

**RESEARCH UNIT FOR EXPERIMENTAL PHONOLOGY**

**UNIVERSITY OF STELLENBOSCH (RUEPUS)**

ANNUAL REPORT: 1997 / 1998

<b>1</b>	<b>BACKGROUND</b>	<b>4</b>
<b>1.1</b>	<b>MISSION AND VISION</b>	<b>4</b>
<b>1.2</b>	<b>LONG TERM GOALS</b>	<b>5</b>
<b>2</b>	<b>PROGRESS REPORT</b>	<b>5</b>
<b>2.1</b>	<b>ACHIEVEMENT OF GOALS</b>	<b>5</b>
<b>2.2</b>	<b>COMPLETED PROJECTS</b>	<b>6</b>
2.2.1	HOMORGANIC NASAL ASSIMILATION IN XHOSA.	6
2.2.2	PHONETIC AND PHONOLOGICAL ASPECTS OF TSWANA-AFRIKAANS/ SPEECH INTERFERENCE	6
<b>2.3</b>	<b>CURRENT PROJECTS</b>	<b>7</b>
2.3.1	PROGRAMME 1: LINGUISTIC PROJECTS	7
2.3.1.1	Lexical disambiguation in Xhosa and Zulu	7
2.3.1.2	The production and perception of statement and question intonation in Xhosa	7
2.3.1.3	The phonetics and phonology of click articulations in Nguni	8
2.3.1.4	Aspects of phonetic and phonological interference in Xhosa speech communication.	9
2.3.1.5	Phonetic and phonological interference in the production of Zulu clicks by Sotho speakers.	9
2.3.1.6	Strengthening processes in Sotho and Nguni	9
2.3.2	PROGRAMME 2: RESOURCE DEVELOPMENT PROJECTS	10
2.3.2.1	Preferred syllable structures in Nguni and Sotho	10
2.3.2.2	SASPEECH	10
2.3.2.3	Development of computer software	11
2.3.3	PROGRAMME 3: LANGUAGE AND SPEECH TECHNOLOGY PROJECTS	11
2.3.3.1	Modelling the intonation of imperatives in a Xhosa text-to-speech system (TTS)	11
2.3.3.2	Speech recognition in African languages (DACST proposal)	12
2.3.4	RESEARCH CAPACITY BUILDING PROJECTS (RCB)	12
2.3.4.1	RCB - University of Venda	12
2.3.4.2	RCB - University of Transkei	12
<b>2.4</b>	<b>PUBLICATIONS AND CONFERENCE PRESENTATIONS</b>	<b>13</b>
2.4.1	PAPERS SUBMITTED FOR PUBLICATION IN ACCREDITED JOURNAL	13
2.4.2	PUBLICATIONS IN NATIONAL AND INTERNATIONAL CONFERENCE PROCEEDINGS	13
2.4.3	CHAPTERS IN BOOKS	14
2.4.4	BOOKS	14
2.4.5	PAPERS READ AT INTERNATIONAL CONFERENCES	14
2.4.6	PAPERS READ AT NATIONAL CONFERENCES	14
<b>2.5</b>	<b>STAFF ACTIVITIES</b>	<b>14</b>
2.5.1	VISITS ABROAD	14
<b>2.6</b>	<b>VISITORS RECEIVED</b>	<b>15</b>
2.6.1	PROF LUDWIG GERHARDT (UNIVERSITY OF HAMBURG, GERMANY)	15
2.6.2	PROF EKKEHARDT WOLFF (INSTITUT FÜR AFRIKANISTIK, UNIVERSITY OF LEIPZIG, GERMANY)	15
2.6.3	DR ANNE CUTLER AND DR JAMES MCQUEEN (MAX PLANCK INSTITUTE, NIJMEGEN, THE NETHERLANDS)	15
<b>2.7</b>	<b>CONFERENCES AND SYMPOSIA PRESENTED</b>	<b>15</b>
2.7.1	COLLOQUIUM ON THE VALUE OF CROSS-LINGUISTIC RESEARCH IN PSYCHOLINGUISTICS, STELLENBOSCH, 20 MAY 1997	15
2.7.2	NATIONAL WORKSHOP ON SPEECH, TEXT AND TECHNOLOGY, UCT, 11 & 12 SEPTEMBER 1997	16
2.7.3	TWO DAY WORKSHOP ON THE TEACHING OF PHONETICS: 9 <sup>TH</sup> INTERNATIONAL CONFERENCE OF THE AFRICAN LANGUAGE ASSOCIATION (ALASA), UNIVERSITY OF NATAL, DURBAN, JULY 1997	16
<b>2.8</b>	<b>IMPLEMENTATION OF RESEARCH</b>	<b>16</b>
2.8.1	TEXT TO SPEECH SYSTEM (VERSION 1: XHOSA)	16
<b>3</b>	<b>INFRASTRUCTURE</b>	<b>16</b>
<b>3.1</b>	<b>STAFF</b>	<b>16</b>
<b>3.2</b>	<b>RESOURCES</b>	<b>17</b>
<b>3.3</b>	<b>EQUIPMENT</b>	<b>17</b>

