Humans as Corpus – Language Learning Strategies in Virtually Mediated Authentic Environments

Daniel Jung

LINGO project, HIT centre,
Allegaten 27
5007 Bergen (Norway)
jung@uib.no

Abstract

This paper deals with the design and presuppositions of a research project yet to be conducted. The notion of "linguistically relevant data" is interpreted as the language resource which is represented by other humans. In a MOO environment (virtually mediated), students' language learning strategies are observed and described, especially the use they make of the communication with co-students and native speakers (authentic environment). *Resource* means thus both *people* and *conserved communication* (transcripts, discussion board messages). Rather than discussing the act of collecting and storing these data technically, the attention is drawn to the *use* the students make of those data, how they learn from each other and their own language production.

1. Introduction

According to the aims declared in www.lrec-conf.org/lrec2002/lrec/aims.html,

"two issues are considered particularly relevant: the availability of language resources and the methods for the evaluation of resources, technologies, products and applications."

In the Age of the Internet, contact between people and cultures in different countries can be easily achieved. This means that language learners can access the target language culture by (1) browsing the internet, and by (2) making personal contact with speakers of the target mother tongue. One can speak of the creation of 'authentic learning environments' not necessarily filtered or prepared by the teacher; environments which are made possible by the new ICT.

This presentation deals thus with a highly fluctuous (but nevertheless available) form of language resource (corpus) for learners: other humans. It will, as well, deal with problems, and possibilities, connected to the evaluation of the learning processes, or strategies, in these environments. It is therefore not focusing on the technology itself, but rather on the use, and effect, of it – and it is covering both of the conference aims cited above.

Due to pending funding, this project has not begun yet. I did count on this funding when I first sent in the paper proposal. Anyway, there are no results to be presented. This presentation will therefore be yet another "this should be done"-paper, in a field where all to many of the publications are presentations of "this is what we have done" and rather vague assessments of "this was good" and "that could have been better". This paper may, however, show again the needs for thourough investigation and theory-based research in this field, and offer an angle how to conduct it. And it might prove interesting for the researching community and the conference participants, since *storage* and *collecting* of data are used in a different way than anticipated.

2. Background

2.1. Curricula

The Curricula for French and German as foreign languages in high schools carried by the Norwegian Ministry of Culture, Education and Church Affairs states that "the contact and cooperation between people across country borders has expanded", and that "the pupil has to develop an independent, critical and constructive way of orientating himself in the information society. [...] He shall discover and examine the language and use it from the beginning". The curricula underline the importance of authentic texts and the use of technology which enables the pupils to participate in living languange communities, thus inviting to autonomous learning.

There is very little, if any, theory or research referenced within the curricula which could justify this need. The research has yet to be carried out, it seems.

2.2. MOOs

For a couple of years now, the LINGO project at the University of Bergen (lingo.uib.no) has made use of, and developed, educational MOOs for the purpose of language learning and teaching (covering, as of today, German, French, Spanish and Italian). The MOOs are used both in long distance education and for campus students, and there has been a few minor projects for school classes as well.

The abbreviation MOO stands for Multi user dungeon, Object Oriented. The MOO is both a (multi user) place and a programming language which is object oriented. The first part (multi user dungeon) is often abbreviated MUD, and is a known telnet-based synchronous game platform from the Eighties, which has been taken to new heights by enhancing them with an integrated web interface (the MOOs run by LINGO are based on the open source encore Xpress software, lingua.utdallas.edu/encore). In other words, everything viewable on the internet can be accessed and shared among users and user groups through the MOO. The MOO becomes thus a portal to the target language world. In addition, it is in

itself a virtual place, where the learners within the virtual world metaphor can create, share, hide and destroy their own objects, be it documents, blackboards, tape recorders, rooms etc. All this is, or can be, done in the target language.

These MOOs offer both synchronous and asynchronous forms and forums of learning and interaction. The point is to create a virtual reality where the language itself is the landscape, thus becoming a living language resource. They are connecting people to people, not people to machines. The response is not a predefined mechanical one, like from a language learning CD-ROM, but a human one. The limits of understanding, and interest, are defined in the partner's language skills (and often, one might add, his imagination), not the circuits in the computer or the bits on a disk.

A first introductions to the MOOs' advantages for the language acquisition and possibilities for the classroom is provided by the old article of Turbee (see references). For further information, please consult Haynes/Holmevik, and Schwienhorst.

