

Oddrun Grønvik:

Report from

ALLEX WORKSHOP, UNIVERSITY OF ZIMBABWE, 20.-24.9.1993

Contents:

- 1 Transport
- 2 Stay
- 3 Purpose
- 4 Equipment
- 5 Workshop Programme
- 6 Evaluation of the first Phaze of the ALLEX project
- 7 Starting phaze two of the ALLEX project
- 8 State of project at the end of ALLEX workshop 1993
- 9 Security
- 10 Summing up

AL/INL  
pb 1001 Blindern

19. October 1993

Report from

ALLEX WORKSHOP, UNIVERSITY OF ZIMBABWE, 21.-25.9.1992

#### 1 Transport

I arrived in Harare 19.9. and left 30.9. The Göteborg participants (Gellerstam, Ridings, Sjögreen) arrived 17.9. and left 28.9. Transport standard with KLM was an improvement on last year's with British Airways - more space between seat rows.

#### 2 Stay

All invited experts - i. e. the Göteborg party, Dr Tsonope from University of Botswana and I - stayed at Visitor's Lodge, UZ, on the campus. This was a great help in connection with daily planning and preparation and an improvement on last year's arrangements, when the Göteborg party stayed at a hotel in the city centre.

#### 3 Purpose

To conduct a one-week workshop in cooperation with the African Languages Department, UZ, in order to evaluate the first phase of the ALLEX project, and work out plans for the next phase, 1993-94.

#### 4 Equipment

The project now has 3 personal computers, which are used by the project secretaries (cf Progress Report). There was full agreement that the computers are for the project only, not for the ordinary work of the department.

AL/INL, UiO, has offered to send down two more of our superfluous machines (IBM XT 286), this time with English software and handbooks. This means that editors will be able to place machines in their offices and work directly on them.

Most of the equipment discussed earlier has been bought and is being used, i. e. tape recorders etc. A scanner is being bought and training will be given in Gothenburg to the researcher due to go there for training, cf. pt 7 c.

There is still need of stronger and more sophisticated computers for processing the corpus. Further equipment needs will be discussed in relation to the project budget at a meeting between GU and UiO, to be arranged fairly soon.

#### 5 Workshop Programme

ALD/UZ had prepared a programme which was adjusted a little before start, primarily to leave more time for working on the style manual. Training sessions in running the concordance programme and encoding for concordancing were run parallel to plenary sessions. The programme was carried out as decided, plus of course a good deal of running informal discussion at breaks.

There were fewer participants this time, about twenty, cf. list of participants (in report to come from ALLEX team, UZ). The workshop included 5 second-year students from ALD/UZ and two students from UB. Staff members were mainly the ones who had devoted time to the project and meant to stay with it. The emphasis this time was on discussing the project and furthering it point by point, while last year we did more general teaching.

#### 6 Evaluation of the first Phase of the ALLEX project

The first project year, 1993-94, was to see the entire corpus computer stored, and the style manual finished. At the time of the workshop, (a) the oral material was collected but not yet encoded; (b)

the literature list for the printed part of the corpus was completed, but no scanning done; (c) the style manual prepared through a series of editorial meetings (minutes taken), but not written out in handbook form.

a) Encoding of oral material

The collection of oral material had been carried out with great energy, cf. the first Progress report from ALD/UZ. An oral corpus of some 2.5 mill words is now stored on tape and in handwritten transcripts, and these materials are properly categorized and cared for. The transcripts are now being encoded prior to concordancing.

As pointed out in the progress report from ALD/UZ, the encoding started late and depends on secretarial assistance. The project team had worked out that at the present rate, the encoding of the oral material only would last until the end of January 1994, at the earliest. Minor problems to do with the ordering and indexing of materials were being dealt with.

The collection of oral material had been almost too successful, and in order to carry the project forward at the planned rate, some rescheduling was required, see below pt 7 a.

b) Printed materials

According to previous plans the scanner was to be bought in Göteborg, and the research scholar who was to visit Göteborg for a three month stay would learn scanning and concordancing there by processing the printed materials selected for the corpus. As there had been a delay in selecting a candidate for this stay, this job was not yet done.

c) Style manual

A draft style manual was completed during the workshop, with decisions on entry format, punctuation etc. On suggestion from Dr Tsonope, we used as a "model" the alphabetical checkword list in Landau: *Lexicography* (1985). The style manual will be polished up immediately and added to the ALLEX report from ALD/UZ as an appendix. For all practical purposes, it was ready for use when the workshop was finished.

## 7 Starting phase two of the ALLEX project

What needs to be done in 1993-94, is to (a) finish encoding the corpus, and run concordances from it, (b) enter all headwords with formal information (tone pattern and part of speech) and semantic information (draft definitions correctly ordered). These tasks must run parallel if there is to be any chance at all of having a draft dictionary ready by August 1994. c) To push the dictionary forward, the two research scholarships for 1993 and 1994 must be filled as soon as possible and candidates selected by the end of the workshop.

a)

In order to get started on phase two without delay, i.e. the selection of headwords and actual drafting of the dictionary text, the encoding of the oral material had to be rescheduled.

It was decided that

- only half the registered oral material, i.e. 1.25 mill words, was needed for this corpus, as the total corpus is planned to 2.5 mill. A plan was set up to achieve even geographical distribution.
- two secretaries would go on encoding oral material full time until the end of October 1993. The encoding of the rest of the oral material needed for this project would continue, but from November 1st headword entering would have priority.
- The encoding of oral material for this project must be finished by February 1st 1994.
- All the oral material should eventually be encoded, for future use, but not as a necessary part of this project.