<b>4</b>	<b>PLANNING FOR NEXT YEAR (1999/2000 - YEAR 2 OF CYCLE 2)</b>	<b>17</b>
<b>4.1</b>	<b>PROGRAMME 1: LINGUISTIC PROJECTS: THE PHONETICS PHONOLOGY INTERFACE CONTINUATION</b>	<b>17</b>
4.1.1	THE PRODUCTION AND PERCEPTION OF STATEMENT AND QUESTION INTONATION IN XHOSA	17
4.1.2	THE PHONETICS AND PHONOLOGY OF CLICK ARTICULATIONS IN NGUNI	17
4.1.3	STRENGTHENING PROCESSES IN SOTHO AND NGUNI	18
<b>4.2</b>	<b>PROGRAMME 2: RESOURCE DEVELOPMENT PROJECTS CONTINUATION:</b>	<b>18</b>
4.2.1	PREFERRED SYLLABLE STRUCTURES IN NGUNI AND SOTHO	18
4.2.2	SASPEECH	18
4.2.3	TEXT CORPORA: PARSING TOOLS FOR THE AFRICAN LANGUAGES	18
<b>4.3</b>	<b>PROGRAMME 3: LANGUAGE AND SPEECH TECHNOLOGY PROJECTS CONTINUATION</b>	<b>18</b>
4.3.1	MODELLING THE INTONATION OF IMPERATIVES IN A XHOSA TEXT-TO-SPEECH SYSTEM (TTS)	18
4.3.2	SPEECH RECOGNITION IN AFRICAN LANGUAGES (DACST PROPOSAL)	19
4.3.3	COCHLEAR IMPLANTS: ACOUSTIC AND PERCEPTUAL FEATURES OF XHOSA (UNCONFIRMED)	19
4.3.4	LANGUAGE ENGINEERING: PRINCIPLES AND PRACTICE	19
<b>4.4</b>	<b>STAFF</b>	<b>20</b>
<b>5</b>	<b>BUDGET</b>	<b>20</b>
<b>6</b>	<b>STATEMENT OF INCOME AND EXPENDITURE FOR THE PREVIOUS FINANCIAL YEAR (1997/8)</b>	<b>20</b>
<b>6.1</b>	<b>BUDGET FOR NEXT FINANCIAL YEAR (1999/2000)</b>	<b>20</b>
<b>7</b>	<b>GENERAL</b>	<b>20</b>

## 1 BACKGROUND

This progress report covers activities of the **fourth and final year** of the first cycle of the Unit, i.e. activities that have taken place in the financial year ending **March 1998**. The report will be presented in the prescribed format for annual progress reports as specified by the CSD.

Following specific suggestions presented to and accepted by the Advisory Board at its annual meeting in 1997 the focus of RUEPUS has been redirected towards research and development in the field of language and speech technology applications with specific reference to the indigenous languages of South Africa. At least three considerations have lead to this change in focus:

- An acute need for the development of re-usable language (text) and speech resources not only for academic linguistic purposes, but especially also for technological applications in African languages.
- A need for the identification and training of mother tongue speakers of the indigenous languages in order to empower them to play a role in the future development of these languages at technological level, specifically to meet expected demands for the development of language and speech based human-machine communication systems in the next millennium.
- A need for attracting new and appropriate funding in order to become self-sustaining in view of the policy of the CSD to fund research units only for an initial period of eight years (two cycles of four years each).

In view of the foregoing the initial mission and vision and long-term goals had to be redefined and this also lead to the formulation of a new long-term strategic plan (please see 4.5).

### 1.1 Mission and vision

The **mission** of RUEPUS may best be described as the active promotion of the official languages of South Africa at **technological level** through research, facilitation and development activities with an explicit focus on the African languages and the speakers of these languages.

The **vision** upheld by RUEPUS is that all languages spoken in South Africa, and more specifically the African languages, be *developed at technological level* to such an extent that

- They will meet the challenges of a technologically driven next millennium, i.e. that they may be implemented in language and speech based human-machine communication systems to the effect that mother tongue speakers may have *access to information* of any kind in the language of his/her own choice as well as in the mode (oral/visual/written) of his or her choice;
- A *new generation of language engineers* (i.e. mother tongue speakers with specific training in language and/or computer science and/or electronic engineering) will be created which will play a central role in the development of language based applications which will infiltrate all walks of life in the next century;

- A technology driven *national resource facility* making provision for at least speech corpora, will be put in place not only serving as a national archive documenting language and speech characteristics of all of the languages (and its dialects) at different points in time, but also to serve as a basis for the development electronic speech based systems for use in education, business and industry etc. (This also entails the replacement of the defunct South African Sound Archives.)

## 1.2 Long term goals

The initial aims of RUEPUS were primarily orientated towards theoretical linguistics, with a spin-off component related to speech technology applications. Given the change of focus that has been necessitated *inter alia* by a number of pragmatic considerations, it will be necessary to phase out projects of an exclusive linguistic theoretical character in favour of research projects of a more applied technological nature. In this process however, a point will inevitably be reached where the focus will be of such a nature that the name and primary aims of the existing research unit may become problematic.

In response to a recommendation by the Advisory Board that a **strategic plan** (Minutes 4/9/1997, p3, 5g) be submitted in this regard, it is suggested that RUEPUS activities will be directed in such a fashion that it may be possible to create a new financially self sustaining institution with exclusive focus on **language engineering** (LE) within the African languages, with RUEPUS eventually becoming part of this new institution in the final year of its second cycle. This new centre, the **African Language Engineering Centre (ALEC)**, will operate on a national scale to the benefit of all potential researchers in LE in the African languages bringing together academic resources from the academic community with financial input from business and industry and government according to a model similar as that implemented by the European Union.

## 2 PROGRESS REPORT<sup>1</sup>

### 2.1 Achievement of goals

The goals set by RUEPUS reflect on

- **Academic activities** (the running of research projects and the dissemination of research results through publications and conference presentations),
- **Networking** (making contact with colleagues at national and international level, *inter alia* by organising conferences/seminars and or participating in other conferences/seminars)
- **Research capacity building** (at graduate and post graduate levels as part of formal academic courses; through dedicated research capacity building projects; through specialised workshops)
- **Fund raising** to create a self-sustainable infrastructure conducive to research and development.