Putting the curricula and the MOOs together (which means putting people in the MOO), one should get a perfect match of supply and demand.

3. Goal

The main goal is to learn more about learners' behavior and strategies in this authentic, thus cultural, room, or environment, and how they make use of one another, and of native, or target language, speakers. The corpus covered by achieving this goal is thus both the living human and the electronic copy of their contact, be it a random chat or a structured discussion. It is a hermeneutic and qualitative approach, where the understanding and the description of strategies and processes stand in the focus. The analysis is carried out with the overall goal to find out about global functions of language.

The research proposed aims at both pupils/students and language didactics teachers, i.e. students training to be teachers for the 'new generation' of pupils following the curriculum cited above. This brings a double meta aspect into the picture: the pupil's language learning, his meta-knowledge, and the teacher's learning and reflection about his own and the pupil's role.

Next, this project covers both German and French, trying to find out if there are differences in the learners' approach to those languages (German being much 'closer' to Norwegian than French is).
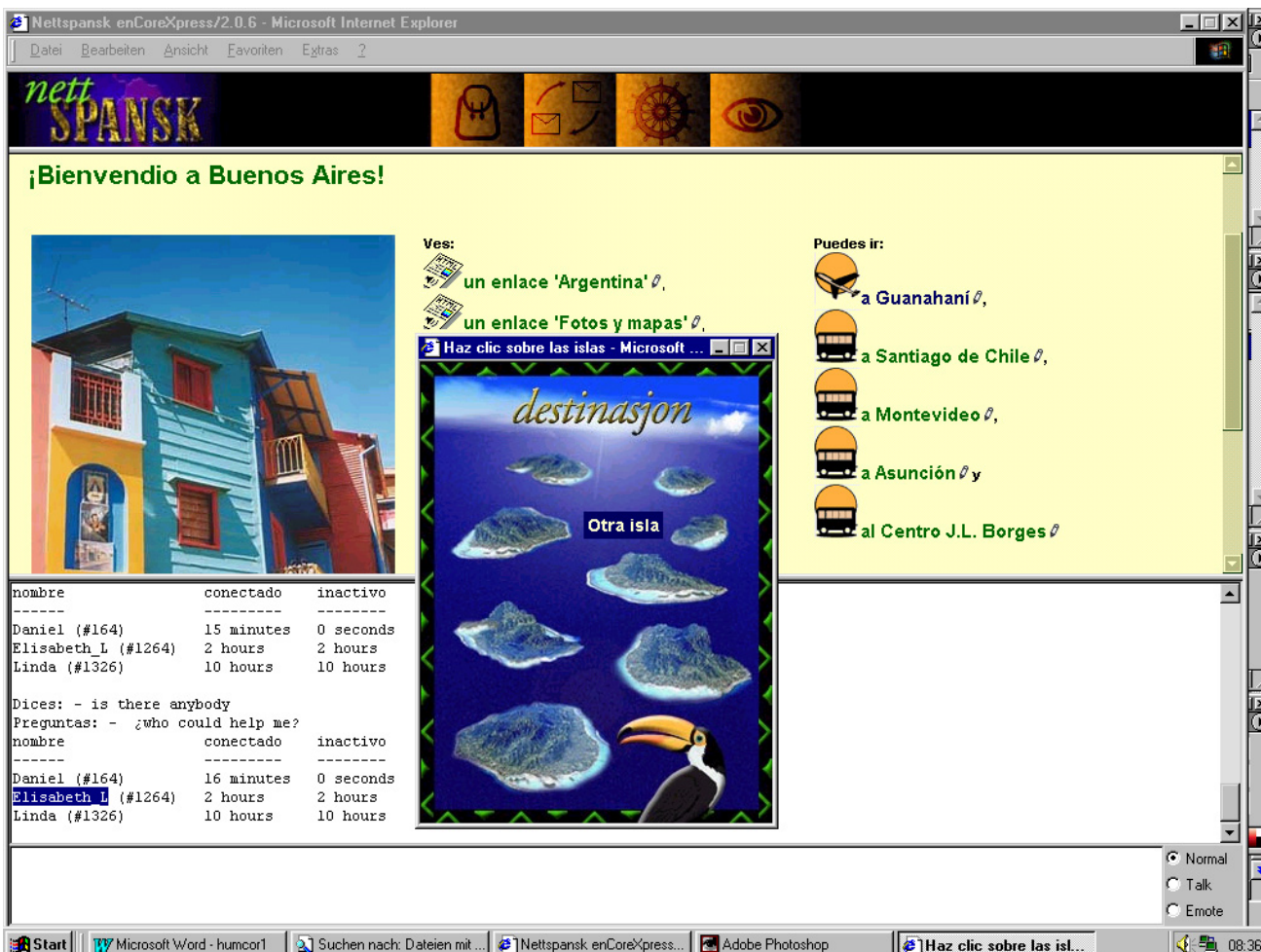

