Chat Logs as a Data Source

Ariel K. Sorensen

The University of Kitakyushu

Many researchers have undertaken research on interlanguage. Standard forms of data collection include recording speech samples and collecting written samples. Recently the logs of synchronous chat rooms have also been used. When we compare the logs resulting from synchronous chat sessions, we see that there are substantial differences between chat discourse and written or spoken samples. This leads us to question the validity of chat logs as an interlanguage data source. In this report I intend to demonstrate that chat logs can be a valid source of interlanguage data. In order to demonstrate this, I examined the effects of weekly online quizzes on four classes. Data was collected in the form of field notes, quiz scores, and chat logs. The resulting data was analyzed for consistency with current language theory. Suggestions are included for maintaining construct validity.

これまで多くの研究者が中間言語研究を行ってきた。収集されるデータは録音されたスピーチか文書のサンプルが標準の形式であったが、最近チャットルームのログも使われるようになってきた。文書サンプルや録音されたサンプルとチャットルームのログを比べてみると、いくつかの相違点があることが分かる。ここで、チャットルームのログは中間言語のデータソースとして果たして妥当性があるかという問題が生じてくる。この論文では、チャットルームのログは中間言語データとして妥当性があるという事が示される。4つのクラスで毎週実施されたオンラインクイズの結果が検証され、データとして、ノート、クイズの点数、チャットルームのログが収集された。それらのデータが現代言語学理論と整合性があるかどうかの分析が行われた。更に、構成概念妥当性保持のためのいくつかの提案もなされている。

Chat logs as a data source

Increasingly researchers are relying on synchronous text chat logs as a source of interlanguage data (Lampe, 1999; Haggerty, 2003; Pellettieri, 2000). There are many reasons for this. Relying on spoken discourse transcriptions is very time-consuming. There is overlapped speech. Some utterances are unintelligible due to background noise or microphone placement. Written discourse is also problematic. A learner has extra processing time when writing. There is also a high possibility of using dictionaries, written work of others, and proofreading by people at a higher level. Each of these problems can invalidate interlanguage data.

While some of these problems remain, most of these problems are solved by using logs of synchronous chat sessions. There is no need for transcriptions, no overlapping speech, no unintelligible utterances due to recording problems, and little extra processing time. In this sense, synchronous chat logs are arguably the closest we can get to raw interlanguage samples. This combined with its natural formatting and digital nature make it simple to globally search, reorganize, count instances, and perform many other functions.

In this paper I will describe a demonstration of the use of text chat logs in gathering interlanguage data. The purpose of this demonstration is not to prove to the world how much the students acquired, but to identify the strengths and weaknesses of using this medium, as well as any issues that must be addressed. To this purpose, there is an experiment within an experiment. The initial experiment concerns use of text chat logs to gain insights into how online quiz activities affect students' grammar accuracy. Much has been written on the subject of tests and how they affect acquisition, creating a base of knowledge from which we can make comparisons. The meta-experiment is an analysis of how the text chat logs performed as a data source. Through this analysis the main purpose of this paper is to demonstrate the usefulness of text chat data, its validity as a data medium, and identify issues of practicality and validity that present themselves through practical application.

It is not my intention to prove that all data generated by chat logs is always valid for all purposes for all language research, rather the simple gathering of valid interlanguage data from text chat logs is possible. To maintain construct validity in any research, a researcher must take care that the data gathered is genuine, and relevant to the research. As a new form of data, it is necessary to demonstrate that chat logs can contain genuine, relevant samples that are not invalidated by their nature, nor by the process of gathering the data. Each form of interlanguage data, be it oral, written, or otherwise has its own strengths and weaknesses. Thus it is up to the researcher to decide which form of data is most appropriate to the research being conducted, and then shore up and maintain data validity.

It would be helpful here to define some of the terms used in this paper at this point. Chat logs are defined as a text file containing all the discourse submitted (posted) by interlocutors in a synchronous chat conversation using computers. Synchronous chat is defined as a conversation using computers for a textual conversation in real time, as opposed to asynchronous chat, commonly referred to as an online bulletin board. Telephone conversations will not be included as they are beyond the scope of this paper. Synchronous text chat interlocutors will be referred to in this paper as "chatters."

