven", "Ich will ja nur helfen", ϵ}

$A_{justification}$ = {"Ich sage das nur, weil du sonst sehr gut Deutsch sprichst", ϵ}

$A_{instructing-later}$ = {"Wenn ich dich damit nicht nerve, habe ich hier noch eine Anmerkung", ϵ}

$A_{express-understanding}$ = {"Es ist sicherlich verwirrend.", "Die Deutschen blicken da selbst manchmal nicht durch.", ϵ}

$A_{encouragement}$ = {"Du machst es aber sonst recht gut.", ϵ}

$A_{emphasis}$ = {"Kein blöder Witz, das heißt wirklich", ϵ}

In a practical computer application, corrections with rich accountings can be generated using only one sort of accountings per correction or combinations of them. Both types are covered by the correction format models in Section 9.2.1. In addition, they can be accompanied by minimal accountings. In order to handle the palette of possible corrections in more detail, I provide more specific correction models. They combine different types of social actions performed by different types of accountings. Some of the accounting types occur in correction-initial positions, some of them were found typically in correction-final positions. If the complete correction is produced in one turn, then they are turn-initial and turn-final, respectively.

Findings discussed in Part II of the dissertation show that it might be a good strategy to negotiate with the user about the acceptance of error corrections. Accountings may help to do this. If the machine finds an error that should be corrected, an exposed correction with specific accountings may be produced. These specific accountings will formally ask for permission to correct and check the acceptance of error correction for the future. Hence, this specific combination of accountings is more appropriate in the beginning of the communication and in the beginning of the "correction history". I use the common notation for regular expressions to denote a set of elements from which only one has to be selected at a particular position $[a_{justification}|a_{intention}]$. Then, the first correction in the communication with a particular user may have the following structure:

$$\mathscr{C}_{init} = a_{request-for-permission} \centerdot e \centerdot F_l \centerdot X \centerdot F_r \centerdot e \centerdot [a_{justification}|a_{intention}] \centerdot e \quad (9.10)$$

As explained above in Section 9.2.1, the same correction may be split in several turns to make the messages shorter.

An exposed correction some time later or after an already produced repair may also start with an excuse followed by an instructing:

$$\mathscr{C}_{later} = [a_{excuse}|a_{excuse-later}] \centerdot a_{instructing} \centerdot e \centerdot f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e \quad (9.11)$$

$$\mathscr{C}_{later} = [a_{excuse}|a_{excuse-later}] \centerdot a_{instructing} \centerdot e \centerdot f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot a_{understanding} \centerdot e \quad (9.12)$$

The contrast in contrasting corrections is achieved by opposing the trouble source and the correction with accounting types $A_{instructing-neg}$ and $A_{instructing}$. A special conjunction may help to create the contrast in the utterance, which I denote by C. In German it is done by the conjunction *nicht ... sondern*. The correction may also be prefaced by an excuse.

$$\mathscr{C} = a_{excuse} \centerdot e \centerdot a_{instructing-neg} \centerdot f_{l,i} \centerdot X^* \centerdot f_{r,i} \centerdot C \centerdot a_{instructing} \centerdot f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e \quad (9.13)$$

Specifically for the integrated corrections, it might be required to add an emphasis after an embedded correction in order to focus on the correction, as discussed in Section 5.3. Embedded and integrated corrections will be subject of the next section. To cover the role of accountings in integrated corrections, I assume that an integrated correction can be produced at some point of the conversation:

T1 Trouble-turn with an error X^*

T2 Integrated correction replacing X^* by X

T3 $a_{emphasis} \centerdot f_{l,i} \centerdot X \centerdot f_{r,i} \centerdot e$

The correction formats listed in this part of the current section are very specific, though they are derived from more general correction formats specified in Equations 9.6 and 9.8. However, they cover a large number of surface variants, because many other ways are possible to express the social actions of requests for permissions, instructing, excuse etc. which were not included into the example sets of respective types of accountings.

Backlinks

Backlinks help to come back to the subject of the talk in delayed exposed corrections. The machine may choose to respond first to the interpersonal trajectory prior to initiating a correction side sequence. In this case, a specific return token may be necessary. I call such tokens *backlinks*. An example set of backlink tokens is BL:

$BL = \{$"PS:", "übrigens", "außerdem", $\epsilon\}$

I found backlink tokens always in the turn-initial position of the correction turns in simple and contrasting corrections with rich accountings.

9.3. Generation of embedded and integrated corrections

Similarly to the conceptual definition of an error, a conceptual definition of a correction appears problematic, although most research publications rely on the intuitive understanding of this term. I choose to restrict my definition of an embedded correction to a description of its properties, namely:

1. They obviously correct an error in form of a replacement of a part of an erroneous utterance.
2. They are fully integrated in syntactic and semantic structure of the correction turn.
3. They do not contain any highlighting of the correction or accountings.

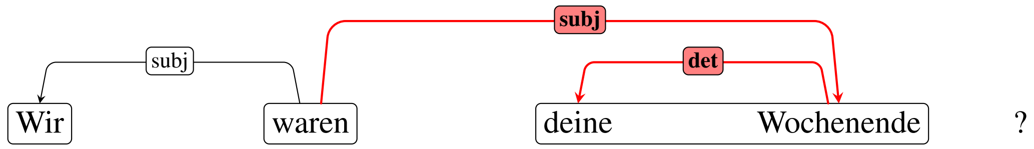
The models of exposed correction described in the preceding section of this chapter cover the broad range of all exposed correction formats found in the reference dataset except of the integrated corrections. Chapter 6 showed that there are also different sequential environments where embedded corrections can occur, however, I could not find any prototypical turn formats for embedded corrections in general. For this reason, I narrowed my focus to embedded corrections for question-answer pairs where the learner produces a question containing one or more errors and the agent is expected to produce an answer which can contain an embedded correction.

Integrated corrections differ from embedded corrections in the last property. They usually contain highlighting and may contain accountings. Both embedded and integrated corrections are semantically and syntactically integrated into the next relevant turn. Therefore, I approach the modelling task for integrated corrections with the assumption that an integrated correction can be formed from an embedded correction by adding accountings and highlighting the replacement.

A detailed analysis of errors in questions produced by the learners and embedded corrections provided by the native speakers and created by the researcher was explained in Section 6.3. This analysis showed that there are types of learner errors which cannot be corrected implicitly. I listed multiple cases of errors which are not implicitly correctable in Section 6.3.4. The main objective of this section is to describe a computational model of embedded corrections for implicitly correctable errors in question-answer pairs. Specifically, I focus on the set of polar questions that (a) are produced by learners, (b) received an answer as a response and (c) contain at least one "real" implicitly correctable error. I explained the composition of this collection in Section 6.3. The set received the name $QAPA$.

I use the error-annotated collection of question-answer pairs described in Section 7.1 for a systematic analysis. The target hypotheses annotated as described in the beginning of this Chapter in Section 9.1 were taken as the basis to create embedded corrections in this phase. As it was shown in Section 6.3.3, repetition-based answers to polar questions offer opportunities for embedded corrections. The erroneous unit is replaced in the response by the assumed target unit. The concept of a minimal correcting replacement was introduced in Section 6.3. I will reuse this concept for the purpose of modelling of embedded corrections in answers to polar questions.

Participants of a conversation have different possibilities to refer to objects and action in conversation after they were already mentioned for the first time. They can repeat the same term, they can use a proterm or they can replace the initially introduced term by a different one (Jefferson, 1983). A correction is a replacement where the erroneous unit is replaced by a correct one addressing the error. For each error type, I need to describe the smallest unit of the utterance that needs to be replaced in order to address the error that needs to be corrected. For a spelling error, it will be the word itself (the wrong spelling is replaced by a correct one), for errors related to determiner-noun congruence (e.g. gender, case and number), the determiner-noun phrase needs to be repeated and the erroneous items in it need to be replaced by correct ones. However, in practice, embedded corrections of some errors are not possible, like for instance errors in interrogative word order because word order changes in the answers.



1	2	3	4	5
wir	sein	dein	Wochenende	?
1 | Pl | _ | Nom	_ | Pl | Praet | Ind	Fem | Adj | Sg	Neut | Nom | Sg	
PPER	VAFIN	PPOSAT	NN	$.
We	*were*	*your*	*weekend*	*?*
Wie	waren	deine	Wochenenden	?
Wie	war	dein	Wochenende	?

Figure 9.1.: A dependency tree for a question produced by L05 (turn 145 on 25.06.12 at 17:18:41). The last two lines correspond to ZH1 and ZH2. Red edges represent ungrammatical subtrees: noun-determiner congruence and verb-subject congruence are problematic.

The maximum number of replacements equals the number of constituents in the question. Taking dependency grammar as an abstract model for syntax, the maximum number of potential embedded corrections for each utterance will be equal to the number of nodes in the syntax tree (including terminal nodes). For the example sentence illustrated in Figure 9.1, the number of potential embedded corrections will be 7: one per word, one for the subtree tokens 3&4, one for the subtree tokens 2&4 and its dependent, one for the subtree tokens 1&2. Two different target hypotheses can be created for the example sentence from Figure 9.1:

ZH1 Wie waren deine Wochenenden?

ZH2 Wie war dein Wochenende?

The first token *Wir* is a result of a typo, however, it is an existing word and it matches to the verb in its morphology, syntax and meaning. However, in this way there are two subjects in the sentence which is problematic. If the typo is corrected first, only one subject is left and problematic congruences can be resolved.

Both ZHs are grammatical, but only ZH2 makes sense in the interaction from the dataset, because only one weekend passed since the previous interaction of this pair of participants. A repetition of the word *deine* alone or of its grammatically correct form *dein* would not repair the congruence problem. Similarly, a repetition of the verb *sein* in its grammatically correct form would not correct the verb-subject congruence error. Instead, the whole subtrees need to be repeated with correction where incorrect forms are replaced by correct ones. This means, for each error type, there must be a minimal unit that needs to be repeated with correction in order to address the error. In order *not to* address the error, the problematic unit must be referenced in a different way.

To describe the mechanism of embedded corrections, I compared how the problematic units are referenced in responses to questions with embedded corrections and without in Section 6.3. The model built upon these findings is explained in Section 9.3.1.

9.3.1. Embedded corrections in answers to polar questions

This section explains models of embedded corrections in question-answer pairs where learners produce polar questions containing an error, and the machine is expected to produce an answer containing an embedded correction. I assume that the question contains exactly one real implicitly correctable error.

Answers to polar questions are classified in the coding scheme as confirming or disconfirming regardless to their form (Stivers and Enfield, 2010). I mentioned also the infamous German *jein* as a possibility. However for simplicity, I assume that only the first two possibilities exist, because the third can be produced from the first two.

A sequence of operations for answer generation to a polar question may look as follows:

1. Locate the question.
2. Determine if the answer should be confirming or disconfirming.
3. Generate responses.

4. Select an answer form the set of all generated answers: interjection-based (marked or unmarked) or repeat-based.

Repeat-based forms of answers support embedded corrections, as it was shown in Section 6.3.2. Because only the last step is in focus of this section, I assume that a language generation function (LGF) generates all possible answers to an identified polar question, as specified in the first three steps.

What is the minimal constituent that needs to be repeated in which a replacement needs to be performed in order to produce an embedded correction? To answer this question, I need an error recognition function (ERF) which finds a target hypothesis and determines the scope of the error, which I refer to as minimal error constituent.

Definition 9.1. *Minimal error constituent (MEC)*
A minimal error constituent is the smallest constituent that allows to identify the error.

Thus, the desired error recognition function identifies the minimal error constituent and the corresponding target hypothesis. Larger parts containing the minimal error constituent will correct the error, if repeated with a replacement in it, that corrects the error. Smaller parts which do not completely contain the MEC will not correct the error, even if they are repeated or replaced. The concept of MEC helps to operationalise the term *embedded corrections* in answers to polar questions. The first property of embedded corrections can be then reformulated as follows:

1. The answer to a polar question contains a MEC for the error.
2. The replacements within the MEC address the error.

As it was discussed in Section 6.3.4, not every unit is repeatable in an answer to a polar question. Repetition-based answers to polar questions contain only repeatable units, however, not every repetition-based answer will correct the error. The first property addresses the problem of repeatability, which in practice is already handled by the LGF. The second property covers the cases where replacements are performed in a way that leaves the error uncorrected. For instance, spelling errors can be only corrected, if the same word is repeated without error. They are not corrected, if the entire word with the error is replaced by a different word. A repair management function (RMF) needs to determine which of the surface variants is the best to be delivered to the user. Figure 9.2 illustrates how the three modules communicate.

From the perspective of correction, all of the response variants that satisfy the two properties would do the correction job. However, different variants of responses with embedded corrections may have different effect on learning an social interaction. The analysis of dependencies between variants of responses with embedded corrections, learning and their interactional import can be addressed to in a future study.

Each format of exposed corrections has been described by a respective regular expression. In contrast, I do not see any possibility to find a general regular expression for all embedded

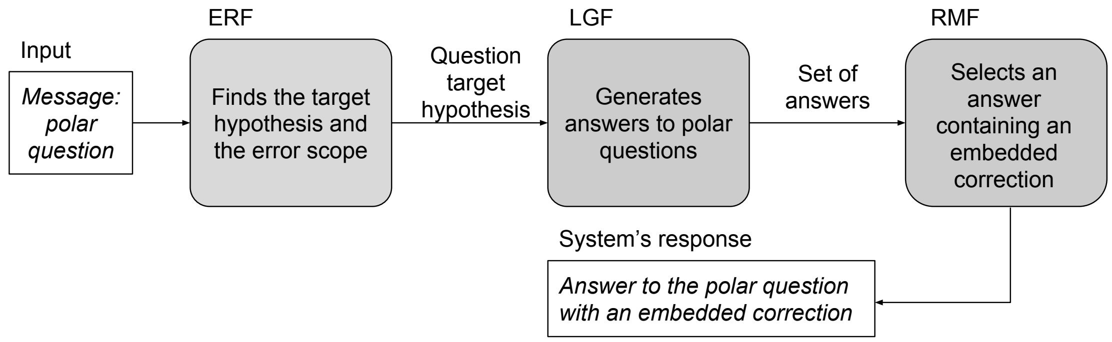


Figure 9.2.: Generation of embedded corrections.

corrections as one correction format. Embedded corrections need to be made an integral part of the response not explicitly formatted as a correction. The minimal error constituents are syntactically not always restricted to single words or neighbouring words, therefore they cannot be in general described just as simple string replacements. For the purpose of illustration of the difference between embedded and integrated corrections, a simplified format of an embedded correction as a string replacement can be described as follows.

Let $Q = q_1...q_n$ a polar question containing a spelling error at position $i, 1 \leq i \leq n$, and $R = r_1...r_m$ a possible repeat based answer. Let q_i the minimal error constituent in the question. The error recognition function ERF finds the target hypothesis for the entire polar question, which includes a target hypothesis for the MEC q_i, $ERF(q_i) = c$.

Then an answer to polar question is an **embedded correction** if the following holds:

$$R = r_1...r_m, \exists i(r_i = c), 1 \leq i \leq m \tag{9.14}$$

Thus, the corrected MEC is placed somewhere in the response turn. The position of the replacement c is determined by the syntax.

9.3.2. Integrated corrections in answers to polar questions

Exposed corrections integrated into next relevant turn are similar to embedded corrections in a way that the corrections are fully syntactically and semantically integrated into the ongoing talk, and do not explicitly initiate a separate sequence, as opposed to other types of exposed corrections. However, in contrast to embedded corrections, they may contain meta-linguistic information and accountings. Therefore, I will build a model for integrated corrections on the model for embedded corrections described in the preceding section.

An **integrated correction** can be produced from an embedded correction by adding highlighting to it:

$$c_{H,j} = \begin{cases} upper(c) & for \quad j = 0 \\ f_{l,j} \centerdot c \centerdot f_{r,j} & for \quad 1 \leq j \leq |F_l| \end{cases} \tag{9.15}$$

$$R_j = r_1 ... r_m \centerdot e, \; \exists i(r_i = c_{H,j}), \; 1 \leq i \leq m, \; 0 \leq j \leq |F_l| \tag{9.16}$$

where $f_{l,j}$ and $f_{r,j}$ are elements from the sets of framing resources and $upper(c)$ the uppercase spelling to highlight the correction and $e \in A_{min}$ is an element from a set of minimal accountings as defined in Section 9.2

In the implementation practice, the possibilities to generate embedded and integrated corrections are bound by the system's capability to generate repetition-based answers to polar questions and to the used language understanding technology in general. I will come back to this question in the next section.

9.4. Applicability of the models for communicative ICALL

This section seeks to analyse practical applicability of the proposed models for exposed and embedded corrections. Using the chatbot technology for the implementation introduced in Section 8.2, the repair manager can be further extended with a correction function. New AIML-categories need to be introduced for each type of correction format. For each variant of the surface, a separate output variant must be specified in the AIML file (specific accountings types, specific smileys). These templates need then only to be filled in with the correct expression. I discuss several possible templates for exposed corrections on one example of an error in Section 9.4.1. For embedded and integrated corrections, in contrast, the ability of the system to find repetition-based answers to polar questions is crucial. I will discuss the chances and the limitations of AIML-based language understanding with regard to embedded corrections in Section 9.4.2.

9.4.1. Templates for exposed corrections

In contrast to embedded and integrated corrections, I did not find any abstract minimal unit that was present in all corrections. For instance, if an error is located in a wrong from of a German participle II as in Example 9.4, different formats of exposed corrections can be generated. Only a simple correction on-the-fly was found in the corpus to correct this error. I discuss below how other types of corrections can be generated using correction templates obtained from the correction models as specified in Section 9.2.

Example 9.4. A sample error and a correction from the corpus.

242	19:19:10	L08	hast du etwas über deutsche Gruppe PUR gehören? *have you something about [*error: zero-article] German band PUR heard [*error: wrong form of the participle]* have you heard anything about the German band PUR?
243	19:19:21	N04	gehört ;-) *heard [*correction] [smile]*
244	19:19:24	N04	ja, kenn ich *yes, I know them*

Although a correction-on-the-fly may be the preferred form of correction in this sequence, other correction formats may be generated to address *the same error*.

Simple and contrasting with minimal accountings

Templates for simple corrections with minimal accountings of the error from Example 9.4 will not differ much from the correction on-the-fly chosen by the native speaker in the dataset. However, the machine will not post a message responding to the interpersonal trajectory immediately after the correction. I will give the user a chance to react to the correction.

Example 9.5. A simple correction with minimal accountings.

i	`User`	hast du etwas über deutsche Gruppe PUR gehören?
i+1	`Bot`	gehört :-)
i+2		User's turn

The place of the [smile] in the correction may be occupied by one of the emoticons from the set A_{min}. Other versions of the correction in turn i+1 may be, for instance:

- "gehört" :-)
- GEHÖRT :-)
- -- gehört -- :-)

A repetition of the error and a negation added to this will form a contrasting correction with minimal accountings:

Example 9.6. Contrasting correction with minimal accountings.

i	`User`	hast du etwas über deutsche Gruppe PUR gehören?
i+1	`Bot`	gehört, nicht gehören ;-)
i+2		User's turn

Even if only this simple correction format should be used, a few variants are available to the machine in order to avoid repetitiveness which is perceived by users as demotivating and disturbing the interaction (Bickmore and Picard, 2005).

Simple and contrasting corrections with rich accountings

Simple corrections with rich accountings in the beginning of the interaction with a conversational agent may have the following form:

- Falls ich das korrigieren darf: "gehört" :-) Ich sage das nur, weil du sonst sehr gut Deutsch sprichst :-)

An exposed correction following a different sequence of repair with linguistic trouble source may be obtained form templates and have the following form:

- Wenn ich dich damit nicht nerve, habe ich noch eine Anmerkung: es sollte heißen GEHÖRT ;-).
- Sorry, aber es heißt "gehört" :-)

A template for a contrasting correction may delifer the following correction form:

- Entschuldigung wenn ich dich damit nerve, aber so was gibt es nicht in der Form "gehören" sondern es sollte heißen "gehört" :-)

Each of the correction templates to generate these corrections need to be specified separately and contain a placeholder for the correct word and the repetition of the trouble source. In addition, they may be split in several turns to make the messages shorter. Other variants can be created with the sets of accountings specified in Section 9.2. The accounting sets may be extended with other expressions, and new sets of accountings may be introduced. This will make the number of potential correction surfaces for each of the formats even larger. However, at most one of the surface variants should be presented to the user. A strategy for this selection is needed. I did not find any factors explaining the selection of one or another variant - neither in the chat logs nor in questionnaires nor in the retrospective interviews. Another study focused on this problem may help to shed light on decision processes and users' preferences. A random selection would be the only way for the implementation at the current stage of the research.

Repair other-initiation-based corrections

Various templates for corrections based on repair other-initiation may be created based on the corresponding formal model.

- Was heißt das, gehören. Meinst du "gehört"? :-).
- Gehören --??? Meinst du 'gehört'? :-)

However, in the dataset this type of correction was used only to correct lexical errors. Corrections of morpho-syntactic errors by this type of exposed correction may sound too teacher-like. Such "known-answer questions" (Sawchuk, 2003) may be "too much teacher" for a Conversation-for-Learning. Native speakers use this type of correction to correct lexical errors when learners used a word non-existing in German or an existing word in a wrong context. Although errors of form can be potentially corrected by repair other-initiation-based formats, such corrections may be misplaced in a Conversation-for-Learning but appropriate in a language classroom.

9.4.2. Generation of embedded and integrated corrections

Possibilities for an implementation of embedded correction in a conversational agent depend on the ability of the system to recognise correctable errors (ERF) and to produce reasonable answers to polar questions (LGF). The chatbot technology used in the implementation case study foresees a very simple mechanism for finding a response to any users' question, namely, select an answer from a predefined list of possible answers. Chatbot's knowledge is available to the programmer from the beginning, and every response where this knowledge may be relevant, can be created in advance by AIML script authors. Error recognition in the baseline chatbot is limited to a predefined list of substitutions, which are applied to the input string in the pre-processing phase in order to find a target hypothesis.

To cover as many potential inputs as possible, a lot of generic responses to questions are part of the AIML-based brain of the chatbot. A typical example of this behaviour is shown below: for every polar question starting with something matching to GEHST DU *, the same response will be delivered to the user.

```
<category>
  <pattern>GEHST DU *</pattern>
  <template>
     Nur wenn mich jemand auf seinem Laptop mitnimmt.
   </template>
</category>
```

This simple question answering mechanism does not require response generation in real time. Therefore, the choice between interjection-based and repetition-based answers to deal with an error needs to be determined in advance. ELIZA-like responses may be used to create repetition-based answers. An additional AIML processor can be introduced to retrieve the information about the error.

In the example below I call such a processor `<contained-error>`. The processor in the example returns either NOERROR, in this case the first response variant will be delivered to the user, or it returns ERROR, then the second variant will be delivered to the user. In this case, the response to user's utterance matching GEHST DU * may be produced by the following template:

```
<category>
  <pattern>GEHST DU *</pattern>
  <template>
    <think>
      <set name="contained-error-tmp">
        <contained-error><star/></contained-error>
      </set>
    </think>
    <condition name="contained-error-tmp">
      <li value="NOERROR">
      Nur wenn mich jemand auf seinem Laptop mitnimmt.</li>
      <li value="ERROR">
      Ich gehe <star/>, aber nur wenn mich jemand auf
      seinem Laptop mitnimmt.</li>
    </condition>
  </template>
</category>
```

The star symbol is used in the patterns to match any input that follows GEHST DU. The same part of the input is then repeated in the answer, if the answer template contains the `<star/>` tag. An embedded correction should repeat the *corrected* version of the input matched by the star symbol.

With regard to error coverage, standard AIML interpreters make use of barely simple replacements for input normalisation to make pattern recognition more stable. This is in general not sufficient for dealing with the variety of learner errors, but can be used to recognise special cases of spelling errors. String replacements are defined in a separate file and performed prior to match the input to the most likely pattern. The AIML interpreter knows all recognisable errors in advance and can re-use this information later for correction formats. Spell checkers can be used to recognise a larger number of errors. In this case, the information about performed changes in the input string needs to be stored and made available for the `<contained-error>` processor.

Integrated corrections usually contain emphasis of the corrected error. In this case, simple ELIZA-like repetitions are in general not sufficient, if the exact error location is not known. The information about the replaced strings can be used to highlight the replacement only on the string level. Additional NLP tools like spellchecker and parser need to be involved to deal with learner language and perform error recognition and highlight the corrected error.

A reasonable alternative has been discussed in Chapter 2 where a more sophisticated language understanding procedure was implemented in an artificial agent Sasha. Sasha was created to support learners of English as a second language and had two jobs. First, it had to produce a recast of the question if the question contained a morpho-syntactic error. Second, it had to generate a semantic response to the question (Petersen, 2010). However, the learner was supposed to ask only questions about pictures that were shown by the system, and many possible questions could be predicted in advance. Learner errors made the understanding of the questions, however, to a challenging task. Several backdoors were implemented to deal with troubles in learner language understanding in form of repair other-initiations produced by the agent.

9.4.3. Dealing with multiple target hypotheses

Different strategies can be used to deal with multiple target hypotheses. Three of them were identified in the dataset. I will therefore restrict the discussion in this section to the following three cases:

1. Test one of the hypotheses in form of an exposed correction.
2. Reduce the number of ZH by reasoning involving for instance information state and knowledge about the user.
3. Demonstrate uncertainty by accountings.

As already discussed in Section 6.3 and illustrated by Example 6.19, conversation partners of language learners have the same problems with interpretation of erroneous learners' utterances as annotators who have to error-annotate learner corpora. Both had to guess which target hypothesis corresponds to learners' intention. Example 6.19 showed that hypothesis testing is a possible strategy. If the first hypothesis tested turns out to be wrong, another one can be tested, or the topic can be changed.

However, there are strategies to narrow down the number of potential target hypotheses for a correction. For instance, only one target hypothesis is correct in Example 6.6 of Section 6.1.1 and in Figure 9.1, because the interlocutors know that only one weekend passed between the previous and the current conversation. In practice, this might be not feasible because it might be difficult or even impossible to determine in advance, what kind of knowledge is relevant for each case.

Accountings may be used to deal with multiple target hypotheses. If a correction is a paraphrase of the trouble source, many variants may exist. Then, a specific type of accountings may be used to express uncertainty:

$A_{uncertainty} = \{$"oder so was in der Art",$\epsilon\}$

A correction may then have the form:

$$\mathscr{C} = a_{excuse} \centerdot \neg f_l \centerdot X^* \centerdot f_r \centerdot f_l \centerdot x \centerdot f_r \centerdot e \centerdot a_{version} \centerdot e \tag{9.17}$$

A successful correction generation may be one of the criteria for the correction decision model. If the machine cannot decide, which of the target hypotheses is correct, it may use one of the strategies chosen by native speakers: initiate repair, let it pass, or simply smile.

9.5. Discussion and preliminary findings

In this chapter I presented and formalised local models for exposed and embedded corrections. I looked at correction models separately from the question of automatic error recognition. I then analysed and presented local correction models independently from the question if a correction of a specific format is appropriate at a particular point in the conversation. Although I analysed integrated corrections as part of exposed corrections in Chapter 5 of this work, it turned out to be more reasonable for computational modelling to group them with embedded corrections. Abstract models for the following types of corrections were presented:

Realised in a side-sequence These are exposed corrections of the following types:

1. On-the-fly
2. Simple with minimal or rich accountings
3. Contrasting-based with minimal or rich accountings
4. Based on repair other-initiations.

Incorporated into the next relevant turn These corrections include the following types:

1. Embedded corrections in answers to polar questions.
2. Integrated corrections in answers to polar questions.

Any of the formats for exposed corrections realised in a side-sequence in a dialogue can be used to correct any of the errors. Template-based responses can be formed according to the models of exposed corrections. The slots in the templates need to be filled in by the corrected version of the trouble source. Even with a technology as simple as chatbots, it may be sufficient enough to implement these correction formats.

Embedded and integrated corrections, however, need to repeat the constituent with the corrected error in the answer to the interpersonal trajectory, without a separate correction side-sequence. Simple ELIZA-like mechanisms like those used in chatbots may be exploited to generate repetition-based answers. However, error recognition tools need to be reliable in finding the correct target hypothesis.

To determine the minimal error constituent, tools for error recognition can be exploited. In Figure 9.1 I provided an example of the recognition of morpho-syntactic errors based on dependency parsing. Language understanding capabilities of the system may vary, depending on what exactly it is, for instance a chatbot or a text-based dialogue system. Different types of such systems were discussed in Section 2.1 from the perspective of learner language understanding, focusing on German as a communication language. Different levels of language understanding, however, influence the number and the types of learner errors that can be recognised. For instance, pattern-based language understanding in chatbots does not support recognition of syntactic errors (De Gasperis and Florio, 2012), as opposed to dialogue systems where parsers are used to find a mapping between user's input with deviations and a correct form in the target language (Petersen, 2010; Amaral et al., 2011).

With regard to the variance within correction formats, the following has been observed. The native speakers did not change their correction variants within one format frequently. Some of the native speakers preferred uppercase for emphasis, other users preferred quotes. However, native speakers preferred not to use the same correction format twice in a sequence. A small number of preferred correction variants for each format may provide the machine with a personal style. Using the full palette of available correction formats would provide the learners with examples of "doing correction" in chat, which would support learning by imitation (Aguado Padilla, 2002).

The selection of different types of correction formats may be related to the interaction history. In the beginning of the interaction, correction formats with rich accountings may be preferred. The accountings may do the additional job of negotiation, whether corrections are accepted by the user. Other types of accountings may help to make visible to the user that it is clear to the machine that corrections might be annoying. Less "intensive" types of corrections, such as simple corrections with minimal accountings are more appropriate when the user and the machine already established a sort of social closeness and the user perceives corrections as helpful. Corrections on-the-fly and embedded corrections may be used from the beginning of the interaction because they do not make the correcting to the matter of the talk. Summarised, there is maybe no dependency between the error and the type of exposed corrections, but there seems to be a dependency between the preferred correction formats and the state of social interaction. This will be taken into account in Chapter 10 where I make a step towards a decision model for corrections.

10. Towards a Decision Model for Corrections

Obviously, the ability to recognise that something went wrong, thus, error recognition, is the first condition for a correction. However, a correction is not a necessary consequence of an error. Various possible alternatives to a correction have been illustrated in Figure 7.1. Moreover, even if the recipient of the trouble talk decides to address the error in interaction, many variants to do a correction are possible, for instance exposed and embedded corrections, and then various formats within these two large categories are available. A computational decision model for corrections in a Conversation-for-Learning needs to select only one of all available alternative and take into account various factors in interaction with a learner on the way to such a decision.

Besides the presence of an error, other factors discussed in preceding chapters influence the presence of a correction. Many factors that might be relevant for this decision were discussed in Chapter 7. Some other factors become important first when computational modelling is required, as discussed in Chapter 9. Summarised, for every recognisable error, the following factors were found important for the decision whether a correction should be produced:

Conventions and agreements mainly determined by the speech exchange system and communication medium or negotiated explicitly. What counts as an error? Are corrections welcome?

Correction produce-ability As it was shown in preceding chapters, it is not always possible to generate all the correction types for each error. Is a target hypothesis projectable? Are multiple target hypotheses available?

Correction or clarification need Are there sequential consequences of non-correcting?

Error properties Error frequency and the number of errors in the turn. Can the error be implicitly corrected?

Interaction history How long does the machine "know" this user? In the first conversation, a different decision might be taken than in the 10th. Other preceding forms of orientation to linguistic identities and self-corrections may have different influence on the decision.

Correction history Has this error or this type of error already been corrected in the past?

Information about the user Is language learning relevant for profession or study? Are any activities planned by the user where language knowledge is important?

Preferences should *let-it-pass* be the preferred strategy or something else?

This chapter seeks to answer the question *how can features in interaction identified in the preceding chapters as influencing the corrections be operationalised for a decision model*? I will provide a basis for further investigations of factors relevant for the correction decision and their analysis with machine-learning-based algorithms. Because there are mutual dependencies among a large variety of features, machine learning approaches may deliver a more stable and more reliable results than a hand-crafted rule-based system.

As shown in Chapter 3 and specifically in Section 3.4.2, every of the participating native speakers and every of the learners were involved in conversation sequences where their linguistic identities have been made relevant. Every native speaker reacted in a very sensitive way to learners preferences and needs, but also to learners language proficiency and interaction management skills (independent from foreign language proficiency). Individual factors such as native speakers' profession and education but also experience in intercultural communication and instant messengers undoubtedly played a role in decisions to correct or to engage in role-play to practice for examinations. To provide an artificial agent with such consistent behaviour, I would need to give the agent a life story and a personality, and I would need to provide the agent with *a* decision model for corrections in order to make it able to behave like *a* language expert. One of them. From millions of possible models for millions of language experts. However, the local models for corrections and and explanations can be re-used for *other* models for corrections, and will give the agent a *different* interaction profile, but still one of a language expert.