In assessing the achievement of the academic aims it is probably true to say that

<sup>1</sup> A detailed description of the contents of all completed and current projects may be found in the Annual Report of 1996/7. Only new and relevant information will be supplied in each case.

although the number of research publications have declined slightly compared to the previous year, (mainly due to staff problems as explained at the previous meeting of the Board) the results may be regarded as satisfactory. As far as networking is concerned the goals have been achieved as a number of conferences were attended and presented which allowed for interaction between RUEPUS staff members and other academics of whom a large number hailed from historically disadvantaged communities. This contributed to the overall visibility of RUEPUS at national as well as at international level. The goals set for research capacity building were partially achieved; it was rather unfortunate that one institution did not display the same commitment as the other (please see 2.3.4). Obtaining funds for research and development proves to be extremely difficult in an environment where there is little understanding for technology development at this level. Some specific suggestions in this regard will be appreciated at the next meeting of the Advisory Board.

A negative assessment may be made of RUEPUS not yet being able to attract researchers from historically disadvantaged communities to work full time in the laboratory of this nature. A large number of these scholars have however been made aware of needs in this field through the seminars, workshops and research capacity building activities. While highly skilled persons with a competence in both languages and the natural sciences are needed to perform basic research, it is still unfortunately so that not many candidates have come through the system, regardless their ethnic status. The reason for this was stated in an earlier report, i.e. the lack of prior (undergraduate) training in either experimental phonetics or computational linguistics. It is the conviction that a financially sound centre such as the proposed ALEC will play an important role in creating opportunities for alternative training, and hence makes a significant contribution to redress.

## **2.2 Completed Projects**

### **2.2.1 Homorganic nasal assimilation in Xhosa.**

#### **Ms Bulelwa Ntlabezo & JC Roux**

Although it was indicated in the Annual Report of 1996/7 that this project had been completed, some detailed information regarding dissemination still need to be reported:

#### **Dissemination of results:**

**Previously reported:** One article in accredited national journal.

#### **New:**

- i. One presentation at international conference in Leipzig: Roux (1997)
- ii. Submission of MA thesis, Ntlabezo (1998)

### **2.2.2 Phonetic and phonological aspects of Tswana-Afrikaans/ Speech interference**

#### **DP Wissing & co-workers, JC Roux**

Although it was indicated in the Annual Report of 1996/7 that this project had been completed, some detailed information regarding dissemination still need to be

reported:

**Dissemination of results:**

**Previously reported:** Four articles in accredited journals; one publication in conference proceedings; one presentation at national conference.

**New:**

- i. One article submitted for publication in the **SA Journal of Linguistics** (Special Supplement on Phonetics and Phonology); Pretorius, Roux & Wissing
- ii. Two presentations at national conferences of respectively, of the Linguistic Society of South Africa and the African Language Association of South Africa (ALASA); Wissing and Roux (1998)

## 2.3 Current projects

In the 1996/7 Annual Report it was indicated that RUEPUS activities would in future focus on three areas expressed in terms of three programmes:

- **Programme 1:** Linguistic projects  
*(The phonetics-phonology interface)*
- **Programme 2:** Resource development projects.  
*(Resource development in African languages)*
- **Programme 3:** Language and speech technology projects  
*(Human-machine interaction through African languages)*

### 2.3.1 Programme 1: Linguistic projects

#### 2.3.1.1 Lexical disambiguation in Xhosa and Zulu

**JC Roux, Joey Haasbroek, A Radebe,**

For a full description of this project - please see the Annual Report 1996/7

In the previous report was indicated that this project had been adversely affected by the resignation of Ms Haasbroek due to staff problems. This situation has further been complicated with Mr Radebe leaving for personal reasons. This project is currently on hold.

**Dissemination of results:**

**Previously reported:** One presentation at national conference.

#### 2.3.1.2 The production and perception of statement and question intonation in Xhosa

**Ms Jackie Jones & JC Roux**

For a full description of this project - please see the Annual Report 1996/7

The researcher is making extremely good progress with this project. Extensive acoustic analyses have been performed which were followed up by the development and presentation of a number of perceptual tests involving mother tongue listeners. The estimated completion date of this project has had to be postponed till April 1999, the reason for this being the vast number of statistical analyses that have had to be performed and following from that, the complex nature of the interpretation of the data.

#### **Dissemination of results:**

##### **Previously reported:**

- i. One presentation at national conference of ALASA, Johannesburg: Jones (1996)
- ii. One presentation at regional workshop of ALASA, Pretoria: Jones (1996)
- iii. One chapter in an academic Festschrift: Jones & Roux (in press)

##### **New:**

- iv. One presentation at an international workshop of ALASA, Durban (1997)
- v. Two academic articles submitted for publication to **SA Journal of Linguistics** (Special Supplement on Phonetics and Phonology); Jones, Louw and Roux

### **2.3.1.3 The phonetics and phonology of click articulations in Nguni**

#### **JC Roux & G Dogil**

For a full description of this project - please see the Annual Report 1996/7

This is an **ongoing study** without any estimated date of completion due to the complexity of the topic. Dr Michael Jessen, a member of the Institute for Computational Linguistics (IMS) at the University of Stuttgart has been awarded a Feodor Lynen Stipendium from the Alexander von Humboldt Stiftung for postdoctoral research in this project. The content of the project has been expanded to include *inter alia* the study of ejectives and implosives in more detail. The University of Stellenbosch has also made a significant financial contribution towards Dr Jessen`s costs through the Stellenbosch 2000 fund.