Traditionally, speech samples have been accepted as valid sources of interlanguage data, as have writing samples. Assuming these are valid sources of data, as questioning this is also beyond the scope of this report, we must compare these forms in order to define what synchronous chat logs are. Using this as a basis, we can then compare features. To make the comparison, I will use a chart originally created by David Crystal (2001, p.26-28), inserting a third column to compare synchronous text chat. Although lengthy, this chart is necessary for an understanding of how chat room logs compare to speech samples and written samples, further helping a researcher choose which genre of data may be most relevant. This chart is included in the appendix at the end of this report (Speech vs. Writing: Crystal, 2001).

As can be seen on this rather lengthy chart, chat has much in common with speech and writing, yet is clearly not either. With this in mind, we must question the validity of using this form of data in interlanguage research. In speech samples, it was stated that code-switching is much more common in EFL classrooms. If a researcher wants more data on the interlanguage in the target language, chat logs may provide that. If a researcher wants to discover how much the student is willing to orally produce in a classroom and compare the mother tongue with instances of target language production, oral samples may better fit the purpose. If a researcher is researching the accuracy of spontaneous production in the target language, as is often the case in interlanguage research, chat logs may have many advantages.

Case study in gathering data through chat logs

In order to demonstrate that valid data can be gathered through chat logs, it was necessary to set up a study that actually made use of this data. One requirement for data to be valid is that it is applicable. Although this study is not intended to prove that all chat log data is valid, it does intend to demonstrate the usefulness of chat log data, as well as support the validity of its use, and make suggestions for maintaining validity. In order to be a complete experiment, proving beyond all doubt that chat logs are a source of valid interlanguage data, this study would necessarily include an equal amount of transcriptions of students' speech, with comparisons made between the mediums.

However, as already discussed, chat logs and speech acts are not identical, and it is arguable that chat logs more accurately represent interlanguage than speech. Although that argument will not be followed up at this point, a comparison chart of written, spoken, and text chat mediums is included in the index of this report which can help researchers decide for themselves if speech samples would add validity to specific research projects using text chat logs. This study focuses on whether or not valid data can be gathered through this medium, and what is uniquely required by this medium to maintain validity. Obviously there are many threats to the validity of a given study not included in this paper, however there are a number of threats that must be avoided in all instances of research using chat logs as a source of interlanguage data, and these are discussed in the conclusions section of this paper.

To verify whether or not valid data can come from chat session logs, and discover issues of maintaining the integrity of the data, I set up a demonstration. I used chat sessions extensively in two classes at two universities for a total of four "Speaking" classes. All participants were informed of the experimental nature of the course and agreed to take part. The participants expressed some enthusiasm for the project as a welcome change. All four classes consisted of second year Japanese students of various majors. Data was gathered in the form of chat logs and field notes and applied to the analysis of the resulting speech patterns before and after administration of online activities. Finally, validity issues were identified, not as much for the purpose of proving that this particular usage demonstration had concrete data, but more to discover threats to chat log data in relation to interlanguage research.

Goals

Although the main purpose of this report deals with issues of chat log data validity, the goal of this case study was to use mainly chat log data to investigate changes that occurred in students' spontaneous production corresponding to online language learning activities. These online study activities are

described below and culminated in a weekly quiz. According to the comparison chart provided in the index, the data should be rich in mistakes similar to the ones observed in spoken samples, yet far easier to analyze and contain fewer instances of code-switching. If the text chat logs contained valid data, analysis of the resulting data should give results predictable by current interlanguage research. The main goal of demonstrating predictable results is to demonstrate a case in which data validity was maintained, and issues of application and maintaining data integrity can be identified as concerning interlanguage research.

Administration

I was interested in the effects of online grammar quizzes on learner's spontaneous production. I created a set of quizzes and preparation exercises and placed them on the internet. The software used to create these exercises is HotPotatoes v.4. The participants were required to prepare for the weekly quiz by first taking the preparatory exercises outside of class time. The weekly quiz, administered during class time, requires knowledge of vocabulary, spelling, and specific discreet grammar points. The quizzes were similar to those administered to students of Hong Kong Polytechnic University (Mak, 1999).