I already put much attention to conventions, agreements, error concepts, target hypotheses and sequential features in preceding chapters. The findings will be taken into account in the attempt to create a decision model for corrections, which I will present in Section 10.4. Because much effort has already been put into error recognition (see Section 2.1.4 for references and discussion), I will assume for the purpose of this work that error recognition techniques discussed in Section 2.1.4 may be employed to detect all errors which count as "real" in chat and are worth being addressed to in a Conversation-for-Learning. To complete the decision model for corrections, I will first specifically focus on the role of the user model, interaction history including correction history and error properties for the final correction decision in this chapter. Section 10.1 will address the influence of the user model and knowledge about the user. The influence of the interaction history will be discussed in Section 10.2. The influence of various error properties such as error frequency, error type, error number and available correction formats will be explained in Section 10.3. I will discuss the findings in Section 10.5.

10.1. User Model

Because every correction is a contribution to social interaction with the user, and because corrections are dispreferred social actions, a model for prediction of user's acceptance of a correction may be helpful. The user may explicitly let the machine know that corrections are expected and perceived as helpful. It may be done in form of a request for correction or in form of an expression of thanks after a correction.

The machine can gather knowledge about the user focusing on the relevance of the target language for user's life, profession or study. If language learning is somehow important, then it is likely, that the user will benefit from corrections and will accept corrections in chat with the machine. Several user models can be provided to the system in the beginning, for instance, stereotype-based (Rich, 1979) covering potential user groups. The models can also be designed for potential users of the artificial agent using, for instance user-centered approach (Petrelli et al., 1999) or personas (Casas et al., 2008).

It cannot be predicted, who will be the final concrete user, therefore the machine will be initialised with a number of user models for potential user categories. Because there are mutual dependencies between the actual user behaviour and the interaction history including interaction process and machine's behaviour, individual-based user models (Pavel et al., 2015) may help to incrementally adjust initial models and adapt to user's needs.

10.1.1. Professional relevance of L2

Language students are used to be corrected by teachers in language classes. They also perceive corrections as helpful to improve, and sometimes the only way to learn. They are familiar with the situation of "being corrected", and probably therefore do not perceive corrections as face-threatening. On the other hand, if L2 knowledge is relevant for profession or study, every chance to improve language skills may motivate to get corrected more frequently. This information can be obtained from the user in the first dialogue, in a similar way as the participants of the study gathered information about their conversation partners.

Dialogue scripts and templates targeting information about professional relevance of L2 can be prepared even for a very simple language understanding technology such as chatbots. The gathered knowledge could trigger the activation of one of the user models.

Because language learning was professionally relevant for all learners in the dataset (and the same problem with the data appears in the majority of all publications in SLA, because their subjects are frequently language students from a university), I cannot make any observations about any other category of users. All other kinds of language learners need to be taken into account in the user modelling phase, for instance people who learn a foreign language just-for-fun or for family reasons. Unfortunately, I cannot make any conclusions about their behaviour or motivation based on my dataset. Therefore, I can observe only features for individual user models for learners who seek to become a foreign language professional. Nonetheless, other user categories may also choose similar strategies to signalise a preference for correction. Additional studies with other categories of participants will help to adjust the models to other user groups and may potentially lead to discovery of new relevant features for the correction decision model.

10.1.2. Learner's requests to correct and uptake

Marques-Schäfer (2013) reports that learners' explicit requests to correct their errors were important for the native speakers in the tutored sessions in their decision to correct. I did not find any such request in my corpus, as I mentioned earlier. This may be explained by the absence of any interaction participant who had a pre-assigned teacher-like role, like the tutors in Marques-Schäfer (2013)'s study. However, such requests may potentially appear in conversations between language learners and artificial conversation partners, and should be properly recognised.

Explicit requests to correct

In my dataset, learners' positive feedback after corrections of linguistic errors encouraged native speakers to continue with corrections (recall Example 3.17 - "es ist gut für mich, dass du meine Fehler korrigierst"). Learners turns in close sequential positions after the correction may contain information about learners' perception of corrections. A positive attitude towards corrections may be expressed, for instance, in words of thanks, explicit requests to correct more and evaluations of the correction activity ("es ist gut für mich..."). Summarised, learners can produce statements equivalent to explicit requests to correct in two places relatively to (potential) correction position:

1. Before the agent starts to correct.
2. Closely after a correction.

I did not find any negatively coloured post-correction work, however, it cannot be completely excluded in conversations with an artificial agent. The agent may need to recognise if there is a kind of post-correction interaction management in urns directly after a correction. If so, their polarity towards the correction will help to adapt the correction decision model for the individual user. Polarity recognition in short messages like Amazon comments and Twitter posts is an area where much research efforts have been put in, see for instance (Thelwall et al., 2010; Davidov et al., 2010; Becker et al., 2013) for English and (Evert et al., 2014) for German.

Implicit signalling and agents initiative

Other forms of learner's orientation to their linguistic identity may be *handled* by the artificial agent to an implicit request to correct. For instance, apology-based face-work may be responded by an encouragement, a positive assessment *or* by a suggestion to correct errors. Such a reaction would show the learner, that this apology-based social action was interpreted as a request for help.

Uptake signalises learner's acceptance of the correction. Correction uptake can be observed during conversations with language learners and included as a feature for the correction model decision. Learners' uptake has been intensively studied from the SLA perspective because it is seen as a measurement for learning progress (Smith, 2005; Panova and Lyster, 2002). Learners' uptake in long-term chat-based dyadic interaction needs to be studied separately from a different perspective in order to find patters related to correction decision models. The study can be performed based on my dataset and other similar corpora, if they become available.

Learners who participated the data collection reported in the retrospective interviews that they desired and expected to receive corrections of their errors more frequently. In order to proactively anticipate this expectations the agent may ask the users whether their linguistic errors should be corrected in conversation. This information can be gathered in a similar way as other kind of knowledge about the user required for user model initialisation, for instance, based on dialogue templates.

10.2. Interaction process and interaction history

As mentioned in Section 3.4.2, both the agent and the user have only a limited influence on their interaction profiles, and the existence of the linguistic identity of a language expert (or a more knowledgeable participant) implies the co-existence of a language novice (or a less knowledgeable participant). Hence, there are mutual dependencies between the user and the agent in their co-construction of the interaction (however, the term co-construction cannot be applied for human-machine interaction in the same way as it is applied in human-human interaction). Every participant of the data collection had her or his own "story", and each new user will have a new "story". However, several features from past interactions and past turns of the ongoing interaction look relevant for the correction decision. I discuss each of the identified features below. Features like social closeness may be relevant, too. Unfortunately, there is no operationalised model for that in CA that would allow to "measure" the grade of the social closeness, however, a few articles made a step towards this problem, for instance "Language of the closeness, language of the distance" (Koch and Oesterreicher, 1985). Social closeness and intimacy is an interactional achievement, as it has been reported in various studies, for instance (Rintel, 2015) and references therein. In particular, the absence of ceremonials and invitations typically index intimacy (Sacks, 1995). Existing models for relationships may also be used, for instance relationship model for relational agents (Bickmore, 2003) and the concept of convergence (Mitchell et al., 2012).

10.2.1. Initialisation of system's beliefs about the user

At some point of every conversation in my dataset, it came to the first, initial construction of linguistic identities of the participants, in form of face work, definition work or other activities described in Chapter 3. Previous research showed that in an equal-power speech exchange system like Conversation-for-Learning, construction of a less powerful linguistic identity of a language novice is normally initiated by the owner of this identity. As a consequence, the construction of a more powerful linguistic identity of a language expert is initiated later than the construction of the novice's identity, normally as a response to it. However, in my dataset a different result appears: native speakers may initiate the construction of their own identity as a language expert earlier than learners display a demand in it. After N04 learned from interactions with one of the learners, that error corrections are not only acceptable, but even desirable and helpful, he inferred that this might be valid for the other learner, too. N04 initiated construction of linguistic identities, including his own as a language expert, although this "first step" is normally reserved for the learner.

This observation makes reasonable to have two alternatives in the correction decision model:

1. The machine is allowed to initiate construction of linguistic identities with the initial assumption that the learner will find it at least acceptable.
2. The machine has to wait for the learner's initiation of the construction of linguistic identities, whatever form they may take. If the learner positions herself as a language novice, then the learner is likely to accept the machine as a language expert.

This decision for the initiation is one of the key criteria for system's initialisation and thus, for the production of the first exposed correction. Embedded corrections do not make correction to the interactional business, therefore, participants' linguistic identities are not oriented to in embedded corrections. However, the dataset provides evidence for participants' *noticing* of implicitly corrected errors (Example 3.16) and thus, noticing of differential language expertise without an explicit orienting to it.

10.2.2. Role of preceding forms of orientation to linguistic identities

Preceding types of participants orientation to their linguistic identities in talk may influence the correction decision positively or negatively. The dataset provides the evidence that other-corrections immediately follow learner's requests for explanations. Learners demonstrate in form of repair other-initiation that they prefer the native speaker as an information source about linguistic matters over a dictionary or a machine translation system.

In contrast, other-corrections of linguistic errors never immediately follow excuses for deficiencies in linguistic knowledge. If a learner produces negative self-assessments or emphasises own linguistic incompetence in a different way, immediately following error corrections would confirm this negative self-assessment and incompetence, and make positive evaluations and encouragement unbelievable. Post-correction face-work may help to deal with these issues, if a correction is absolutely necessary after excuses and negative self-assessments.

Machine learning approaches based on further similar datasets would help to find further mutual dependencies among various forms of orientation to linguistic identities in conversation.

10.2.3. Correction history and format selection

Every exposed correction for each pair of participants corrected each error only once. However, there are pairs of participants and pairs of embedded corrections for them, which correct the same error more than once. I also rarely found a sequence of two corrections where the same exposed correction format was applied. Native speakers preferred to provide correction using different variants. However, the better the participants knew each other, the lower restraints were between them in conversation, the less interactional work native speakers involved in exposed corrections. Time passed after the last correction, however, seems not to play a big role. I found cases where two exposed corrections were produced close one after the other for a pair where in total only a few corrections were found.

To gather this kind of information for each user, all corrections of all errors and all applied correction formats need to be stored. Because corrections do not occur frequently, this data structure is not expected to grow quickly. Information about learner's uptake might also be stored in the same data structure and used for the individual user model, as already noticed earlier in Section 10.1.2.

I observed in the dataset two distinct cases of correction format selection:

1. Sensitive to interaction history: formats that gave the learner less chances to focus on corrections were preferred for the first correction (embedded or on-the-fly). If it was necessary to make the correction to the interactional business, formats with rich accountings were selected (requests for permission to correct, justifications of correction).

2. Insensitive to interaction history: native speaker started to correct in the first meeting without any learner's orientation to own linguistic identity in form of an exposed correction of a morpho-syntactic error that occurred for the first time. In this case, the native speaker selected a simple delayed correction with rich accountings (request for permission and instructing) and backlinks ("PS:"). Thus, more interactional work

was required to deal with the face-threatening effect of the correction. Making the correction delayed makes the error less severe.

These two cases may be taken as a model for the systems' initialisation in according to the initial state of systems beliefs about user's needs, as explained in Section 10.2.1. A preference scale for initial correction may be needed, for instance:

1. Embedded correction should be preferred over exposed.
2. On-the-fly correction should be preferred over exposed with rich accountings if there is an uncertainty that corrections are welcome.
3. Repair-initiation-based correction formats should be preferred for correction of lexical errors that have sequential consequences.

In the continuation of the talk, corrections with less accountings may be more appropriate if there is some evidence for increase of social closeness between the user and the machine.

Because of many features with multiple mutual dependencies, it is easy for a human system designer to lose the overview over all the rules and their interdependencies. If new features are detected, all interdependencies have to be checked for consistency in a hand-crafted rule-based system. Machine learning approaches may deliver a more consistent model.

10.2.4. Learners' self-correction

Normally no correction was provided by native speakers after learners' self-corrections (self-initiated self-repair). It was not important if the self-correction targeted the errors which could be addressed by the other-correction. Therefore, the machine needs to deal with self-corrections, too. Even if a correction was successfully generated before it is presented to the user, the machine needs to check user's input again. If a self-correction is posted by the user, the machine should cancel the correction and continue the talk on the interpersonal trajectory. The self-correction must then be taken into account for the production of the semantic response.

Influence of the self-correction to the semantic response was subject of academic publications discussed in Chapter 2. Previous research focused mainly on self-corrections in speech recognition applications, specifically on same-turn self-corrections. As opposed to spoken interaction, such self-corrections are normally produced in chat *before* a message is posted. Nonetheless, transition-space self-corrections remain for the analysis in chat protocols. Due to specific characteristics of responsiveness in chat, the time slot where a self-correction can be expected depends on many factors, such as type of trouble source, length of the correction turn, length of the trouble-turn. For the machine, an approximation is needed how long is it reasonable to wait for user's self-correction. The dependencies in responsiveness values in the dataset are presented in Table 10.1.

The majority of all self-corrections occur in the first 30 seconds after the trouble turn (85%). The longest responsiveness values (15%, between 34 and 58 seconds) appeared

r	ctl	ttl	ttl_{av}	Trouble source	Correction type
1-4	1-3	10-66	33	typo: question mark	One symbol
5-11	2-31	14-56	29	typo, case, missing subject, word order	One word
14-20	3-22	27-176	103	typo, spelling, missing or wrong words	One word, two-word phrases, *ich meine*+correction
23-30	6-15	15-188	90	verb forms, gender, spelling, typos	rejection token+verb
34-58	7-55	55-481	222	gender in DET+ADJ, case, wrong word, missing *zu*	repeat utterance, repeat verb + object, 1 word

Table 10.1.: Responsiveness values: responsiveness r in seconds, correction turn length ctl, trouble turn length ttl, average trouble turn length $ttl_{av)}$ in symbols.

if the learner chose to repeat the whole turn or the error was hard to find because the trouble-turn was very long, longer than 300 symbols. However, long responsiveness was also found in two cases where the trouble turn was relatively short and the produced self-correction consisted of only one token. In these two cases the errors were morpho-syntactic (wrong case of a noun and wrong form of the auxiliary verb "sein"). In both cases it is only one missing letter which lead to the erroneous form, which probably were result of a typo, and it may have been unlikely for the learners that they made these errors. While longer responsiveness values can be predicted after longer turns, the length of learner's self-correction cannot be predicted from the small number of examples. However, the probability to receive a learner's self-correction after 30 seconds is quite low.

10.3. Error properties

Because embedded corrections do not focus on the errors in talk, they do not have the same face-threatening effect, and can be produced more frequently and repeatedly. In addition, embedded corrections have the function of documenting what the speaker is actually talking about (for instance, what exactly is the speaker answering to, if there was something in the question that may have caused misunderstandings or might have been imprecise from the perspective of the implicitly correcting speaker). Therefore, embedded corrections seem to be preferred in a Conversation-for-Learning and should be preferred by an artificial agent. However, only 18,7% of all potentially produceable embedded corrections were factually produced by native speakers. A qualitative analysis of embedded corrections in answers to polar questions shows that embedded corrections appear most of all if there are sequential consequences, a possible difference in understanding of an answer without a correction. If many readings of learners utterance are possible (not necessarily caused by an error), an embedded correction helps to channel further development of the dialogue. Thus, an embedded correction is also a strategy to deal with multiple target hypotheses.

Domain knowledge is required to detect multiple target hypotheses in order to detect the preposition error like in Example 10.1. The machine needs to know that the teams play *against* each other and the players of one team *with* each other if they play football.

Example 10.1. Embedded correction clarifies the meaning

220	20:09:32	L02	Und mit welchem Land spielt Deutschland nächstes Mal? *And with [* error: wrong meaning] which country does Germany play the next time?*
221	20:09:56	N01	Gegen Niederlande am Mittwoch. *Against the Netherlands on Wednesday.*

All types of errors (orthography, morpho-syntax and lexicon) may imply multiple target hypotheses not only in understanding of the utterance, but also in correction detection. However, only one target hypothesis was normally selected by native speakers in corrections. Machines certainty in target hypothesis selection may be included as a feature into the correction decision model. There is no need to produce as many corrections as possible, but if a correction is produced, it has to be accurate. Otherwise, the machine will not be accepted as a language expert. The user will not believe machines opinion, if "mis-corrections" are produced. I discussed three strategies that may be used to deal with multiple target hypotheses: hypothesis test, reasoning and uncertainty demonstration. If still in the end too many variants exist, a correction should probably be omitted.

As discussed in Chapter 7, native speakers see error frequencies as a criterium for correction (Marques-Schäfer, 2013). However, there are two problems with this criterium. On one hand, there is a difference between what people say they do and what they actually do. I observed in my dataset that native speakers did not wait until an error starts to occur "frequently", they corrected the first occurrence, thus there is a contradiction. On the other hand, participants intuition of error frequency may be very subjective and cannot be expected to reflect the reality. Specifically, native speakers in chat do not have the same capability to track and to catalogue all the errors like a machine. In addition, it needs to be specified more precisely, how frequent is frequent. A number is expected, not an intuition. Moreover, if an error occurs too frequently, a different strategy of dealing with it may be preferred, such as let it pass. For these reasons, error frequency may be included into correction decision model as a weak feature, and concrete frequencies need to be specified for concrete decisions. To track error frequency, a special data structure or a database is required, where all automatically detectable learner errors need to be stored.

Every correction in the dataset addressed only one error. If too many errors were in learner's turn, it was difficult for the native speakers to provide a correction, and therefore, a different strategy was selected. Hence, the number of errors in learners' turn needs to be taken into account in the correction decision model.

Different error types are handled differently in chat-based Conversations-for-Learning. I found the following regularities:

- Typos were classified as spelling errors and were corrected either if several cases of the same error occurred in a small window and the reading of the word was the same as with the correct spelling OR if several cases of the same error occurred and it was a "hard" case where native speakers make mistakes, too. "Hard" cases can be part of the knowledge base. A phonetic index such as Kölner Phonetik (Postel, 1969; von Reth and Schek, 1977) for German or Soundex for English (Russell and Odell, 1918), can be used to find the other cases.
- Morpho-syntactic errors have been corrected even after the first occurrence, however, errors occurring constantly have not been addressed. Dependencies between the frequency of morpho-syntactic errors and their corrections may exist.
- Lexical errors have been corrected even after the first occurrence, because wrong lexical choice and lexical creations cause misunderstandings (= sequential consequences) or even make comprehension impossible.

Learner language understanding is specifically challenging because learners produce deviations from language standard, and because statistical NLP tools have been normally trained on native speaker language samples. Additional re-training is required to prepare NLP tools for dealing with learner language. These additionally trained tools help to detect errors and to guess the intended utterance. Some error types lead to machines inability to understand the intended meaning. Petersen (2010) deals with this problem quite radically: every misspelled lexeme triggers a repair initiation and every unsuccessful parsing triggers a repair initiation. However, chat conventions allow many variations in orthography and do not allow to much focusing on these variations, thus, Petersen (2010)'s strategy cannot be adopted. Other types of errors, such as missing verb, may lead to problems with parsing. Example 10.2 illustrates one such case: the answer of N01 in turn 280 becomes an embedded correction because of the verb insertion.

Example 10.2. L01 omits a verb in her question, N01 inserts a verb in his answer.

277	21:08:00	L01	Wie deine Arbeit? *How [* error: missing verb] your work?*
280	21:09:57	N01	Meine Arbeit war ganz ok. [simplified] *My work was totally ok.*

Other verbs or verb forms could potentially be placed instead of *war* in learners question, for instance *ist* (is), *läuft* (runs), *verkauft sich* (is being sold) and *endet* (ends). Each of them leads to a different meaning of the question and therefore, to a different space for potential responses. The embedded correction in the response in turn 280 clearly documents, which of the possible verbs N01 actually took for the interpretation of the question. Hence, corrections are a device to deal with sequential consequences of an error. If the machine can detect similar cases and can produce a correction, it should be done.

Summarised, the following error properties need to be taken into account in a correction decision model:

1. Correctability of the error by an embedded correction;
2. The number of target hypotheses;
3. Error frequency (weak feature);
4. Error number in learner's turn.
5. Error type;
6. Sequential consequences.

Error tracking data (e.g. number and types) may then be taken into account for the individual user model. If errors occur only rarely, a strategy for increasing the language complexity may be needed to make the conversation more challenging for the learner and maybe trigger repair other-initiations, like it was observed in conversations with N01. There are also mutual dependencies with correction history. Every corrected error for a learner was never corrected again explicitly, but embedded corrections were observed several times for the same error produced by the same learner.

Further comparative analysis of differences between repetition-based answers to polar questions and other answer forms may help to find further features that should be included in the correction decision model. Machine learning approaches may then be used to detect mutual dependencies among the features. These features should then be included into the decision model.

10.4. Correction decision model

From what has been discussed in the preceding sections of this chapter, the following factors need to be taken into account in a correction decision model:

User model the system may be initialised with for instance a number of stereotypes or personas, and needs to be incrementally adapted to the user forming an individual user model for each user. The initial selection of the user model may be done according to the relevant knowledge about the user

1. Knowledge about the user: purpose of language learning, experience with language classroom or self-directed learning etc.
2. Learner's requests to correct (explicit and implicit) and learner's uptake.

Because the individual user model needs to be incrementally updated according to what happened in interaction, there are mutual dependencies between interaction history and user model.

Interaction process and interaction history a set of rules for the beginning of the interaction and for the continuation of the interaction is required.

1. Initial state of beliefs. Who is allowed to initiate orienting to linguistic identities, only the user or both user and machine?
2. Dependencies from preceding forms of orientation to linguistic identities. Some of them negatively influence the occurrence of corrections, some of them provide opportunities for correction.
3. Correction history plays a role in correction format selection. Repetitiveness should be avoided: do not explicitly correct the same error twice, use different formats for subsequent corrections. Embedded corrections of the same error are allowed.
4. A data-driven model for social closeness may be helpful but needs to be developed first. Relationships models can be taken as a start.
5. Learner's self-correction leads to correction cancelation even if the correction decision model decides to generate a correction.

All the important features need to be documented and will be used to update the user model (e.g. detected errors and corrected errors).

Error properties are relevant for the correction decision:

1. Frequency may be a factor but needs to be specified more exactly or learned from the data.
2. Number: many errors in a turn make it too difficult to correct.
3. Type: different rules for different error types.
4. Sequential consequences.

Information about the recognised errors may be useful for the user model and will influence the interaction process indirectly.

A set of features can be defined and observed for each discussed factor. A machine learning model for each of the feature sets may help to find proper weights for each of the factors and to calculate dependencies. Each of the parts may generate own decision based on internal factors using, for instance, a decision three generation method, such as C4.5 (Quinlan, 2014). The results of the separate decisions may be composed to a feature vector. Another machine learning model will need to generate a (pre-)final decision from this vector.

Each of the values in the vector may have its own weight which, in turn, needs to be estimated from training data. Mutual dependencies between the features must be covered by the machine learning model. Finally, if the user still does not produce a self-correction, the machine may correct learner's linguistic error using a specific correction format, which needs to be selected as part of the decision model. Figure 10.1 illustrates the feature sets and their interdependencies on a high level.

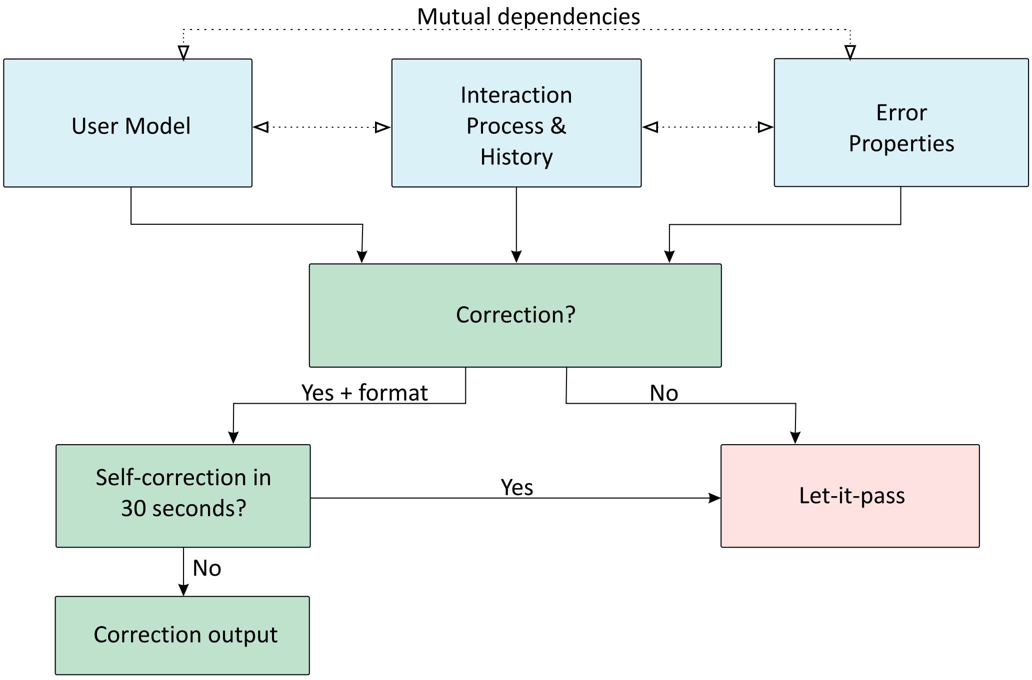


Figure 10.1.: Schema of the correction decision model. Blue blocks – the feature sets, green blocks – correction path, red block – path without correction; black arrows - information flow in the decision model, dotted arrows – mutual dependencies.

10.5. Preliminary findings and discussion

In this chapter I made the first attempt to formulate a computational decision model for corrections. The features relevant for this decision which I identified in preceding chapters were analysed from the perspective of the decision modelling. I discussed the role of each of the features and listed several mutual dependencies among them. The presence of an error does not necessarily trigger a correction. Other factors such as user professional interests, user's explicit requests to correct, the grade of the social closeness, preceding corrections, other forms of orientation to linguistic identities, eventual learner's self-corrections and different error properties may be crucial for this decision. Each of these factors can be formulated as a feature for a feature-based decision model. I categorised the features into three groups:

1. Related to user modelling;
2. Related to interaction process and interaction history;

3. Related to error properties.

Each of the features in each of the groups may be more or less important for the final decision, however, more examples are required to generate a reliable machine learning model and to estimate the importance (weight) of each of them. The final decision of the model should be not just yes or no, but a specific correction format that will be then presented to the user. However, responsiveness in chat needs to be considered before the output is presented to the user, because the user may produce a self-correction. Hence, a human-like responsiveness in chat may be an advantage. I made an attempt to find patterns in responsiveness in chat in an earlier article (Danilava et al., 2013b). However, there are many unobservable and uncontrollable elements influencing responsiveness values, such as parallel activities of the users and network connection values. A different, more detailed recording of the chat production process is required in order to make reliable conclusions about the patterns in responsiveness.

The dataset used for this work was created with participants for whom language learning has a high professional relevance: they all were DaF students and were going to work as DaF teachers and translators. All findings and conclusions made in this work can be only applied to this group of potential users. Further datasets are required to cover different groups of language learners, such as people who learn additional foreign languages just as a leisure activity without any professional interests, and people who want to go for vacation to a foreign country and need conversational skills in the language of the country.

An additional study focused on operationalisation of the social closeness in interaction may help to find a computational model which would allow to measure the grade of the social closeness, and may be included into the decision model for corrections. Because every correction is a social action, this model may be crucial for the correction decision. However, the data collection used for this work was not designed to contain this kind of information, therefore, participants perception of the social closeness is not explicitly reflected neither in the questionnaires nor in the retrospective interviews, and cannot be connected to the chat protocols. An additional data collection and a study based on it may help to close this gap.

Error properties reported as important for the correction decision by the tutors in the study by Marques-Schäfer (2013), such as error frequency, can be hardly operationalised for the computational model of corrections without further investigations. A important question is, how many errors the learners actually made and how frequent these errors actually have been corrected in that dataset. If an error occurs too frequently, it is rather ignored. Even the first occurrence of errors has been corrected in my dataset. In contrast, other error properties like error type, number of errors in the turn and sequential consequences seem to be more important for the correction decision. For instance, spelling errors are normally not addressed in chat because of a high probability of mis-typing due to high typing pace and conscious deviations in orthography, as discussed earlier in Chapter 7.1.2. However, if it is likely that a deviation in typing is caused by lack of knowledge, even spelling errors may be addressed in a chat-based Conversation-for-Learning.

A stronger connection between the chat protocols and native speaker's motivation to correct may be achieved by explicit focusing on error corrections in retrospective interviews, as it was done in other studies for which the relation between the dialogue data and participants' perception was of interest, for instance (Mitchell et al., 2012). Because error corrections were not the specific focus of the work in the beginning, only a few insights can be made into this issue based on the existing dataset. Focusing on corrections in questionnaires, however, is not desirable because questions about them during the data collection would influence the behaviour of the chat participants.

As noted by Schegloff (1993), "... in examining large amounts of data, we are studying *multiples or aggregates of single instances*". From the relatively small number of correction examples in the dataset (22 exposed and 37 embedded of approximately 4800 messages), it was possible to identify important features for correction decision and to propose a correction decision model. Participants of the chats decided to correct not based on any statistical metrics, but when it was relevant. I made here an attempt to express the relevance as a set of features. The importance of each of them can be expressed as weights that need to be estimated from more data or specified manually for the small number of examples.

Several issues need to be solved for a practical implementation of the model. First of all, a quantification of mutual dependencies and calculation of weights of each feature need to be approached. Because still not enough examples are available for machine learning, manual calculations may be approached by creating of linguistic variables and assigning fuzzy numbers to them (Zadeh, 1983). Hence, the model seeks to be seen as close to the rules of the speech exchange system of a Conversation-for-Learning. A quantitative study of dependencies among the identified features and, maybe, finding new features, would not decline the results of this qualitative study, but rather help researchers to deal with complexity and to preserve consistency within a complex system of rules and mutual dependencies.

Part IV.

Results

11. Findings and Discussion

I emphasised the multidisciplinary nature of communicative ICALL research in the first chapter and throughout this work. Findings in Software Engineering, Second Language Acquisition (SLA) and Natural Language Processing (NLP), Psychology, Linguistics, Sociology, Pedagogy, Computer-Mediated Communication (CMC), Foreign Language Teaching and Human-Computer Interaction (HCI) already established a well-founded basis to create meaningful and useful ICALL applications. I sought to show in this work how communicative ICALL would benefit from including Conversation Analysis (CA), specifically Conversation Analysis for Second Language Acquisition (CA-for-SLA) into the circle of related disciplines.

Inspired by CA-driven research on native/non-native speaker computer-mediated communication, I turned my attention to models for conversational agents in communicative ICALL in roles other than teachers or tutors. Starting with the idea of creating a machine that behaves "like a language expert", I defined two research objectives:

I Find and describe interactional practices in native/non-native speaker chat-based Conversation-for-Learning where chat participants orient to their linguistic identities of language experts and language novices.

II Create computational models of those practices and analyse technical requirements and limitations to implement the resulting models in a communicative ICALL application.

To get a first intuition about the complexity of the problem and to work toward a functional specification, I applied CA methods in my longitudinal study of instant messaging dialogues between German native speakers and advanced learners of German as a foreign language (DaF). I discussed the applicability of CA methods to the study of CMC and SLA with support of academic literature in Chapters 1 and 2. While CA methods have been already widely used to find patterns for human-robot interaction and to evaluate interaction between people and robots, only selected topics from CA were successfully taken up to approach challenges in dialogue design and evaluation with disembodied conversational agents, such as turn-taking and repair. The state of the art was discussed in Section 2.3. Despite all efforts, I could not find any academic publication about CA-informed communicative ICALL studies at the time of writing this dissertation. Therefore, this work makes a step towards creating new and strengthen existing multidisciplinary connections between Conversation Analysis, AI and communicative ICALL.

The research presented in this dissertation was approached in three phases: unmotivated looking at data from longitudinal native/non-native speaker interactions, detailed micro-analysis of three types of sequences in the data and computational modelling of patterns found in these sequences. During the first phase of "unmotivated looking" at the data, I found various types of sub-dialogues where language learners position themselves as language novices at the same moment creating the novice-expert dichotomy and putting the native speakers in the role of language experts. Theoretical and methodological framework for this analysis was discussed in Sections 1.2 and 2.2.4. The first insights from this research phase have been presented in (Danilava et al., 2013a). I discuss the outcome from the phase of unmotivated looking in Section 11.1.

Dealing with repair in dialogues with users is seen as a problem in speech recognition research. User's same-turn and transition space self-initiated self-repair is problematic for speech recognition applications (Grote et al., 1997; Purver and Hough, 2014). Same-turn self-corrections are not present in chat logs because they are completed by the participants *before* the messages have been posted. Transition space self-corrections, in contrast, are recorded in the chat logs, because they are produced immediately after the trouble-turns by the same speakers. This type of self-corrections produced by learners have impact on the correction decision model and were discussed in Section 10.2.4. Second position repair initiations have been handled in AI for the situations when speech recognition was successful but the next steps were still unclear for the machine, for instance with a robot (Kruijff et al., 2008). This corresponds to other-initiated self-repair when the user is the trouble speaker and the machine needs clarification. Such types of repair (with non-native speakers as trouble speakers) were identified in the dataset in the function of error corrections. Native speakers corrected learners' errors employing OISR-based correction formats.