#### **Dissemination of results:**

##### **Previously reported:**

- i. Presentations at international conferences in, respectively, Vienna (Dogil, 1996), Budapest (Roux, 1996), München (Dogil, 1996), Adelaide (Roux, 1996), as well as at national and regional conferences in South Africa: (Roux, 1996a, 1996b).
- ii. Two publications as proceedings of international conferences: Dogil et al (1996), Dogil & Roux (1996)



- iii. One publication in an accredited journal: Roux & Lewis (1996)

**New:**

- iv. Presentation at 16<sup>th</sup> International Conference of Linguists, Paris (Roux & Dogil, 1997)
- v. Two publications as proceedings of international conferences (Roux & Dogil 1998, Dogil, Mayer & Roux 1998)

#### **2.3.1.4 Aspects of phonetic and phonological interference in Xhosa speech communication.**

**PW Lewis, JC Roux & G Dogil**

For a full description of this project - please see the Annual Report 1996/7

This project has made significant progress to the extent that the researcher has just presented his findings in the form of a doctoral thesis, which has been approved by the University of Stellenbosch. This project makes a significant contribution to an understanding of the so-called "difficult" and "easy" sounds of a language such as Xhosa, both from the views of non-mother tongue speakers and mother tongue listeners. The results have implications for intercultural communication, language learning and the teaching of phonetics, as well as for speech technology as sets of very valuable acoustic and perceptual data have been made available for technology development.

**Dissemination of results:**

**New:**

- i. One paper submitted for publication to the **SA Journal for Linguistics** (Special Supplement on Phonetics and Phonology); Lewis and Roux.

#### **2.3.1.5 Phonetic and phonological interference in the production of Zulu clicks by Sotho speakers.**

**AM Radebe & JC Roux**

For a full description of this project - please see the Annual Report 1996/7

This project had to be abandoned due to the fact that the researcher returned to Kwa-Zulu-Natal due to personal problems.

#### **2.3.1.6 Strengthening processes in Sotho and Nguni**

**I. Olivier & JC Roux**

For a full description of this project - please see the Annual Report 1996/7

This project is making slow progress.

**Dissemination of results:**

None thus far.

## 2.3.2 Programme 2: Resource development projects

### 2.3.2.1 Preferred syllable structures in Nguni and Sotho

**JC Roux, A Jones, J Louw, Dorothy Calata & Bongiwe Fingo**

For a full description of this project - please see the Annual Report 1996/7

This **ongoing study** with an expected completion date of **December 1999** is making good progress although some specific problems occurred with regard to securing text databases with the assistance of the Bible Society of SA. Due to the large text corpora, which need to be analysed, it has in the mean time become necessary for Mr Louw to redevelop the initially developed computer programme ("Patana" - pattern analysis).

#### **Dissemination of results:**

- i. One presentations at an international conference: Vienna (Roux, 1996)

#### **New:**

- i. One presentation at national workshop on Speech, Text and Technology, UCT, (A Jones, 1997)
- ii. One presentation at workshop at 9<sup>th</sup> International conference, of ALASA, Durban (A Jones, 1997)
- iii. One paper submitted for publication to the **SA Journal for Linguistics** (Special Supplement on Phonetics and Phonology); Roux and Louw

### 2.3.2.2 SASPEECH

This project may be regarded as pivotal for any real significant development within RUEPUS in view of its declared change in focus and movement towards language and speech technology applications. For this reason extremely much attention was devoted by the Director to promote this project with a budget (which may be implemented in phases) of R7.5 million and which will eventually secure at least 15 new job opportunities for mother tongue speakers of African languages, and reinstate a lost national facility (the South African Sound Archives). High-level contact was kept (through discussions, correspondence etc) with a number of key figures in this process:

Dr Neville Alexander, former executive member of PANSALB

Dr Cynthia Marivate, CEO, PANSALB

Dr Wally Serote, Chairman of the Parliamentary Committee on Arts, Culture, Science and Technology

Dr Rob Adam, Assistant Director-General, DACST (Technology)

Dr Hans Martin Schmid, German Technical Co-operation Office (GTZ), Pretoria

Dr. N Mkhulisi, Director: Language Services

Mr K Makhudu, Liaison officer reporting to Mr Enoch Sithole, CEO, SABC

The reaction has up till this point in time been quite diverse (some surprisingly so), and may possibly be discussed at the meeting of the Advisory Board. However, it is clear that more high profile discussions will be necessary and that RUEPUS will approach the CSD for possible intervention and/or mediation in this regard.

### 2.3.2.3 Development of computer software

Software programmes are developed continuously as resources for applications in African languages. The following programmes were developed by Mr Jan Louw and are in principle available to any researcher in this field.

Software	Description
Perception testing	Automatically presents sound stimuli for up to twenty listeners at a time. The programme also collects and processes the listeners' responses.
Phonetic transcription	The program transcribes orthographic text as phonetic text according to a set of language specific letter-to-sound rules. Available for the Sotho and Nguni languages
Phonetic symbols	The program provides a user-friendly interface for inserting phonetic text into a document.
Phonetic pattern matching	The program searches for specified sequences in a phonetic transcription and returns various statistics on the occurrence of the sequences found.
Pitch calculation	The program calculates the pitch period markers, which are used in calculating the pitch contour of an utterance or in pitch synchronous prosodic modification of speech.
Extraction and analysis of prosodic features	The program extracts the duration, pitch and loudness features from a set of words spoken by a number of speakers and calculates various statistic parameters on these features.
Prosodic modification of speech	The program modifies the duration and pitch of an utterance according to user specified values using a pitch synchronous overlap-add method (PSOLA).
Speech database management	Various smaller program aids in managing speech databases used in current ongoing projects.