Figure 1: Inside the MOO (here: the Spanish one). The screen is divided into a webframe (where html-content is shown) and a telnet frame (where the chat is going on). The world and the communication is controlled from the command line below. On top: A navigator to get to the main islands by mouseclick. Users can access and alter different imaginative and factual "rooms", to construct their own learning environment. Other students, and random guests, can join them.

The goal then is not only research and finding out about methods. The findings are intended to give the learners a better understanding of what they are doing, and to point out possibilities for (even) better learning. Consequently, the project results should provide a basis for the educating system (i.e., the language departments which use the MOOs) to redesign their education.

The goal could be described as *finding out what is going on in the language learners head when communicating on the MOO, and when recollecting the data gathered from this communication.*

4. Method

This goal should be achieved by observing triggered and untriggered language production and thought process. (The notion of triggering has yet to be carefully investigated and described.) The material will thus consist of:

- messages in a threaded discussion board. This is available and used by the MOO students. They are encouraged to use the target language, though the access is password protected and reserved to the courses, thus not available to random [target language] users. (This might be changed if needed)
- transcripts from MOO sessions, be it the offical teacher driven class or the semi-organized colloquium or the random chat (where the transcripts are available), see figure 2. Transcripts show a "faked oral" conversation put down in keyboard-typed letters. Things have to go fast, and you have to quickly get the meaning across rather than to file on elaborate sentences. This opens for spontanous language production, including errors, misunderstanding, but also "positive" strategies
- collection of surveys (see below), self assessments and interviews. This project being hermeneutic, the form of the surveys etc. will vary and be modeled along the line.

The material includes, of course, all the observations and findings gathered from the items above.

One challange is to create an ethically clear basis for the research. Since language production is an intellectual achievement, the users should be aware of that they are being "probed", and they should give their permission to their material being used. The act of observing could then affect what is observed. But since this whole project is about eye opening, this might not represent a problem.

4.1.1. Collection and storage

In this approach, the learner is responsible for keeping track of what he or she is doing. One of the main purposes of this project is to find out how the learner accesses the human material, i.e. makes use of other people's, preferably native speakers', resources. This includes finding out how he or she collects and stores the material, and what he or she does to access and "decypher" it. One question is to what level language learning is involved – or if the stored transcripts are merely the source of non-linguistic knowledge. In other words, do the students read the transcripts to recollect factual information about, e.g., Marcel Proust, only, or do they have an eye for what is happening linguistically with their own and others participation and means of expression?

I have interviewed one student who copies the html transcripts into his text editor, cuts the Good-afternoon-is-everybody-here-yadayada, corrects some typos, but also structural linguistical errors, prints it out and keeps it for his private exam preparation. The main purpose for him is, however, of course, the non-linguistical content. His

E. exclama:	– Me parece bien, ¡che!
E. pregunta:	– ¿Sabés que significa che?
D. dice:	– no, lo siento...
D. pregunta:	– ¿que dirás que significa?
E. dice:	– Una manera de decir vos o tú. Se lo usa entre amigos.
D. dice:	– somos amigos :–)
E. exclama:	– ¡y si!
D. pregunta:	– ¿pero no sos latina tu tampoco?
E. pregunta:	– Soy noruega. De Kristiansand. ¿Y vos?
D. dice:	– una historia longa... mi padre es alemán, mi madre es francés, mi esposa es noruega
D. dice:	– habitamos en Bergen ahora
E. pregunta:	– Y te fuiste a Noruega por tu mujer...? ¿O ya estabas en Noruega cuando la conociste?
D. dice:	– pero he pasado mis años de joven (?) en Alemania
E. pregunta:	– ¿Sprichst du Deutsch denn...?
D. dice:	– jepp
E. dice:	– Disculpa, una vez sabía hablar alemán yo también, pero ya no me sale nada... Lástima..
D. dice:	– venía para estudiar, y conoció (?) mi esposa futura
E. dice:	– Conocí...
D. dice:	– ok
D. dice:	– han muchos años
E. pregunta:	– Bueno, ¿más adelante tal vezme puedas ayudar a hablar aleman de vuelta?
D. dice:	– lo pongo corectamente con un diccionario, pero...
E. exclama:	– Está bien igual! ¡No te preocupes!
D. dice:	– no, pero quiero aprender más
E. exclama:	– Bueno. ¡Entonces nos ayudamos el uno al otro y vamos a aprender mas los dos!
D. pregunta:	– ¿Podrás conectarte todas las tardes?
E. dice:	– No todas... Pero casi. Por lo menos tres, cuatro veces por semana..
D. dice:	– estás en tu casa ahora
E. pregunta:	– Me parece que en poco tiempo se me va a cortar la conexion. ¿Pero si querés, seguimos chateando mañana..?
D. exclama:	– ¡si!