Till present, researchers mostly assumed that the user understands everything what the machine says, because human users are clearly superior in their ability to understand human languages. Only one exception which I found in the literature is the assumption that the user may have troubles with understanding caused by, for instance noisy environment or machine's troubles with understanding and therefore initiate repair (Gehle et al., 2014). This assumption is not completely valid in communication with language learners. Language learners may need a clarification of the meaning of some lexical material or grammatical constructions because they communicate using a language that they have not yet fully mastered. Subsequently, the machine will need to produce a relevant explanation. This explanation may differ in format and linguistic resources as compared to repairs dealing with for instance environmental issues. Hence, other-initiated self-repair when the machine is the trouble-speaker was put in focus of this dissertation. I discuss the results and the contribution of this work on other-initiated self-repair when the other (the native speaker or a conversational agent) is the trouble-speaker in Section 11.2.

One of the ways to orient to linguistic identities in talk are corrections of linguistic errors. Other-corrections of linguistic errors are recognised in SLA research as important tool to support learning. Teacher in a language classroom and classifications of corrective feedback obtained from classroom data (Lyster et al., 2013) were taken as a model for corrections in communicative ICALL applications (Petersen, 2010). However, these models are not transferrable into Conversations-for-Learning because of differences in the speech exchange system, as discussed in Section 2.2.1. In addition, different sets of interactional resources are available in text-based chat and face-to-face interaction in a traditional language classroom. I approached the problem of computational modelling of other-corrections by a detailed analysis of correction formats. This resulted in local models of error corrections and will be discussed in Section 11.3.

Not all error are corrected in a language classroom, and not all errors are corrected by ICALL applications that play a role of an artificial tutor or teacher (Amaral and Meurers, 2011). Even less error corrections can be found in informal communication with language learners. Factors relevant for a decision whether a correction should be produced have been analysed in empirical studies of native/non-native speaker text chat communication in (Marques-Schäfer, 2013). The tutors in the study by Marques-Schäfer (2013) and the ICALL system in the work by Amaral and Meurers (2011) have in common that they own the role of a language expert and have the right to correct errors because of this role. In contrast, in an ordinary conversation and in a Conversation-for-Learning as defined by Kasper (2004), the linguistic identity of a language expert may but does not have to emerge as a product of the interaction process. A good decision model for corrections needs to determine when error correction is appropriate. In this work I made a first attempt to specify a decision model for corrections in a Conversation-for-Learning based on observable features in interaction. I discuss the results in Section 11.4.

11.1. Participants' orientation to linguistic identities

Participants orient to their linguistic identities chat in different sequences of talk. The following types of orientations can be listed:

1. Face-work and evaluation:
   a) Learners make "self-deprecating remarks" toward their language knowledge (*My German is not so good*). Native speakers respond with encouragement and positive evaluation (*You are doing a good job*).
   b) Learners excuse for their "bad" language knowledge (*Sorry for my all little mistakes*). Native speakers respond with encouragement and evaluation (*For now, I did not see any mistake*).
   c) Express anxiety toward potential difficulties in comprehension (*I hope I understand everything, otherwise you will have to explain*). Native speakers position themselves as willing to help and provide responses designed to reduce anxiety.
   d) Native speakers evaluate learner's specific achievements in form of compliments and praise.

2. Meta-talk about learning and collaborative learning:
    a) Comparison of grammar of shared languages.
    b) Discussion of (potential) difficulties in a particular language.
    c) Discussion of potential exam topics for exam preparation.
    d) Role play to rehearse exams.
3. Repair with linguistic trouble source.

In addition, participants may present themselves as members of the social category of language experts throughout the entire communication. The forms of presentation include participants' choice of linguistic resources such as orthography, vocabulary and syntax. Similar types of orientation have been described in previous publications (Kasper, 2004; Hosoda, 2006; Tudini, 2010; Dings, 2012; Vandergriff, 2013).

Meta-talk about language learning as form of orientation to linguistic identities was also found by Dings (2012). I did not find any leaner error correction initiated by non-native speakers, neither could I observe changes in error correction which have been reported by Dings (2012) who observed changes from more form-focused to more meaning-focused repairs. This may be explained by the study duration and the number of corrections. The present study was much shorter than the study described by Dings (2012). It took several weeks as opposed to one year in (Dings, 2012). Some of the native speakers in my corpus corrected only once, therefore changes of correction forms are not observable. However, I found changes in correction formats from more implicit and more polite to more explicit but with less accountings. Native speakers tend to reduce the interaction management required for corrections with rich accountings when they become more familiar with the non-native speakers. This, in turn, confirms the dependencies between the selection of correction formats and social proximity.

Vandergriff (2013) points out that "the most prevalent way of indexing identity is by evaluating something."(p. 395). She found that stance-taking in native/non-native speaker chat dialogues may have the form of accountings (learners critically evaluate their linguistic competences), explicit labelling of interactional roles, evaluations and taking certain interactional structures (typographic markers, lexical items). The orientations to linguistic identities in Vandergriff (2013)'s data were primarily invoked by non-native speakers. Native speakers, in contrast, made differential language expertise relevant only in response to non-native speakers "self-deprecating remarks". With regard to this, Vandergriff (2013)'s findings go in line with those reported by Tudini (2010) for chat and Kasper (2004) for face-to-face interactions. My research completes these findings in the following way.

- Native speakers oriented to differential language expertise in form of compliments where they took a positive stance toward learner's linguistic knowledge (praise for extraordinary knowledge of language and literature).
- Learners made "self-deprecating remarks" primarily in the beginning of the interaction (first 10 minutes of the first session).

- Negative assessment of own L2 proficiency in the continuation of the interaction was usually triggered by other troubles with production or comprehension.
- There are huge differences in pair-by-pair analysis. While pairs of participants involving N02 and N03 only rarely engaged in teacher-student-like sub-dialogues of linguistic repair, N01 and N04 did it more frequently. In dialogues between N04 and L08, teacher-student-like sequences of linguistic repair can be found more frequently than the other two types of orientation to linguistic identities.

However, my data do not confirm the role of the interaction managers dedicated to native speakers described by Vandergriff (2013). Non-native speakers showed themselves as equally competent in initiating and closing sequences as native speakers.

While the first two classes of sub-dialogues can be transferred to a communicative ICALL application in a relatively simple way by creating small dialogue scripts and using existing task-based dialogue models, the third class of sub-dialogues requires new models for dealing with different kind of trouble in conversations with learners. By "relatively simple" I mean that local models of sub-dialogues belonging to these two classes already exist or can be covered even by pattern-based language understanding techniques such as Artificial Intelligence Markup Language (AIML) (Wallace, 2003). Recognition of user-initiated face-work as face-work (e.g. distinguishing between excuse-based face work and real excuses) and a decision when meta-talk about learning and collaborative learning are appropriate remain a challenge.

11.2. Dealing with troubles in comprehension

I analysed a collection of sequences of other-initiated self-repair when the native speaker is the trouble-speaker in Chapter 4 to create an empirical basis for computational modelling of this repair type. The resulting computational models and their applicability were discussed in Chapter 8. This section synthesises the results with regard to the following two research questions:

RQ1 Which interactional resources do language learners use in a chat-based Conversation-for-Learning with native speakers to initiate repair in order to deal with troubles in comprehension and how do native speakers deal with these repair initiations?

RQ2 How can other-initiated self-repair when the machine is the trouble-speaker be handled in a chat-based communicative ICALL system?

I specifically looked at interactional resources which are used to tell the trouble-speaker *that* there is a problem with a previous utterance and to show *what kind of* trouble occurs. In this way, this work is close to studies described by Egbert (2009) for German oral naturally occurring interaction and a cross-linguistic study described by Dingemanse et al. (2014). I discuss the results and the contribution of the work to CALL and ICALL by a comparison of the findings of this study with studies of repair in native/non-native speaker chat (Tudini, 2010; Marques-Schäfer, 2013). The findings and the contribution are discussed in Section 11.2.1. The applicability of the models and the contribution of this work to communicative ICALL and NLP in a broader sense are discussed in Section 11.2.2.

11.2.1. Repair initiation and resolution in chat with language learners

In my analysis of repair initiation formats, I distinguish between resources used to reference the trouble source, and resources used to signal trouble in comprehension. This was a necessary abstraction to prepare a basis for computational recognition of repair initiations and trouble source extraction. I will argue that the selection of a specific format for a repair initiation may be influenced by the type of trouble, communication medium, turn adjacency in chat and learner's level of language proficiency. Repair carry-outs depend on the repair initiations in terms of format, referencing trouble source, type of information required to solve the problem and interactional resources that may be employed for that. The information provided in the repair projects the expectations the trouble-speaker toward what is known to the repair initiating speaker.

I focused in my work only on repair initiations with linguistic trouble source in text chat-based dialogues The following formats of repair initiations produced by the learners to point to a linguistic trouble source were found:

1. Open-class repair initiations realised through symbolic and lexical means, for instance *unklar* (unclear), *ich verstehe nicht* (I don't understand) or simply *???*.
2. Demonstratives-based repair initiations, such as *Was bedeutet das?* (What does it mean?).
3. Repetition-based repair initiations.

I chose to use the concept of an *action ladder* to start a discussion about the repair initiation formats in text chat. Repair initiations signal that there is a problem with a previous utterance, and they allow to understand what kind of trouble the recipient of the trouble-talk may have with it. Dingemanse et al. (2014) describe the problem space building on previous work by Austin (1962) and Clark (1996) in form of a so called *action ladder*.

Level	Speaker A's actions	Addressee B's actions
4.	A is proposing joint project w to B	B is considering A's proposal of w
3.	A is signalling that p for B	B is recognising that p from A
2.	A is presenting signal s to B	B is identifying signal s from A
1.	A is executing behaviour t for B	B is attending to behaviour t from A

Table 11.1.: The Austin/Clark action ladder helps to understand the problem space. Adopted from (Dingemanse et al., 2014, p. 8) .

In the action ladder presented in Table 11.1, higher levels of comprehension can only be reached if the lower levels are passed: one needs to attend to speaker's talk, correctly recognise speaker's words, find an interpretation of speaker's words and recognise the intended social action. As Dingemanse et al. (2014) note, "all four levels are involved in building mutual understanding, and each of them can be a locus of trouble" (p. 9). Moreover, the levels of troubles determine the selection of formats for repair initiation.

> Repetition is a commonly used device in the other-initiation of repair, but if something was imperfectly produced by A or not attended to by B (level 1), repetition-based formats will not be available. Conversely, the format selected by B can be inspected by A for its downward evidence. An interjection like *huh?* entails at least that some expressive behaviour was perceived (level 1), but not much more than that, and therefore indicates that there was likely a low-level problem. A question word like *who?* entails not only that some words were perceived, but also that they were identified by B as a person reference, and therefore indicates that the problem likely lies at the level of signalling and recognition. (Dingemanse et al., 2014, p. 9)

While a repetition of the trouble turn is an appropriate repair proper after an open-class repair initiation in oral talk (the problem is located on level 1), I have not found any example of such repairs in the chat protocols. Probably for the same reasons, repair initiation formats based on question words are not present in the dataset. However, I found *WHAT?s* in chat to express surprise. This might be explained by the influence of the medium giving the possibility to re-read and making repetitions to repair mis-hearings unnecessary. Similarly, the formulaic German *bitte?* or *wie bitte?* does not occur in the dataset, because the relevant next action after this repair initiation is a repetition of the trouble-turn. Repair initiation formats based on *Wie + repetition* described by Egbert (2009) were not found in my dataset, either. This may be explained by language classes not covering such repair initiations, which in turn, can be only acquired in interaction with native speakers. Sometimes language learners fail in producing an appropriate repair initiation. This may be caused by, for instance, their inability to identify the unit where the trouble source is located and selecting a wrong reference to it (Example 4.7). This, in turn, may lead to troubles in trouble source identification by the native speaker and, consequently, in repair carry-out. Hence, leaner's interactional and linguistic competence influence the selection of a repair initiation format and its successful recognition.

In contrast to oral communication analysed in (Dingemanse et al., 2014; Egbert, 2009), repetition-based repair initiation formats obscured that the source of trouble was on a lower level than normally required in oral talk to produce a repetition-based repair initiation. Namely, in a repetition-based repair initiation the repeated trouble-source was repeated with an error: *Frage machen* instead of *Fliege machen*. The original idiom used by the native speaker seems to be misread, although text chat allows to re-read all previous turns. There are for sure explanations in psycho-linguistics or cognitive linguistics, how such errors may occur. For the study of computer-mediated communication it might be important *that* problems on the lowest level of attention may occur in text chat, too, and that they can be detected through other means than in oral communication.

Another difference as compared to Dingemanse et al. (2014)'s and Egbert (2009)'s results is that candidate understandings are formatted differently in repair initiations with linguistic trouble source. Meaning check is a typical repair initiation format based on a candidate understanding: the trouble source is repeated (usually left-hand side), followed by a comparison token and the candidate understanding is presented (usually right-hand side). Candidate understandings are usually formatted as polar questions which require a confirmation or a disconfirmation of the hypothesis. *You mean*-based candidate understandings were produced by learners in my dataset only as an additional resource for repair initiation, which follows another, less specific repair initiation (Example 4.3). Candidate understandings for linguistic trouble sources seem to require additional resources in oder to mark the trouble source as a *linguistic* trouble source.

Open class repair initiations such as lexicalised equivalent of an oral *hä?* are not used by learners, but only by natives. This may obscure learner's lack of familiarity with oral open-class repair initiations in German, and therefore, the level of communicative competence. However, learners used multiple question marks *???* to initiate unspecific repair. The majority of all repair other-initiations produced by the learners are repetition-based. These repair initiations always entailed repair resolutions presenting explanations of the meaning of the trouble source. Dingemanse et al. (2014) explain a similar phenomena in native speaker data as follows:

> A difference in the type of repair solution provided in response to a repeat-formatted repair initiation is not always directly linked to a difference in formatting, but may also be linked to expectations about what is known. [...] Partial repetitions of terms clearly known to both speakers never result in clarifications of the terms, but are treated as taking an epistemic position that calls for another type of response, for instance a justification. (p. 24)

In conversations with language learners, not a conceptual clarification of the terms is required, but an explanation of some lexical material in a language-to-be-learned. This need may be satisfied by paraphrasing, synonyms, examples, and translations. The latter are rather untypical for native-speaker-only talk.

Further, I found differences in repair initiations formats caused by the timing of repair initiations relatively to the trouble source. Although all repair initiations with linguistic trouble source correspond to second-position repair initiations, they may occur immediately after the trouble-source turn in the timeline (adjacent position) or with one or more turns after the trouble source (virtual adjacency (Tudini, 2010)). Virtual adjacency provides a constraint for open-class repair initiations. A more explicit reference to the trouble source is needed if a repair initiation does not immediately follow the trouble turn.

With regard to interactional resources used to signal troubles in comprehension, question mark is an important and effective device, it is involved in the majority of repair initiations in chat. Dingemanse et al. (2014) explain it as follows:

> At the most general level, questions are next-speaker selection devices (Moerman & Sacks 1988). They are well-fitted to the other-initiation of repair because they put the ball in the court of the trouble-source producer. (p. 21)

Besides the practice of questioning, declarative utterances labelling the preceding turn explicitly as unclear were found.

The types of trouble sources addressed in sequences of other-initiated self-repair with linguistic trouble source were classified by length and unit integrity:

1. One unit: a single word, an abbreviation, an idiomatic expression or an utterance.
2. Copy-paste of a part of the trouble-turn containing the trouble source regardless of unit boundaries.

Copies of turn parts regardless of unit boundaries may imply that learners sometimes are not able to identify the problematic unit and copy a random part of the native speaker's turn which contains the trouble source. Such repair initiations may be placed between open class and restricted class repair initiations, because they restrict the search space, however do not completely specify the problem.

Repair carry-outs contain references to the trouble source. The selection of a proper referring expression depends on the reference to the trouble source in the repair initiation, type of the trouble source and timing of the repair carry-out. For instance, abbreviations are usually repeated in both repair initiation and repair carry-out.

Marques-Schäfer (2013, Ch. 8) does not explicitly analyse repair sequences in chat, but she classifies *questions with linguistic matter* (orig.: *sprachbezogene Fragen*). She classifies question content as form-related and meaning-related. She distinguishes between triggers for questions which are related to the interaction and which are not related to the interaction. Responses to questions with linguistic matters are classified as zero-response, direct response, translation, synonym, paraphrasing, example and hyperlink to a web page. With regard to this classification, repair other-initiations in response to NS's turn with linguistic trouble source in my dataset correspond to questions related to interaction. All of question-based repair-initiations were meaning-related. Form-focused repair initiations are not present in my dataset. This may be explained by the absence of a pre-assigned role of a tutor who is responsible for all linguistic issues and by the learners' advanced level of L2 proficiency. With regard to repair carry-outs, I found the same types as Marques-Schäfer (2013). Both translations in a different shared language or machine translation in leaner's native language belong to interactional resources for repair.

Tudini (2010) applied CA methods to analyse different repair types in native/non-native speaker chat. She concludes that learners have various opportunities to improve their foreign language skills in a Conversation-for-Learning. However, interactional resources for dealing with trouble in chat as they are made available by participants were not in focus of her study. In this sense, the present study continues the CA-informed analysis of native/non-native speaker chat started by Tudini (2010) and Vandergriff (2013). Since language earners are expected to benefit from a conversation with an artificial chat partner, aspects of learning are an important issue for communicative ICALL. It needs a further investigation whether language learners will make use of the same opportunities to engage in repair with linguistic trouble source with an artificial conversation partner as they do it with a native speaker.

Relying on the SLA-theoretic results on influence of specific interactional routines on language acquisition, Marques-Schäfer (2013) argues that negotiations with regard to questions with linguistic matters support language learning and help the learners to improve their knowledge of L2. Tudini (2010) concludes more general that learners' engagement in repair sequences supports learning. However, repair initiation is one option to deal with troubles in comprehension, but not the only one. The participants of the present study reported that they used online dictionaries and machine translation if they had difficulties in comprehension. Such strategies belong to learning strategies, too. The selection of a strategy to deal with trouble depends on many factors, and leaner's choice to position herself in a weaker position of a language novice may be one of them. Strategies of dealing with trouble in comprehension not visible in the chat protocols need to be targeted in the research design phase in order to assess their influence on learning.

11.2.2. Computational models of other-initiated self-repair

In this dissertation I made the first step towards closing the gap in modelling repair initiated by the user specifically targeting language learners. Related academic literature has been discussed in Section 2.3.2.

The new model of other-initiated self-repair when the machine is the trouble-speaker (OISR$_M$) allows to recognise learner repair initiations and to extract the trouble source based on a description of language-specific and medium-specific resources for repair initiation. The model is created on a necessary level of abstraction to be applicable for text chat interaction in languages other than German. This assumption builds on Dingemanse et al. (2014)'s finding that similar repair initiation formats exist across languages. Therefore, when provided a set of language-specific devices for repair initiation (such as lexicalised unspecific signalling resources and demonstratives), it can be implemented for other languages. The extraction of the trouble source is based on abstract features like repetition of parts of the trouble-turn and adjacent position. These features are language independent.

The problem of the trouble source extraction is related to referring expression recognition or reference resolution described in NLP textbooks (Martin and Jurafsky, 2009, Ch. 21), which is addressed in a large number of scientific publications (Dahan et al., 2002; Iida et al., 2010). Usually only noun phrases or their pronominalised alternatives are considered for reference resolution in NLP. These are usually definite and indefinite noun phrases, pronouns, demonstratives and names. The analysis of repair initiations shows that verbs or parts of utterances may be used to refer to the trouble source. The model of OISR$_M$ implicitly includes a local *discourse model* which "contains representations of entities which have been referred to in the discourse" (Martin and Jurafsky, 2009, p. 730). The local discourse model in OISR$_M$ sequences is restricted to the possible representations of the trouble source.

Compared to the model of clarification requests proposed in (Purver, 2004), the model introduced in this work has the following advantages. First, the inconsistencies form CA perspective found in (Purver, 2004)'s classification (see Section 2.3.2 for critiques) do not exist in the model presented in this work because of a close cross-disciplinary connection with CA. The model for repair initiations presented here strictly differentiates next-turn repair other-initiations from all other types of repair and describes only these repair initiations. Second, (Purver, 2004) introduced the model for clarification requests in a strong connection to the HPSG formalism. He lists technological requirements regarding language understanding capabilities, which include:

- The representation of utterances must include information at phonological, syntactic and semantic levels.
- This representation must have an appropriate semantic structure: it must be made con- textually dependent, with words and certain phrases contributing elements which must be contextually identified during grounding.
- Both user and system utterances must share this representation, as both may be subject to clarification (Purver, 2004, p. 236) .

As opposed to these strong requirements, the model presented in this work is already implementable with such simple language understanding technology as chatbots. The separation between resources for signalling trouble and resources for referencing trouble source allows creating a rule-based grammar which can be implemented in dialogue systems and conversational agents with different levels of complexity in language understanding.

With regard to the analysis of causes of troubles in understanding introduces in (Schlangen, 2004) and discussed in Section 2.3.2, manly problems on the third level (Meaning and understanding) were subject of learner's repair initiations. Consequently, the modelling was approached in this work with the assumption that the required kind of clarification is mainly determined by the user model targeting language learners. Similarly to the (Schlangen, 2004)'s approach to map the variance in form to a small number of readings, repair initiations in this work are mapped either to a content question *What does X mean?* or to a polar question *Does X mean Y?* where X is the trouble source and Y is the candidate understanding. In this way, the two approaches to modelling repair initiations are similar.

Models of repair covering repair initiations proposed in (Purver, 2004) and (Schlangen, 2004) and extended in follow-up work (Purver, 2006; Ginzburg et al., 2007; Ginzburg, 2012) were motivated by Conversation Analysis research. However, other approaches for modelling were preferred because of the insufficient operationalisation of CA findings for computational modelling. As an implication, the factors influencing the interaction that have been identified as important in CA studies and building a *system* did not become part of the baseline models in (Purver, 2004) and (Schlangen, 2004). Such factors include repair, turn taking, membership categorisation, adjacency pairs and preference organisation. In contrast to the previous models of repair (Purver, 2004; Schlangen, 2004) this work analyses repair initiations in a system of interconnected factors in conversation. More specifically, the proposed model of repair initiations takes turn taking and sequential organisation of interaction explicitly into account by distinguishing between immediate and delayed repair initiations and respective options for trouble source extraction. In addition, the new model takes virtual adjacency in chat into account. It explicitly differentiates repair initiated by the user from repair initiated by the system taking the sequential organisation into account. Finally, the preference organisation and recipient design were taken into account by the user model. Based on the findings from the dataset discussed in Chapter 4, the user model assumes that language learners will request a special kind of clarification

While recognition of repair initiations and trouble source extraction can be implemented using the simplest type of language understanding, namely, pattern-based language understanding, most repair carry-outs require more sophisticated linguistic capabilities. I will go through all of them based on the list of identified repair carry-out types.

Definitions provide an explanation of the trouble source. Existing online dictionaries such as Wiktionary or Wikipedia may be used to create linguistic knowledge bases. Because one term may have multiple meanings, a linking to the correct meaning may be required. This problem has been approached in NLP mainly in the area of lexical ambiguity resolution also known as meaning resolution (Small et al., 1987) and is part of a larger area of computational lexical semantics (Martin and Jurafsky, 2009, Ch. 20).

Paraphrasing provide a reformulation of the trouble source. A lot of efforts have been put in automatic paraphrase generation and recognition. Several recent publications are (Metzler et al., 2011; Regneri and Wang, 2012; Marton, 2013).

Synonyms provide usually a short reformulation of the trouble source. Existing language resources such as WordNet (Fellbaum, 2010) and GermaNet (Hamp et al., 1997). Similarly to definitions, multiple meanings of a word may need to be resolved.

Translations may be generated by using existing machine translation systems (Avramidis et al., 2015; Burchardt et al., 2014). Open source statistical machine translation systems such as Moses[1] make experimental implementations feasible. Commercial machine translation API can be integrated into the dialogue manager, for instance Google Translate API[2].

Demonstrations include hyperlinks to websites containing relevant information or as a way to show one example of an object referenced by the trouble source. For instance, objects and concepts related to web technologies such as *life ticker* and *life stream*

[1] http://www.statmt.org/moses/
[2] https://cloud.google.com/translate/docs

may be explained by an example of such objects. Using search engines to find relevant websites may be one of the ways to find relevant information in real time. However, only one of the found documents can be presented to the user. For semi-automatically created databases of linguistic knowledge, such information may be included into examples. Wikipedia articles sometimes also contain links to example websites and pictures, which may be used as examples of concepts described in the article.

The implementation case study covered only a simplified version of all possible repair carry-outs. Repairs have been generated from a pre-processed Wiktionary dump called ExplanationDB. The database contained synonyms, meaning explanations, examples and notes on pragmatics. If multiple meanings of a trouble source were found in the database, the machine generated an explanation for each of the meanings based on a response template. If examples were available, each meaning was accompanied by an example. More sophisticated language technology than a simple AIML-based chatbot is required to cover the broad range of repair carry-out formats found in the empirical data.

Because everything may become a trouble source in conversation, repair initiations may come after *every* utterance and even after silence. It means for human-machine dialogues, that the machine needs the ability to distinguish the action of repair initiation from all other actions which may come routinely after system's utterance. This dissertation made a contribution to the computational dialogue and conversational agent research by proposing a local model for OISR$_M$. Explicit handling of repairs targeted for language learners allows an implementation in a communicative ICALL system mimicking Conversations-for-Learning. In this way, this dissertation advances state-of-the-art in ICALL research described in Section 2.1 and strengthens multidisciplinary connections to related disciplines, such as Conversation Analysis and NLP. Other types of tutorial dialogues where a clarification of the terminology may be necessary, would also benefit from the presented model.

11.3. Other-corrections of linguistic errors

This research continued the work towards disengagement of trouble in talk and practices of dealing with it started in Conversation Analysis (Schegloff, 1987). In particular, I distinguish between local practices of learner error correction and the choice of error correction as one of possible practices to deal with learner errors.

With regard to dependencies between error types and correction types, this study confirms findings reported in earlier academic publications showing that types of errors produced and types of correction format applied are not mutually dependent (Schegloff, 1987). However, other factors in the trouble turn may influence the preference in correction formats. These factors include the availability and certainty of a target hypothesis, sequential environment where the trouble turn occurs, and whether form or meaning is the focus of the correction. The analysis of embedded corrections shows additional restrictions. Because only repeatable units from the trouble source can be made part of an embedded corrections, some errors of person, number, gender and congruence cannot be repaired implicitly.

Tudini suggests to analyse exposed and embedded corrections as different options "in the continuum of explicitness of exposed correction in online text chat" (Tudini, 2010, p. 101). My study of corrections provides additional empirical support for this idea: exposed correction with rich accountings and contrasting exposed corrections as one extreme and embedded, implicit corrections as another extreme, and various cases in between. In addition, further dimensions for a classification have been found in this work.

Previous academic literature distinguished between exposed an embedded corrections. Exposed corrections are delivered in a separate side sequence and may include accountings. Embedded corrections are delivered as part of the relevant next action to the trouble turn and do not permit for accountings or any other form of focusing on correcting. In this work I found a type of corrections not previously described as a separate correction type, which I named *integrated*. I argue that integrated corrections are a distinct type of corrections combining features of exposed and embedded corrections. Integrated corrections do initiate a separate side sequence but they are fully semantically and syntactically integrated into the next action. They may contain accountings and usually emphasise the correction, and become an exposed correction through accountings and/or emphasis. The existence of integrated corrections extends the understanding of the correction types stepping away from the dichotomy of embedded as opposed to exposed correction types.

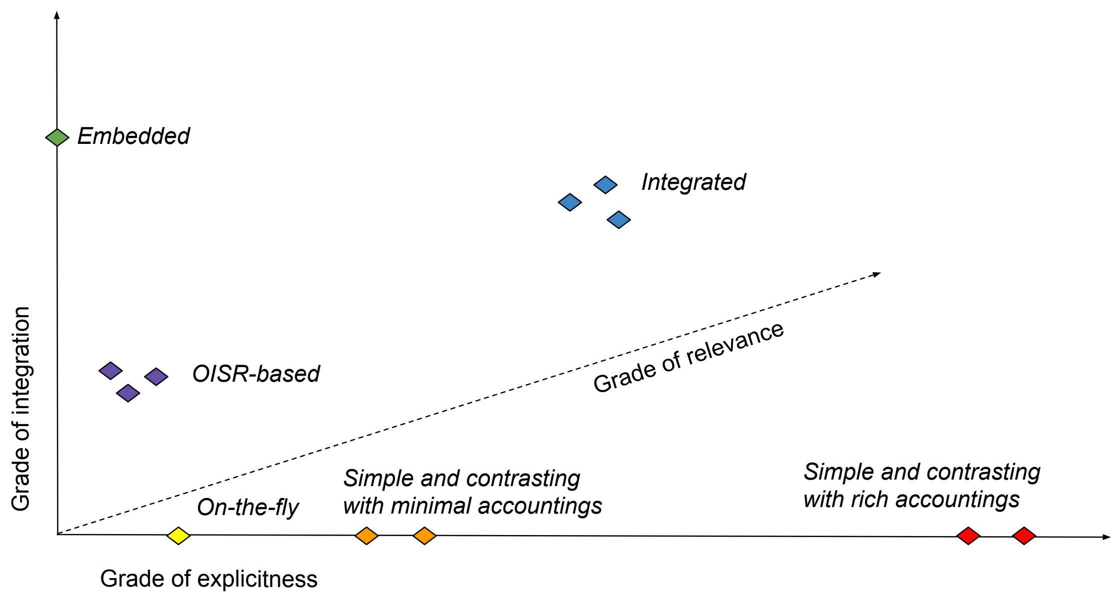


Figure 11.1.: The space determined by the explicitness, integration and relevance of correction formats.

OISR-based error correction formats have been previously described in classroom research (e.g. initiations of meaning negotiations (Varonis and Gass, 1985)) and in CA-for-SLA literature (Kurhila, 2006). Such correction formats explicitly initiate a repair side sequence to deliver a correction, however, they formally provide the learner the opportunity to self-correct. In addition, repair initiations are potential relevant nexts after each turn. On the other hand, the correction in this case is not integrated into a preferred or sequentially projected response, as opposed to embedded or integrated corrections. Therefore, in my understanding they also have features from both exposed and embedded corrections. They hide their face-threatening effect behind the preference for self-correction. However they are exposed, because they clearly do correction and are taken up as exposed corrections. Consequently, this type of corrections opens a third dimension in the classification of correction formats, that is grade of relevance.

Figure 11.1 illustrates the correction space based on the three dimensions. Though Tudini (2010)'s idea to analyse corrections in the continuum of explicitness is very intuitive and supported by the socio-linguistic data, it is hardly operationalisable with respect to computational modelling of corrections. The mapping from a correction format type to the grade of explicitness is rather subjective and makes a value assignment for the moment problematic. The grade of relevance buries the same difficulties. In contrast, the grade of integration into the next relevant action may be expressed by means of syntax. Therefore, a separation of correction formats according to their syntactic integration into the next projected turn was chosen for computational modelling. OISR-based corrections were classified as not integrated into the next projected turn, because repair initiations are not projected, though they are possible after each utterance.

Learner Type/NS	L01 N01	L02 N01	L03 N02	L04 N02	L05 N02	L06 N03	L07 N03	L08 N04	L09 N04	 Sum
Exposed	0	1	1	1	0	1	4	9	5	22
Embedded	3	7	2	3	4	3	4	10	1	37

Table 11.2.: The number of all corrections of linguistic errors in the corpus: distribution over all pairs of participant labeled by the learner code: L01 - learner 01 to L09 - learner 09.

The majority of all other-corrections of linguistic errors in my dataset were embedded corrections. This supports the finding by Kurhila (2001) who found embedded corrections (called corrections 'in a repetition slot' in the original work) to be the largest class of all corrections. This shows the opposite to the findings by Tudini (2010) as well as by Marques-Schäfer (2013) who report that the majority of all corrections in their datasets are explicit corrections. Table 11.2 the number of exposed and embedded corrections in my dataset. A comparative analysis of the datasets may help to explain why such different results were achieved.

The numbers in Table 11.2 give an intuition, how much additional data may be needed to build a collection of error corrections that is sufficient to train machine learning models.

Interaction recordings of hundreds of native/non-native speaker pairs may be needed. For now, the small numbers of occurrences of error corrections was sufficient to find prototypical structure in corrections and to define rules for machines behaviour based on these prototypes. This shows that Conversation Analysis offers a powerful tool to create computational models of rare phenomena in conversation, such as error corrections.

11.3.1. Error recognition and learner corpus research

While multiple research projects put efforts into learner corpus annotation for conceptually and medially written learner language, only few conceptually oral learner corpora exist till now (see the overview for German learner corpora in Section 2.1.6). The corpus created for the purpose of this research may satisfy the needs of many other research endeavours and can be obtained for free from the European Language Resource Association (ELRA, 2015). To ensure result comparability and better evaluation, all examples discussed in this dissertation can be found in the original corpus by using turn number and participants codes as a reference. With the discussion of the problem with instant messaging data provided in Section 2.2.2, the corpus offers a great opportunity to perform further studies on native/non-native speaker instant messaging communication.