### 2.3.3 Programme 3: Language and Speech Technology Projects

#### 2.3.3.1 Modelling the intonation of imperatives in a Xhosa text-to-speech system (TTS)

**P Swart, J Louw, L Botha, & JC Roux**

For a full description of this project - please see the Annual Report 1996/7

This project is making extremely good progress. Apart from setting specific guidelines of how to develop a TTS system in an African tone language, an important methodological point has been addressed regarding the nature of tagged speech when used as a basis for concatenation in speech synthesis. Research is now at a critical point regarding the manner in which monotonously produced tagged speech may be used to resynthesize intelligible speech in Xhosa.

#### Dissemination of results:

- i. Presentation at two day workshop at 9<sup>th</sup> International ALASA conference,

Durban (Swart, 1997)

### 2.3.3.2 Speech recognition in African languages (DACST proposal)

RUEPUS entered into a consortium proposal agreement with the departments of Electric and Electronic Engineering at respectively, the Universities of Pretoria and Stellenbosch which was presented as an Innovation Fund proposal to the Departments of Arts, Culture, Science and Technology. The topic of this project is **Development of African Language Speech Recognition Software and Voice based Information System Prototypes**. This proposal of approximately R8 million spread over a period of three years includes an amount of R1.1 million for RUEPUS for speech data base development and management in the first year of the project. The outcome of this proposal will be known in November 1998.

### 2.3.4 Research capacity building Projects (RCB)

#### 2.3.4.1 RCB - University of Venda

Although this project was completed in 1997 a special request was made to CSD to make some surplus funds available for follow up consultations and monitoring of the process. This was agreed upon and the Director made a special visit to the University of Venda for the period 16-19 November 1997. It appeared as if the system was functioning fine and that the members of staff trained in the process were quite at ease in their use of it. During the training phase, the participants were supplied with a number of possible projects as well as the methodology needed for conducting such projects. The idea was that having a system set up, and having been trained in the use of the system, and having been informed about the vast number of phonetic and phonological problems to be solved in the different languages, that staff members would proceed on their own to implement the facilities for research and teaching purposes. Unfortunately, this has not taken place. It will be difficult to speculate on all possible reasons, however, it remains disappointing that the staff are seemingly not using the system. On a personal level, the Director has tried to involve some individuals in a joint descriptive project, however, with little response. On the other hand, it is good to know that there are mother tongue speakers around who are qualified and who may assist in data collection and processing in the event of the SASPEECH or DACST project materialising.

#### 2.3.4.2 RCB - University of Transkei

The research capacity building project with the University of Transkei has since its initial stages taken a more positive turn than the one reported above. This is mainly due to the commitment of Dr Mbulelo Jokweni and his fellow staff members. The project was executed in different phases.

**Phase 1:** Dr Jokweni visited RUEPUS for a week in September/October 1997 and was informed by the local team about all the aims, procedures etc for setting up work station for phonetic research. At that point in time the Dean: Research of Unitra also visited RUEPUS and he acquired first hand knowledge of the situation. These two persons then conferred with the authorities and succeeded in obtaining funds to acquire the necessary hardware as prescribed by RUEPUS. Dr Jokweni then proceeded to introduce two staff members informally to the functioning of the system through impromptu exercises.

**Phase 2:** The Unitra team consisting of Dr Jokweni, Ms KN Qamata and Mr AT Sotashe attended an intensive 6-day training course in speech acoustics, perception and methods and techniques applicable to experimental phonetic research presented by staff of RUEPUS. They were supplied with appropriate literature and specific tasks for exercise at Unitra.

**Phase 3:** The RUEPUS team consisting of the Director, a phonetician (Ms P Swart) and an electronic engineer (Mr J Louw) then paid Unitra a three day visit to ensure that the technical set-up was correct and functioning and that the staff could cope in their use of the system. An introductory seminar was held and it was well attended by staff members from a number of departments within the Arts faculty as well as by language students. Approximately 50 people attended this seminar in which the emphasis was on emerging technologies and how the Humanities could/should respond to it.

This exercise has been a great success up till this point in time and it has created considerable interest also on the campus of Unitra. Following from this training a new interest developed amongst the trainees for some specific projects to be undertaken. The Director and Dr Jokweni are now also in a process of applying for funds from the Research committee of Unitra as well as from PANSALB to embark on a systematic experimental phonetic analysis of the sound systems of respectively, Rharhabe and Gcaleka.