The books selected for the corpus were to be ordered immediately and brought to Göteborg for scanning and concordancing by the research scholar going to Göteborg (see pt 7 c below).

During the workshop, the College Press publishing house in Harare donated textbook files in Shona equalling ca 1000 pages of text, which are ready to use once they've been formatted for

concordancing. The College Press is interested in publishing the dictionary, and expects to receive a finished manuscript in the autumn of 1995.

#### b) Drafting the dictionary

By the end of August 1994, the dictionary should be drafted, with headwords, tone pattern, part of speech information and draft definitions correctly ordered.

It was decided that

- the entering of headwords with formal information would start immediately, building on existing Shona dictionaries, and last until January 1st 1994. To achieve 24 000 headwords in three months, 400 headwords must be entered each working day. Entering this information must be a fulltime job for a good and careful secretary. At least two editors must start immediately selecting headwords. The files containing the headwords with formal information, set up in the pattern decided on in the style manual, must be copied so that the research scholar coming to Oslo can bring them with him.
- the drafting and ordering of definitions will be a full time job for three editors from eight months from January 1st 1994. If ca 25 000 words are to be defined in eight months, each editor will have to finish ca 50 entries per working day. This is a tall order, even with the simplified entry format selected for derivatives/compounds, and will require very hard work from the editorial team.

If the project is to have any chance of succeeding, the editors must have their other obligations limited in the project period, especially their administrative duties. The department must consider ways immediately to ease the load of general departmental tasks for the core of its editorial team.

#### c) Research scholars 1993 and 1994

Three candidates were put forward for the research scholarships, in priority order. It was decided that nr 1, Edgar Mberi, go to Göteborg as soon as possible, preferably in the beginning of November 1993, to learn and do scanning and concordancing. Nr 2, Bernhard Batidzirai, should come to Oslo as early in 1994 as possible to start drafting his part of the dictionary, while learning about general lexicographical practice. The Scandinavian participants would start preparations for these visits immediately on return to Göteborg and Oslo.

Both candidates were asked and it was arranged that they can come on short notice as soon as their host institutions are ready to receive them.

#### 8 State of project at the end of ALLEX workshop 1993

The project definition is unchanged, i. e.:

- The Shona monolingual dictionary is to have 20-25 000 headwords in the space of ca 500 pages. It is to be a general dictionary suitable for secondary school and adult users.
- Headword selection will be based on previous Shona lexicography, with additions from all relevant (secondary school) subjects and topics, and it is to be representative of Shona as it is spoken and written now.
- The emphasis is to be on definitions and examples (from the corpus) while information on etymology, pronunciation etc will be severely limited (i. e. to abbreviation, no discussion of variants or doubtful issues etc)
- The definition style will build on the Cobuild model. Definitions will be in plain and simple language and refer to general experience and understanding of concepts among expected users. The lexicographical precision level is to conform with recognized international lexicographical practice. (A sample exercise of definition practice was gained in a final session with Oddrun Grønvik, at the first editorial meeting after the workshop (30.9.1993), when ten words selected for defining during the workshop, were analysed and defined.)
- examples will be taken from the corpus if at all possible. Minor adjustments and abbreviations of quotes (without altering their sense or essential wording) is acceptable.

#### 9 Security

It was decided that the editorial team must make arrangements immediately for safe storage of copies of all encoded material, either within the university or with the College Press.

#### 10 Summing up

The workshop was successful and necessary. The project team at ALD/UiZ are committed to making a go of it, and they are well aware that they are working on a pioneer project and that their progress is being followed with a great deal of interest in many places. They also know that they have taken on a heavy load of work, and that other engagements will have to wait until this project is finished.

A great encouragement was the news that NUFU had allowed the project additional funding for 1993-94, with good chances for further added funds next year if the project progresses as planned.

It should be noted that the ALLEX project has the full support of its own university, the Vice Chancellor taking a great personal interest in following its progress. The Vice Chancellor's opening speech showed that he fully appreciates the perspectives which could open up both institutionally and professionally if this project is as successful in the end as it now looks. This in turn could mean a great step forward for the cultivation of Shona and ultimately other indigenous languages of Southern Africa.

AL/INL  
Postboks 1001 Blindern  
0315 Oslo

19. oktober 1993

Til

Stipendiat frå University of Zimbabwe - ALLEX-prosjektet 1993 og 1994.

Etter planane og budsjettet for ALLEX-prosjektet skal Universiteta i Oslo og Göteborg ha kvar sin stipendiat på tre-månaders opphald i 1993/1994 for å få opplæring i ulike aspekt ved leksikografi og utføre arbeid for prosjektet.

ALL/UZ la under workshopen 20.-24.9.93 fram eit prioritert framlegg på tre personar som er knytte til prosjektet, jf vedlegg. I samråd med deltakarane frå Oslo og Göteborg vart d3et avgjort at nr 1, Edgar Mberi, skal dra til Göteborg så fort som råd, ikkje seinare enn 1.11.93, og arbeide med skanning og konkordansar for prosjektet. Nr 2, Bernhard Batidzirai, skal koma til Oslo så tidleg i 1994 som mogleg, og arbeide med definisjonsskriving. Begge har forplikta seg til å lære opp kolleger når dei kjem heim, og halde fram ved prosjektet til det er fullført.