The corpus was published as an untagged and as an annotated version. Besides the creation of the corpus itself, this work contributes to leaner corpus research in the following way:

1. Conceptual work on learner errors in chat, analysing what counts as error and what is worth correcting.
2. Learner language annotation focusing on dialogue moves in other-initiated self-repair and exposed corrections, partial error annotation.

Though it is still necessary to recognise as many as possible learner errors for the purpose of language understanding in ICALL, some of the deviations from the language standard are excluded from error corrections due to preference organisation in chat. The interplay between the variations in orthography, language standard and social closeness was discussed in Section 7.1.2.

Tudini (2010, p.19) notes that the Italian native speakers in her data set tend to "tone down" chat jargon in conversations with language leaners and use mainly informal colloquial Italian. From this observation can be concluded that native speakers modify their chat language and design their turns to make them understandable for not-yet-fully proficient language users. However, changes in the chat language discussed in Section 7.1.2 show that making the utterances comprehendable for non-native speakers is only one aspect in the selection of the interactional resources. There is at least one more function performed through deviations from linguistic standard, that is regulation of the social proximity.

While this finding may play a less important role for conceptually written learner language corpora, the interactional function of deviations from standard is important for conceptually oral learner corpora. Specifically, the difference between "real" or potentially correctable linguistic errors and those which do not count as errors in chat is important. In addition, further conceptual work on error annotation in chat data is required.

The majority of learner corpora focus on error annotation and only a few of the projects focus on linguistic markup. This work addresses an additional dimension in learner language annotation, namely dialogue moves markup. SLA-inspired first corpus version contains annotated dialogue moves in sequences of corrective feedback and meaning negotiation. Annotation of the interactional aspects of learner language may be specifically of interest for communicative ICALL research, because it makes information related to sequences of repair with linguistic trouble source accessible for the machine. Corpus user's guide provides a detailed information about the annotation. The relevant part of the manual can be fond in Appendix A.

11.3.2. Data-driven models of side-sequence exposed corrections

I argued in Section 2.1.5 with support from CA-for-SLA literature discussed in Section 2.2, that an empirical base for the specific speech exchange system needs to be created in order to obtain reasonable models for communicative ICALL systems mimicking Conversations-for-Learning. I analysed a collection of exposed corrections in Chapter 5 to prepare an empirical basis for further computational modelling. Computational modelling of exposed corrections was then approached in the second step and documented in Section 9.2. In this section I summarise the research results in the field of exposed corrections answering the following research questions:

RQ3b. Which types of exposed corrections of linguistic errors exist in the dataset representing a chat-based Conversation-for-Learning?

RQ4b. How can these types of exposed corrections of linguistic errors be modelled in order to be implemented in a chat-based communicative ICALL system?

To make the contribution of this work to CA-for-SLA and CALL clear, I compare the findings with the results described by Markee (2000), Kurhila (2006) and Tudini (2010) and the study of German native/non-native speaker chat described in (Marques-Schäfer, 2013). In addition, I compare the types of exposed corrections with the classification of corrective feedback obtained from language classroom data (Lyster et al., 2013).

How this work advances the state-of-the-art in ICALL and NLP will be made clear by a comparison of the findings to automatic corrections and automatic feedback generation in literature discussed in Section 2.1.5. Implementation issues and required NLP technology will be also discussed.

Exposed corrections in native/non-native speaker chat

Marques-Schäfer (2013) found that the most corrected errors are either corrections of orthographical errors or corrections of morpho-syntactic errors. She concludes that the tutors and the chat participants perceive the corrections of orthographic and morphosyntactic errors as more important than correction in the "way of expression" (orig.: *Ausdrucksform*, p. 195). My dataset does not confirm this finding. The majority of all corrected errors were lexical errors in my dataset. It can be explained by the speech exchange system. Due to the study setup, the obtained Conversation-for-Learning appears less formal and less focused on learning (more focused on communication) than the group chat in *JETZT Deutsch lernen*. The chat participants in the study by Marques-Schäfer (2013) must have oriented to the didactic purpose of the learning platform through error correction.

Corrections *en passant* identified in oral talk (Kurhila, 2001) also called *on-the-fly* occurred in native/non-native speaker chat-based Conversation-for-Learning (Tudini, 2010) allow to correct explicitly without making correcting a big interactional deal. I found in my dataset the same two types of corrections on-the-fly as described by Tudini (2010). Either correction and interpersonal trajectory are contained in one turn, or correction turn produced by the native speaker is immediately followed by another turn continuing the interpersonal trajectory.

The organisation of corrections has been seen in CA literature as turn-by-turn organisation, for instance (Jefferson, 1983). The prototypical structure of a correction is then the following:

1. Speaker A produces an item X;
2. Speaker B produces an alternative item Y;
3. Speaker A accepts X or rejects X by using Y again.

How the turns are formatted was not that important. However, in addition to X and Y, corrections may contain various types of accountings and refer to earlier points in interaction in some way. In order to provide a conversational agent with a detailed instruction how to use all these parts to generate a correction of a particular type, a detailed analysis of correction turn formats was required, which I did in Chapter 5.

Table 11.3 shows the result of classification of all exposed corrections from my dataset according to Lyster et al. (2013)'s classification of corrective feedback. I did not find any repetitions, metalinguistic cues, paralinguistic signals or elicitations in my dataset. These types of corrective feedback seem to be typical for classroom interaction (typical teacher's expressions) but too teacher-like for a Conversation-for-Learning. Paralinguistic signals do not occur in the data set to signalise an error. This result confirms the hypothesis, that classifications obtained from a different speech exchange system may be not applicable for the Conversation-for-Learning.

Speaker	repairs by type				
	conversational recast	didactic re-cast	clarifica-tion request	explicit correction	explicit with MLE
N01				1	
N02		1	1		
N03		2	1	1	1
N04	2	2	1	9	
TOTAL	2	5	3	11	1

Table 11.3.: Classification of exposed corrections according to the classification of corrective feedback proposed by Lyster et al. (2013, p.4)

The classes of corrective feedback are not disjoint. This has already been criticised in CA literature (Markee, 2000; Kurhila, 2006). Especially recasts were frequently selected for analysis in communicative ICALL. However, recasts as form of correction covers several *different* types of corrections with *different interactional import* found in the dataset and described in earlier CA-driven SLA studies. Recasts may have a form of an embedded correction and of an exposed correction. Subsequently, different next actions may be expected after them. This research showed how CA-for-SLA may be an effective methodology on the way to create a consistent classification of disjoint categories of exposed corrections. The proposed classification of error correction formats contains only disjoint classes which are based on interactional recourses employed in their construction.

While contrastive stress and laughters were found in oral data as devices for accomplishing corrections (Hauser, 2010), other medium-specific resources are on participants' disposal in chat to perform a correction. These resources are uppercase writing, quotes, dashes and emoticons. Repetitions of the errors and explanations are common to oral and chat data as devices for correction. The role of the specific correction devices for construction of different correction formats in chat needs to be taking into account for design of chat-based communicative ICALL applications. The proposed correction format classification serves this purpose.

Exposed correction formats for communicative ICALL

The classification of corrective feedback discussed just above in this section was taken up by communicative ICALL (Petersen, 2010; Wilske, 2014). In contrast, this research proposed a different classification of error corrections grounded in CA research. A formal model for correction formats was created based on the new classification. Devices such as accountings and backlinks are part of the correction formats. To my knowledge, this is the first model of corrections that explicitly takes them into account.

As discussed in Section 9.2, the proposed model may be implemented using simple template techniques for each of the correction types. However, automatic error recognition remains a challenging prerequisite. Various approaches have been used for automatic error recognition and were discussed in Section 2.1. With its further development, a broader range of learner errors will be recognised automatically and potentially corrected by an artificial conversation partner. However, at present there is a tradeoff between the need to correct more errors focused on form and the ability to recognise such errors automatically.

Wilske (2014, Sec. 5.4) discusses the amount of information and the complexity of language understanding necessary to extract the required data in order to produce different types of corrections in ICALL systems. For instance, detailed error explanations require a very detailed error checking while open-class (unspecific) repair initiations such as *I did not understand* only needs information *that* an error occurred. As argued in Section 11.2, such types of corrective feedback correspond to open-class repair initiations. In oral native/native speaker communication, they project a repetition as a response. In chat-based native/non-native communication, I found such repair initiation only produced by non-native speakers to deal with troubles in comprehension. Native speakers used open class repair initiations to initiate conversational repair, but not as a tool for error corrections. Even if such repair initiations are easy to produce automatically, their use as a format for error correction in native/non-native speaker chat is not supported by empirical data. However, they might be appropriate in a different speech exchange system.

Because the majority of error correction formats found in the dataset do not contain meta-linguistic information, detailed explanations of the errors may be unnecessary for some forms of communicative ICALL, such as conversational agents that help to practice conversation. The availability of at least one target hypothesis is sufficient in oder to generate the correction formats from the models proposed in this work. However, different error correction formats have different import into interaction. Correction formats identified and modelled in this work reflect the view on error corrections in talk as a social action. Accountings, and not meta-linguistic information are included (or not included) in all of them. Therefore, the right choice of accountings (or the decision to omit them) may be a new challenge. I addressed this problem of the correction format selection in Chapter 10 where I make a step towards a correction decision model. To sum up, a generation of the correction formats introduced in this work are quite simple from the computational perspective, if the information about the target hypothesis and the decision to use a particular correction format are available. The latter two tasks are very challenging from the computational perspective.

11.3.3. Data-driven models of embedded and integrated corrections

In this section I discuss the research results in the field of embedded corrections to answer the corresponding research questions. The respective empirical and computational findings have been presented in Chapters 6 and 9. The research questions are:

RQ3a. Which types of embedded corrections of linguistic errors exist in the dataset representing a chat-based Conversation-for-Learning?

RQ4a. How can these types of embedded corrections of linguistic errors be modelled in order to be implemented in a chat-based Communicative ICALL system?

The present study confirms previous observations that embedded corrections may be used to modify previous speaker's talk in order to correct linguistic errors (Section 6.1.2) but also - as in L1 talk - to deal with imprecise formulations, for example, in style and expressivity (Section 6.2).

For the purpose of communicative ICALL, I focused specifically on embedded corrections of linguistic errors. In particular, I analysed sequential environments where embedded corrections on linguistic errors occur. In oder to prepare a data-driven basis for computational models of embedded corrections, I analysed the mechanism of embedded corrections in pairs of polar questions and answers. I discuss the results in the remainder of this section.

Embedded corrections of linguistic errors in chat

In contrast to the results described by Tudini (2010) saying that "dyadic online intercultural chat favours exposed correction to deal with pedagogical repair", the present study shows that embedded corrections may occur even more frequently than exposed corrections to deal with learner errors. This confirms the result reported by Brouwer et al. (2004) and Kurhila (2001) according to which embedded corrections (corrections "in a repetition slot", (Kurhila, 2001)) are typical for native/non-native speaker interaction. However, as Table 11.2 shows, the number of occurrences of embedded corrections varies for each pair of participants. The analysis of pairs of polar questions and answers in Section 6.3 showed that native speakers made use of embedded corrections only in a small number of all opportunities. From 91 answers where an embedded correction could have been provided, only 17 contained an embedded correction. An explanation for this finding may be found in an additional study.

Embedded corrections were found in the following sequential positions:

1. In a second pair part if an error occurs in the first pair part.
2. In a post-expansion if an error occurs in the second pair part.

This confirms the finding by Kurhila (2006) and extends the findings by Brouwer et al. (2004). Brouwer et al. (2004) found only corrections in second pair parts. The majority of all embedded corrections in the corpus were found in responses to questions. Therefore, question-answer pairs were chosen for modelling. This type of embedded corrections was then restricted to only embedded corrections in answers to polar questions as specified in the coding scheme by Stivers and Enfield (2010). Repetition-based answers to polar questions were found to provide opportunities for embedded corrections. Thus, a communicative ICALL system needs to be made aware of differences between repetition-based answers and other types of responses to polar questions (answers and non-answers).

As opposed to findings described by Brouwer et al. (2004), I found embedded corrections of linguistic errors not only in the second pair parts after the first pair parts with deviations, but also in counter-questions and post-expansions (minimal and not-minimal) after errors occurring in second pair parts. Table 11.4 summarises the findings with regard to sequential positions of the error and the relative position and the action type of correction.

Error location	Correction location	Example
First pair part	Second pair part	Greeting - greeting Question - answer
Second pair part	Counter question	Answer - counter question
	Minimal post-expansion	Assessment Evaluation Acknowledgement
	Non-minimal post-expansion	Surprise Topicalisation realised through an information request

Table 11.4.: Location of embedded corrections relatively to the error location

With regard to observations by Brouwer et al. (2004) on embedded corrections in the second pair part, my study confirms that embedded corrections frequently occur in the second pair parts and correct errors from the first pair parts. However, as opposed to the collection analysed by Brouwer et al. (2004), the collection of embedded corrections found in my corpus contains errors in the second pair parts and corrections in minimal and non-minimal post-expansions. "Sequence closing thirds" are examples of minimal post-expansions (Schegloff, 2007). Embedded corrections with this function in form of an evaluation, acknowledgement and assessment are found in my collection. Non-minimal post-expansions described by Schegloff (2007) are for instance, repair initiations and topicalisation. Expressions of surprise and information requests which contained embedded corrections in my dataset were found in both types of non-minimal post-expansion.

Brouwer et al. (2004) report that they did not find any embedded correction in an answer to a polar question. In light of the analysis of question-response-pairs based on question coding scheme (Stivers and Enfield, 2010) presented in Section 6.3, answers to polar questions can be designed in a way which supports or allows to avoid embedded corrections. Repetition-based answers to polar questions provide opportunities to correct implicitly. Interjection-based (marked or unmarked) answers to polar questions do not embed a correction. In my collection of question-answer pairs, repetition-based answers to polar questions clearly do more than simply giving a confirmation or a disconfirmation. The same is reported by Brouwer et al. (2004). In addition, this confirms the result described by Lerner and Kitzinger (2007) showing that what is done by an embedded correction is action specific to the trouble source.

The concept of the *minimal correcting replacement* reflects findings by Schegloff (1979) where the author analysed self-correction and the unit boundaries looking at what is repeated in a self-correction. Schegloff (1979) found that no repetition starts in the middle of a unit, the whole unit is renewed (whatever the unit is). Similarly, the minimal correcting replacement repeats the smallest unit necessary to locate the error, and replaces the erroneous part in it by a correct one. A repetition of a part of such a unit would not correct the error.

Jefferson (1974) describes correction devices which show that current talk is a quote of directly prior talk. These devices employ a repetition of the erroneous item, which was not completely produced but could be guessed. With this regard they are very similar to acknowledgement-based embedded corrections. The difference is that there is no cut-off in the directly prior talk. The learners completely produce the erroneous item but the native speakers use the same quotation format as a base. Due to a replacement within the quotation-based format a correction can be produced.

With regard to tracking learning progress, embedded corrections present a particular difficulty because they allow learners not to respond to it, as Kurhila (2006) observed. Confirming this finding, it was illustrated in Example 3.16 how a postponed uptake after an embedded correction may take place. In this example, the correction was taken up in the subsequent chat session, and not as a direct response to the correction. Postponed uptake and imitation are very difficult to find in the data for human annotators, but might be approached by computer-assisted qualitative data analysis (Wickham and Woods, 2005) and taken as a base for evaluation in ICALL.

Embedded corrections in ICALL

Because ICALL research mainly builds on results reported in SLA literature, the concept of embedded corrections was not explicitly included in ICALL till present. Implicit handling of embedded corrections was performed by using recasts as form of implicit feedback in ICALL applications, such as (Morton et al., 2008; Petersen, 2010; Wilske, 2014). However, as argued earlier in this chapter, recasts include embedded and exposed corrections. For instance, recasts presented by Petersen (2010) mainly correspond to corrections on-the-fly. This research is the first attempt to explicitly operationalise embedded corrections for communicative ICALL.

Because of the complete syntactic and semantic integration of embedded and integrated corrections into system's response *not formatted as a correction*, generation of such corrections may present a challenge for a computer system. An exact instruction needs to be given to the machine, how to construct a turn with and without an embedded correction.

In this work I approached this challenge by focusing only on embedded corrections in answers to polar questions. The coding scheme for classification of questions and responses introduced by Stivers and Enfield (2010) was taken as a basis to create a collection of examples of question-answer pairs. Questions produced by learners and containing "real" errors and responses to them have been selected for the analysis. The focus on pairs of polar questions and answers allows to restrict the potential search space for potentially relevant next actions. This space is termed by Schegloff (1996) as *contingency*.

Because of the contingency, polar questions (as any action) may be responded by a number of relevant next actions, such as answers which deal with the questions as put, and responses which include indirect dealing with the questions, idontknows and all other kinds of relevant next actions to a polar question. This variation is reflected in the coding scheme proposed by Stivers and Enfield (2010). From all this variety of possible responses, repetition-based answers represent the class of responses allowing to embed corrections of linguistic errors.

A repetition-based answer corrects an error in a polar question if some parts of the question were repeated, and some other parts were replaced or changed in a way that something wrong in the question became correct in the answer. I express this process in the concept of a *minimal correcting replacement* in Section 6.3.

A minimal correcting replacement is given if:

1. There is constituent in the learners turn which completely contains the error,
2. The response to this turn contains the same constituent, but with changes, so that the error is corrected, and
3. Any smaller constituent will not correct the error, and any larger constituent will be not minimal.

For spelling errors in separate words, the minimal correcting constituent will be a repetition of the word with replaced symbols. To correct errors in relationships between parts of an utterance (e.g. congruence errors), all parts of the utterance involved in the erroneous relationship need to be repeated in order to make the correct relationship visible. Sometimes, the whole utterance needs to be repeated in order to correct errors in, for instance, subject-verb congruence. All correcting replacements for an error correct this error, even if they are not minimal. This is how repetition-based answers to polar questions would work: they frequently repeat the whole question content. To generate an integrated correction, an embedded correction can be produced and the unit where the error was located needs to be marked by symbolic devices (framing and highlighting) and may be accompanied by accountings. Therefore, integrated corrections may be seen as a sub-class of embedded corrections.

Even simple, pattern-based language understanding techniques allow to make use of embedded corrections in dialogues with learners. For instance, AIML-based chatbots already implement matching of a number of surface strings to one pattern in order to cover at least the most prominent variants of spelling. This can be done either by string replacements during the preprocessing phase or by `srai` operator. The `srai` operator allows to forward the processing of a matched pattern from one category to another one in order to handle synonyms or paraphrases. Replacements allow to recognise, for instance, small variations in spelling. Special cases of spelling errors can be targeted by such replacements. A repetition of the part of the utterance after the replacements in machines response would provide an embedded correction. As argued in Section 9.4.2, deeper language understanding techniques would be helpful to produce integrated corrections. Simple cases of integrated corrections can be based on the same embedded corrections as those handled by the AIML chatbot. In addition, they need to include information which tokens have been change. This can be done by highlighting the correct spelling with uppercase.

Because this is the first attempt to propose a model of embedded corrections for communicative ICALL documented in the academic literature of which I am aware, pioneer work in understanding of the working mechanism of embedded corrections in talk was done for polar question-answer pairs. This was only partially successful due to the task complexity. This work produced more questions than answers in the field of computational modelling of embedded corrections. Even more pioneer work still needs to be done to cover other types of question-answer pairs and other sequential environments where embedded corrections can occur. I will discuss some of the identified questions in Chapter 12 and outline future research directions.

11.4. To correct or not to correct

In this section I discuss the research results with regard to a correction decision model addressed by the following research questions:

RQ5 Which factors besides an occurrence of an error are relevant for an occurrence of a correction of a linguistic error in native/non-native speaker chat-based Conversation-for-Learning?

RQ6 How can these factors be modelled for in order to be implemented in a chat-based communicative ICALL system?

Because this question was not on the research radar at the time of the data collection, only partial information about participant's decisions to correct or not to correct is available. Nonetheless, all the local correction formats and their model are useless for a communicative ICALL system, if it cannot decide, when they are applicable. Therefore, this research faced this challenge. The respective empirical findings have been presented in Chapter 7. The computational model for correction decision based on these findings has been proposed in Chapter 10.

11.4.1. Interactional relevance of a correction

Studies on native/non-native speaker and non-native/non-native speaker chat interaction show different and sometimes contradicting results with regard to frequencies of exposed corrections. The conclusions of the studies mostly connect the presence or non-presence of correction to participants' linguistic identities or the communication medium. For instance Tudini (2007) explains that native/non-native speaker chat does not support corrections as much as non-native/non-native speaker chat does. Tudini (2010) concludes that intercultural chat does not support embedded corrections. I see such conclusions as problematic, because they do not address reasons for non-present or present corrections in conversation in terms of discourse type, sequential consequences and interactional import of corrections and their alternatives. Other studies discussed in Section 2.2 describe efforts made towards an explanation of presence or non-presence of correction, for instance (Kurhila, 2001, 2004; Hosoda, 2006; Kasper, 2004). Differences between language classroom interaction and less formal types of native/non-native speaker interaction need to be taken into account in terms of decisions to correct or not to correct and which form of correction should be preferred.

As analysed by Schegloff (Schegloff, 1987) and observed in the present work, a correction is not the only way to deal with learner errors. As it was shown in multiple CA-driven SLA studies, interactional import of specific correction formats is relevant for speakers decision to correct and speaker's selection of a specific correction form. As Kurhila (2006) argues:

> the prevailing features of correction in non-pedgagogic native/non-native speaker conversation: first, (grammatical) deficiencies are subject to outright repair but not to repair initiations and, second, (linguistic) correction is done so as to diminish the interactional prominence of the activity. (p. 43)

As it is reflected in the number of corrections in Table 11.2, there is a huge variation in the number of both exposed and embedded corrections over different native speaker-learner pairs. A detailed analysis of all corrections was helpful to understand which factors may be relevant for a decision to correct or not to correct. The let-it-pass strategy was preferred by some of the pairs while other pairs frequently engaged in discussions of linguistic matters.

A presence or a non-presence of a correction is determined not barely by the statuses of the participants and the medium of the communication. It is also determined by other integral parts of the conversation such as

- conventions and agreements,
- turn properties (e.g. length and present or non-present narrative character),
- error complexity and availability of a target hypothesis allowing or not allowing for a quick and clear correction,
- engagement with the topic and
- sequential consequences of error.

In this sense, this study continues work on the reasons for corrections started by (Kurhila, 2001, 2004; Hosoda, 2006; Kasper, 2004). Mutual dependencies among these factors make both empirical analysis and computational modelling to a very challenging research endeavour.

11.4.2. Computational decision model for corrections in chat

Because previous ICALL application usually build upon existing SLA theories, the selection of a correction format (or type of corrective feedback) is mostly driven by the supposed effectiveness of particular correction types to promote or facilitate learning. In contrast, this work presented the first attempt to define a feature-based correction decision model in a Conversation-for-Learning. Till present, mainly two factors have been considered in communicative ICALL as relevant for correction decision: first, the occurrence of a recognisable error, and second, the ability of the machine to provide meaningful and accurate feedback, see for instance the discussion in (Wilske, 2014, Sec. 5.4). Other factors identified in the empirical part of the present dissertation as relevant have not yet been taken into account. These factors are related to interaction history going beyond corrections in the past, user model including but not limited to linguistic features and error properties.

The disengagement between errors, corrections and decisions to correct was realised in a separation of local models for correction and a correction decision model. Local models for corrections are responsible for the generation of a particular correction format. The correction decision model determines if a correction should be produced at a particular given point of the interaction. As a whole it can be represented as a product of two regulatory formalisms, from which one is responsible for the activation of a particular type from the other one. Figure 11.2 visualises the interaction of the two regulatory mechanisms in the correction framework. Figure 11.3 shows the place of the decision mechanism in the repair manager.

ICALL research addressed the issue of correction decision from a different perspective. Factors like the number of errors, error types in user's language and activity types were considered. For instance, Amaral et al. (2011) describe the following priority for feedback messages:

> Feedback messages for reading, listening and description activities should prioritise meaning over form. [...] If multiple errors are diagnosed, meaning-based errors will be displayed first for these types of activities. Feedback messages for rephrasing activities, on the other hand, can focus on syntactic errors at the sentence level." (p. 8).

Such activities are not specified explicitly in a Conversation-for-Learning, however they may be introduced by participants in the course of social interaction and made *part of* interpersonal trajectory, for instance, in sub-dialogues of collaborative learning. Because focus on learning is emphasised in such sub-dialogues, context-sensitive modifications in correction decisions may be required and can be addressed in a future study. However,

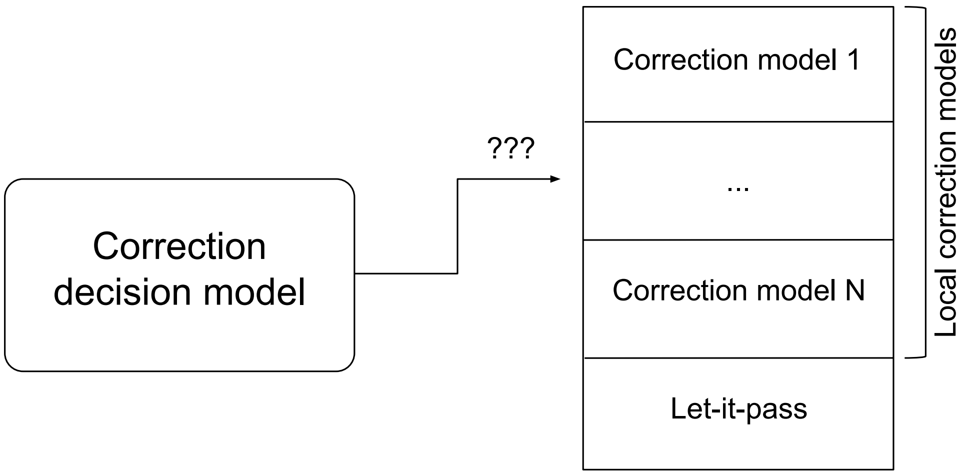


Figure 11.2.: Local and global regulation in the correction model

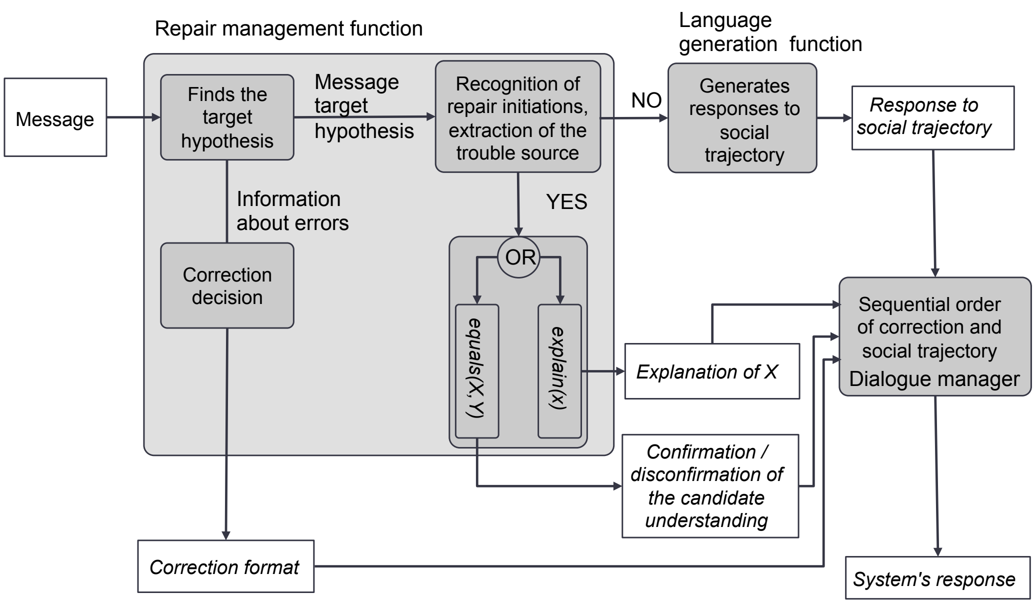


Figure 11.3.: The repair manager and the place of the correction decision mechanism

activities listed by Amaral et al. (2011) such as reading and writing are involved in chat throughout the entire conversation.

Social interaction with the user influenced the decision to correct in the work by Petersen (2010) in the way that the social trajectory was followed immediately after corrections, forming every time the same type of correction. However, empirically grounded research in this and other studies show that correction formats and correction focus change over time Dings (2012). Choosing a different types of correction formats over time will also allow to avoid repetitiveness which has been seen as problematic in conversations with artificial agents (Bickmore and Picard, 2005).

Because every correction in talk is also a social action, observable features of social interaction were included into the correction decision model, such as interaction history and user model. Amaral et al. (2011) integrate information from the activity model and the user model (e.g. which errors the learner typically makes) to improve learner language understanding. Several academic publications discussed in Section 2.1.5 take information from the user model and interaction context into account for providing corrective feedback (Heift, 2003; Amaral and Meurers, 2008, 2011). However, the decisions that non-communicative ICALL applications have to take and the features that need to be observed are different. While context understanding for a non-communicative ICALL system may include information about the activity and the task, interactional history needs to be tracked in addition to the actual activity understanding. The actual activity in a Conversation-for-Learning may range from teacher-student simulation (collaborative learning) over chit-chat about the work day to telling jokes and narrations. The sub-dialogues of collaborative learning make corrections less dispreferred. In addition, preceding forms of participants orientations to their linguistic identities influence the decision to correct. Specifically, preceding learners negative self-assessments make a correction less preferred as opposed to learner-initiated explanations (other-initiated self-repair). In contrast, focus on meaning in immediately preceding explanation sequences initiated by language learners (other-initiated self-repair) supports subsequent focus on form, as discussed in Section 4.1.1.

Other types of self-correction are important for correction decision in chat than as compared to spoken communication. While same-turn self-corrections have been intensively studied in the speech recognition context, they do not play a role for language understanding in text chat. Such self-corrections are completed before the message is posted. However, transition-space self-corrections do influence correction behaviour in human-to-human chat and are taken into account for the decision model.

Issues like dealing with multiple target hypotheses and selection of one surface string from many possible are already known from previous ICALL publications. In addition, in communicative ICALL applications where corrections need to be provided within the ongoing dialogue, the interactional import of the correction *besides of the error correction by itself* needs to be taken into account. This interactional import is accomplished by various types of rich or minimal accountings, backlinks, giving the learner the opportunity to respond to the correction or "keeping the floor" (correction on-the-fly) as well as the primary decision to make the correction to an interactional business or not (exposed vs.

embedded correction). Previous academic publications in communicative ICALL did not focus on this issue.

Though this research made a clear contribution to communicative ICALL, it has several limitations with regard to correction decision model. Motivation to perform the learning task is listed by human tutors as an important factor for a correction (Amaral et al., 2011). The concept of *motivation* was operationalised as professional relevance in this work. All of the learners participating the study were students from same university and studied German as a second language to become teachers and translators. Therefore, the dependencies between the motivation and the learning behaviour are only partially analysable from my dataset. However, further adjustments based on more observations are required to cover other user groups than students of foreign languages at the university level.

Although the group of learners in the corpus was homogeneous in terms of their affiliation, variations in motivation and consequent changes in interaction were observed. Specifically, exam preparation was made relevant in several chat sessions, because the data collection took place as the learners had their exam time at the university. In this time, the learners were more motivated to engage in pedagogical activities (e.g. exam preparation) and to chat at all (Section 3.2 Example 3.7).

Because of multiple mutual dependencies among the factors, machine learning approaches may be more useful in finding priorities over different factors and calculating the importance of each of them. Due to a small number of examples in the dataset the creation of a machine learning model was left for a future study. Several ideas are discussed in Section 12.4.2 how the data issue may be solved.

12. Future Research Directions

I described various types of participant's orientations to their linguistic identities in Chapter 3. All of them are important for computational modelling of corresponding types of sub-dialogues for conversational agents in Communicative ICALL. However, only a small part of them could be made focus of this dissertation. Each of the remaining types of positioning described in Chapter 3 is a potential research challenge and can be approached by methods used for data analysis and modelling in this dissertation. The types of positioning handled in the present work only sparingly include:

1. Face-work and evaluation
2. Meta-talk about learning and collaborative learning
3. Repair types other than exposed error corrections and embedded corrections in answers to polar questions, as well as other-initiated self-repair initiated by the learner.

In addition, the repair types that became the main subject of this work triggered other, follow-up questions while I was working on them. In this chapter, I outline future research directions for which I made a start and identified in this dissertation. Some of the problems may be approached by further analysis of the dataset that I created for the purpose of this work. More or different types of data might be required for the others. In this case I make suggestions, how the necessary data can be obtained.

12.1. Further work on repair for Communicative ICALL

In Section 3.3 I described various types of repair with linguistic trouble source as they allow to make participants' linguistic identities relevant in conversation. Although this work is dedicated to models of repair to great extent, several repair types are left for future studies, namely:

1. Other-initiated self-repair when the learner is the trouble speaker.
2. Self-initiated other-repair when the learner is the trouble speaker.
3. Embedded corrections in utterances other than answers to polar questions.

Moreover, only a preliminary, initial decision model for corrections in a Conversation-for-Learning was specified in this work. The idea of intensity-based correction format classification introduced by Tudini (2010) is also worth further attention and may be promising for communicative ICALL research. I discuss some open problems and suggest how they can be approached in this section.