**Request:** That the surplus amount of R6 379 be kept available for further monitoring and the ensuing SASPEECH version of Xhosa (please see 4.2.2, as well as Appendix B)

## 2.4 Publications and conference presentations

### 2.4.1 Papers submitted for publication in accredited journal

- **Roux, JC, Wissing, DP & R Pretorius**, Front vowel raising in Tswana. Submitted for publication in special supplement on Phonetics and Phonology of **South African Journal of Linguistics** (to appear 1998)
- **Lewis, PW & Roux, JC**. Xhosa L2 speech intelligibility. Submitted for publication in special supplement on Phonetics and Phonology of **South African Journal of Linguistics** (to appear 1998)
- **Jones, J, Louw, J & Roux, JC**. Acoustic aspects of Queclaratives in Xhosa. Submitted for publication in special supplement on Phonetics and Phonology of **South African Journal of Linguistics** (to appear 1998)
- **Jones, J, Louw, J & Roux, JC**. Queclaratives in Xhosa: a perceptual study. Submitted for publication in special supplement on Phonetics and Phonology of **South African Journal of Linguistics** (to appear 1998)
- **Roux, JC & Louw, J**. Preferred syllable structures in Xhosa: A first approximation. Submitted for publication in special supplement on Phonetics and Phonology of **South African Journal of Linguistics** (to appear 1998)

### 2.4.2 Publications in national and international conference proceedings

- 1998**      **Dogil, G, Mayer, J & Roux, JC**. Syllables and unencoded speech: Clicks

and their accompaniments in Xhosa. In **PHONOLOGICA 96: Syllables?!** Eds. J Rennison & K Kuhnhammer. The Hague: Holland Academic Graphics. 49-60.

**I.p:** **Roux, JC & Dogil, G.** On the phonetic representation of clicks: Some experimental phonetic considerations. In **Proceedings of the 16<sup>th</sup> International Congress of Linguists**, Paris (ed B Caron) Appearing in CD-ROM format

#### 2.4.3 Chapters in books

**I.p:** **Jones, J & Roux JC.** Queclaratives in Xhosa revisited. In **Festschrift for Prof JA Louw**, (ed R Finlayson)

#### 2.4.4 Books

**1998** **Gerhardt, D & Roux, JC. Otto Dempwolff: Induktiver Aufbau des Urbantu.** Archiv afrikanistischer Manuskripte, Nr 3. R.Köppe Verlag, Köln. (In press)

#### 2.4.5 Papers read at international conferences

**1997** **JC Roux.** Phonetic explanations for phonological phenomena: Labiodental assimilations in Xhosa. **2<sup>nd</sup> World Congress of African Linguists (WOCAL 2)**, Leipzig, Germany.

**1997** **JC Roux.** On the phonetic representation of clicks: Some experimental phonetic considerations. **16<sup>th</sup> International Congress of Linguists, Paris, France.**

#### 2.4.6 Papers read at national conferences

**1997** **JC Roux.** Human-machine interaction in South Africa. **National Workshop on Speech, Text and Technology**, University of Cape Town

See also workshops and seminars presented in 2.7

### 2.5 Staff activities

#### 2.5.1 Visits abroad

The Director made a visit to France and then to Leipzig in July - August 1997 in order to participate in the 16<sup>th</sup> International Conference of Linguists in Paris as well as the 2<sup>nd</sup> World Congress of African Linguists in Leipzig, Germany. The visit to Leipzig served to strengthen academic ties with the Institut fuer Afrikanistik to the effect that an academic exchange programme between the two universities was set up for staff and student exchange. One postgraduate student in phonetics/phonology (Mr J Mokgophi) as well as one postgraduate student in socio-linguistics (Mr J Masalesa) was invited to follow advanced courses in Leipzig for one semester (October 1998 - March 1999) in accordance with this programme<sup>2</sup>.

<sup>2</sup> A member of staff, Mr MW Jadezweni is also currently teaching a literature course at Leipzig in terms of this agreement.

## **2.6 Visitors received**

### **2.6.1 Prof Ludwig Gerhardt (University of Hamburg, Germany)**

Prof Gerhardt head of the Institut fuer Afrikanistik at the University of Hamburg visited RUEPUS for a period of one month in January 1997 in order to finalise a joint publication between himself and the Director. This publication focuses on historical aspects of the sound system of Bantu and is a critical revision of the work of Dempwolff, one of the students of Carl Meinhoff. This book publication is due to appear late in 1998.

A public lecture was held by Prof Gerhardt, which was well attended by staff and students of UCT and UWC.

### **2.6.2 Prof Ekkehardt Wolff (Institut für Afrikanistik, University of Leipzig, Germany)**

Although Prof Wolff primarily visited the Department of African Languages, he had lengthy discussions with staff of RUEPUS and with the Director in view of some specific needs in the area of phonetics and phonology. A tentative arrangement has been reached in terms of which the Director and/or staff members of RUEPUS will assist in the development of a new phonetics laboratory in Leipzig in 2000, and simultaneously supply them with available acoustic speech databases for the study of the South African Bantu languages at phonetic level.

### **2.6.3 Dr Anne Cutler and Dr James McQueen (Max Planck Institute, Nijmegen, The Netherlands)**

Following discussions at a meeting in Australia in 1996 between the Director and Dr Cutler (Scientific Director, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands), she and a colleague Dr James McQueen visited RUEPUS in order to check and discuss some of their tentative findings on children's speech gathered on a field trip to Lesotho. Acoustic analyses were made of some of their data and some interesting points were raised regarding cross-linguistic findings in psycholinguistics. Some of these discussions had direct impact on the current study of Ms Jackie Jones, especially with respect to the application of some "gating" experiments.

## **2.7 Conferences and symposia presented**

### **2.7.1 Colloquium on the value of cross-linguistic research in psycholinguistics, Stellenbosch, 20 May 1997**

A colloquium was organised by RUEPUS on 20 May 1997 to coincide with the visit of Dr Cutler and Dr McQueen. These two academics were the main presenters and the Director the respondent. The meeting was attended by 25 student and staff members from Stellenbosch, UCT and UWC, not only from African language departments, but also from psychology and logopaedics. This meeting contributed a lot to international networking as a number of postgraduate students and staff continued their contact with the Max Planck Institute through the two visiting academics.