The exercises were given in the following order. First there was a vocabulary exercise designed to familiarize the learners with the target vocabulary items and how to spell them. This was in the form of online flashcards. The flashcards are similar to the well-known version. A student sees on the monitor a vocabulary item or its definition. The student must recall what the definition of the shown item is, or if the definition is shown, what the vocabulary word is, and how it is spelled. The student then clicks a button to see the other side of the virtual card for feedback. The cards are randomly selected by the computer to avoid memorizing the order of the answers. The student can do this as many times as necessary until mastering the vocabulary items, including definitions and spelling. In the majority of these items the definitions were in English.

Following this was an exercise in which the learners must spell the target vocabulary correctly. This was done by means of a common input form. The student sees the definition of the target word, then must type the word in a given input space. Clicking on a button, the student can see if the word was properly spelled, and may receive hints in case of mistakes. Again, the order of the vocabulary was randomly selected by the computer. A score is given automatically, and the student can try again until the student is satisfied with the score. This score is not collected, and is only seen by the student as instant feedback.

After completing these exercises until they felt they had mastered the vocabulary and spelling, they were to do exercises designed to reinforce a discreet grammar point, which also use the target vocabulary extensively. This was done by means of cloze tests or short answer quizzes, or both. The cloze tests had ten sample sentences that required use of the target structures. The words that comprised the target structures were omitted, and the students were to fill them in. Cloze test forms were administered intact with no random elements. Clues were also available by click of a button in case the students could not guess. Again scores were provided only for instant feedback, not for collection purposes.

The short answer quizzes were similar to the cloze tests. A question or other prompt was provided in an attempt to elicit a specific response. The input was compared to all possible

correct responses the author of the quiz could guess and scored accordingly. A list of twenty questions was compiled and the computer randomly selects ten of these questions to be answered, so each quiz is different in order and repetition. Again hints were available and the scores were not collected.

The final preparatory exercise in this sequence was a quiz with identical questions to the final quiz. This quiz contained questions or prompts requiring longer answers. The prompts were designed to elicit specific answers, and the answers were compared to a list of all possible correct answers the author could guess. The responses included as much of the target vocabulary and structures as possible. As with the short answer quizzes, twenty prompts were prepared with a list of all possible correct answers. Ten of these are randomly selected by the computer and administered one by one.

Participants were to do all these preparatory exercises outside of class time, freeing up class time for the chat sessions and other speaking activities. This enabled the students to practice as much or as little as they felt was necessary before each weekly quiz. This was in an attempt to counter the problems posed by good or bad test takers, and ensure that all students had actually studied the material and felt they had a mastery of the target vocabulary and structures sufficient to pass the weekly tests.

In the first twenty minutes of class time, the students complete the weekly quiz within the fifteen-minute time limit allotted by the quiz program. Failure to finish the quiz within the time limit redirects the participants to another page, disallowing a score to be given. Upon completion of the quiz, the participant is instantly given a score calculated by the computer and sent to a text file for later analysis and course assessment. Using this score as one means of assessment encouraged the students to study for the weekly quizzes more carefully. The entries in the text file included the participant's name, identification number, class identification, quiz number, score, and time and date of taking the quiz. This allowed a before – after comparison by class and by individual. In the cases of a student being unable to finish a weekly quiz, some time was set aside for make-up quizzes in the last two weeks of the course.