12.1.1. Other-initiated learner's self-repair

I explained the relevance of other-initiated self-repair when the learner is the trouble speaker for communicative ICALL in Section 3.3.2. With regard to SLA theory it falls under the concept of *meaning negotiation* which is seen important for language learning (Varonis and Gass, 1985). Although it was partially handled as a type of error correction (repair initiation-based correction formats), a systematic analysis and modelling of all such repair sequences remains for a future study. In particular, a description of differences between repair with linguistic trouble source and all other kinds of trouble may be needed to formulate models for communicative ICALL.

This type of repair initiation caught HCI researchers attention, as discussed in Section 2.3.1. Because different formats of repair initiation are designed to signal troubles of different levels in terms of Austin/Clark action ladder, a model systematically taking these differences into account may help to generate more natural repair initiations in order to deal with problems in speech recognition and language understanding.

12.1.2. Learner's self-initiated other-repair

In Section 3.3.3 I illustrated by a number of examples how word searches may be composed. Similar requirements for computational modelling are valid for word searches as for OISR. The machine will need to recognise that the learner initiated a word search and to extract the trouble source. However, the trouble in this case is the inability to find the proper word or expression. Therefore, a different way will be selected by the user to describe what she is looking for. To find abstract models for them might be a great challenge. This research question is related to such hard NLP problems as paraphrase recognition and generation (Regneri and Wang, 2012; Marton, 2013) and word sense disambiguation (Martin and Jurafsky, 2009, Ch. 20).

12.1.3. Embedded corrections not in answers to polar questions

Only a small part of all turn types where embedded corrections were identified could be expressed in models for communicative ICALL in this work. Many questions remained unsolved because of the problem complexity. Specifically, following research questions and objectives may be approached in future studies of embedded corrections:

1. Comparative analysis of embedded corrections in L1 and L2 talk: are there differences in environments where the corrections occur and their function?
2. Computational modelling of embedded corrections in non-answers to polar questions and in responses to non-polar questions are worth research efforts to complete the work that was started in this dissertation. As a preliminary step, it might be required to complete the question coding scheme initiated by Stivers and Enfield (2010). Responses to content questions and non-answers to polar question need to be classified first in order to complete the scheme. Systematic analysis of necessary modifications of the coding scheme to code text chat data may be needed.
3. A separate study focusing on responses with and without corrections may help to discover further reasons for non-present embedded corrections besides the error location and error types described in Section 6.3.4. This will allow to make the correction decision model more stable.
4. Embedded corrections are based on repetitions, however, we need to keep in mind that repetitions have a function in talk. Repeating a reference to an object or an action in an answer to a question needs to be analysed through the lens of research on references in conversation (Enfield N.J., 2013; Enfield, 2007; Schegloff, 1972) and referring expression generation (Janarthanam and Lemon, 2009, 2010).
5. Further analysis and modelling of embedded corrections on post-expansions and sequence closings after errors in second pair parts would extend the current understanding of their role in interaction. This is a necessary step to prepare computational models for such embedded corrections, not only for ICALL.

12.1.4. Exposed corrections of language form

Correction of language form in a Conversation-for-Learning is very rare, as my dataset shows and as it has already been reported in academic literature, for instance (Hauser, 2010). The majority of all linguistic errors corrected in my dataset were focused on meaning (both exposed and embedded). As Hauser (2010) shows in his analysis of other-corrections of language form in a Conversation-for-Learning, language learners may focus on form later than an error occurs so that a repair sequence is placed between the error and the correction. In such cases the corrections are exposed and the errors are usually repeated in order to renew the context. The difference between focus on form and focus on meaning may be important for the correction decision, however, their sequential position plays a role, too. Further analysis of error corrections with the focus on form as opposed to focus on meaning is required in order to understand participants' decisions and the influence of the interaction process, and with the purpose to make better distinctions for correction decision models in communicative ICALL.

12.1.5. Creating a better correction decision model

A formal decision model for correction based on the preliminary model proposed in Chapter 10 needs further investigations. The following steps may be helpful on the way to a formal definition of a correction decision model. As explained in Chapters 7 and 10, finding a correction decision model was not the focus of the work in the early stage as the data were collected. Therefore, little information is available about participant's motivation to correct and learner's correction perception. Direct connections between participant's feelings about corrections and the data logs cannot be reconstructed from my dataset. A different study focused on this problem may be designed in a way that allows to grasp the important factors in conversation that make a correction relevant. More stable and accurate models may follow from such a study.

Simmons-Mackie and Damico (2008) examined the use of exposed and embedded corrections in individual and group aphasia therapy sessions. They found that exposed corrections occur more frequently in sessions focusing on repairing deficits, as opposed to embedded corrections which dominated in sessions focusing on natural communication events. It will be interesting to see, if similar regularities can be found in Conversations-for-Learning. Because there is a variation within this speech exchange system with regard to distributions of social and pedagogical talk, as discussed in (Tudini, 2010) and present dissertation, a comparative analysis of different types of "sessions" may help to find more features to define a more accurate correction decision model.

A decision to correct is not only a binary yes/no answer, but a selection of a specific correction format. For instance, as argued in Chapter 10, there is a preference for embedded corrections, but they are not always produceable. The correction decision model introduced in Chapter 10 suggests to select a correction format according to the state of social interaction among other factors. A already noticed by Tudini (2010) and supported by this study, all error corrections are intuitively perceived as more intensive or less intensive. An intensity-based model for error corrections might have an advantage of an intuitive mapping from the state of social interaction to a correction format selection. I will share my ideas regarding modelling long-term interaction and social closeness with CA methods in Section 12.3.

12.2. Dialogue Processing and Affective Computing

In this work I focused specifically on text chat interaction with language learners. The same methodology (CA-informed data-driven design and modelling) may be used to approach open problems Dialogue Processing and Affective Computing, such as creation of individual user models, topic detection and tracking and emotion recognition and generation. As argued in Section 2.3.1, the relationships between CA and NLP (Computer Linguistics) was difficult from the beginning. However, optimistic voices have seen a happy common future for these two disciplines. One of the objectives of this work was to strengthen the multidisciplinary connections between the two disciplines. This section suggests further topics of collaborative investigation.

12.2.1. Dialogue topics

Different definitions of "topic" can be found in conversational agent literature, linguistics and social sciences. Academic literature frequently appeals to an intuitive understanding of "topic" when attempts to an automatic detection of what people talk or write about are made, for instance (Dong et al., 2006).

Building on conversational maxims by Grice (1975), Brown and Yule (1983) uses the notion of *speaking topically*, which means "making your contribution relevant in terms of the existing topic framework" (p. 84). The mechanism of *speaking topically* may be noticed in conversations where participants pick-up elements from others' contributions and incorporate them into own contributions (Brown and Yule, 1983, p. 84). A similar conceptual understanding of topic is frequently used in academic literature on conversational agents. For instance, Breuing et al. (2011) defines a topic as *an independent, self-selected category superordinate to a co-constructed sequence of dialog contributions*.

While conversational agent research is mainly concerned with topic detection, classification and selection (don't talk about dispreferred topics!), conversation analysis investigates how to talk on-topic or to change a topic, thus how a talk on-topic is sequentially organised. Schegloff (2007) describes two in principle different organisations for talk on-topic, which are both sequentially post-expansions:

1. Preferred responses are relevant for sequence closure and dispreferred responses are relevant for post-expansions.
2. Preferred responses lead to expansions and dispreferred responses imply sequence closure.

With regard to *speaking topically* (Brown and Yule, 1983), the contributions are relevant in both response versions for each of the cases, however, interactional consequences are different in each case.

To improve the abilities of conversational agents to *speak topically*, both understanding of what the talk is about and the ability to perform relevant next action with regard to topical talk (e.g. sequence closure) are important for conversational agent research. Methods and data used in this dissertation may be applied to create data-driven models of sequence openings and closings as well as topic-proffering utterances.

Action recognition (what the user is doing with a particular utterance) is as important as recognition of what the user wants to talk about. For instance, the adjacency pairs in apology-based face work sequences described in Chapter 3 (learners apologise for their errors) should not be understood by the agent as user's wish to talk about language errors. The preferred responses in the corresponding examples were sequence closures in form of encouragement (e.g. Examples 3.1 and 3.2 in Section 3.1). Such adjacency pairs can be taken as templates for conversational agents communicating with language learners. Other typical adjacency pairs may be found, too.

12.2.2. Emotions in text chat

Symbolic and lexicalised representations of emotions are available in chat to express participants' emotional state. Various attempts have been made to describe and to classify representations of emotions in chat, for instance for German (Orthmann, 2004). However, display of affective states is not the only function of emotions in text chat.

I showed in this work how emoticons may be used as accountings in error correction turns. Vandergriff (2014) analysed pragmatic functions of emoticons in chat between native and non-native speakers of English. The identified functions of emotions are markers of affective stance (joy, happiness, or relief), but also as keying markers or contextualisation cues. Two functions of the latter were identified: (1) orientation to a dispreferred action and mitigate the face-threatening potential of requests and (2) cueing conversational humour, irony and contextual inappropriateness. A similar study may be performed on the corpus used for this work to continue work on pragmatic functions of emotions in native/non-native speaker chat and to detect possible differences in use of emoticons by language learners who are native speakers of different languages. A comparative analysis of multiple existing corpora will be an advantage (e.g. Russian native speakers who learn German, this work, (ELRA, 2015), Swedish native speakers who learn English (Sauro, 2009) and Australian English native speakers who learn Italian (Tudini, 2003)). Differences in types of discourses need to be considered.

Such qualitative studies have the potential to provide empirically grounded classifications and data-driven models for automatic emotion recognition and sentiment understanding for different types of applications, such as sentiment analysis of short text messages (for example product evaluations, Twitter news and Facebook comments), artificial companions for daily care (Wilks et al., 2015; Ren et al., 2014) and artificial coaches (Yasavur et al., 2014).

Vandergriff (2014) outlined possible connections between non-native speaker's use of emoticons and language proficiency. If there are differences in use of emoticons by native and non-native speakers, they may be part of interactional competence in CMC, and need to be acquired as part of the foreign language. Communicative ICALL applications may help to approach this learning goal.

12.2.3. Improving conversational agents with CA

Various problems have been investigated in statistical NLP and CA separately, such as references in conversation, temporal and spatial expressions, paraphrases and formulaic expressions. From my perspective, only insufficient multidisciplinary collaboration can be observed between these two disciplines. However, NLP and Computer Linguistics may greatly benefit from CA results and methods, specifically in dialogue-focused research.

In this work I only covered a small part of the variety of structures in chat conversations, namely two types of repair with linguistic trouble source. However, in order to implement the findings in a real-life application more efforts need to be put into the opposite of them, thus "normal talk". I touched in my research such problems as references in conversation for the purpose of trouble source recognition and correction format generation. A lot of work has been done on references in conversation outside repair sequences. Research results from CA on references in conversation may help to improve existing reference resolution techniques, for instance by more specific models for references to places and persons (Enfield, 2007; Enfield N.J., 2013; Schegloff, 1972) and expert-novice talk (Isaacs and Clark, 1987).

As it was discussed on the example of emoticons in chat in Section 12.2.2, classifications and models for recognition of pragmatic functions of specific structures in conversation could help to improve language understanding models. I suggest several areas where I see a potential knowledge transfer from CA:

1. Work on epistemic stances and states (Heritage, 2012).
2. Questions and answers (Stivers and Enfield, 2010; Heritage, 2012; Schegloff, 1988).
3. The role and pragmatic function of repeats (Perrin et al., 2003; Stivers, 2005).
4. The role of figurative idiomatic expressions in topical changes (Holt and Drew, 2001, 2005).

12.2.4. Chatbots as minimum viable products

Although chatbots have been criticised for their limited capabilities in language understanding (Klüwer, 2011), several attempts have been made to adapt chatbots for the purpose of language learning, for instance (Jia, 2009; Zakos and Capper, 2008) and this dissertation. I discussed in Part II of this work, which types of repair can be implemented even with the simple chatbot technology. This investigation can be continued with regard to other structures in conversation, such as sequence initiations and closings, making appointments and topicalisation.

Specific application scenarios may benefit from employing chatbots for language learning. For instance, acquisition of the recent orthography standard for native and non-native speakers. German orthography has changed several time, language users have difficulties in decisions, what the current standard is (e.g. *am Mittwoch Nachmittag* vs. am *Mittwochnachmittag*). The chatbot technology could be useful to retrieve and learn the recent orthography standard. Similarly, information about specific structures in the native language of the user may be provided to the chatbot (*ogo!* as a news marker and surprise token in the dataset) and the chatbot could help the learner to acquire the corresponding structures in the foreign language.

Several instant messenger service providers announced recently their plans to integrate chatbots for specific purposes into their messengers, first of all in the role of personal assistants (Beer, 2015). The models proposed in this work can be implemented in such a service chatbot to offer conversation training to specific topics in a foreign language.

12.3. Long-term human-machine interaction

Artificial companions (Wilks, 2005), relational agents (Bickmore and Picard, 2005) and companion technologies (Wendemuth and Biundo, 2012) are in focus of multidisciplinary research projects related for instance, to health, sports & fitness (Ståhl et al., 2009), ambient assisted living (Dorr et al., 2015; Caire and van der Torre, 2009), coaching & stress relief (Pulman et al., 2010; Wilks et al., 2015) and second language acquisition (Danilava et al., 2013a). The interpretation of the term *companionship* is for each of them slightly different, but they all have in common the expectation that the companion interacts with the user for a prolonged period of time and adapts its behaviour to user's needs. In an earlier publication, my co-authors and I argued that analysis and modelling of long-term interaction may be a promising approach to companionship (Danilava et al., 2012). My co-authors and I outlined how the challenge of modelling long-term interaction may be approached by methods of Conversation Analysis in our attempt to transfer the concept of *interaction profiles* into the field of Artificial Intelligence (Höhn et al., 2015).

From the classical computer science perspective, interaction profiles are defined by static set of attributes to which values from a specified interval or set of classes can be assigned. For instance (de Alencar et al., 2014) specify user's interaction profile as a set of attribute-value pairs where the attributes are input and output method (e.g. keyboard), physical and cognitive characteristics of the person (e.g. visual impairment and attention deficit), interests, literacy level and age. The interaction profile of each interaction participant as defined by Spranz-Fogasy (2002) emerges during the interaction and is influenced by all interaction participants and the interaction process. I argue that interaction profiles as understood by Spranz-Fogasy (2002) may be a useful concept for long-term human-machine interaction design. In particular, interactional practices of dealing with specific events in interaction can be modelled in a similar way as I formulated local models of correction and explanations. Another model (or maybe many of them) decides, under which circumstances which of the local models may be activated. In this way, a different interaction profile emerges during the interaction with the user. I illustrated this perspective on the example of different profiles of "language experts" in Section 3.4.2. Conversational agents for roles other than a more knowledgeable language user in an equal-power speech exchange system may be designed in a similar way.

I started the discussion on the role of the variations in orthography for the regulation of the social closeness in Section 7.1.2. Further work on medium-specific variations in orthography and its relationship to social closeness on one hand and language standard on the other hand needs to be done. I argued in Section 10.2 that a model for social closeness derived form CA results could be helpful for different applications where long-term human-machine interaction is desired, including Communicative ICALL. However, operationalisation efforts are needed to translate "Language of the closeness, language of the distance" (Koch and Oesterreicher, 1985) into data-driven computational models of social closeness in interaction. Because social closeness and intimacy are found to be interactional achievements (Rintel, 2015; Sacks, 1995), features indexing them can be translated into formal models by methods used in this dissertation.

12.4. Effects on learning and learner corpus research

This work did not focus directly on the effect on language learning, however, it prepared a computational basis to implement longitudinal language learning methodology initiated and elaborated in CA-for-SLA (Kasper, 2004; Markee, 2008). With the findings of this work, recommendations for further directions in learner corpus research can be made. Because the non-presence of data is a hug problem for this kind of research, I make suggestions how to overcome this problem. Finally, I share my vision how the work started here may help to improve the world educational system.

12.4.1. Evidence of language learning

ICALL and communicative ICALL research has been criticised for not paying enough attention to the aspects of learning and not-present evaluation of a learning effect (Petersen, 2010; Wilske, 2014). The evaluation of learning effect by ICALL systems has been approached by methods grounded in SLA theory such as experimental comparison of specific types of corrective feedback provided by human tutors vs. ICALL system (Petersen, 2010; Wilske, 2014). Such experimental methods are not feasible when evaluating a Conversation-for-Learning, because specific moments of learning such as error correction and repair initiations should not be elicited in a free conversation, as opposed to experiments in laboratory conditions. CA-for-SLA suggests a different method to track the effects on learning by tracking learning objects (specific expressions or words) and learning behaviour (specific interactional practices) in longitudinal studies (Markee, 2008). From my perspective, this approach is more suitable for evaluation of effects on learning in communicative ICALL applications.

I suggest two specific directions in learning behaviour and learning object tracking which may be tracked automatically by an ICALL application:

- An analysis of postponed uptake after embedded corrections would help to get insights into noticing issues. Because embedded corrections do not make correcting to the interactional business, there is no direct evidence of noticing of the correction by the learner.
- Immediate uptake after an exposed correction but postponed error repetition will help to track the stability of interlanguage modifications. Usually, this issue is approached by a comparison of results of immediate and delayed post-tests (Wen and Mota, 2015).

Learners' interactional competence may be measured in addition by metrics focused on their ability to perform specific social actions, such as formulating questions, initiating repair and providing appropriate responses to compliments. The latter are very difficult even for proficient speakers (Hauser, 2010) citing (Golato, 2002, 2005). Classifications and typologies of repair other-initiations obtained from native-speaker talk offer a valuable base for comparison, for instance those described in (Drew, 1997; Dingemanse et al., 2014; Benjamin, 2013; Egbert, 2009; Enfield et al., 2013). The present work provides the necessary computational models to track such development and do detect differences automatically in an ICALL application or in a traditional language classroom involving computer-mediated communication.

The *imitation hypothesis* formulated by Aguado Padilla (2002) suggests that second language learning may happen by imitation, similarly to the first language acquisition. On the other hand, conversation participants adapt their language to the language of the co-participants. This phenomenon is referred to as *convergence* (Mitchell et al., 2012). It may be an interesting research question, how these two concepts correlate in communication with language learners.

12.4.2. The data issue

As motivated in Section 2.2.2, data of high quality is an essential part of any data-driven research. Because several datasets of chat Conversations-for-Learning exist, studies requiring additional datasets could be potentially started without waiting for more data. However, the majority of the datasets of interest are not publicly available (Fredriksson, 2012; Tudini, 2010; Marques-Schäfer, 2013). An "open data initiative" for CALL may be beneficial for all CALL and ICALL researchers, but also for learner corpus research and Natural Language Technology.

Given that such an open data initiative is far from being a reality for now, it may be interesting, how to get more examples, for instance, of error corrections, with less data. In Section 11.3.2 I made a comparison between types of corrective feedback in classroom and chat-based Conversation-for-Learning. More examples of correction types present in both chat corpus and classroom data may be obtained from classroom datasets to improve local models of corrections.

Current language learning platforms and social networks supporting instant messaging between language learners and native speakers have large amounts of data of desired type. For instance, Busuu (Busuu, 2008) targeted from the beginning tandem language learners and their community of tandem learners is growing. Text chat is one of the communication options for the learners of different languages who benefit mutually from other's native language skills. Collaboration with such platforms may be beneficial for data-driven research. Facebook is a place, too, where language learning communities meet. New language learning groups may be created with the purpose of data collection.

Wizard-of-Oz methodology is a popular solution to do research about technology which does not yet exist or is not mature enough for tests with users. Wizards normally receive instructions how to behave like a machine, and the users interact with a machine operated by the Wizard. In the beginning of my study I did not know what instructions a Wizard should receive, how to "behave like a language expert". As a result of the present research it can be specified how to "do being a language expert. Some of the ways to "behave like a language expert" can even be implemented. However, we still have limitations given by language technology with regard to talk on topic, error recognition, emotion recognition and generation and so on. This is a state where a wizarded data collection might be reasonably approached in oder to test the models, to collect more data and to see if a machine might be accepted as a language expert by language learners.

With regard to repair annotation addressed in Section 11.3.1, a new version of the corpus is being prepared where all repair types discussed in this dissertation will be annotated. A new annotation scheme needs to be worked out which takes turn-taking, virtual adjacency, repair structure and repair formats into account in order to encode all the information found during this work. This will be a step towards preparing a repair-annotated corpus for supervised machine learning. I am not aware of any such linguistic resource till present.

12.4.3. Language classroom without grades

Markee (2008) proposed a longitudinal methodology in CA-for-SLA to capture learning called learning behaviour tracking. This methodology includes learning process tracking and learning object tracking. What Markee suggests is quite cumbersome for teachers, however, ICALL and other computer-assisted learning tools (not only for languages) may be good teacher's assistants in tracking.

How should this change the world? Grades in school and in higher education change motivation. Students are motivated to get good grades, and optimise their learning strategies according to this goal. Not the learning is important, not the creativity, but the marks. The curiosity, that children inherently have from their birth, and the passion to learn is being killed as soon as they are confronted with grades. I recently read a non-scientific article where the following example of grading was given: one student A writes a text with 20 errors and gets a bad grade, the other student B writes a text with 2 errors and gets a good grade. One semester later, B writes again a text with 1 error and gets again a good grade, a little bit progress is there. But A writes a text with only 8 errors, and gets again a bad grade, however, A's progress is huge.

All people are different, all have different interests and need individualised approach. Teachers are overwhelmed by the amount of work they already have. They do not have time to write a qualitative assessment of learning achievements for each student, therefore, both A and B will just have to be satisfied with the marks, and A will probably be very demotivated, because the efforts and the progress were not noticed by the teacher, and the feedback was generic. Intelligent Tutoring Systems and ICALL may help the teachers not only to track learning behaviour, but also to prepare qualitative assessments as feedback, and to get rid of demotivating marks and frustration in learning.

13. Conclusions

As motivated in Chapter 1, this research aimed at exploring how participants of native/non-native speaker dyadic chat orient to their linguistic identities with the purpose of computational modelling of conversational agents for communicative ICALL in roles other than teachers or tutors. Driven by the initial idea of creating a machine for practicing conversation in chat, the study had the following objectives:

I Find and describe interactional practices in native/non-native speaker chat-based Conversation-for-Learning where chat participants orient to their linguistic identities of language experts and language novices.

II Create computational models of those practices and analyse technical requirements and limitations to implement the resulting models in a communicative ICALL application.

This work shows that methods of Conversation Analysis (CA) can be successfully applied when facing such open research objectives in computational modelling. Specifically, it shows that CA methods can be applied to find typical structures in longitudinal dyadic chat-based Conversations-for-Learning between German native speakers and advanced learners of German as a foreign language. Further, the study shows that the identified structures can be used a basis to build data-driven computational models for conversational agents in roles other than teachers or tutors communicating with users who are foreign language learners.

Repair with linguistic trouble source is one of the identified structures. Because of the emphasised importance of various types of repair for language learning in SLA literature, two types of repair with linguistic trouble source have been selected for a more detailed analysis. The academic publications discussed in Chapter 2 are inconclusive on several vital questions within the discourse on various types of repair in a Conversation-for-Learning. The study sought to answer six of these questions, as specified in Section 1.3. The results of these research endeavour have been discussed in Chapter 11 by comparing the findings with the state of the art.

The methodological novelty of this research is mainly determined by the inclusion of Conversation Analysis into the multidisciplinary research paradigm of communicative ICALL. This chapter evaluates the advantages and limitations of the methodological innovation, it synthesises the results and discusses the implications of the contribution.

I provide a concise summary of the contribution of this research with regard to research objectives and research questions in Section 13.1. I evaluate the presented research and its implications for communicative ICALL in Section 13.2. I make my final conclusions in Section 13.3.

13.1. Findings, conclusions and theoretical implications

The main empirical findings have been presented within the corresponding Chapters 3, 4, 5, 6 and 7 in Part II of this work. Computational models have been discussed in Part III in Chapters 8, 9 and 10. The findings to each of the research questions and research objectives have been discussed in light of the relevant academic literature in Chapter 11. This section will synthesise the findings and the contribution of this research to communicative ICALL and discuss theoretical implications for communicative ICALL and related disciplines.

13.1.1. Findings and conclusions

This research shows that looking at close-to-natural conversations through the lens of CA helps to discover new scenarios for interaction with users taking communicative ICALL as application area. In addition, it has been shown that a translation of findings from Conversation Analysis to "serve computational interests" (Schegloff, 1996) requires conceptual work and is sometimes methodologically tricky. However, the study has shown that computational modelling of specific sequences in conversation can be successfully approached by micro-analytic methods inspired by CA. In particular, it appeared a feasible approach to computational modelling of such rare phenomena as error corrections in informal talk.

Conversational agents for communicative ICALL in roles other than teachers

Chapter 3 has shown that all of the native speakers position themselves as language experts in conversations with non-native speakers. However, all of them do it in a different way, demonstrating that there is no such thing like *the* language expert. The variety of practices where participants' differential language expertise is made relevant has been classified as face work and evaluation, meta-talk about learning and collaborative learning, and repair with linguistic trouble source.

The concept of *interaction profiles* introduced by Spranz-Fogasy (2002) has been used to build a bridge between the worlds of socio-linguistics and computational systems. Because each of the native speakers had only a partial influence on the formulation of the own linguistic identity in conversation, computational modelling of artificial agents that behave like language experts can be approached in the following way. The artificial agent needs access to a pool of computational models for specific practices. Another decision mechanism will then determine, when the agent should make use of particular local models in conversation. This conceptual approach has been applied in this research to analyse and model repair with linguistic trouble source. It resulted in local computational models for explanations and corrections, and a decision model for corrections of linguistic errors.

In light of the discussion started in Section 1.2 and continued in Chapter 3 regarding participants' orientation to their linguistic identities, the following conclusions can be made:

- Modelling interactional practices where participants of an interaction orient to their linguistic identities can be approach independently from the question of the "real" expertise.
- Although it depends largely on the user (language learner), which of the local models of those practices will be finally activated, the machine needs the ability to recognise them and to react appropriately.
- The activation of specific local models will be different for each user, and will lead to a different interaction profile emerging in each case.

Practices and local models of other-initiated self-repair

Chapter 4 has shown that learners make use of a variety of repair initiation formats to display difficulties in comprehension, however, this variety does not cover all repair initiation formats used by the native speakers. In addition, the study has shown that repair initiations may appear immediately after the trouble turn or with a few turns delay. Both types correspond interactionally to second-position repair initiations in oral interaction. Non-native speakers employ lexical and symbolic means to signal troubles. The information about the kind of trouble is included into repair initiations and is made part of repair initiation format in form of references to trouble source. The references to trouble source include adjacent position, demonstrative expressions and repeats. Types of trouble addressed by repair initiations are single words, idioms, abbreviations, and sometimes longer utterances or their parts. With regard to repair carry-outs, Chapter 4 has shown that native speakers may use paraphrases, synonyms and metalinguistic information, but also machine translation into learners native language, translation into other shared languages and examples in form of hyperlinks to other electronic resources in order to deal with learners trouble in comprehension.

Chapter 8 has shown that all identified formats can be divided in two abstract classes. First, they can mark something as unclear and mapped to the content question *What does X mean?* for a trouble source X. Second, they can present candidate understandings and can then be mapped to the polar question *Does X mean Y?* with a trouble source X and the candidate understanding Y. Repair carry-outs need to be designed in a way that provides a projected answer to either the content question or the polar question. The computational model for other-initiated self-repair reflects these findings and allows recognising a repair initiation, extracting the trouble source and generating a repair proper.

In order to specify technical requirements from the computational perspective to simulate such repair sequences, an implementation case study was set-up with an AIML-based chatbot. Chapter 8 has shown that even such simple pattern-based language understanding technology allows implementing recognition of repair other-initiations using the model presented in this work. Linguistic resources open for research have been helpful to retrieve the necessary paraphrases and explanations from hand-crafted large knowledge bases such as Wiktionary. The conversation program simply selected relevant records from the database to generate an explanation. However, more sophisticated NLP techniques are required to cover the whole palette of repairs as described in the empirical chapters of this work. The required NLP technologies include paraphrase generation, word sense disambiguation and machine translation. Extending the linguistic database with hyperlinks to examples of particular concepts will be an advantage.

The study of other-initiated self-repair lead to the following concluons:

- Because repair initiations to deal with troubles in comprehension is a frequent form of orientation to linguistic identities, it has to be an integral part of each machine designed to interact with language learners.
- Repair initiations are closely connected with the system of other factors in interaction, such as turn taking, adjacency pairs, virtual adjacency in chat, membership categorisation and preferences. Therefore, repair initiations need to be modelled as part of this system.
- Responses to learners' repair initiations reflect the assumption that difficulties in comprehension caused the repair initiation. However, it cannot be excluded that learners also may initiate conversational repair.

The model can also be applied in other domains where professional talk is mimicked by an artificial agent in conversations with novices, for instance, to acquire professional terminology.

Practices and local models of corrections of learner errors

Chapter 5 is dedicated to exposed correction formats. The chapter has shown that various factors in interaction influence participants choice of a specific correction format regardless to error type. Such factors include the availability and certainty of a target hypothesis, sequential environment where the trouble turn occurs, and whether form or meaning is the focus of the correction. In addition, Chapter 5 has shown that there are types of corrections which combine characteristics of exposed and embedded corrections. I named them integrated corrections, because they are fully syntactically integrated into the next relevant turn, however, emphasise the action of correcting.

With regard to changes of correction formats over time, this research found other types of changes than previously documented in academic literature. Synthesised with findings reported in literature, the following changes may take place with the time in conversations with language learners:

1. Changes in the use of exposed correction formats from more polite (accompanied by rich accountings) and discreet (on-the-fly) to correction formats requiring less interactional management (simple with minimal accountings).
2. Changes in correction focus: form focus on form to focus on meaning (Dings, 2012).

Changes in the use of embedded corrections have not been observed yet. Chapter 6 has shown that the same embedded correction types as in native-speaker interaction are also present in native/non-native speaker chat interaction. I named these types *conceptual* and *stylistic* embedded corrections. In addition, there are embedded corrections of linguistic errors. The fact that embedded corrections occur more frequently than exposed corrections does not imply that it is due to the speech exchange system or participants' linguistic identities. The occurrence of embedded corrections is determined by the place of the trouble turn within the larger sequential environment, but also by the presence of a clear target hypothesis. Embedded corrections occur in second pair parts to correct errors in the first pair parts. They can occur in responses to second pair parts to correct an error in the second pair parts. Embedded corrections are only analysable in pair with the trouble turn. The embedded correction formats employ special interactional actions requiring repetitions of linguistic material from the preceding utterance. Such actions include repeat-based answers to questions, counter questions, evaluations, acknowledgements, requesting information, initiating repair and expressing surprise. Consequently, embedded corrections cannot be produced if the part of the trouble turn cannot be repeated, in which a correction is needed.

Chapter 9 has demonstrated how the identified correction formats can been classified according to three attributes: grade of explicitness, grade of integration and grade of relevance. In addition, Section 9.4.1 has shown that simple pattern-based language understanding technology may be sufficient to implement the identified exposed correction types in form of correction format templates. In contrast, covering the entire palette of embedded corrections by computational models requires more sophisticated language understanding technologies.

Because embedded corrections in answers to questions are the most frequent type in my dataset, I focused on embedded corrections in answers to polar questions in the modelling phase. Section 9.4.2 shows that simple pattern-based language understanding can be used to correct spelling errors and some types of morphology errors implicitly. Embedded corrections of errors in congruences and lexical choice present a particular challenge for modelling. Their computational models will depend on the abstract language model such as syntax and speech acts.

Integrated corrections in answers to polar questions can be produced when the unit where the correction is located can be highlighted. In addition, accountings may be made part of the correction turn. Section 9.4.2 shows that embedded corrections can be produced even by pattern-based language understanding technology which are not aware of any syntactic units. Error recognition with the AIML technology, without extensions by parsers and taggers, can work on the level of preprocessing where multiple string surfaces are matched to one pattern.

The following conclusions can be made with regard to the study of error corrections:

- The separation in local models for specific practices of correction and a central decision mechanism allows for independent modelling covering a large palette of various correction formats. Modelling other practices in conversation may be approached in a similar way.
- Although the majority of correction formats can be realised as utterance templates, this type of repair is still interconnected with sequential organisation and turn taking (correction on-the-fly) and adjacency pairs (embedded corrections).
- Because some correction formats modify the syntax of utterances (embedded and integrated corrections), it might be reasonable to look at another direction of CA research, namely *syntax for conversation* (Schegloff, 1979) with the purpose to understand, how repair modifies syntax.

Appropriateness of error correction in Conversations-for-Learning

The empirical analysis of factors in interaction which may make an error correction to a relevant action in a Conversation-for-Learning have been discussed in Chapter 7. This analysis shows that the decision to correct is influenced by various parameters in the speech exchange system, communication medium, participants' linguistic expertise, participants' engagement, participants' negotiations and agreements, social proximity, error properties and interaction history. In addition the study showed that all these parameters and their importance for the participants may change over time.

The factors identified in the empirical part of the work have been translated into a feature-based computational decision model for correction. Chapter 10 makes clear that only hand-crafted rules for the correction decision are available for now, however, the model and the communicative ICALL system employing it would benefit from machine learning models. Specifically, machine learning models display in general a more stable behaviour in identification of mutual dependencies among features and re-calculating weights for particular features while keeping and updating the dependencies.