### **2.7.2 National Workshop on Speech, Text and Technology, UCT, 11 & 12 September 1997**

RUEPUS was co-organiser (with the Department of Applied Mathematics, at UCT) of this two-day workshop that attracted staff members from departments of African languages, computer science and electronic engineering from local universities, Pentech and the University of Pretoria. The basic problems surrounding effective development in the field were identified and critically discussed. The crucial roles of government as well as the private sector were emphasised in most of the issues raised. Fifteen people attended this workshop.

### **2.7.3 Two day workshop on the teaching of phonetics: 9<sup>th</sup> International Conference of the African Language Association (ALASA), University of Natal, Durban, July 1997**

RUEPUS presented a two day workshop with the topic: Why teach phonetics and phonology?" It was primarily directed towards teachers and university lecturers focusing on innovative approaches to the teaching of the subject. Many participants were from historically disadvantaged institutions. The discussions were brisk, innovative and to the point. An informal interest group was formed, which is e-mail co-ordinated by RUEPUS. Since this meeting RUEPUS has been approached by different institutions *inter alia* to assist in the acquisition of computer hardware for a laboratory (in Lesotho) as well as for assistance in curriculum design.

## **2.8 Implementation of research**

### **2.8.1 Text to Speech System (version 1: Xhosa)**

Good progress has been made in designing the basic architecture of the system. The applied research of Ms Swart (2.3.3.1) has given new impetus to this project. An interesting deviation from standard practice of speech synthesis, i.e. the modulation of monotonous speech, holds promising results for the eventual synthesis of tone languages.

## **3 INFRASTRUCTURE**

### **3.1 Staff**

The following persons were appointed for the periods specified:

Ms A Jones (MA)	1/2/1997 - 31/12/1997
Mr Jan Louw (M Eng.)	Ad hoc (1997)
Ms Pippa Swart (BA Hons)	1/4/1997 - 31/3/1998
Ms D Calata	Ad hoc (1998)
Ms B Fingo	Ad hoc (1998)
Mr P Mtwendinga	Ad hoc (1998)

It needs to be mentioned here that although RUEPUS is totally committed to affirmative action it is still extremely difficult to find suitably qualified staff from previously



disadvantaged communities to. A potential candidate and current MA student Ms N Shoba declined a full time position as research assistant at a very late stage on the grounds of personal circumstances regarding her moving from Alice to Stellenbosch for a limited period of time. An appointment in this regard has been negotiated with Mr J Masalesa who is currently involved in the Stellenbosch- Leipzig exchange programme. He will join RUEPUS in March 1999 on his return to take up a research position.

As has been proven elsewhere in phonetics laboratories world wide, it is absolutely necessary to have at least one electronic engineer/phonetician on the staff on a full time basis. This will be even more necessary when RUEPUS redirects its focus to language and speech technology. The technical nature of experimentation necessitates expertise of this kind.

In order to maintain and run a phonetics laboratory of this interdisciplinary nature it is necessary to have at least one electronic engineer, one general/applied phonetician and one mother tongue researcher. These positions will be included in the budget of the second year of the second cycle (1999/2000) (Appendix B).

### **3.2 Resources**

Due to the fact that RUEPUS had recourse to the ad hoc services of an electronic engineer it was not necessary to make use of the services of the Institute for Electronics at the University of Stellenbosch to assist in setting up certain configurations and programs. Resources for the study of both technology and phonetics are readily available in the different libraries. RUEPUS furthermore has unlimited access to the Internet.

### **3.3 Equipment**

RUEPUS had to upgrade some of its computer equipment as routine practice. This is reflected in the financial report. At this point in time the laboratory is adequately equipped for running its own projects. Any new project sponsored from an external source will have to include equipment in its budget.

## **4 PLANNING FOR NEXT YEAR (1999/2000 - YEAR 2 OF CYCLE 2)**

The planning below is in accordance with the proposed strategic plan as presented in **Appendix A.**

### **Projects continued as well as new projects**

#### **4.1 Programme 1: Linguistic projects: The phonetics phonology interface Continuation**

##### **4.1.1 The production and perception of statement and question intonation in Xhosa**

Ms Jackie Jones & JC Roux

##### **4.1.2 The phonetics and phonology of click articulations in Nguni**

JC Roux, G Dogil, M Jessen and D Calata

This is an ongoing study with estimated date of completion in 2001. Dr Michael Jessen, a member of the Institute for Computational Linguistics (IMS) joins the team together with Ms D Calata a postgraduate student of Stellenbosch. The contents of the project have been expanded to include *inter alia* the study of ejectives and implosives in more detail. The relationship between phonetics and phonology will be explored with reference to these sound types.

#### **4.1.3 Strengthening processes in Sotho and Nguni**

I. Olivier & JC Roux

### **4.2 Programme 2: Resource development projects Continuation:**

#### **4.2.1 Preferred syllable structures in Nguni and Sotho**

JC Roux, J Louw, Dorothy Calata & James Masalesa

The projects has reached a stage where results reflecting the differences between the nature of syllable structures in the Nguni and Sotho languages are to be compared and explained in terms of either acoustic or perceptual parameters. Mr Masalesa will join the team.

#### **4.2.2 SASPEECH**

**JC Roux, J Louw, P Swart, M Jokweni, K Qamata, A Sotashe**

Although no specific funding is available yet, co-researchers from Unitra will conduct the following sub-project: **The acoustic phonetic differences between Gcaleka and Rharhabe**. This will entail a process of systematic data acquisition and acoustic analysis on the full sound system of these two varieties of Xhosa. The results of this study will be of extreme importance for differentiating in the design of either a speech recognition system or a text to speech system. Application has also been made to the Research Committee of Unitra as well as to PANSALB.