Chat sessions

The chat sessions were held for part of the class time each week. Each class was given ninety minutes. The first twenty minutes of each class consisted of calling roll, during which time the students logged on to their computers and began the weekly quiz. A sufficient number of computers were available so each participant had one available during class time. Weekly quizzes were administered on ten consecutive weeks. The first week of the course was an introduction to chat sessions, how to navigate the website, assignment of special chat names for anonymity, and signing the permission list. During the second week of the course the students did some practice quizzes to learn the functional processes of taking the quizzes and more practice in the chat rooms. The first weekly quiz was administered the third week of the course, and the tenth quiz being the final quiz was administered on the twelfth week. The thirteenth through fifteenth weeks also included a chat session. This guaranteed some data before and after administering the weekly quizzes for analysis. None of the activities were available for more than one and a half weeks before administration of the corresponding weekly quiz in order to maintain separation between before and after quiz production. In an attempt to make the quizzes more relevant to their actual language production skills, the target grammar was based on frequent mistakes collected in class, and from written and spoken samples from previous semesters.

After the weekly quiz was finished, we had an activity which required a considerable amount of language production. For this, often they were divided into two groups. One group performed the activity in a chat room. The other worked on the activity in face to face spoken discourse. This allowed me to observe differences in their speech compared to their chat samples, to monitor for problems, and to take field notes on frequent mistakes. When nearly half the remaining time was up the groups switched their medium. With some activities this was impractical, so the entire class used the chat medium, or the spoken medium. In each session the participants were to use only what they know, with no outside help, and were fully allowed to be creative. During this time I listened and took notes of their speech, including code-switching, frequently occurring grammar mistakes, and communication breakdowns. I also monitored the students to ensure that they were not using dictionaries or notes. They were told that they would not be penalized for mistakes, but would have a portion of their final assessment on the quantity of their chat session participation. This resulted in hundreds of pages of chat logs. A rated sample page of the chat logs is included at the end of this report in the index section.

Analysis

When all of the data was collected it was analyzed. The field notes were compiled, the quiz scores were assessed. The chat logs were then carefully scrutinized using an obligatory occasion analysis similar to the one carried out by Pica (1983). In this analysis two people rated the samples for instances of usage of target structures. The ratings were based on correct usage, incorrect usage, and omission of obligatory usage. What structures would be analyzed, and how it would be rated was agreed upon verbally. The same data were rated without reference to the ratings given by the other rater. Instances of target usage attempts were recorded by date and compared with the date of the quiz that contained the target structures for a comparison of before and after. As one example of a target structure, I was able to gather data from all four classes on the use of *a*, *an*, t*he*, and zero article. Each class had different quantities of usage of different target structures, so I was not able to gather data on all target structures from all four classes.

After discussion of the few instances of rating discrepancies, both raters came to an agreement of the data ratings on each point. The frequency of usage before the quiz was administered, compared to after administration, was graphed. The graphed grammar categories included mistaken use, correct use, and omission of obligatory use. Total word count and frequency of target structure usage was also included. The results of each class as a whole were charted, as well as four individuals from each class, giving a cross-sectional view of sixteen students. The sample page in the index is scored for reference. Although there were numerous instances of grammar mistakes, only the target structures were rated. In two of these classes I was able to gather enough data to compare the effects of nine lessons and in the other two classes only six due to lack of spontaneous usage of the target structures in the chat logs. Using the scores from the weekly quizzes it was easy to identify the students who had regularly completed the required preparatory activities on time and those that had not. The sixteen students in the cross section had all completed the activities on time, and thus were appropriate for pre and post-quiz analysis. I also included analysis and graphs of grammar usage by whole class as well, and the results were further scrutinized by comparison with the field notes I took during administration.

Findings

When the field notes were added, I found that there were many instances of regression that can be explained by the capacity model (Ellis, 2001). Chat log samples reflected the assertations of Loschkey and Bley-Vroman (1993) about making contexts obligatory for discreet grammar point usage and automaticity described by Ellis (2001). They also demonstrate the characteristics of overlearning described by Lightbown (1983). Following learning of specific targets, the participants reverted back to their natural order of acquisition as predicted by Lee & Vanpatten (1995) and Pica (1983). I found no instances that contradicted current second language theory.

I discovered that the mistakes I found in the text logs closely resembled the mistakes I observed in their spontaneous speech. I found that the students were much more likely to use the target language in the text chat than in spontaneous speech acts with their peers. Length of turns seemed to be similar in speech acts and in text chat; however, I did not collect enough complete speech samples to statistically prove this similarity. I can state that when students spoke with their peers I often had to ask them to speak in the target language, while that was completely unnecessary in text chat.