The model differs from other correction decision models in ICALL in the following aspects:

1. Features relevant for the correction decision are not restricted to the occurrence of an error and system's ability to recognise the error correctly, as compared with other communicative ICALL publications (Petersen, 2010; Wilske, 2014).

2. Interaction history and social proximity are included into the decision model in addition to the error properties and user model. The latter two are already part of ICALL applications (Amaral and Meurers, 2011), however they include a different set of knowledge and parameters.

3. The presented model for correction decision works in a specific speech exchange system which has not been previously explicitly considered, as compared to communicative ICALL. The presented models acts between the two extremes, teacher-like behaviour one one hand (Wik and Hjalmarsson, 2009; Petersen, 2010; Wilske, 2014) and zero-correction on the other hand (Wik and Hjalmarsson, 2009).

The study of interactional factors relevant for a correction decision and their computational modelling lead to the following conclusions:

- Micro-analytic methods of Conversation Analysis can be effectively used in order to identify and describe complex mutual dependencies in speakers selection of a particular practice to deal with a particular inout in interaction. This shows in general that the problem of *contingency* (Schegloff, 1996) in interaction "to serve computational interests" can be effectively approached by such methods.

- Because the action of correction has a pedagogical and a social import in interaction, both need to be studied. From the analysis of factors it became clear that social import of corrections is more important in a Conversation-for-Learning than pedagogical import. Therefore the correction decision is mainly determined by social factors. This may change in a different speech exchange system.

13.1.2. Theoretical implications

From what has been discussed so far, I see the following theoretical implications for communicative ICALL research inferred from this research:

- Including Conversation Analysis into the circle of disciplines related to communicative ICALL has several practical advantages reported in this work. However, it also has the disadvantage that a rethinking of the mainstream theoretical paradigm may be needed.

- Because orientation to linguistic identities in chat Conversations-for-Learning is mainly initiated by the learner, the primary focus in computational modelling of such sequences of talk needs to be put on recognition of learner's orientations to linguistic identities and appropriate reactions to them. Active positioning as a language expert by an artificial conversation partner is only rarely appropriate, but also needs attention.

- There are places in conversation where learners' orientations to their linguistic identities are more likely, that is in the beginning of the first talk and after sequences of repair with linguistic trouble source. This locational information may be made part of the recognition model.

- A model to capture social proximity in conversation is urgently needed. Incremental changes of the social proximity need to be covered by the model since there are dependencies between social proximity and error corrections.
- Because only a small number of errors are corrected in a Conversation-for-Learning, the current state of the art in error recognition may provide sufficient information for automatic correction generation in communicative ICALL. There is a tradeoff in the current state of the art in automatic recognition of lexical errors and errors in pragmatics and the empirical evidence that the majority of corrected errors in a Conversation-for-Learning are focused on meaning.
- With regard to the differences in complexity between exposed and embedded corrections, the following can be observed. Embedded corrections are more frequent and more preferred but they are harder to operationalise and to model and to implement. Regular expressions are sufficient to describe exposed corrections while embedded corrections require syntax-based modelling (with a few exceptions). Therefore, more theoretical investigations are needed to reach practical applicability of the entire findings of this work.

I see the following theoretical and methodological implications of this research for SLA:

- This research supports findings of other CA-for-SLA such as (Markee, 2000; Kasper, 2004, 2006; Markee, 2008; Vandergriff, 2013) stating that the existing conceptual definition of learning and the experimental approach as the only research instrument are not sufficient to analyse naturally occurring opportunities for language acquisition. Specifically, occurrence of deviations from language standard, error corrections and their interactional import have no place in the traditional SLA and require a cardinal rethinking of the underlying concepts of errors and corrections.
- Orthography, vocabulary and syntax are interactional resources used to regulate social closeness and update participant's identities, and not barely a display of language proficiency or result of production pace. A closer interdisciplinary collaboration between CA and SLA may help to explain *why* some deviations from a language standard occur in chat. Simply to register *that* deviations are present and conclude that they are caused by lack of knowledge is not sufficient.
- Uptake after embedded corrections is documented by socio-linguistic data, therefore, embedded corrections and their effect on learning should be put under the loupe of SLA research.
- Postponed uptake documented in this research may make it difficult to provide empirical evidence of learning. Automatised tracking of learning objects may offer an additional value for both, SLA and CA-for-SLA.

After all, it is important to bear in mind that reacting to a repair initiation with a repair proper is the preferred response, but not the only possible. That reacting to an error with a correction is one, but not the only acceptable way to continue the interaction. This study did not aim at simplification of the complexity of the talk, but it tried to break down the complexity to manageable pieces. Therefore, I see the following theoretical implications for NLP, specifically dialogue research:

- Natural Language Understanding in dialogue implies understanding of the social action encoded in an utterance. Speech acts are *the* model used in NLP to manage this issue (Schiffrin, 2005). However, CA researchers found that there are social actions not covered by the speech act classifications, such as pre-offers, pre-tellings and other pre- sequences (Schegloff, 1996). However, an operationalisation of the new concepts for the computational purposes with CA methods is costly.
- Disengagement of specific triggers in talk and practices of dealing with these triggers can be handled in a similar way as it has been done for errors and corrections in this work. This disengagement helps to manage the complexity in form of concise descriptions of *practices*, that is potentially eligible forms of producing specific actions. Another mechanism can then decide when each of them can be activated.

Furthermore, I see the following implications of this work for leaner language and learner corpus research:

- Learner corpus research needs to go beyond error annotation, especially for conceptually oral corpora.
- Distinguishing between the objective deviation form the written language norm and conventions applicable for a specific speech exchange system and communication medium.
- Taking the radically emic perspective in error annotation in conceptually oral corpora will allow to deal with potentially multiple target hypotheses.
- Because data of good quality, that is, replicating the speech exchange system of interest, are crucial for CA-informed modelling, the methodology of data collection is always important and should be described in detail.

13.2. Recommendations for future communicative ICALL

In this section I evaluate the advantages and disadvantages of the method applied for this research. I analyse opportunities and risks of the method for future studies in communicative ICALL. I make use of a structured planning method known as SWOT (Strengths, Weaknesses, Opportunities and Threat (Mintzberg, 1994)) to identify strategic implications of this research for communicative ICALL. In Section 13.2.1 I analyse internal and external positive and negative factors. In Section 13.2.2 I make conclusions on strategic implications following from the analysis.

13.2.1. Strengths, weaknesses, opportunities and threats

Table 13.1 presents the SWOT canvas to evaluate the applicability of the used method for communicative ICALL research and to identify strategies for future development of communicative ICALL based on this study. The evaluation is determined by a detailed analysis of internal and external factors. Internal factors include strengths and weaknesses of the research method itself. External factors include opportunities and risks given by the environment.

The top half of the SWOT canvas in Table 13.1 describes the situation with regard to advantages and limitations of the method used in this research. The bottom half lists advantageous and disadvantageous factors determined by state of-the art, language complexity, factors related to potential users, existing machine learning models and efforts required for obtaining a larger number of examples for machine learning.

I see the *contingency* as described by Schegloff (1996) as both risk and opportunity for AI, and specifically communicative ICALL. Contingency is the reason why such primitive conversation machines like chatbots work. Contingency is also a reason why chatbot users experience conversations with chatbots as unpleasant. The most important strategic implication is the need to handle contingency, to make it manageable.

13.2.2. Strategic implications and recommendations

The SWOT canvas is rather a static representation of the situation not including any recommendations to actions. However, different strategies for a future development can be elaborated based on the description of the situation. The *matching strategy* builds on the combination of strengths and opportunities and searches for new chances. The *transformation strategy* combines weaknesses and opportunities in order to transform weaknesses to new strengths. The *neutralisation strategy* combines strengths and threats with the purpose of risk prevention. Finally, the *defence strategy* combines weaknesses and threats to prevent identified weaknesses from becoming a target for the risks. I identify the strategies and make recommendations for future research in reminder of this section.

The strategies should always be developed towards a concrete goal. The goal in my case is to create new and to improve existing communicative ICALL systems by CA-informed computational modelling. I formulate the strategies in form of DOs and DONTs.

Strengths	Weaknesses
• Hypothesis formation with regard to new user and system models. • Structures and prototypes from naturally-occurring data. • Takes special characteristics of the speech exchange system into account. • Micro-analysis allows to detect new features and dependencies. • Separation in local and global models allows to extend both easily. • Long-term changes can be observed.	• Voluntary participants from only one population group. • Only one language and medium. • Micro-analysis is time-expensive and exhausting. • Sparse information about the correction decision and social import of corrections due to study design. • A very small number of examples may be problematic for modelling. • Manual creation of rules of mutual dependencies is tricky. • Comparability to other methods, quantitative validation is tricky.
Opportunities	**Threats**
• Contingency. • Dealing with complexity in form of role play. • Current NLP tools allow to make good models and applications for well-defined short tasks. • Integration in an ICALL system as an additional service. • Big data, deep learning. • More data for various other populations may be found in current ICALL applications.	• Contingency. • Interaction with a machine may work completely different. • Learners may not accept the machine as a language expert. • More data does not imply more examples. • Trade-off in automatically detectable error types and most frequently corrected error types.

Table 13.1.: SWOT-analysis

New chances for development

To find a strategy for identification of new chances for development, strengths and opportunities are combined.

DO Integrate well defined short dialogue models to practice specific situations (e.g. mentioned in Section 12.4.1) into existing work-book-like ICALL systems.

DO Develop short role-play-based situations for specific speech exchange systems where participants' identities (not only linguistic identities) are relevant, for instance a job interview in the language-to-be-learned.

DO Develop methods for long-term tracking of learning objects and learning behaviour.

DO Create partnerships with data holders from which communicative ICALL can learn.

DO Use qualitative approaches to evaluating learners' performance.

DONT Rely on grades.

Transform weaknesses to strengths using opportunities

The *transformation strategy* combines weaknesses and opportunities in order to transform weaknesses to new strengths.

DO Create collaborations to obtain sufficient data to cover user groups and languages not covered in this study.

DO Look even at very rare phenomena, because a small number of examples helps to identify at least the phenomenon in focus and to understand their role and structure. Then try to find more examples in other datasets, where such phenomena are more likely to be found.

DO Apply machine learning approaches to dependency identification and weight estimation.

Neutralise risks by strengths

The *neutralisation strategy* combines strengths and threats with the purpose of risk prevention.

DO Apply qualitative evaluation methods to dialogue assessment.

DO Setup a wizarded study. Provide the Wizard with instructions identified here.

DO Identify most frequent errors with focus on meaning and target them in modelling.

DONT Work with artificially created examples.

Defend against risks targeting the weaknesses

The combinations of weaknesses and threats allows preventing identified weaknesses from becoming a target for the risks.

- DO Apply contingency to deal with contingency. Develop "back doors" to deal with inputs that the system does not cover explicitly on the necessary level of detail. Rely on similar strategies observed in naturally occurring data.
- DO Use existing findings on structures of interest to set-up new data collections (specifically design of interviews and questionnaires) in order to get sufficient information from the data. Set-up a new study targeting the correction decision.
- DO Take the speech exchange system into account.
- DONT Promise practicing free conversation at the current state of research.
- DO Approach modelling "free conversation" by practicing specific actions, for instance, variants of possible reactions to a news telling.

13.3. Final conclusion

This work continues the research efforts bringing together Conversation Analysis and Artificial Intelligence which have been started in Human-Robot Interaction but received not enough attention in communicative ICALL. In this way, innovative software to improve users' experience can be created. Models of specific interactional practices and their pragmatic function will provide the agents the ability to react appropriately to specific social actions. The same can be performed for various speech exchange systems and interaction modalities in the same or a similar way as it was shown in this dissertation. Conversational agents for specific purposes in specific interactional roles such as personal assistant, fitness coach and housekeeper can get closer to users expectations.

Although the study shows that specific phenomena in interaction may be discovered, described and modelled for computational purposes with CA-informed methods, it should not be understood as a simplification of the complexity of the interaction. I choose to start closing this work with a quotation which is applicable to this research as much as it is an integral part of the original article:

> The several themes to which I have called attention all involve a major challenge to computational interests in discourse. [...] The challenge is *contingency*. Although the organisation of talk-in-interaction is orderly [...], it is characterised by contingency at virtually every point. [...]
>
> One underlying "burning" issue for computational interests in discourse analysis is how to come to terms with the full range of contingency which talk-in-interaction allows and channels. [...]

> The problem of contingency [...] poses truly formidable obstacles to computational approaches. But if some useful interchange between these modalities of work is to be realised, it is most likely to come not from transforming the object from which you would like to learn, but from taking it seriously in its own terms. In the end, it will be the computationalists who will have to figure out how to do this. (Schegloff, 1996, pp. 21-22, 29)

Being a "CA-informed computationalist", I see CA-informed modelling as a sustainable approach of dealing with *contingency*. This research was not an attempt to simplify the object of studies, namely conversation, but an attempt to break down the complexity to manageable pieces. I took the object of studies seriously, and I approached contingency by distinguishing between local models of corrections and a decision mechanism which is able to select correction as one of the options in the space of contingency. I showed how this approach can work on the example of chat conversations with language learners.

Part V.

Supplements

A. Data

In this supplement chapter I describe the process and the results of the data collection. In an attempt to resolve the problems that exist in investigating chat conversations, this Chapter presents the data capturing approach that was chosen to overcome the limitations of instant messaging dialogue datasets.

A.1. Data collection

This section documents the study design process, the decisions that had to be taken and the issues. Prior to set up the data collection experiment, the data collection tools and three interaction scenarios were tested in a pilot study. Two candidates for the data collection method were considered:

Wizard-of-Oz experiment: human operators simulate technology, usually performed when the research is focused on user's behaviour while the desired technology does not exist (WOE).

Qualitative experiment: natural interaction modified in a controlled way (QE).

Based on the results of the pilot study, the decision was taken, which of these methods should be used in the longitudinal study. Both methods are used for data collection where human behaviour in a near-to-natural environment is in focus, however the former is more frequently used in HCI domain (Gould et al., 1983; Bradley et al., 2009), and the later is more commonly used in social science (Lamnek, 2010).

A.1.1. Study preparation

During the pilot study the following scenarios have been tested:

$S1$ Interaction flow for the WOE setting where a German native speaker plays the role of an artificial conversation partner pretending to be a chatbot.

$S2$ Interaction flow for the QE setting without any additional instructions.

$S3$ Interaction flow for the QE setting with an additional instruction for the native speaker of German to correct language mistakes of the learner.

These scenarios have been tested on five pairs of voluntary participants. The Wizard participant in $S1$ setting was told to behave like a computer program without further instructions. The participants in settings $S2$ and $S3$ had to communicate over a forwarding chatbot system, they had an appointment for the chat made by the researcher, and the length of the chat was given. In a natural chat conversation, these restrictions usually do not exist. After each interaction, the researcher took a semi-structured phone interview with each of the interaction parties. The goal of the interviews was to find out how the participants feel while talking with a person who they have never met before, what was perceived as pleasant, what was disturbing or annoying. The interviewer made notes for questions and answers The interviews have not been recorded. All interviews were based of the following frame questions:

1. Was there something difficult, unpleasant or annoying?
2. Was there something outstandingly pleasant?
3. Was it difficult to find a topic to talk about?
4. If the other party were a chatbot, would you like to chat with such a chatbot? (Only for $S2$ and $S3$.)

To parametrise a "pleasant" interaction, the researcher performed a set of chat-interviews and a questionnaire survey with two questions:

$Q1$ What is for you a pleasant interaction or a pleasant dialogue?

$Q2$ How can you describe unpleasant interaction or an unpleasant dialogue?

The non-native speaker participants were all Russian native speakers. They were allowed to give their answers either in German or in Russian, they could select the language in which they could better express themselves. German native speakers gave the answers in German.

Results of the Pilot Study

The interaction became unnatural if one of the parties thought that she/he talked to a computer program, like it was in $S1$. For example, the participants tended to test conversational abilities of the machine in order to find out, what the machine does not understand. They also tried to hide some personal information or used grammatical forms that they thought would be better understood by the machine. In addition, it was unclear to the researcher which specific instructions might be necessary for the Wizard. It was clear as a result of this part of the pilot study, that an answer to an open question "What exactly is *doing being a language expert* in text chat?" is important in order to be able to formulate a functional specification for the computer program.

For $S3$ setting, it was difficult for the participants to do mistake correction, because they did not want to be impolite or annoying. For $S2$ setting, the participants reported that they have had a nice conversation, but the situation itself ("let's chat") was unnatural, they did not know what to say at the beginning.

The researcher selected the $S2$ setting for the final study, but we were aware of this risk. Furthermore, other risks for the data collection were identified and are described in the next section.

In order to create the connection between the data and participants' personal perception of the interaction, the researcher created a questionnaire based on the interview and survey outcome that needed to be filled in after each interaction by each of the participants of the longitudinal study. The questionnaire contained three open questions concerning a general description of the chat and pleasant and unpleasant moments of it, and two quantitative questions where the participants need to quantify their feeling of the interaction (how pleasant was it in general, did the participant feel well understood by the partner, were there interesting topics etc.).

The risks for data collection

There are several risks for data collections and data quality that had to be taken into account for the longitudinal study. The major risk for data collection in this free chat interaction scenario was that study participants may not chat at all. I discuss here other risks that have to be kept in mind when targeting longitudinal data.

Topics: it is not so easy to find topics for communication with persons whom one has never seen or talked to before. Therefore, it was not granted that the participants can "just chat" for period of time long enough for the analysis. It was possible that the participants would lose their interest or would not know what to say. For this case, the researcher planned to suggest some topics for the parties.

Motivation and relationship: the participants could lose their motivation if the interaction was unpleasant or boring; if they could not understand their partner or if they would not match at the interpersonal level. For this case the researcher planned to change the partners.

Language: despite the instruction to use the target language for communication, it is not granted that the participants would not interact in a different language. This risk can be reduced by choosing a population which is not multilingual. In my data collection, both native speakers and non-native speakers of German speak other languages and thus, this risk could not be avoided.

Learner language: although all learners were already at an advanced DaF level, it was still a foreign language for them. I could not assume that they would use the lexical resources correctly. Misinterpretations of native speakers' utterances (e.g. social signals) are always possible. However, it is another important aspect for user modelling in ICALL applications.

A.1.2. The longitudinal instant messaging dialogues

CMC studies show that unlike face-to-face conversation where participants can spontaneously speak to each other, chat interaction presents difficulties related to the non-natural aspect of electronic devices (Loewen and Reissner, 2009). In order to make the IM data as natural as possible, I emphasised freedom of the participants during their selection and throughout the experiment. Hence, all participants were volunteers to the experiment and the only instruction they received was “just chat!”. As noted in the preceding section, the principal risk of this method is that participants might not chat at all. The communication has been set up as explained below.

Participants, instructions and software

Participants were four native speakers of German (G1) and nine advanced learners of German as a foreign language (G2). NS were all from my private circle and learners were contacted through their university in the Republic of Belarus where they studied German as a foreign language in order to become German teachers. The communication started in May and ended in August 2012. Prior to starting the communication, participants had to sign the agreement that their IM talk will be recorded, and that an anonymised version of it will be used later on in a publicly available corpus.

Participants needed only a Jabber or a Google account on their electronic devices, and they could use every IM client of their preference or chat via Google Talk web site. Participants were allowed to chat using any device (PC, laptop, smartphone or tablet) and they could switch these devices at their convenience. There were no time constraints and participants were also free to chat whenever they wanted during the weeks of data collection. There was no other instruction given to participants than "just chat!".

Connection between participants occurred through the chatbot because participants did not have the direct private address of their partners, rather the address alias registered by the chatbot. Figure A.1 illustrates how the users were connected, how the communication was designed, and what the role of the researcher was. If the NS sent a message to the NNS, technically, the message arrived at the NNS alias address of the forwarding chatbot, and the original message was immediately forwarded to the addressee. A copy of the message was automatically saved in the database. The forwarding chatbot was not visible for the participants who can chat as if they were connected directly. However, the participants were informed, that there was a special system, which connected them with their partners. The forwarding chatbot was implemented using Google Web Toolkit and hosted on Google App Engine. These computer tools are offered by Google for free and are easy to use.

Participant	Age	Gender	Native language	Other languages	Occupation	Experience with IM
L01	22	F	ru	de, en, be	Student	No
L02	22	F	ru	de, en, be	Student	No
L03	22	F	ru	de, en, be	Student	Yes
L04	23	F	ru	de, en, be, it	Student	Yes
L05	20	F	ru	de, en, be	Student	Yes
L06	22	F	ru	de, en, be	Student	Yes
L07	22	F	ru	de, en, be	Student	Yes
L08	22	F	ru	de, en, be	Student	Yes
L09	22	F	ru	de, en, be	Student	Yes
N01	27	M	de	en, la	Student	Yes
N02	27	F	de	en, es	Teacher	Yes
N03	22	M	de	en, fr, lb	PhD student	No
N04	22	M	de	en, fr, es	Physicist	Yes

Table A.1.: Participants' demographics in the longitudinal study.

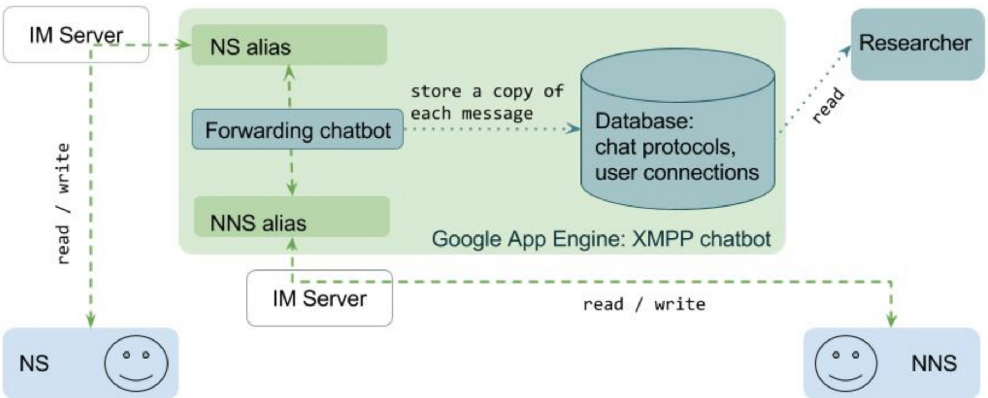


Figure A.1.: Connections between participants using the forwarding chatbot.

Design of the study

Connections between the participants did not overlap: each participant from G1 communicated with exactly one participant from G2, whereas each participant from G2 communicated with two or three participants from G1. The communication was established over the forwarding chatbot described in the previous section. Though the difficulties that occurred at connecting the parties, participants could communicate with the same partner for the complete duration of the experiment. Since the participants communicated voluntarily in their leisure time, the total amount of time that they could invest in the experiment was subject to personal agreement with the researcher. Nevertheless, the lowest bound of time per connection was set to two sessions per week, each 30 minutes. The experiment was planned for 4 weeks, which corresponds to at least 8 sessions of chat per pair.

For the first chat, the researcher coordinated the interaction between the parties. The pairs were built only according to participants' availability for the first interaction since the parties did not know anything about their partners at the beginning. After each chat session, each of the parties had to fill in the questionnaire (online feedback form) described in Section A.1.4. The participants were allowed to answer the open questions in German or in Russian. This bilingual-mode of filling in the questionnaire was more convenient for the Belorussian participants because sometimes they could better express themselves in their native language.

After eight sessions were completed, the researcher interviewed each participant separately via IM. Considering the advantages and disadvantages of IM interviews (Voida et al., 2004), I looked at learners' motivation to communicate and to engage in, native speakers' motivation for "doing being a teacher", e.g. in providing error corrections and explanations of linguistic matters. All interviews were semi-structured, based on a set of frame questions provided in SectionA.1.3 for completeness.

A.1.3. Questions for retrospective interviews

Frame questions for the retrospective IM interviews:

1. Was there something difficult, unpleasant or annoying?
2. Was there something outstandingly pleasant?
3. Was it difficult to find a topic to talk about?
4. If the other party were a chatbot, would you like to chat with such a chatbot?

Other questions were included into the interviews when they were relevant, for instance "Why did you decide not to correct errors?" to the NS and "Did you feel like the chat helps you with the language?"

A.1.4. Questionnaires

1. Wie angenehm war das heutige Chatgespräch insgesamt? (Scale)
2. Gib an, inwieweit die folgenden Aussagen für den heutigen Chat zutreffen:
    a) Wir haben interessante Themen / ein interessantes Thema besprochen.
    b) Das Thema des Gesprächs war mir unangenehm.
    c) Wir haben Spaß miteinander gehabt.
    d) Es war schwer, ein Gesprächsthema zu finden.
    e) Ich habe mit meinem Chatpartner nur gesprochen, weil man das von mir erwartet hat. Eigentlich habe ich besseres zu tun.
    f) Ich verstehe meinen Chatpartner gut.
    g) Mein Chatpartner versteht mich gut.
    h) Wir reden oft aneinander vorbei.
    i) Wir sind auf derselben Wellenlänge.
    j) Ich war erleichtert, als die 30 Minuten um waren.
3. Gab es etwas im heutigen Chatgespräch, was Dir unangenehm aufgefallen ist oder Dich gestört hat?
4. Was fandest Du bei dem heutigen Chatgespräch besonders angenehm?
5. Wie würdest Du den heutigen Chat beschreiben?

A.1.5. Technical issues

Technical support was provided by the conducting researcher in the study and participants were invited to contact her in case of any difficulty. The following issues were identified during the experiment:

1. Connection: geographical moves during a chat created problems with the Internet connection and interrupted conversations. In order to minimise these problems, participants informed their partners about their moves and specific locational circumstances (e.g. entering a tunnel while travelling by train).

2. Awareness of presence: unlike classical IM-clients, which offer the possibility to sent an automatic status update to all contacts as an interactional resource (is somebody online or offline?), the forwarding chatbot always occurs online and participants cannot see each other directly. However, participants quickly learned how to deal with this inconvenience by sending a presence request like “Are you there?”, for instance.

3. IM service reliability: when IM provider services were not available, the connection was disrupted. This might lead to misunderstanding between the participants who did not know that the problem was due to the technical system rather than to their partners.

4. Connection between users in the database. One of the NS specified for the connection his private account at a free Jabber server. However, problems with IM service availability and reliability of this provider caused an account change. It was necessary to change the connection between this NS and all his partners in the DB. An additional unforeseeable issue was faced as Google offered all German users to change from *googlemail* to *gmail* domain. One of the NS made use of this offer and did not think on the relevance of this change for her connection with her partners. The researcher had to clarify why the message exchange with this participant was disturbed.

A.1.6. Participant-related issues

- Missing appointments: it refers to situations where participants could not come to their appointments for chat and did not inform their partners.
- Time investment: participants might underestimated the time that they needed to invest in the experiment and would not be able to fulfil their commitments. Their partner reported that they felt uncomfortable.
- Offline messages: the majority of participants was familiar with instant messaging in Skype and Facebook, but new to Jabber and Google Talk. Hence, they could not find messages that they had received while offline and would ask for help although they were provided with a help-document at the beginning of the experiment. Problems related to offline messages created technical support that was very time-consuming for the researcher due to the wide variety of web browsers and IM clients and hardware.
- Duration: the experiment took longer than the planned 4 weeks. Some sessions were longer than 30 minutes (up to 90 minutes) and this might be an issue for specific research questions that could be approached based on the presented corpus.
- Code-switching: the instruction to the participants was to chat in German as the target language in the study. Nevertheless, code-switching occurs in English and Russian within the dialogues in German. Both NS and NNS used repair in word

search and meaning negotiation sequences. NNS used English translation as an interactional resource to explain meaning of unknown German words to the NNS. One NS used Russian in greetings, farewells and repair sequences. In addition, NNS paired not with the Russian learner used Russian words for greetings and farewell in order to teach their partners a bit Russian.

- Responsiveness: it refers to uncontrolled elements that lead to longer time intervals between messages, and which influenced the interaction. For instance, participants might be busy while chatting or the network delays may slow typing in a different language than the usually used language in the same device.

A.1.7. Ethical issues

At the beginning of the experiment, participants were informed about the purpose of the study, the recordings and the use of the produced chat protocols for research endeavour. Only those people who signed an agreement could participate in the experiment. Participants' privacies was protected and all email addresses and other identification possibilities were replaced after the data collection. The researchers only have access to the anonymised copy of the chat protocols. The demographic information about the participants was encoded in the corpus as provided in Table A.1.

A.2. Corpus creation

The attributes of each message stored in the database are timestamp, sender ID, receiver ID and message body (UTF-8 encoding). The log files for each pair of participants consist of eight dialogic sessions and some single messages that were not answered (non-dialogic). The corpus consists of 72 dialogues (8 dialogues by each of 6 pairs, 9 dialogues by each of 2 other pairs, and 7 dialogues by the 9th pair), which correspond to ca. 2500 minutes of instant messaging interaction, 4548 messages with ca. 52000 tokens in total, and ca. 6100 single tokens both calculated with wordpunkt_tokenizer of Python Natural Language Toolkit (NLTK) (Bird et al., 2009). Three aspects are specifically important for longitudinal CMC studies for SLA:

1. Data quality: the chatbot-in-the-middle approach allows for the instant capture of messaging dialogues and at the same time preserves the natural way of IM communication already adopted by the users, i.e. the freedom of participants to chose IM client software and hardware, time and location for communication is not disturbed. Moreover, participants deal quite competently with the limiting effects of technology-mediated communication like network delays and connection interruptions .

2. SLA aspects: the chatbot-in-the-middle approach allows for the analysis of learning during instant messaging between NS and NNS of German. Hence, corrective feedback (Lyster et al., 2013), meaning negotiations (Varonis and Gass, 1985), embedded corrections (Jefferson, 1983) were analysed as described in detail in this dissertation. Following the imitation hypothesis (Aguado Padilla, 2002), I observed that learning takes place in chat also by imitation.

3. Long-term interaction: having a nice conversation and curiosity is the most important motivation for long-term interaction. "Being taught" by the NS is only a secondary aspect for the NNS. However, it was perceived as helpful by NNS if NS and NNS engaged in talk about linguistic matters. According to Loewen and Reissner (2009), participants may have difficulties in CMC scenarios. Our analysis of CMC with the chatbot-in-the-middle shows, however, that instant messaging participants may deploy different intelligent strategies in order to deal with technology-mediated communication.

The dataset was annotated in the beginning in the following way:

1. Dialogue moves in sequences of corrective feedback and meaning negotiation.

2. Partial error annotation.

Two versions of the corpus were published, with and without annotations (ELRA, 2015). A new repair-annotated corpus version is planed for publication.

The annotation was performed by two independent annotators in August-September 2013. The first annotator was German non-native speaker with a near-native fluency in German and Russian as native language with strong professional background in linguistics, natural language processing and language teaching. The second annotator was German native speaker with no professional background in linguistics but with strong knowledge of the German language. In August 2014 and in February 2015, the annotation was rechecked by the first annotator, several annotation errors and ambiguous cases were corrected.

A.2.1. TEI-P5 modules and customisation

TEI-P5 standard already contains several customisations for spoken interaction data, poetry, linguistic corpora and drama. Chat or instant messaging data have similar features with some of them, but none of the existing customisations could be used without modification and with the appropriate semantics of the tags. Therefore, a decision was taken to create a new customised annotation scheme using the existing TEI customisation for linguistic corpora as a basis. This allows for continuous extension of the annotation of linguistic phenomena in the dataset.

For the purpose of the related PhD project only annotation of repair sequences in chat has been performed, the TEI schema was customised according to the requirements annotation (see Section A.2.4 below).

The corpus is provided as a set of 10 files: one root file containing the description of the corpus and information about participants (TEI header), and 9 files with chat logs produced by 9 pairs of participants, one file per pair. Each file containing chat logs includes all dialogues produced by one pair of participants. The root file contains links to each of those files.

A.2.2. Text replacements

In general, the original spelling and textual symbols used are kept as produced by the used. However, there are a few exceptions made for the purpose of storing the data in XML format and data analysis. All the replacements are summarised in Table 2.

Original	Replaced by
&	&
All posted hyperlinks	HYPERLINK
Facebook ID	FACEBOOK_ID_{LXX,NXX}
Email address	EMAIL

Table A.2.: Text replacements

A.2.3. Chat structure in TEI-P5 XML

The goal of this annotation was to provide a TEI-P5 conform encoding for chat data. Two new tags were introduced:

- `<message>` contains the text of one instant message produced by a chat participant OR more than one non-empty message line (`<ml>` tag). Message lines were introduces for the cases where chat participants insert line breaks in their messages. Different lines may relate to different previous messages of the partners, and need to be linked separately. Important attributes of a message are sender, timestamp and id. The sender is specified by the standard TEI attribute `who` and is linked to a chat participant listed in the root file. The `timestamp` attribute specifies the server time when the message arrived at the server (time zone GMT+0, needs to be recalculated to determine factual time in the time zones of the chat participants - Germany and Belarus).
- `<ml>` contains a message line if and only if the sender of the message inserted breaklines in the message.

The corresponding schema is contained in the file `tei_corpus_chat.rng` which is provided with the corpus.