Figure 2:
A chat transcript with lots of starting points for research

thoughts and methods when correcting the transcript deserve, nevertheless, a closer look.

On the other side, there is the researcher. The material consists of primary and secondary sources. Primary sources are, both for the learner and the researcher, the (virgin) transcripts and forum messages. The secondary sources are the students' washed, or edited, form of the transcripts and messages, but also observation made by close reading of the texts, and survey plus interview results.

Collection and storage are as sophisticated as being based on ordinary text files, stashed away somewhere on the users' (or the researcher's) computers. The focus is on what is in them, and how to get it out mentally, not technically.

4.1.2. Survey

The first survey should ask the learners the following questions:

- How long have you been using the MOO? Why?
- Do you feel comfortable using the MOO?
- What introduction did you get to the possibilities on a MOO?
- What is, in your opinion, the main purpose of the MOO in general? Why does the language department you attend use it?
- What do *you* use it for?
- Have you read transcripts of classes you attended and/or have missed? What did you do with these?
- Have you met native speakers of your target language on the MOO? If any: Were these meetings prearranged? What happened? Did you get to know them?
- Have you used the target language with other non-native speakers? When and how?
- Did you experience language or comprehension problems talking/writing to either of the groups? What did you do then?
- Did anyone make transcripts of the meetings? Did you access them afterwards? How?

This first survey should help to see where the land lies. I have described the goal as *qualitative*. Within this preparing part of the project, there could, or should, be, however, a *quantitative* part involved. And although the project mainly aims at learners of German and French, it would be unwise not to make use of the learners of Spanish and Italian and the resource they represent, at least in the preparing part of the project. They share the same educational setting, and they can provide vital points to where to look further. This is why I have made a screenshot of our spanish MOO (figure 1) and included an authentic, only slightly shortened, chat transcript excerpt from there (figure 2) (and, of course, because the conference takes place in Las Palmas).

The survey should in itself be a trigger for the participants to think about their own behavior, and it could very well lead to new ways for them to assess their own and the others' language skills. Therefore, the survey should be repeated in a similar form a few months from the first one.

The next step should be an in-depth interview with the most prominent participants (as shown by the survey). Simultanously, their behavior and experiences on the MOO should be evaluated and be compared to their own evaluation. The details of this procedure, however, have yet to be found.

5. Conclusion

As I pointed out in the introduction, there are no results, or findings, to be presented in this version of the paper. By the time of the conference, I hope that I will, however, be able to present a more detailed description of the elicitation and evaluation process, and supply with preliminary findings and observations.

6. References

Haynes, Cynthia & Jan Rune Holmevik 1998: High wired: On the design, use, and theory of educational MOOs. Ann Arbor [=The University of Michigan Press]

Little, David 1991: Learner autonomy 1: Definitions, Issues, and Problems. Dublin [=Authentik]

Schwienhorst, Klaus 1997: Talking on the MOO: Learner autonomy and language learning in tandem. Paper presented at the CALLMOO: Enhancing Language Learning Through Internet Technologies, Bergen, Norway. Available at: www.tcd.ie/CLCS/assistants/kschwien/Publications/CALLMOOtalk.htm

Turbee, Lonnie 1996: MOOing in a Foreign Language: how why and who? Paper presented at the Information Technology Education Connection's International Virtual Conference/Exhibition on Schooling and the Information Superhighway. Available at: web.syr.edu/~lmturbee/itechtm.html