Fund raising efforts for this project will continue.

**New:**

#### **4.2.3 Text Corpora: Parsing Tools for the African Languages**

**JC Roux, Marianna Visser, Kathy Demuth (Brown University, USA) and Annie Zaenen (Xerox, France)**

High-level negotiations will continue and preliminary resource analyses will be made so as to draft a joint proposal for implementation as soon as ALEC has been stabilised

### **4.3 Programme 3: Language and Speech Technology Projects Continuation**

#### **4.3.1 Modelling the intonation of imperatives in a Xhosa text-to-speech system (TTS)**

P Swart, J Louw, L Botha, & JC Roux

**New:**

#### **4.3.2 Speech recognition in African languages (DACST proposal)**

**J Louw, JC Roux, J Masalesa, P Swart, D Calata and co-workers**

If this proposal is successful (to be known at the end of November 1998) dedicated speech databases will be developed for Xhosa, Zulu, Northern Sotho and Tswana focusing on speech communication in a banking environment in order to develop a prototype query system for the banking industry. Hardware and software have been included in the budget of R1.1 million.

#### **4.3.3 Cochlear implants: Acoustic and perceptual features of Xhosa (Unconfirmed)**

**L Muller, JC Roux, J Louw**

An international company Cochlear, Australia has donated two cochlear implant devices for implantation in deaf mother tongue speakers of Xhosa. The acoustic features of some speech sounds in Xhosa however pose a severe problem for the calibration of the hearing aid device in such a system. The nature and extent of this problem will be investigated and guidelines will be set for future implantations for speakers of African languages. This project, when confirmed will take place in co-operation with the Department of Speech Therapy at the Medical School of Stellenbosch on their request. RUEPUS has in the past been involved with some research projects of students in Logopaedics.

#### **4.3.4 Language Engineering: Principles and Practice**

**JC Roux, J Louw, P Swart, J Masalesa**

This project will entail the following:

- i. A continuous study of developments regarding LE technologies as well as of models of language and speech processing, and principles and policies underlying implementation.
- ii. The implementation of an LE awareness campaign which includes
  - Two high level seminars for business and industry, government and academia with international participation from the European Union (limited to decision makers)
  - Presentation of a paper and a two day workshop on LE during the 10<sup>th</sup> International Conference of the African Language Association of Southern Africa at UNISA, Pretoria, July 1999
  - Specific fund raising initiatives with national and international role players

#### **International conferences**

The following two international conferences will be attended by one or more members of RUEPUS where papers will be delivered and workshops presented.

- 10<sup>th</sup> International Biennial conference of ALASA, UNISA, Pretoria, 5-9 July 1999
- 14<sup>th</sup> International Congress of Phonetic Sciences (ICPhS 99), San Francisco, 1-7 August 1999

#### 4.4 Staff

In order to create a stable environment not only for existing projects to be conducted, but also for the intended transition to take place, it is extremely important to keep the current core staff employed as well as to introduce a new researcher. This new researcher is **Mr James Masalesa**, a brilliant young man, completing his BA Honours study in the department of African Languages at Stellenbosch, who is currently in Leipzig in terms of the Stellenbosch-Leipzig exchange programme. He will be appointed according to current payment schedules, i.e. R100/day x 22 days x 11 months.

As has been explained it is furthermore extremely important to have the services of a person with good technical and computer skills in the phonetics laboratory. **Mr Jan Louw** has shown his value in the past; apart from his direct contributions to publications (cf. 2.4), he has been of enormous support for doctoral students working on projects of RUEPUS (cf. the projects of Lewis and of Jones). Mr Louw has a Masters Degree in Electronic Engineering (cum laude) from Stellenbosch and is showing extreme interest in the field of language and speech processing. In order to retain his highly sought expertise, it will be necessary to offer him an adequate remuneration package. The proposal is that he be appointed initially for one year on a contract basis as Researcher with a salary equivalent to that of the position of lecturer (R74 000 - with a subsidy of R4 000 towards a medical assistance plan: this amounts to a total of R78 000).

**Ms Pippa Swart** is currently doing excellent work as a young applied phonetician in African languages. She will be completing her MA study early in 1999. In order to keep her services, and at the same time balance the books of RUEPUS, the Department of African Languages will make a significant contribution towards her remuneration package of one year of R43 220 (without benefits). The department will contribute R33 220.

## 5 BUDGET

## 6 STATEMENT OF INCOME AND EXPENDITURE FOR THE PREVIOUS FINANCIAL YEAR (1997/8)

Please see **APPENDIX B** (as well as the special requests)

### 6.1 Budget for next financial year (1999/2000)

Please see **APPENDIX C**

## 7 GENERAL

It is an understatement to say that without the financial support of the CSD and the University of Stellenbosch none of these activities would have been possible. I would like to express my sincerest thanks and appreciation to both these institutions for their generous support.

Thank you to the representatives of the private sector, i.e. to Gencor and Datafusion for insightful comments and support not only expressed at the previous Board meeting but also during private conversations. I sincerely hope that with the new focus of RUEPUS in its transformation to ALEC it would be possible for closer and even more tangible co-operation to the benefit of all concerned.

Thank you to the staff and students of RUEPUS and their dedication to quality work, and finally also to the members of the Advisory Board sharing your expertise with us, not only at official meetings, but also in collegial interactions.

**PROF. JC ROUX**  
**DIRECTOR: RUEPUS**

**STELLENBOSCH**  
**19 NOVEMBER 1998**