More importantly for this report, during this process I discovered that there were some points to keep in mind when attempting to use data from chat sessions for interlanguage research. Although these points are not sufficient of themselves to guarantee validity in every case, they seemed to me to be the most important findings of this research. Failure to follow these guidelines would invalidate any interlanguage research based on chat log data.

Conclusions

As has already been pointed out, students in language classes tend to attempt to use more complete sentences, more grammatically accurate utterances than in ordinary discourse, as this is how most language teachers teach (Beauvois, 1994; 1997; Kelm, 1992; Kern, 1995; Chun, 1999). Given this we can see some advantages of using a physical language classroom as an environment for gathering interlanguage data from chatroom sessions. For interlanguage research, we want to maximize the instances of spontaneous speech that the student feels is correctly formed. This assists the production of a larger quantity of quality data, and it also assists in guarding against some forms of data invalidation.

In order to maintain the validity of the data from chat sessions, a researcher should physically monitor the learners for data invalidation. These forms of invalidation have counterparts in speaking and writing and would also invalidate those samples. Examples of data invalidation would include such instances as using the copy and paste functions of their computers to insert the words of others into their posts, as if they were their own. This would be the equivalent in speaking of a student reading out loud something written by another. This would not be valid as a source of interlanguage, just as its written equivalent plagiarism would not be.

Another instance of data invalidation would be *spoofing*. Spoofing is when a person logs in to a chat room under a character name used by another person. In writing samples this would be considered cheating, such as when one student asks another to do her homework or take her test for her. It could also be considered a form of forgery.

Most instances of data invalidation can be easily prevented by using personal passwords to gain access to the chatroom, and by holding the chat sessions in a physical room carefully monitored by the course instructor. Physically monitoring the participants of a chat session in a language course is a requirement for maintaining validity of data. This creates a special genre of chat session I labeled "language classroom chat session" in previous research (Sorensen, 2003). Due to the possible anonymity of digital discourse, a researcher must take extra measures to ensure that the data came from the source that is claimed. It is possible for one of the participating interlocutors to not even exist outside of the chat environment, commonly called a *bot*.

Practical application of text chat data revealed it to be far easier to analyze than common written or spoken samples. It is textually intact, it is in a form complete with data and time of production, and can easily be imported into most database programs. In this case study I used Microsoft Excel. This allows global searching and easy reorganization. Due to its inherent organization, rating is far easier and can be easily double-checked. The amount of analyzed production produced would have taken far longer if it required transcription from speech and did not allow such easy reorganization or global searching. It would likely have resulted in far more inter-rater discrepancies as well.

In summary, there are always considerations relevant to the validity of any data collected for the purpose of interlanguage research. Text chat logs can provide data of higher quantity and quality than speech samples and, with care, have the potential to be equally as valid. They are also far easier to collect accurately, rate, and analyze.

References

Beauvois, M. H. (1992a). Computer-assisted classroom discourse in French using networked computers. Unpublished doctoral dissertation. In M. Bush & R. Terry (Eds.), *Technology-enhanced language learning*, (pp 172). Illinois: National Textbook Company.

Beauvois, M. H. (1992b). Computer-assisted classroom discussion in the classroom: Conversation in slow motion. *Foreign Language Annals.* 25, 455-464.

Beauvois, M. H. (1994). "E-talk: attitudes and motivation in computer-assisted classroom discussion. *Computers and the Humanities,* 28, 177-190.

Beauvois, M. H. & Eledge, J. (1996). Personality types and megabytes: student attitudes toward computer mediated communication (CMC) in the language classroom *CALICO Journal* 13(2,3), 27-45.

Beauvois, M. H. (1997). Computer-mediated communication (CMC): Technology for improving speaking and writing. In M. Bush & R. Terry (Eds.). *Technology-enhanced language learning* (pp. 168). Illinois National Textbook Company

Chun, D. (1999). Real-time class discussion on a LAN. In M. Warschauer (Ed.). *Virtual connections*. University of Hawai'i Press: Honolulu Hawaii.