A.2.4. Chat log annotation

For the purpose of the PhD project, the data have been annotated according to the two classifications:

1. Corrective Feedback as explained by Lyster, Sato Saito (2013). The sequences containing CF usually consist of three types of interactional moves described in classroom research literature: error, correction, uptake.
2. Meaning Negotiation as introduced by Gass and Varonis (1985). According to this model, a MN sequence is composed of 4 moves: trouble source, indicator, response, reaction to response. The messages were labelled according to these types of moves.

All moves in these sequences may consist of several turns (messages). The TEI `note` tag us used for explanations of some complicated cases of annotation (ambiguities, complex sequences).

All of these sequences are repair sequences from the point of view of Conversation Analysis. However, in the beginning of the related PhD project the data analysis was influenced by the language classroom research and Second Language Acquisition theory.

Corrective Feedback

Types of Corrective Feedback (CF) adopted for the PhD work are:

- conversational recast,
- repetition,
- clarification request,
- explicit correction,
- explicit with metalinguistic explanation (MLE),
- didactic recast,
- metalinguistic clue,
- elicitation,
- paralinguistic signal.

The types of the corrections are explained in the article by Lyster, Sato, Saito (2013). Not all of them occur in the dataset because of the text based nature of the chat and due to an informal conversation contrasting to a teacher-fronted classroom where the original classification was obtained from. Metalinguistic clues, elicitations, repetitions and paralinguistic signals were not found in CF-sequences in this dataset.

An example of an annotated CF-sequence is provided below:

```
<im:message xml:id="L06N0320120710-281" who="deL1L2IM-root.xml#L06"
timestamp="2012-07-10T08:44:09">
man kann versuchen </im:message>
<im:message xml:id="L06N0320120710-282" who="deL1L2IM-root.xml#N03"
timestamp="2012-07-10T08:44:29">
[[wir koennen es versuchen]] </im:message>
<im:message xml:id="L06N0320120710-283" who="deL1L2IM-root.xml#N03"
timestamp="2012-07-10T08:44:32"> :-) </im:message>
<im:cfseq>
<im:cfturn turntype="ts" corresp="L06N0320120710-281"/>
<im:cfturn turntype="cf" corresp="L06N0320120710-282"
cftype="explicit_correction"/>
</im:cfseq>
```

No uptake was produced in this sequence.

Partial error annotation

The error annotation has been performed in place (the error is tagged where it occurs). If an item is missing (missing main verb or missing prefix), the content of the tag is empty. The following error types have been annotated in the corpus (all occurrences of them, not only the corrected ones).

1. Morpho-syntactic errors:
   a) Missing main verb in Futurum 1: missing_main_verb_futur1
   b) Wrong word order in a sentence. Only the following types have been annotated:
      - Wrong position of the main verb in the main or subordinate close, or missing main verb.
        Possible types:
        missing_finiite_verb_main_clause
        missing_finiite_verb_subordinate_clause
        position_finiite_verb_main_clause
        position_finiite_verb_subordinate_clause
      - Wrong posistion of or missing separable prefix in verbs.
        Possible types:
        missing_verb_prefix_main_clause
        missing_verb_prefix_subordinate_clause
        position_verb_prefix_main_clause
        position_verb_prefix_subordinate_clause.
2. Lexical erros: only lexical errors in collocations. Definition, classification and types for locations: s. Paper "Towards a Motivated Annotation Schema of Collocation Errors in Learner Corpora". Examples of collocations: frei haben, leicht fallen, Fliege machen.
   Possible types:
   substitution | creation | synthesis | analysis | different_sense
   Possible locations:
   base | collocate | collocation

```
<im:error errtype="position\_finite\_verb\_subordinate\_clause"
target="man kann darüber recherchieren" corrected="NO">
man darüber rescherschieren kann</im:error>.
```

NOTE: The error annotated above can be analysed and corrected in a different way (multiple target hypotheses). An alternative annotation of the same error is shown below.

```
<im:message xml:id="L07N0320120716-602" who="deL1L2IM-root.xml#L07"
 timestamp="2012-07-16T19:29:15"> [content hidden for the example]
 und das wichtigste ich müsse so spät wie möglich heiraten,
<altGrp>
 <alt>
  <im:error errtype="substitution" location="collocate"
   target="sonst geht alles kaputt" corrected="NO"/>
  </alt>
 <alt>
  <im:error errtype="missing_main_verb_futur1"
   target="sonst wird alles kaputt gehen" corrected="NO"/>
 </alt>
sonst wird alles kaputt
</altGrp>
 ...)))))) xD </im:message>
```

Meaning Negotiation

According to the model of a Meaning Negotiation (MN) sequence suggested by Varonis & Gass (1985), the classification of the dialogue moves includes the following classes:

- Trouble-source: the problematic item that is not clear and needs a clarification;
- An indicator that something previously said is not clear;
- A response to the indicator (normally a clarification or an explanation);
- A reaction to response (for example an acceptance or a rejection of a term).

This model is very simple, MN sequences can be very complex, the indicator for example can also be a trouble source and trigger a nested MN sequence or a cascade of MN sequences. To tag such sequences I still used this basic model and handled each nested or cascading sequence as part of a large sequence.

- `mnseq`: Contains a Meaning Negotiation sequence.
- `mnturn`: Contains a reference to a turn that is part of a specific `mnseq`. Turn types must be specified in the attribute `turntype` wich contains the type of the turn that it part of this Meaning Negotiation sequence. Typical turn types are trouble-source (ts), indicator (ind), response (resp), reaction to response (rr). Types are not predefined in the schema because different types not fitting in the basic classification are possible. Cascading and nested MN sequences can be also part of `mnseq` (see turns from L08N0420120531-226 to L08N0420120531-237 of the corpus and their annotation).

Example of a Meaning Negotiation sequence:

```
<im:message xml:id="L01N0120120618-268" who="deL1L2IM-root.xml#N01"
   timestamp="2012-06-18T21:02:44"> Das schaffst du mit links :-)
</im:message>
<im:message xml:id="L01N0120120618-269" who="deL1L2IM-root.xml#N01"
   timestamp="2012-06-18T21:02:55"> (Verstehst du die Bedeutung?)
</im:message>
<im:message xml:id="L01N0120120618-270" who="deL1L2IM-root.xml#L01"
   timestamp="2012-06-18T21:03:42"> Und was bedeutet das?
</im:message>
<im:message xml:id="L01N0120120618-271" who="deL1L2IM-root.xml#N01"
   timestamp="2012-06-18T21:04:37"> mit links schaffen/erledigen/
   erreichen = ohne Probleme schaffen/erledigen/erreichen
</im:message>
<im:message xml:id="L01N0120120618-272" who="deL1L2IM-root.xml#L01"
   timestamp="2012-06-18T21:05:46"> Vielen Dank für Erklärung :)
</im:message>
<im:message xml:id="L01N0120120618-273" who="deL1L2IM-root.xml#N01"
   timestamp="2012-06-18T21:06:36"> Kein Problem :-)
</im:message>

<im:mnseq>
  <im:mnturn turntype="ts" corresp="L01N0120120618-268"/>
  <im:mnturn turntype="ind" corresp="L01N0120120618-269"/>
  <im:mnturn turntype="ind" corresp="L01N0120120618-270"/>
  <im:mnturn turntype="resp" corresp="L01N0120120618-271"/>
  <im:mnturn turntype="rr" corresp="L01N0120120618-272"/>
</im:mnseq>
```

B. Coding Scheme for Polar Questions

Stivers and Enfield (2010) suggest a data-driven coding scheme for questions. The study was performed for questions collected from naturally occurring dyadic and multiparty conversations in ten languages: Akhoe Haillom (Namibia), Danish, Dutch, English (US), Italian, Japanese, Korean, Lao, Tzeltal (Mexico), Yélî Dnye (Papua New Guinea).

The coding scheme proposed by Stivers and Enfield (2010) was used for the analysis of question-answer pairs. Because this coding scheme is relevant for the description of embedded correction in the empirical part, and for the formulation of the computational model of embedded corrections in the modelling part, I provide here a detailed description of the part of the coding scheme which polar questions and responses to polar questions are handled. Because neither German nor text chat language were part of the study presented by Stivers and Enfield (Stivers and Enfield, 2010), it was required to modify the scheme according to the question properties found in the dataset. I explicitly declare all modifications made.

What utterances qualify as questions need to be determined first. I use the basic set of criteria as suggested by (Stivers and Enfield, 2010, p. 2621) but with adjustments for chat:

1. A question had to be either (or both) a formal question (i.e., it had to rely on lexico-morpho-syntactic or prosodic interrogative marking) or a functional question (i.e., it had to effectively seek to elicit information, confirmation or agreement whether or not they made use of an interrogative sentence type).
2. News markers such as *wirklich?* were coded as functional questions. Under the categorisation of question suggested by Stivers and Enfield (2010), news markers qualify because they are routinely treated as seeking confirmation.

Stivers and Enfield (2010) note that "it was difficult to attain validity or reliability in the coding of some aspects of sequential position" (p. 2620). As a result, information about sequential position was not included into coding scheme. For the purpose of this work and in general for the analysis of chat data, it is necessary to add at least the following inclusion criterion for questions with regard to sequential organisation of chat:

3. One question may be delivered in one turn (message) and in multiple turns (messages).

Following the question coding scheme by Stivers and Enfield (2010), all questions were classified as polar, content, alternative and through-produced multi-questions (two or more questions produced as a chunk). Questions delivered in multiple turns are not necessarily through-produced multi-questions. Because only polar questions and responses to polar questions were in focus of this work, I restrict myself to polar questions in further explanation of the coding scheme. Readers interested in the remaining question types are invited to see the original publication by Stivers and Enfield (2010).

According to the coding scheme, responses to polar questions can be classified as answers (deal directly with the question "as put") and non-answers (all the I'don't-knows, maybes, indirect responses). Answers can be further classified by the polarity (confirming or disconfirming) and by the form they take. A confirmation for a negative answer in German could be done by a *nein* (no). Sometimes it is not possible to clearly say yes or no, the infamous German *jein* might be an additional class in the coding of polar questions. With regard to the form of the answers to polar questions, Stivers and Enfield classify the answers as repetitional answers (include full, partial or modified repeats of the question), interjection answers (*ja* or *nein* and their variations) and marked interjection answers (*absolut*, *total* and similar).

Form and aspects of polar questions

Questions can be formed as a declarative question, as an assertion followed by a confirmation request, or as a full question with interrogative word order. The following forms of polar questions were found in the dataset:

1. Declarative word order. Example:

   8 17:13:19 L08 wir müssen heute um 21-00(belorussischer Zeit) chaten, oder wir können jetz anfangen?)

2. Declarative word order with turn-final element. Example:

   15 18:44:44 L07 einen bechlor bekkommt mat nach der Absolvierung der Uni.... stimmt das–?

3. Interrogative word order. Example:

   154 20:35:28 L02 Habt ihr zusammen mit Freunden etwas Tolles untergenommen?

Polar questions can be marked negatively and dubitatively. Stivers and Enfield (2010) suggest to use these characteristics as binary features (yes/no). For simplicity, I used the following coding:

0 Unmarked polar questions.

1 Negatively marked polar questions (*nicht*, *keine*, *un-...*).

2 Dubitatively marked polar questions (*vielleicht*, *wahrscheinlich*).

If both, negative and dubitative marking was present, the question received both codes in this field.

Responses to polar questions

Stivers and Enfield (2010) suggest the following classification of responses to polar questions:

0 No response.

1 Answer, a response dealing directly with the question "as put".

2 Non-answer, a response dealing indirectly with the question, all clarifications, I-dont-knows etc.

Only answers were considered for further analysis in this thesis, therefore I provide the part of coding scheme for answer responses below. Moreover, non-answers were not further coded in the original coding scheme by Stivers and Enfield (2010). A consistent classification of types of non-answer responses to polar questions remains an open question for a future study.

Answers

The coding scheme by Stivers and Enfield (2010) distinguishes between confirming and disconfirming answers with regard to answer polarity and regardless of form. I found several examples where responses to polar questions deal with the question as put and should be classified as "answers", however, do neither confirm nor disconfirm. Such answers contain partial confirmations (*nur teils*) or both, confirmation and disconfirmation which may count as partial confirmation or partial disconfirmation (*jein*) and disambiguations in order to deal with learner errors (*leicht fallen* vs. *gfallen* as discussed in Example 9.2). Therefore, I suggest to extend the classification of answers to polar questions as follows:

1. Confirming answer.

2. Disconfirming answer.
3. Partial confirmation (e.g. *jein* or *nur teils*). This category has been added to the coding scheme because it deals with the question directly as put, but is neither confirming nor disconfirming.
4. Disambiguation (corrective feedback). This category has been introduced for responses dealing with the question as put where native speakers were not able to directly confirm or disconfirm due to learner's linguistic errors. It may contain both, confirmation and disconfirmation.

With regard to answer form, Stivers and Enfield (2010) distinguish among interjection-based answers, marked interjection-based answers and repetition-based answers. Repetitions with replacements (embedded corrections) were covered by this category.

Summary question coding scheme

Modifications were necessary due to communication medium (text chat, multi-turn questions) and learner language (linguistic errors in the question made disambiguation in the answer necessary). Partial confirmation may be needed in other languages, too, not only in German. Table B.1 summarises the scheme with modifications (bold).

Unit	Features	
Question	Form	Declarative word order with a question mark (including questions without verb
		Declarative word order with followed by turn-final element (e.g. *ok?*, *stimmt das? ist es so?*)
		Interrogative word order
	Marked:	Negatively (e.g. *nicht*, *keine*)
		dubitatively (e.g. *vielleicht*, *wahrscheinlich*)
	Location	**Delivered in one turn or multiple turns**
Answer	Form	Repeat-based
		Interjection-based
		Marked interjection-based
	Polarity	Confirming
		Disconfirming
		Partially confirming
		Disambuguation

Table B.1.: Coding scheme for question-answer pairs with polar questions with modifications (bold)

Bibliography

Abel, A., Glaznieks, A., Nicolas, L., and Stemle, E. (2014). Koko: an l1 learner corpus for german. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*, pages 2414–2421.

Abrams, Z. I. (2003). The effect of synchronous and asynchronous cmc on oral performance in german. *Modern Language Journal*, pages 157–167.

Aguado Padilla, K. (2002). *Imitation als Erwerbsstrategie. Interaktive und kognitive Dimensionen des Fremdsprachenerwerbs*. Habilitationsschrift, Universität Bielefeld.

Ai, R., Krause, S., Kasper, W., Xu, F., and Uszkoreit, H. (2015). Semi-automatic generation of multiple-choice tests from mentions of semantic relations. In *2nd Workshop on Natural Language Processing Techniques for Educational Applications (NLP-TEA-2) at ACL-IJCNLP 2015*. ACL.

Ai, R. and Xu, F. (2015). A system demonstration of a framework for computer assisted pronunciation training. *ACL-IJCNLP 2015*, page 1.

Al-Zubaide, H., Issa, A., et al. (2011). Ontbot: Ontology based chatbot. In *Innovation in Information & Communication Technology (ISIICT), 2011 Fourth International Symposium on*, pages 7–12. IEEE.

Alemi, M., Meghdari, A., Basiri, N. M., and Taheri, A. (2015). The effect of applying humanoid robots as teacher assistants to help iranian autistic pupils learn english as a foreign language. In *Social Robotics*, pages 1–10. Springer.

Allen, J. F. and Small, S. L. (1982). Discourse understanding at the u. of rochester. *SIGART Bull.*, pages 83–85.

Allwood, J. (1995). An activity based approach to pragmatics. *Gothenburg papers in theoretical linguistics*, (76):1–38.

Amaral, L. and Meurers, D. (2007). Conceptualizing student models for icall. In *User Modeling 2007*, pages 340–344. Springer.

Amaral, L. and Meurers, D. (2008). From recording linguistic competence to supporting inferences about language acquisition in context. *Computer Assisted Language Learning*, 21(4):323–338.

Amaral, L., Meurers, D., and Ziai, R. (2011). Analyzing learner language: towards a flexible natural language processing architecture for intelligent language tutors. *Computer Assisted Language Learning*, 24(1):1–16.

Amaral, L. A. (2011). Revisiting current paradigms in computer assisted language learning research and development doi: 10.5007/2175-8026.2011 n60p365. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies*, (60):365–390.

Amaral, L. A. and Meurers, D. (2011). On using intelligent computer-assisted language learning in real-life foreign language teaching and learning. *ReCALL*, 23(01):4–24.

Amaral, L. A. and Meurers, W. D. (2009). Little things with big effects: On the identification and interpretation of tokens for error diagnosis in icall. *CALICO Journal*, 26(3):580–591.

Amoia, M., Brétaudiere, T., Denis, A., Gardent, C., and Perez-Beltrachini, L. (2012). A serious game for second language acquisition in a virtual environment. *Journal on Systemics, Cybernetics and Informatics (JSCI)*, 10(1):24–34.

Anderson, J. N., Davidson, N., Morton, H., and Jack, M. A. (2008). Language learning with interactive virtual agent scenarios and speech recognition: Lessons learned. *Computer Animation and Virtual Worlds*, 19(5):605–619.

Ariel Chatbot, Spanish (2011). Ariel. http://www.sustentum.com/sustentum/ariel.asp.

Auer, P. (1999). *Sprachliche Interaktion - Eine Einführung anhand von 22 Klassikern*. Tübingen: Niemeyer.

Avramidis, E., Popovic, M., and Burchardt, A. (2015). Dfki's experimental hybrid mt system for wmt 2015. In *Proceedings of the Tenth Workshop on Statistical Machine Translation*, pages 66–73. Association for Computational Linguistics.

Bartz, T., Beißwenger, M., and Storrer, A. (2014). Optimierung des stuttgart-tübingen-tagset für die linguistische annotation von korpora zur internetbasierten kommunikation: Phänomene, herausforderungen, erweiterungsvorschläge. *Zeitschrift für germanistische Linguistik*, 28(1):157–198.

Becker, L., Erhart, G., Skiba, D., and Matula, V. (2013). Avaya: Sentiment analysis on twitter with self-training and polarity lexicon expansion. In *Second Joint Conference on Lexical and Computational Semantics (* SEM)*, volume 2, pages 333–340.

Becker, M., Bredenkamp, A., Crysmann, B., and Klein, J. (2003). Annotation of error types for german newsgroup corpus. In *Treebanks*, pages 89–100. Springer.

Beer, K. (2015). Google will whatsapp und co. mit neuem messenger konkurrenz machen. http://m.heise.de/newsticker/meldung/Google-will-WhatsApp-und-Co-mit-neuem-Messenger-Konkurrenz-machen-3055331.html.

Beißwenger, M. (2002). Getippte» gespräche «und ihre trägermediale bedingtheit. In Schröder, I. W. and Voell, S., editors, *Moderne Oralität.*, CURUPIRA. Philipps-Universität Marburg.

Bender, E. M., Flickinger, D., Oepen, S., Walsh, A., and Baldwin, T. (2004). Arboretum: Using a precision grammar for grammar checking in call. In *Instil/icall symposium 2004*.

Benjamin, T. (2013). *Signaling trouble On the linguistic design of other-initiation of repair in English conversation*. Phd, University of Groningen.

Bickmore, T. and Picard, R. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer Human Interaction*, 12:293–327.

Bickmore, T. W. (2003). *Relational Agents: Effecting Change through Human-Computer Relationships*. PhD thesis, MIT.

Bird, S., Klein, E., and Loper, E. (2009). *Natural Language Processing with Python*. O'Reilly Media.

Blake, R. (2008). *Brave new digital classroom*. Washington, DC: Georgetown University Press.

Blake, R. and Delforge, A. M. (2004). Language learning at a distance: Spanish without walls. In *Selected papers from the 2004 NFLRC Symposium*.

Bolden, G. B. (2009). Beyond answering: Repeat-prefaced responses in conversation. *Communication Monographs*, 76(2):121–143.

Bonneau, A. and Colotte, V. (2011). Automatic feedback for l2 prosody learning. In *Speech and Language Technologies*, pages 55–70. Intech.

Bono, M., Ogata, H., Takanashi, K., and Joh, A. (2014). The practice of showing 'who i am': A multimodal analysis of encounters between science communicator and visitors at science museum. In *Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge*, pages 650–661. Springer.

Boyd, A. (2010). EAGLE: an Error-Annotated Corpus of Beginning Learner German. In *Proc. of LREC*. ELRA.

Bradley, J., Mival, O., and Benyon, D. (2009). Wizard of Oz Experiments for Companions. In *Proc. of the 23rd British HCI Group Annual Conf. on People and Computers: Celebrating People and Technology*, pages 313–317. British Computer Society.

Brandt, A. (2011). *The Maintenance of Mutual Understanding in Online Second Language Talk*. PhD thesis, Newcastle University, School of Education, Communication and Language Sciences.

Breckle, M. and Zinsmeister, H. (2010). Zur lernersprachlichen generierung referierender ausdrücke in argumentativen texten. *Textmuster: schulisch-universitär-kulturkontrastiv.*, pages 79–101.

Breuing, A. and Wachsmuth, I. (2012). Let's talk topically with artificial agents! providing agents with humanlike topic awareness in everyday dialog situations. In *Proceedings of the ICAART'12*, pages 62–71.

Breuing, A., Waltinger, U., and Wachsmuth, I. (2011). Harvesting wikipedia knowledge to identify topics in ongoing natural language dialogs. In *Proc. of the 2011 IEEE/WIC/ACM Int. Conf. on Web Intelligence and Intelligent Agent Technology*, pages 445–450. IEEE Computer Society.

Brouwer, C. E., Rasmussen, G., and Wagner, J. (2004). Embedded corrections in second language talk. In Gardner, R. and Wagner, J., editors, *Second Language Conversations: Studies of Communication in Everyday Settings*, pages 86–103. A&C Black - Verlag.

Brown, G. and Yule, G. (1983). *Discourse analysis*. Cambridge University Press.

Burchardt, A., Lommel, A. R., Rehm, G., Sasaki, F., van Genabith, J., and Uszkoreit, H. (2014). Language technology drives quality translation. *MultiLingual*, 143:33–39.

Bush, N. (2006). Program D. http://www.aitools.org/Program_D.

Buß, O. and Schlangen, D. (2011). DIUM – An Incremental Dialogue Manager That Can Produce Self-Corrections. In *SEMDIAL*.

Busuu (2008). Busuu. https://www.busuu.com/about/.

Cahill, A. (2015). Parsing learner text: to shoehorn or not to shoehorn. In *The 9th Linguistic Annotation Workshop held in conjuncion with NAACL 2015*, page 144.

Caire, P. and van der Torre, L. (2009). Convivial ambient technologies: Requirements, ontology and design. *The Computer Journal*, 52.

Canale, M. and Swain, M. (1980). Theoretical bases of com-municative approaches to second language teaching and testing. *Applied linguistics*, 1(1):1–47.

Casas, R., Marín, R. B., Robinet, A., Delgado, A. R., Yarza, A. R., Mcginn, J., Picking, R., and Grout, V. (2008). User modelling in ambient intelligence for elderly and disabled people. In *11th International Conference on Computers Helping People with Special Needs*, ICCHP 2008, pages 114–122. Springer.

Chang, C.-W., Lee, J.-H., Chao, P.-Y., Wang, C.-Y., and Chen, G.-D. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Journal of Educational Technology & Society*, 13(2):13–24.

Chat, C. (2008). Chatbot mike representing cypris chat in second life. http://maps.secondlife.com/secondlife/Wellston/167/88/23, http://cyprischat.org.

Chee, B. T. T., Wong, A. H. Y., Limbu, D. K., Tay, A. H. J., Tan, Y. K., and Park, T. (2010). Understanding communication patterns for designing robot receptionist. In *Social Robotics*, pages 345–354. Springer.

Chrysafiadi, K. and Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, 40(11):4715–4729.

Clark, H. H. (1996). *Using language*. Cambridge university press.

Dahan, D., Tanenhaus, M. K., and Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47(2):292–314.

Danilava, S., Busemann, S., and Schommer, C. (2012). Artificial Conversational Companion: a Requirement Analysis. In *Proc. of ICAART'12*, pages 282–289.

Danilava, S., Busemann, S., Schommer, C., and Ziegler, G. (2013a). Towards Computational Models for a Long-term Interaction with an Artificial Conversational Companion. In *Proc. of ICAART'13*.

Danilava, S., Busemann, S., Schommer, C., and Ziegler, G. (2013b). Why are you silent?-towards responsiveness in chatbots. In *Avec le Temps! Time, Tempo, and Turns in Human-Computer Interaction". Workshop at CHI 2013, Paris, France*.

Darhower, M. A. (2008). The role of linguistic affordances in telecollaborative chat. *Calico Journal*, 26(1):48–69.

Davidov, D., Tsur, O., and Rappoport, A. (2010). Semi-supervised recognition of sarcastic sentences in twitter and amazon. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning*, pages 107–116. Association for Computational Linguistics.

de Alencar, T. S., Machado, L. R., de Oliveira Neris, L., and de Almeida Neris, V. P. (2014). Addressing the users' diversity in ubiquitous environments through a low cost architecture. In *Universal Access in Human-Computer Interaction. Aging and Assistive Environments*, pages 439–450. Springer.

De Gasperis, G. and Florio, N. (2012). Learning to read/type a second language in a chatbot enhanced environment. In *International Workshop on Evidence-Based Technology Enhanced Learning*, pages 47–56. Springer.

De Marco, A. and Leone, P. (2012). Computer mediated conversation for mutual learning: acknowledgement and agreement/assessment signals in italian as l2. In *Proceedings of the EUROCALL Conference*, pages 70–75.

Delmonte, R. (2003). Linguistic knowledge and reasoning for error diagnosis and feedback generation. *Calico Journal*, 20(3):513–532.

DeSmedt, W. H. (1995). Herr Kommissar: An ICALL conversation simulator for intermediate german. In Holland, V. M., Sams, M. R., and Kaplan, J. D., editors, *Intelligent language tutors: Theory shaping technology*. Routledge.

Díaz-Negrillo, A. and Domínguez, J. F. (2006). Error tagging systems for learner corpora. *Revista española de lingüística aplicada*, 19:83–102.

Díaz-Negrillo, A., Meurers, D., Valera, S., and Wunsch, H. (2010). Towards interlanguage pos annotation for effective learner corpora in sla and flt. In *Language Forum*, volume 36, pages 139–154.

Diego Chatbot, Spanish (2011). Diego. http://www.soydiego.com/.

Dingemanse, M., Blythe, J., and Dirksmeyer, T. (2014). Formats for other-initiation of repair across languages: An exercise in pragmatic typology. *Studies in Languag*, 3(81):5–43.

Dings, A. (2012). Native speaker/nonnative speaker interaction and orientation to novice/expert identity. *Journal of Pragmatics*, 44(11):1503–1518.

Dodigovic, M. (2005). *Artificial intelligence in second language learning: Raising error awareness*, volume 13. Multilingual Matters.

Dong, H., Cheung Hui, S., and He, Y. (2006). Structural analysis of chat messages for topic detection. *Online Information Review*, 30(5):496–516.

Dorr, B., Galescu, L., Golob, E., Venable, K. B., and Wilks, Y. (2015). Companion-based ambient robust intelligence (caring). In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Drew, P. (1997). Open class repair initiations in response to sequential sources of troubles in conversation. *Journal of Pragmatics*, 28:69–101.

Droßmann, C. (2005). German AIML set. http://www.drossmann.de/wordpress/alicebot/.

Dzikovska, M. O., Callaway, C. B., Farrow, E., Moore, J. D., Steinhauser, N., and Campbell, G. (2009). Dealing with interpretation errors in tutorial dialogue. In *Proceedings of the SIGDIAL 2009 Conference: The 10th annual meeting of the special interest group on discourse and dialogue*, pages 38–45. Association for Computational Linguistics.

Edlund, J., Heldner, M., and Gustafson, J. (2005). Utterance segmentation and turn-taking in spoken dialogue systems. *Sprachtechnologie, mobile Kommunikation und linguistische Ressourcen*, pages 576–587.

Egbert, M. (2009). *Der Reparatur-Mechanismus in deutschen Gesprächen*. Verlag für Gesprächsforschung, Mannheim.

ELRA (2015). deL1L2IM: Corpus of long-term instant messaging NS-NNS conversations. http://islrn.org/resources/339-799-085-669-8/.

Enfield, N. J. (2007). Meanings of the unmarked: How'default'person reference does more than just refer. In *Person reference in interaction: Linguistic, cultural, and social perspectives*, pages 97–120. Cambridge University Press.

Enfield, N. J., Dingemanse, M., Baranova, J., Blythe, J., Brown, P., Dirksmeyer, T., Drew, P., Floyd, S., Gipper, S., Gisladottir, R. S., Hoymann, G., Kendrick, K. H., Levinson, S. C., Magyari, L., Manrique, E., Rossi, G., Roque, L. S., and Torreira, F. (2013). Huh? what? - a first survey in twenty-one languages. In Hayashi, M., Raymond, G., and Sidnell, J., editors, *Conversational Repair and Human Understanding*, chapter 12. Cambridge University Press.

Enfield N.J. (2013). Reference in Conversation. In J. Sidnell and Stivers, T., editors, *The Handbook of Conversation Analysis*, chapter 21, pages 433–454. MA: Wiley-Blackwell.

Evert, S., Proisl, T., Greiner, P., and Kabashi, B. (2014). Sentiklue: Updating a polarity classifier in 48 hours. In *Proceedings of SemEval 2014*, page 551.

Fellbaum, C. (2010). Princeton university: About WordNet. http://wordnet.princeton.edu. WordNet.

Firth, A. and Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in sla research. *The Modern Language Journal*, 81(3):285–300.

Fredriksson, C. (2012). About collaboration, interaction, and the negotiation of meaning in synchronous written chats in l2-german. In Bradley, L. and Thouësny, S., editors, *CALL: Using, Learning, Knowing, EUROCALL Conference, Gothenburg, Sweden, 22-25 August 2012, Proceedings*, pages 88–92. Research-publishing.net.

Fredriksson, C. (2013). Learning german by chatting–a study of social interaction and language production in an academic german online course. In *ICT for Language Learning, Florence, Italy 14-15 November 2013*, pages 218–222. Pixel.

Fryer, L. and Carpenter, R. (2006). Emerging technologies - bots as language learning tools. *Language Learning & Technology*, 3(10):8 – 14.

Gabsdil, M. (2003). Clarification in Spoken Dialogue Systems. In *AAAI Technical Report*. AAAI.

Gardent, C., Lorenzo, A., Perez-Beltrachini, L., and Rojas-Barahona, L. (2013). Weakly and strongly constrained dialogues for language learning. In *the 14th annual SIGdial Meeting on Discourse and Dialogue SIGDIAL 2013*, pages 357–359.

Gehle, R., Pitsch, K., and Wrede, S. (2014). Signaling trouble in robot-to-group interaction. emerging visitor dynamics with a museum guide robot. In *Proceedings of the second international conference on Human-agent interaction*, pages 361–368. ACM.

Georgeff, M., Pell, B., Pollack, M., Tambe, M., and Wooldridge, M. (1999). The belief-desire-intention model of agency. In *Intelligent Agents V: Agents Theories, Architectures, and Languages*, pages 1–10. Springer.

Gilmartin, E., Bonin, F., Vogel, C., and Campbell, N. (2013). Laughter and topic transition in multiparty conversation. In *Proceedings of SIGDIAL 2013 Conference*, pages 304–308.

Ginzburg, J. (2012). *The interactive stance*. Oxford University Press.

Ginzburg, J., Fernández, R., and Schlangen, D. (2007). Unifying self-and other-repair. In *Proceeding of DECALOG, the 11th International Workshop on the Semantics and Pragmatics of Dialogue (SemDial07)*.

Golato, A. (2002). German compliment responses. *Journal of pragmatics*, 34(5):547–571.

Golato, A. (2005). *Compliments and compliment responses: Grammatical structure and sequential organization*, volume 15. John Benjamins Publishing.

González-Lloret, M. (2011). Conversation analysis of computer-mediated communication. *Calico Journal*, 28(2):308–325.

Gould, J. D., Conti, J., and Hovanyecz, T. (1983). Composing letters with a simulated listening typewriter. *Commun. ACM*, 26(4):295–308.

Gray, E. F. (1992). Interactive language learning: "a la rencontre de philippe". *The French Review*, 65(3):499–507.

Greene, C. E., Keogh, K., Koller, T., Wagner, J., Ward, M., and Van Genabith, J. (2004). Using nlp technology in call. In *InSTIL/ICALL 2004 Symposium on Computer Assisted Learning*. International Speech Communication Association.

Grice, H. P. (1975). Logic and conversation. *Syntax and semantics*, 3, Speech Acts:41–58.

Grote, B., Hagen, E., Stein, A., Teich, E., and Magdeburg (1997). Speech Production in Human-Machine Dialogue : A Natural Language Generation Perspective ? In Meier, E., Mast, M., and Luperfoy, S., editors, *Dialogue Processing in Spoken Language Systems*, volume 1236 of *Lecture Notes in Computer Science*, pages 70–85. Springer.

Gut, U. (2009). *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German*, volume 9 of *English Corpus Linguistics*. Peter Lang.

Hallili, A. (2014). Toward an ontology-based chatbot endowed with natural language processing and generation. In *26th European Summer School in Logic, Language & Information*.