Cononelos, T., & Oliva, M. (1993). Using computer networks to enhance foreign language / culture education. *Foreign Language Annals,* 26, 525-534.

Crystal, D.(2001). *Language and the internet*. UK: Cambridge University Press

Ellis, R. (2001). Non-reciprocal tasks, comprehension and second language acquisition. In M. Bygate, P. Skehan & M. Swain, (Eds), *Researching pedagogic tasks, second language learning* (pp. 49-74).*Teaching and Testing*. Harlow: Longman

Falsetti, J. (1999) SchMOOze university: A moo for ESL/EFL students. In Warschauer, M. (Ed.), *Virtual connections*. University of Hawaii: University of Hawai'i Press

Haggerty, K. (2003, September 10)*Conversational analysis: a study of on-line chatrooms vs. oral communication*. Western Connecticut State University. Online thesis retrieved September 10, 2003, from http://people.wcsu.edu/mccarneyh/acad/HaggertyThesis.html

Kelm, O. R.(1992). The use of synchronous computer networks in second language instruction: a preliminary report. *Foreign Language Annals*. 25, 441-445.

Kern, R. G. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. *Modern Language Journal* 79(4), 457-476.

Lafford, P., & Lafford, B. (1997). Learning language and culture with internet technologies. In M. Bush & R. Terry (Eds.). *Technology enhanced language learning*. National Textbook Company. Illinois: USA.

Lampe, M. (1999). *BabelChat: Structured chatroom usage in first-year German classes*. Retrieved September 10, 2003, from http://lampe.net/lampe/thesis/thesisdefense_lampe.pdf

Lee, J. & Vanpatten, B. (1995). *Making communicative language teaching happen*. USA: McGraw-Hill.

Lightbown, P. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. Seliger & M. Long (Eds.). *Classroom-oriented research in second language acquisition* (pp. 217-243). Rowley MA: Newbury House

Loscky, L. & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S.M. Gass (Eds.).*Tasks and language learning.* Cleveland, UK: Multilingual Matters.

Mak, S. (1999). Designing exercises and tests on the web. In M.Warschauer (1999). *Virtual connections* (pp. 351-353). University of Hawai'i Press: Honolulu Hawaii.

Pellettieri, J. (2000). Negotiation in cyberspace. In Warschauer, M. & Kern, R.(Eds.) *Network-based language teaching: Concepts and practice*. USA: Cambridge University Press

Pica, T.(1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning* 33(4), 465-497.

Sorensen, A. (2003). Synchronous chat logs as a source of interlanguage data. Unpublished masters degree thesis. Anaheim University, New Port Beach, California.

Acknowledgements

I would like to thank the people that participated in my classes and Professor David Nunan for helping me make this something worth reading.

Appendix A. Speech, writing, chat characteristics chart

Speech vs. Writing (Crystal, 2001, p.26-28)

Speech	Writing	Chat Rooms (Author)
1. Speech is time-bound, dynamic, transient. It is part of an interaction in which both participants are usually present, and the speaker has a particular addressee (or several addressees) in mind.	Writing is space-bound, static, permanent. It is the result of a situation in which the writer is usually distant from the reader, and often does not know who the reader is going to be (except in a very vague sense, as in poetry).	Chat is time-bound, dynamic, transient (Beauvois, 1997), and space bound as well. It is usually a result of multiple interlocutors being attentive at the same time (Beauvois, 1997), and the chatter has a particular addressee (or several addressees) in mind, although they are usually distant from the reader (Chun, 1999).

2. There is no time-lag between production and reception, unless one is deliberately introduced by the recipient (and thus, is available for further reaction on the part of the speaker). The spontaneity and speed exchanges make it difficult to engage in complex advance planning. The pressure to think while talking promotes looser construction, repetition, rephrasing, and comment clauses (e.g. you know, you see, mind you). Intonation and pause divide long utterances into manageable chunks, but sentence boundaries are often unclear.	There is always a time-lag between production and reception. Writers must anticipate its effects, as well as the problems posed by having their language read and interpreted by many recipients in diverse settings. Writing allows repeated reading and close analysis, and promotes the development of careful organization and compact expression, with often intricate sentence structure. Units of discourse (sentences, paragraphs) are usually easy to identify through punctuation and layout.	From production to reception there is a very short time lag, from fractions of a second to several seconds, imposed by the connection and processing speeds of the intervening equipment. Chatters tend to expect a real-time discussion. The spontaneity and speed of exchanges make it difficult to engage in complex advance planning. The pressure to think while chatting promotes looser construction, repetition, rephrasing, and signals meaning an extended turn will be taken (Beauvois 1997). Sentence boundaries are usually clearer than in ordinary speaking, but may be divided between conjoining posts by posts from other chatters. Each post will arrive uninterrupted on the receivers' screens. The post may be an entire turn, or a piece of one turn. Reflection is normally reduced by screen size more than time. Pressure to complete the reflection process before the text is no longer visible can be intense (Beauvois, 1997; Kern, 1995).

3. Because participants are typically in face-to-face interaction, they can rely on such extralinguistic cues as facial expression and gesture to aid meaning (feedback). The lexicon of speech is often characteristically vague, using words which refer directly to the situation (deictic expressions, such as that one, in here, right now).	Lack of visual contact means that participants cannot rely on context to make their meaning clear; nor is there any immediate feedback. Most writing therefore avoids the use of deictic expressions, which are likely to be ambiguous.	Lack of visual contact means that participants cannot rely on physical context to make their meaning clear. However, there is immediate feedback. The context is strictly textual, supplemented possibly by emoticons, which take the place of facial expressions, and in some cases can be more expressive (Crystal, 2001). Given that the context is textual, with near instant feedback (Pellettieri, 2000), deictic expressions are allowed, yet must refer to textual content. In a language classroom the usage of the target language is more likely to be in complete form, as this is how most language teachers teach (Beauvois, 1994, 1997; Kelm, 1992; Kern, 1995; Chun, 1999).

4. Many words and constructions are characteristic of (especially informal) speech, such as contracted forms (isn't, he's). Lengthy co-ordinate sentences are normal, and are often of considerable complexity. There is nonsense vocabulary (e.g. thingamajig), obscenity and slang, some of which does not appear in writing, or occurs only as a graphic euphemism (e.g. f***).	Some words and constructions are characteristic of writing, such as multiple instances of subordination in the same sentence, elaborately balanced syntactic patterns, and the long (often multi-page) sentences found in some legal documents. Certain items of vocabulary are never spoken such as the longer names of chemical compounds.	Many words and constructions are characteristic of informal speech, such as contracted forms. Different genres of chat lead to wide variations of grammar usage and correctness. The context being textual, yet time-bound, allows for highly complex sentences that require greater processing time, yet are likely to be less wordy, or are broken into chunks, as opposed to the long, uninterrupted paragraphs that can be found in writing. There is nonsense vocabulary, obscenity, slang, ASCII and graphic emoticons, spelling aberrations, and acronyms representing entire clauses. Omission of words considered unimportant is standard, although grammatically not correct. Some single letters replace whole words. Code switching in language learning settings is rare compared to written or spoken discourse (Beauvois, 1994; 1997, Kelm 1992, Kern, 1995; Chun, 1999).