Hamp, B., Feldweg, H., et al. (1997). Germanet-a lexical-semantic net for german. In *Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, pages 9–15. Citeseer.

Han, J. (2012). Emerging technologies:“robot assisted language learning”. *Language and Learning Technology*, 16(3):1–9.

Hauser, E. (2010). Other-correction of language form following a repair sequence. *Pragmatics and Language Learning*, 12:277–296.

Hayes, P. and Reddy, R. (1979). Graceful interaction in man-machine communication. In *Proceedings of the 6th International Joint Conference on Artificial Intelligence - Volume 1*, IJCAI’79, pages 372–374. Morgan Kaufmann Publishers Inc.

Heift, T. (2002). Learner control and error correction in icall: Browsers, peekers, and adamants. *Calico Journal*, 19(2):295–313.

Heift, T. (2003). Multiple learner errors and meaningful feedback: A challenge for icall systems. *CALICO journal*, 20(3):533–548.

Heift, T. (2007). Learner personas in call. *CALICO Journal*, 25(1):1–10.

Heift, T. (2015). Web delivery of adaptive and interactive language tutoring: Revisited. *International Journal of Artificial Intelligence in Education*, pages 1–15.

Heift, T. and Schulze, M. (2007). *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. Routledge.

Heine, L., Witte, M., and Holl, T. (2007). Babbel! https://www.babbel.com/.

Heritage, J. (2012). The epistemic engine: Sequence organization and territories of knowledge. *Research on Language & Social Interaction*, 45(1):30–52.

Heritage, J. and Goodwin, C. (1990). Conversation analysis. *Annual review of anthropology*, 19:283–307.

Hirst, G. (1991). Does conversation analysis have a role in computational linguistics? *Computational Linguistics*, 17(2):211–227.

Hjalmarsson, A., Wik, P., and Brusk, J. (2007). Dealing with deal: a dialogue system for conversation training. In *Proceedings of SIGDIAL*, pages 132–135.

Höhn, S. (2015). Dealing with trouble: A data-driven model of a repair type for a conversational agent. In *Proceedings of the twenty-ninth AAAI conference on Artificial Intelligence*. AAAI.

Höhn, S., Busemann, S., Max, C., Schommer, C., and Ziegler, G. (2015). Interaction profiles for an artificial conversational companion. In *Proceedings of the first International Symposium on Companion Technology*, ISTC, pages 73–78.

Höhn, S., Ouafo, A., Ziegler, G., and Schommer, C. (2016). Instant data capturing from chat conversations for learning. *Language Learning and Technology (under review).*, 21(1).

Holt, E. (2010). The last laugh: Shared laughter and topic termination. *Journal of Pragmatics*, 42(6):1513–1525.

Holt, E. and Drew, P. (2001). Idiomatic expressions as topical pivots. In *Proceedings of Interdisciplinary Workshop on Corpus-Based & Processing Approaches to Figurative Language (part of Corpus Linguistics Conference 2001)*. Lancaster University.

Holt, E. and Drew, P. (2005). Figurative pivots: The use of figurative expressions in pivotal topic transitions. *Research on Language and Social Interaction*, 38(1):35–61.

Hosoda, Y. (2006). Repair and relevance of differential language expertise in second language conversations. *Applied Linguistics*, 27(1):25–50.

Hutchby, I. (1995). Aspects of recipient design in expert advice-giving on call-in radio. *Discourse processes*, 19(2):219–238.

Hymes, D. (1972). On communicative competence. In J.B.Pride and J.Holmes, editors, *Sociolinguistics*, pages 269–293. Harmondsworth: Penguin.

Hynninen, N. (2012). Icl at the micro level: L2 speakers taking on the role of language experts. *AILA Review*, 25:13–29.

Hynninen, N. (2013). *Language Regulation in English as a Lingua Franca: Exploring language-regulatory practices in academic spoken discourse*. PhD thesis, University of Helsinki.

Iida, R., Kobayashi, S., and Tokunaga, T. (2010). Incorporating extra-linguistic information into reference resolution in collaborative task dialogue. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1259–1267. Association for Computational Linguistics.

Isaacs, E. a. and Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General*, 116(1):26–37.

Ivanov, A. V., Ramanarayanan, V., Suendermann-Oeft, D., Lopez, M., Evanini, K., and Tao, J. (2015). Automated speech recognition technology for dialogue interaction with non-native interlocutors. In *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, page 134.

Janarthanam, S. and Lemon, O. (2009). A wizard-of-oz environment to study referring expression generation in a situated spoken dialogue task. In *Proceedings of the 12th European Workshop on Natural Language Generation*, pages 94–97. Association for Computational Linguistics.

Janarthanam, S. and Lemon, O. (2010). Learning to adapt to unknown users: referring expression generation in spoken dialogue systems. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 69–78. Association for Computational Linguistics.

Jefferson, G. (1974). Error correction as an interactional resource. *Language in society*, 3(02):181–199.

Jefferson, G. (1983). On exposed and embedded correction. *Studium Linquistik*, 14:58–68.

Jia, J. (2004). The Study of the Application of a Web-based Chatbot System on the Teaching of Foreign Languages. *Proc. of Society for Inf. Tech. and Teacher Education Int. Conf.*, pages 1201–1207.

Jia, J. (2009). CSIEC: A Computer Assisted English Learning Chatbot Based on Textual Knowledge and Reasoning. *Know.-Based Syst.*, 22(4):249–255.

Jian, C., Zhekova, D., Shi, H., and Bateman, J. (2010). Deep Reasoning in Clarification Dialogues with Mobile Robots. In *Proceedings of the 19th European Conference on Artificial Intelligence*, pages 177–182. IOS Press.

Jin, L. (2013). Language development and scaffolding in a sino-american telecollaborative project. *Language Learning & Technology*, 17(2):193–219.

Kasper, G. (1985). Repair in foreign language teaching. *Studies in Second Language Acquisition*, 7:200–215.

Kasper, G. (2004). Participant Orientations in German Conversation-for-Learning. *The Modern Language Journal*, 88:551–567.

Kasper, G. (2006). Beyond repair: Conversation analysis as an approach to sla. *AILA Review*, 19(1):83–99.

Katagiri, Y., Bono, M., and Suzuki, N. (2004). Capturing conversational participation in a ubiquitous sensor environment. In *Pervasive 2004 Workshop on Memory and Sharing of Experiences*, pages 101–106.

Kaur, J. (2011). Doing being a language expert": The case of the ELF speaker. In *Latest trends in ELF research*, pages 53–75. Cambridge Scholars Publishing.

Kim, J. (2003). A study on negotiation of meaning in NNS-NNS interactions - Focusing on synchronous CMC. In *Proceedings of the 8th Conference of Pan-Pacific Association of Applied Linguistics*, pages 190–210. PPAL.

Kitade, K. (2000). L2 learners' discourse and sla theories in cmc: Collaborative interaction in internet chat. *Computer assisted language learning*, 13(2):143–166.

Klüwer, T. (2009). RMRSBot – using linguistic information to enrich a chatbot. In *IVA*, pages 515–516. Springer.

Klüwer, T. (2011). From chatbots to dialogue systems. In Perez-Marton, D. and Pascual-Nieto, I., editors, *Conversational Agents and Natural Language Interaction: Techniques and Effective Practices*, pages 1–22. IGI Global Publishing Group.

Koch, P. (1994). Schriftlichkeit und sprache. In *Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung*, pages 587–604. Walter de Gruyter.

Koch, P. and Oesterreicher, W. (1985). Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. In *Romanistisches Jahrbuch*, volume 36. Walter de Gruyter.

Kost, C. R. (2004). *An investigation of the effects of synchronous computer-mediated communication (CMC) on interlanguage development in beginning learners of German: Accuracy, proficiency, and communication strategies*. PhD thesis, The University of Arizona.

Kramer, M., Yaghoubzadeh, R., Kopp, S., and Pitsch, K. (2013). A conversational virtual human as autonomous assistant for elderly and cognitively impaired users? social acceptability and design considerations. *Lecture Notes in Informatics (LNI)*.

Krashen, S. (1981). *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.

Kruijff, G.-J. M., Brenner, M., and Hawes, N. (2008). Continual planning for cross-modal situated clarification in human-robot interaction. In *The 17th IEEE International Symposium on Robot and Human Interactive Communication*, pages 592–597. IEEE.

Kruijff, G.-J. M., Zender, H., Jensfelt, P., and Christensen, H. I. (2006). Clarification dialogues in human-augmented mapping. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 282–289. ACM.

Krummes, C. and Ensslin, A. (2014). What's hard in german? whig: a british learner corpus of german. *Corpora*, 9(2):191–205.

Kurhila, S. (2001). Correction in talk between native and non-native speaker. *Journal of Pragmatics*, 33(7):1083–1110.

Kurhila, S. (2004). Clients or language learners–being a second language speaker in institutional interaction. *Second language conversations*, pages 58–74.

Kurhila, S. (2006). *Second language interaction*, volume 145. John Benjamins Publishing.

Kwok, V. H. (2015). Robot vs. human teacher: Instruction in the digital age for esl learners. *English Language Teaching*, 8(7):p157.

Lamnek, S. (2010). *Qualitative Sozialforschung*. BELTZ.

Lavolette, E., Polio, C., and Kahng, J. (2015). The accuracy of computer-assisted feedback and students' responses to it. *Language Learning and Technology*, 19(2):50–68.

Lee, C. K. (2007). Text-making practices beyond the classroom context: Private instant messaging in hong kong. *Computers and Composition*, 24(3):285–301.

Lee, M. K., Kiesler, S., and Forlizzi, J. (2010). Receptionist or information kiosk: How do people talk with a robot? In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, CSCW '10, pages 31–40, New York, NY, USA. ACM.

Lee, M. K. and Makatchev, M. (2009). How do people talk with a robot?: an analysis of human-robot dialogues in the real world. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pages 3769–3774. ACM.

Lerner, G. H. and Kitzinger, C. (2007). Extraction and aggregation in the repair of individual and collective self-reference. *Discourse Studies*, 9(4):526–557.

L'Haire, S. (2011). *Traitement automatique des langues et apprentissage des langues assisté par ordinateur: bilan, résultats et perspectives*. PhD thesis, University of Geneva.

Liddicoat, A. J. (2011). *An Introduction to Conversation Analysis*. Continuum.

Loewen, S. and Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, 19(1):1–14.

Loewen, S. and Reissner, S. (2009). A comparison of incidental focus on form in the second language classroom and chatroom. *Computer Assisted Language Learning*, 22(2):101–114.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. *Handbook of second language acquisition*, 2(2):413–468.

Lüdeling, A., Briskina, E., Hantschel, J., Krüger, J., Sigrist, S., and Spieler, U. (2010). *LeKo Lernerkorpus Handbuch*. Humboldt-Universität zu Berlin, Institut für Deutsche Sprache und Linguistik, Philosophische Fakultät II.

Lüdeling, A., Walter, M., Kroymann, E., and Adolphs, P. (2005). Multi-level error annotation in learner corpora. *Proceedings of corpus linguistics 2005*, pages 15–17.

Luff, P., Frohlich, D., and Gilbert, N., editors (1990). *Computers and Conversation*. London: Academic Press.

Lyster, R. and Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in second language acquisition*, 19(01):37–66.

Lyster, R., Saito, K., and Sato, M. (2013). Language teaching. *Oral corrective feedback in second language classrooms.*, 46.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied linguistics*, 27(3):405–430.

Maier, E. (1997). Clarification Dialogues in VERBMOBIL. In *ACL/EACL Workshop on Interactive Spoken Dialog Systems: Bringing Speech and NLP Together in Real Applications*. ACL.

Makkonen, J., Ahonen-Myka, H., and Salmenkivi, M. (2004). Simple semantics in topic detection and tracking. *Information Retrieval*, 7(3-4):347–368.

Markee, N. (2000). *Conversation Analysis*. Mahwah, N.J.: Lawrence Erlbaum.

Markee, N. (2008). Toward a learning behavior tracking methodology for ca-for-sla. *Applied Linguistics*, 29(3):404–427.

Marques-Schäfer, G. (2013). *Deutsch lernen online. Eine Analyse interkultureller Aktionen im Chat*. Gunter Narr Verlag.

Martin, J. H. and Jurafsky, D. (2009). *Speech and language processing: an introduction to natural language processing, computational linguistics and speech resognition*. Pearson International Edition, second edition edition.

Martinez, M. (2013). Adapting for a personalized learning experience. In Huang, R., Kinshuk, and Spector, J., editors, *Reshaping Learning*, New Frontiers of Educational Research, pages 139–174. Springer Berlin Heidelberg.

Marton, Y. (2013). Distributional phrasal paraphrase generation for statistical machine translation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4(3):39.

Matt Chatbot, English, iOS (2011). Matt. http://www.speaktoit.com/products.html.

Matthews, C. (1993). Grammar frameworks in intelligent call. *CALICO Journal*, 11(1):5–27.

McGinn, J. J. and Kotamraju, N. (2008). Data-driven persona development. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1521–1524. ACM.

Meena, R., Skantze, J. L. G., and Gustafson, J. (2015). Automatic detection of miscommunication in spoken dialogue systems. In *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, page 354.

Metzler, D., Hovy, E., and Zhang, C. (2011). An empirical evaluation of data-driven paraphrase generation techniques. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pages 546–551. Association for Computational Linguistics.

Meurers, D. (2009). On the automatic analysis of learner language: Introduction to the special issue. *CALICO Journal*, 26(3):469–473.

Meurers, D. (2012). Natural language processing and language learning. *The Encyclopedia of Applied Linguistics*.

Mintzberg, H. (1994). The fall and rise of strategic planning. *Harvard business review*, 72(1):107–114.

Mitchell, C. M., Boyer, K. E., and Lester, J. C. (2012). From strangers to partners: Examining convergence within a longitudinal study of task-oriented dialogue. In *Proceedings of SIGDIAL 2012 Conference*, pages 94–98. ACL.

Morton, H., Davidson, N., and Jack, M. (2008). Evaluation of a speech interactive call system. *Handbook of research on computer-enhanced language acquisition and learning*, page 13.

Mubin, O., Shahid, S., and Bartneck, C. (2013). Robot assisted language learning through games: A comparison of two case studies. *Australian Journal of Intelligent Information Processing Systems*, 13(3):9–14.

Nagata, N. (2009). Robo-sensei's nlp-based error detection and feedback generation. *CALICO*, 26(3):562–579.

Orthmann, C. (2004). *Strukturen der Chat-Kommunikation: konversationsanalytische Untersuchung eines Kinder- und Jugendchats.* PhD thesis, Freie Universität Berlin.

Oxford, R. L. (1993). Intelligent computers for learning languages: The view for language acquisition and instructional methodology. *Computer Assisted Language Learning*, 6(2):173–188.

Panova, I. and Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult esl classroom. *TESOL Quarterly*, 36(4):573–595.

Pavel, M., Jimison, H. B., Korhonen, I., Gordon, C. M., and Saranummi, N. (2015). Behavioral informatics and computational modeling in support of proactive health management and care. *Biomedical Engineering, IEEE Transactions on*, 62(12):2763–2775.

Peres, F. and Meira, L. (2003). Educational software evaluation centered on dialogue: Interface, collaboration and scientific concepts. In *Proceedings of the Latin American Conference on Human-computer Interaction*, CLIHC '03, pages 97–106, New York, NY, USA. ACM.

Perrin, L., Deshaies, D., and Paradis, C. (2003). Pragmatic functions of local diaphonic repetitions in conversation. *Journal of pragmatics*, 35(12):1843–1860.

Petersen, K. A. (2010). *Implicit Corrective Feedback in Computer-Guided Interaction: Does Mode Matter?* PhD thesis, Georgetown University.

Petrelli, D., De Angeli, A., and Convertino, G. (1999). A user-centered approach to user modeling. In *Proceedings of the 7th International Conference on User Modeling*, UM99, pages 255–264. Springer.

Pikkarainen, M. (2015). *Finnish and Russian as lingua francas. Joint activity in conversations*. PhD thesis, University of Helsinki, Faculty of Arts, Department of Modern Languages, Russian language and literature.

Pillet-Shore, D. (2012). Greeting: Displaying stance through prosodic recipient design. *Research on Language & Social Interaction*, 45(4):375–398.

Pitsch, K., Gehle, R., and Wrede, S. (2013a). Addressing multiple participants: A museum robot's gaze shapes visitor participation. In Hermann, G., editor, *Proceedings of ICSR*, volume 8239 of *LNAI*, pages 587–588.

Pitsch, K., Gehle, R., and Wrede, S. (2013b). A museum guide robot: Dealing with multiple participants in the real-world. In *Workshop "Robots in public spaces. Towards multi-party, short-term, dynamic human-robot interaction" at ICSR 2013*.

Pitsch, K., Kuzuoka, H., Suzuki, Y., Süssenbach, L., Luff, P., and Heath, C. (2009). "the first five seconds": Contingent stepwise entry into an interaction as a means to secure sustained engagement in hri. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 985–991. IEEE.

Pitsch, K., Neumann, A., Schnier, C., and Hermann, T. (2013c). Augmented reality as a tool for linguistic research: Intercepting and manipulating multimodal interaction. In *Multimodal corpora: Beyond audio and video (IVA 2013 workshop)*.

Pitsch, K. and Wrede, S. (2014). When a robot orients visitors to an exhibit. referential practices and interactional dynamics in real world hri. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication, 2014 RO-MAN*, pages 36–42.

Plurkowski, L., Chu, M., and Vinkhuyzen, E. (2011). The implications of interactional" repair" for human-robot interaction design. In *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2011 IEEE/WIC/ACM International Conference on*, volume 3, pages 61–65. IEEE.

Postel, H. J. (1969). Die kölner phonetik–ein verfahren zur identifizierung von personennamen auf der grundlage der gestaltanalyse. *IBM-Nachrichten*, 19:925–931.

Pulman, S. G., Boye, J., Cavazza, M., Smith, C., and de la Cámara, R. S. (2010). "How was your day?". In *Proc. of the 2010 Workshop on CDS*, pages 37–42. ACL.

Purver, M. (2004). *The theory and use of clarification requests in dialogue*. PhD thesis, King's College, University of London.

Purver, M. (2006). CLARIE: Handling clarification requests in a dialogue system. *Research on Language and Computation*, 4(2-3):259–288.

Purver, M. and Hough, J. (2014). Strongly Incremental Repair Detection. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 78–89.

Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.

Quintano, L. and Rodrigues, I. P. (2008). Question / Answering Clarification Dialogues. In *MICAI2008*, pages 155–164. Springer.

Ragheb, M. and Dickinson, M. (2013). Inter-annotator agreement for dependency annotation of learner language. *NAACL/HLT 2013*, page 169.

Raux, A. and Eskenazi, M. (2004a). Non-native users in the let's go!! spoken dialogue system: Dealing with linguistic mismatch. In *HLT-NAACL*, pages 217–224. Citeseer.

Raux, A. and Eskenazi, M. (2004b). Using task-oriented spoken dialogue systems for language learning: potential, practical applications and challenges. In *InSTIL/ICALL Symposium 2004*.

Read, T. and Bárcena, E. (2014). Modelling user linguistic communicative competences for individual and collaborative learning. In *From Animals to Robots and Back: Reflections on Hard Problems in the Study of Cognition*, pages 205–218. Springer.

Regneri, M. and Wang, R. (2012). Using discourse information for paraphrase extraction. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 916–927. Association for Computational Linguistics.

Ren, J., Schulman, D., Jack, B., and Bickmore, T. (2014). Supporting longitudinal change in many health behaviors. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 17:1–17:7. CHI.

Reznicek, M., Lüdeling, A., and Hirschmann, H. (2013). Competing target hypotheses in the falko corpus: A flexible multi-layer corpus architecture. In Díaz-Negrillo, Ana, N. B. and Thompson, P., editors, *Automatic Treatment and Analysis of Learner Corpus Data*, pages 101–123. Amsterdam: John Benjamins.

Reznicek, M., Lüdeling, A., Krummes, C., Schwantuschke, F., Walter, M., Schmidt, K., Hirschmann, H., and Andreas, T. (2012). *Das Falko-Handbuch. Korpusaufbau und Annotationen*. Humboldt Universität zu Berlin, 2.01 edition.

Rich, C., Ponsleur, B., Holroyd, A., and Sidner, C. L. (2010). Recognizing engagement in human-robot interaction. In *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction*, HRI '10, pages 375–382, Piscataway, NJ, USA. IEEE Press.

Rich, E. (1979). User modeling via stereotypes. *Cognitive science*, 3(4):329–354.

Rieger, C. L. (2003). Repetitions as self-repair strategies in english and german conversations. *Journal of Pragmatics*, 35(1):47–69.

Rintel, S. (2015). Omnirelevance in technologised interaction. In Fitzgerald, R. and Housley, W., editors, *Advances in Membership Categorisation Analysis*. SAGE.

Robins, B., Dickerson, P., Stribling, P., and Dautenhahn, K. (2004). Robot-mediated joint attention in children with autism: A case study in robot-human interaction. *Interaction studies*, 5(2):161–198.

Rodríguez, K. J. and Schlangen, D. (2004). Form, intonation and function of clarification requests in german task-oriented spoken dialogues. In *Proceedings of Catalog (the 8th workshop on the semantics and pragmatics of dialogue; SemDial04)*, pages 101–108.

Russell, R. and Odell, M. (1918). Soundex. US Patent.

Sabelli, A. M., Kanda, T., and Hagita, N. (2011). A conversational robot in an elderly care center: an ethnographic study. In *Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on*, pages 37–44. ACM.

Sacks, H. (1995). *Lectures on Conversation*, volume 1 and 2. Wiley Online Library.

Sacks, H., Schegloff, E. A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *language*, pages 696–735.

Sagae, A., Johnson, W. L., and Valente, A. (2011). *Conversational Agents in Language and Culture Training*, pages 358–377. IGI Global.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of l2 grammar. *Language Learning and Technology*, 13(1):96–120.

Sawchuk, P. H. (2003). Informal learning as a speech-exchange system: Implications for knowledge production, power and social transformation. *Discourse & Society*, 14(3):291–307.

Schäfer, U. (2008). Shallow, deep and hybrid processing with UIMA and Heart of Gold. In *Proc. of the LREC-2008 Workshop Towards Enhanced Interoperability for Large HLT Systems: UIMA for NLP, 6th Int. Conf. on Language Resources and Evaluation*, pages 43–50.

Schegloff, E. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes*, 23(3):499–545.

Schegloff, E. A. (1972). Notes on a conversational practice: Formulating place. *Studies in social interaction*, 75:75–119.

Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation in discourse and syntax. *Syntax and Semantics Ann Arbor, Mich.*, 12:261–286.

Schegloff, E. A. (1987). Some sources of misunderstanding in talk-in-interaction. *Linguistics*, 25(1):201–218.

Schegloff, E. A. (1988). Presequences and indirection: Applying speech act theory to ordinary conversation. *Journal of Pragmatics*, 12(1):55–62.

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 97(5):pp. 1295–1345.

Schegloff, E. A. (1993). Reflections on quantification in the study of conversaiton. *Research on Language and Social Interaciton*, 26(1):99–128.

Schegloff, E. A. (1996). Issues of relevance for discourse analysis: Contingency in action, interaction and co-participant context. In *Computational and conversational discourse: Burning Issues – An Interdisciplinary Account*, pages 3–35. Springer-Verlag Berlin Heidelberg.

Schegloff, E. A. (2000). When 'others' initiate repair. *Applied Linguistic*, 21(2):205–243.

Schegloff, E. A. (2007). *Sequence Organization in Interaction: Volume 1: A Primer in Conversation Analysis*. Cambridge University Press, 1 edition.

Schegloff, E. A., Jefferson, G., and Sacks, H. (1977). PreferenceSelfCorrectSm.pdf. *Language*, 53(2):361–382.

Schiffrin, A. (2005). *Modelling Speech Acts in Conversational Discourse*. PhD thesis, The University of Leeds.

Schlangen, D. (2004). Causes and Strategies for Requesting Clarification in Dialogue. In *5th Workshop of the ACL SIG on Discourse and Dialogue*.

Schlobinski, P. (1997). *Syntax des gesprochenen Deutsch*. Westdeutscher Verlag.

Schmidt, R. W. (1990). The role of consciousness in second language learning1. *Applied linguistics*, 11(2):129–158.

Schnier, C., Pitsch, K., Dierker, A., and Hermann, T. (2011). Adaptability of communicative resources in ar-based cooperation. In *Proceedings of the 2nd Workshop on Gesture and Speech in Interaction (GESPIN 2011)*.

Schulze, M. (2008). Ai in call–artificially inflated or almost imminent? *Calico Journal*, 25(3):510–527.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*, volume 626. Cambridge university press.

Shriberg, E. E. (1994). *Preliminaries to a Theory of Speech Disfluencies*. Phd, University of California at Berkeley.

Simmons-Mackie, N. and Damico, J. S. (2008). Exposed and embedded corrections in aphasia therapy: issues of voice and identity. *International Journal of Language & Communication Disorders*, 43(sup1):5–17.

Skantze, G. and Hjalmarsson, A. (2010). Towards Incremental Speech Generation in Dialogue Systems. In *SIGDIAL*, pages 1–8.

Small, S., Cottrell, G., and Tanenhaus, M. (1987). *Lexical ambiguity resolution*. Morgan Kaufman Publishers, Inc., Los Altos, CA.

Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *Tesol Quarterly*, 39(1):33–58.

Smith, B. (2008). Methodological hurdles in capturing CMC data: The case of the missing self-repair. *Language Learning and Technology*, 12(1):85–103.

Spranz-Fogasy, T. (2002). *Interaktionsprofile: Die Herausbildung individueller Handlungstypik in Gesprächen.* Radolfzell: Verlag für Gesprächsforschung.

Ståhl, O., Gambäck, B., Turunen, M., and Hakulinen, J. (2009). A mobile health and fitness companion demonstrator. In *EACL '09*, pages 65–68.

Stewart, I. A. D. and File, P. (2007). Let's chat: A conversational dialogue system for second language practice. *Computer Assisted Language Learning*, 20(2):97–116.

Stivers, T. (2005). Modified repeats: One method for asserting primary rights from second position. *Research on Language and Social Interaction*, 38(2):131–158.

Stivers, T. and Enfield, N. J. (2010). A coding scheme for question–response sequences in conversation. *Journal of Pragmatics*, 42(10):2620–2626.

Stoyanchev, S., Liu, A., and Hirschberg, J. (2013). Modelling human clarification strategies. In *Procedings of the SIGDIAL 2013 Conference*, pages 137–141.

Suchman, L. A. (1985). *Plans and situated actions: the problem of human-machine communication*. XEROX.

Sussenbach, L., Riether, N., Schneider, S., Berger, I., Kummert, F., Lutkebohle, I., and Pitsch, K. (2014). A robot as fitness companion: towards an interactive action-based motivation model. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, pages 286–293. IEEE.

Tannen, D. (1987). Repetition in conversation: Toward a poetics of talk. *Language*, pages 574–605.

ten Have, P. (2007). *Doing Conversation Analysis:2nd (Second) edition*. SAGE Publications, 2 edition.

Tetreault, J. R. and Chodorow, M. (2008). Native judgments of non-native usage: Experiments in preposition error detection. In *Proceedings of the Workshop on Human Judgements in Computational Linguistics*, pages 24–32. Association for Computational Linguistics.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558.

Thórisson, K. R. (2002). Natural turn-taking needs no manual: Computational theory and model, from perception to action. In *Multimodality in language and speech systems*, pages 173–207. Springer.

Tudini, V. (2003). Using native speakers in chat. *Language Learning and Technology*, 7(3):141–159.

Tudini, V. (2007). Negotiation and intercultural learning in italian native speaker chat rooms. *The Modern Language Journal*, 91(4):577–601.

Tudini, V. (2010). *Online Second Language Acquisition: Conversation Analysis of Online Chat*. Continuum.

Vandergriff, I. (2013). " my major is english, belive it or not:)"-participant orientations in nonnative/native text chat. *CALICO Journal*, 30(3):393.

Vandergriff, I. (2014). A pragmatic investigation of emoticon use in nonnative/native speaker text chat. *Language@Internet*, 11(4).

Vandewaetere, M. and Clarebout, G. (2014). Advanced technologies for personalized learning, instruction, and performance. In *Handbook of Research on Educational Communications and Technology*, pages 425–437. Springer.

Varonis, E. M. and Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics*, 6(1):71–90.

Vicente-Rasoamalala, L. (2009). *Teachers' reactions to foreign language learner output*. PhD thesis, Universitat de Barcelona.

Vickers, C. H. (2010). Language competence and the construction of expert–novice in ns–nns interaction. *Journal of Pragmatics*, 42(1):116 – 138.

Vöge, M. (2008). *All You Need is Laugh*. PhD thesis, Syddansk Universitet.

Voida, A., Mynatt, E. D., Erickson, T., and Kellogg, W. A. (2004). Interviewing over instant messaging. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1344–1347.

von Ahn, L. and Hacker, S. (2012). Duolingo.

von Reth, H. and Schek, H. (1977). *Eine Zugriffsmethode für die phonetische Ähnlichkeitssuche*. Heidelberg Scientific Center Technical reports. Wiss. Zentrum.

Wagner, J. and Gardner, R. (2004). Introduction. In Gardner, R. and Wagner, J., editors, *Second Language Conversations*, pages 1–17. Continuum, New York.

Wahlster, W. and Kobsa, A. (1986). Dialogue-based user models. *Proceedings of the IEEE*, 74(7):948–960.

Wallace, R. S. (2003). *The Elements of AIML Style*. Alice A.I. Foundation, Inc.

Waterworth, J. (1986). Conversational analysis and human-computer dialog design. *ACM SIGCHI Bulletin*, 18(2):54–55.

Wen, Z. E. and Mota, M. B. (2015). *Working Memory in Second Language Acquisition and Processing*, volume 87. Multilingual Matters.

Wendemuth, A. and Biundo, S. (2012). A companion technology for cognitive technical systems. In *Cognitive Behavioural Systems*, pages 89–103. Springer.

Wickham, M. and Woods, M. (2005). Reflecting on the strategic use of caqdas to manage and report on the qualitative research process. *The qualitative report*, 10(4):687–702.

Wik, P. and Hjalmarsson, A. (2009). Embodied conversational agents in computer assisted language learning. *Speech communication*, 51(10):1024–1037.

Wik, P., Hjalmarsson, A., and Brusk, J. (2007). Computer assisted conversation training for second language learners. In *Fonetik 2007, 30 maj-1 juni 2007, Stockholm*, volume 50, pages 57–60.

Wilkinson, S. and Kitzinger, C. (2006). Surprise as an interactional achievement: Reaction tokens in conversation. *Social psychology quarterly*, 69(2):150–182.

Wilks, Y. (2005). Artificial Companions. *Interdisc. Science Rev.*, 30(2):145–152.

Wilks, Y., Jasiewicz, J., Catizone, R., Galescu, L., Martinez, K., and Rugs, D. (2015). CALONIS: An artificial companion within a smart home for the care of cognitively impaired patients. In Bodine, C., Helal, S., Gu, T., and Mokhtari, M., editors, *Smart Homes and Health Telematics*, volume 8456 of *Lecture Notes in Computer Science*, pages 255–260. Springer International Publishing.

Wilske, S. (2014). *Form and Meaning in Dialog-Based Computer-Assisted Language Learning*. PhD thesis, University of Saarland.

Yaghoubzadeh, R., Kramer, M., Pitsch, K., and Kopp, S. (2013). Virtual agents as daily assistants for elderly or cognitively impaired people. In *Intelligent Virtual Agents*, pages 79–91. Springer.

Yamazaki, A., Yamazaki, K., Kuno, Y., Burdelski, M., Kawashima, M., and Kuzuoka, H. (2008). Precision timing in human-robot interaction: Coordination of head movement and utterance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 131–140, New York, NY, USA. ACM.

Yang, H.-C. and Zapata-Rivera, D. (2010). Interlanguage pragmatics with a pedagogical agent: the request game. *Computer Assisted Language Learning*, 23(5):395–412.

Yasavur, U., Lisetti, C., and Rishe, N. (2014). Let's talk! speaking virtual counselor offers you a brief intervention. *Journal on Multimodal User Interfaces*, 8(4):381–398.

Zadeh, L. A. (1983). A computational approach to fuzzy quantifiers in natural languages. *Computers & Mathematics with applications*, 9(1):149–184.

Zakos, J. and Capper, L. (2008). CLIVE - an artificially intelligent chat robot for conversational language practice. In *Proceedings of the 5th Hellenic conference on Artificial Intelligence: Theories, Models and Applications*, pages 437 – 442. Springer-Verlag.

Zinsmeister, H. and Breckle, M. (2012). The alesko learner corpus: design–annotation–quantitative analyses. *Multilingual Corpora and Multilingual Corpus Analysis. Amsterdam: John Benjamins*, pages 71–96.

Zourou, K. (2012). Towards a typology of corrective feedback moves in an asynchronous distance language learning environment. *Media in Foreign Language Teaching and Learning*, 217242.

Zwarts, S., Johnson, M., and Dale, R. (2010). Detecting Speech Repairs Incrementally Using a Noisy Channel Approach. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, number August, pages 1371–1378.