5. Speech is very suited to social or 'phatic' functions, such as passing the time of day, or any situation where casual and unplanned discourse is desirable. It is also good at expressing social relationships, and personal opinions and attitudes, due to the vast range of nuances which can be expressed by the prosody and accompanying non-verbal features.	Writing is very suited to the recording of facts and the communication of ideas, and to tasks of memory and learning. Written records are easier to keep and scan, tables demonstrate relationships between things, notes and lists provide mnemonics, and text can be read at speeds which suit a person's ability to learn.	Chat is very suited to social or 'phatic' functions, such as passing the time of day, or any situation where casual and unplanned discourse is desirable between multiple interlocutors in an anonymous environment, or at an unknown distance. It is also good at creating social relationships between people that have similar interests that would probably never meet otherwise. These relationships begin between two characters, and have the possibility of evolving into relationships between the humans that they represent (Falsetti, 1999; Crystal, 2001).
6. There is an opportunity to rethink an utterance while the other person is listening (starting again, adding a qualification). However, errors, once spoken, cannot be withdrawn; the speaker must live with the consequences. Interruptions and overlapping speech are normal and highly audible.	Errors and other perceived inadequacies in our writing can be eliminated in later drafts without the reader ever knowing they were there. Interruptions, if they have occurred while writing, are also invisible in the final product.	Errors and other perceived inadequacies in chat can not be eliminated (with exceptions). While most errors are ignored, and in fact outside a moderated chat setting, may be intentional, errors requiring fixing may be recast (Chun, 1999) with an apology. Overlapping text is not possible (Beauvois, 1997; Crystal, 2001).

7. Unique features of speech include most of the prosody. The many nuances of intonation, as well as contrasts of loudness, tempo, rhythm, pause, and other tones of voice cannot be written down with much efficiency.	Unique features of writing include pages, lines, capitalization, spatial organization, and other aspects of punctuation. Only a very few graphic conventions relate to prosody, such as question marks and italics (for emphasis). Several written genres (e.g. timetables, graphs, complex formulae) cannot be read aloud efficiently, but have to be assimilated visually.	Unique features of chat include the potential to converse with unlimited numbers of people in near-real time across wide distances and political boundaries (Beauvois, 1997), the ability to remain completely anonymous, and to converse with other anonymous characters with similar interests. The limits on production time create an environment, in which chatters will normally use only words and grammar they know, except in repeating text. This form is seen by many as the closest form to direct interlanguage (Crystal, 2001; Pellettieri, 2000). The text format allows a log to be kept of the text entries in a format that is by default in a database form (Lafford & Lafford, 1997). Some words are omitted or shortened to speed up typing time. Dialectic forms are created and forgotten at an increased speed (Crystal, 2001). In language learning environments there is a marked increase in use of the target language spontaneously compared to spoken discourse in a classroom (Beauvois, 1992a; 1992b; Beauvois & Eledge, 1996; Cononelos & Oliva, 1994; Kern, 1995)

Appendix B. Sample of chat log with scores

Thu Oct 17 06:54:09 2002	secretcave	Naoya says	I taked mistake. I wanted to write I feel it difficult to use the words a and the. I am sorry.	8c	m0 tc	21
Thu Oct 17 07:05:42 2002	mainauditorium	naoki says	English is too difficult. Because front of vocabulary (a,the,some,etc)		m0	12
Thu Oct 17 07:05:46 2002	language	firefly says	I see...it's difficult. I have not been English teacher...		m0	10
Thu Oct 17 07:07:19 2002	mainauditorium	Hiriyuki says	I like both see and play the baseball. which do you like see or play?	4c	tm	15
Thu Oct 17 07:08:41 2002	mainauditorium	ET says	Hello.Teacher said "what grammer do you think difficult for you." I think we should talk about difficult grammer.	4c 7c		19
Thu Oct 17 07:08:49 2002	sports	Kazu says	Hello,H.R. My name is Kazu. I like soccer,too. Please tell me your favorite player?			16
Thu Oct 17 07:09:24 2002	language	nao says	A difficult grammar for me is what we use "if"and "would".In japanese "kateihou"and"kateihou kako"	4m	am	18
Thu Oct 17 07:11:04 2002	secretcave	Naoya says	I also feel difficult to use the words, to, for, in, on, and so on. When I speak in Japanese, which I should use te ni wo ha, I decide in my sense of words. Peaple speaks in English also do so, I think.	4c 6c 6m 6c	tc	44

Scoring Guide

Column 1: Lesson number + mistake or correct usage.
(Ex. 3c = Lesson 3 grammar point, correct usage)

Column 2: Mistaken use of Zero Article, A / An correct or mistaken use,
The / Some correct or mistaken usage
(Ex. m0 = Mistaken use of zero article, ac = A / an correct)

Column 3: Total word count in this turn.