languagematters@UKZN

University Language Planning and Development Office
Second Quarterly Report: 2016

INSPIRING GREATNESS

Langa Khumalo, PhD Linguistics.
Director: Language Planning and Development Office

The period under review saw the official launch of the *Bua Lenna* 'Let's Talk' Programme. This is in consonant with the letter and spirit of the University's Language Policy and Plan, which seeks *inter alia* to "[…] preserve and promote respect for, and proficiency in, the languages referred to in the Constitution, and other languages, including the heritage languages, that facilitate potentially valuable cultural, scientific and economic ties […]". Multilingualism is thus the hallmark of the UKZN language policy. This is vitally important for the University since this programme cultivates social cohesion within the University community. The former Dean and Head of School of Arts, Professor NS Zulu, the Acting Dean and Head of School of Arts Professor D McCracken, and the Head of African Languages Discipline Professor SE Ngubane graced the occasion.

Second Quarterly Report
April - June 2016.

Contents

1. Introduction

The highlight of this period has been without doubt the launch of the *Bua Lenna* 'Let's Talk' Programme. The launch is a culmination of the pilot sessions that have been taking place at the two residences, Ansell May and John Bews since 2014. At the pilot sessions the University Language Planning and Development Office (ULPDO) and its Sesotho language champions have provided a platform for non-Sesotho speaking students to learn conversational Sesotho in a relaxed environment devoid of any strictures. The programme has grown in leaps and bounds since its inception, and is very popular with students in these residences. ULPDO has been reporting regularly in its quarterly and annual reports on the development of this programme, and how this programme has fostered multilingualism and social cohesion.

The Terminology development programme has also progressed well. The Law discipline has made the most progress in developing isiZulu terminology for its discipline. For the third time, Law experts presented their terminology for consultation, at a workshop that was well attended by language experts and the discipline experts. The office also hosted a verification workshop for psychology terminology, which was also well attended. Terminology development as a process has given rise to the scholarship of intellectualization through terminology development, a theme that ULPDO is taking to the Afrilex International Conference in July this year.

The development of isiZulu National Corpus has continued, albeit aided by our partner institutions that produce text materials in isiZulu. Shutter & Shooter has been the most productive institution. Killie Campbell Library and the UKZN Press have provided us with a lot of corpus materials in the period under review. This has seen the growth of the isiZulu National Corpus from 10 million at the beginning of the year to 13.5 million running words at the end of May 2016.

Langa Khumalo
Director: ULPDO

1. Psychology Terminology Verification Workshop

The ULPDO hosted the Psychology Terminology Verification Workshop on the 19-21 April 2016 at Coastlands Hotel.

In the picture are the language and discipline experts. The language experts include Prof AM Maphumulo, Dr N Sibiya, Mr. Z Xala, Mr. M Mbatha. The subject specialists are Prof S Magwaza, Ms N Buthelezi and Dr T Sifunda. The PanSALB office was represented by Mr. J Simelane and Ms J Ngcobo.

This workshop was convened in order to present psychology terminology to the verification committee in conformity to the PanSALB terminology development processes. Psychology discipline experts, Prof S Magwaza, Dr T Sifunda and Ms N Buthelezi provided the requisite expertise as the committee of linguists and terminologists studiously considered each of the 500 terms. The workshop was a tremendous success. The 500 terms are now ready for the final authentication stage.

2. Law Terminology Consultative Workshop

The Law Terminology Consultative Workshop was hosted on the 17-19 May 2016 at Coastlands Hotel by the ULPDO. This Law Workshop was the third in a series, which has successfully taken Law terminology through the statutory terminology development processes. The Law discipline has distinctly shown its seriousness in advancing the transformation agenda of UKZN particularly through the implementation of Language Policy and Plan. The terms were too technical and required more time than expected. The office evaluation processes confirmed that the workshop went very well, hence one of the participants shared the following sentiments "this was an interesting and well organized workshop, we are really grateful to have been part of it." A total of 468

terms were considered at this consultative workshop, and are now ready for the second stage, which is Verification.

The collaborators for this workshop included, Prof AM Maphumulo, Dr N Sibiya, Dr M Nzuza, Mr. Z Xala, Dr G Mazibuko, Dr L Makhubu, Dr S Ntuli, Dr M Ndlovu, Mr. B Nxumalo, Mr. SW Sibeko, Mr S Ndlovu, Mrs. M Ngongoma, Mr B Mdluli, HS Gumede, Ms N Dladla, and Ms G Mkhize, The subject specialists included Advocate S Phungula, Advocate S Ngcobo, and Mr K Zondi. Students included Ms C Kunene and Mr M Busane, and the ULPDO staff members.

3. IsiZulu Orthography Workshop

The ULPDO hosted the isiZulu Orthography Workshop which, was held on the 4th and 5th of May 2016 at UKZN Unite Building.

The main objective of this workshop was to share knowledge about the history of isiZulu orthography and how it has evolved over the years. This workshop was facilitated by two renowned experts in the field of isiZulu orthography, with over 30 years of experience in writing and teaching of isiZulu language, Professor OEHM Nxumalo and Mr S Ngobese.

The workshop was attended by people from different institutions, namely, UKZN academics, DUT, EThekwini municipality, National DAC, KZN DAC and Ukhozi FM. It is worth noting that novice language practitioners immensely benefited from the practical experiences of the two facilitators. Imparting such knowledge to them will not only benefit UKZN but also contributes to the development of isiZulu and the entire nation.

As part of monitoring our operations the workshop evaluations were distributed to all participants and their overall feedback suggested that all went very well. They also indicated that they gained more insight on how to write isiZulu. Below is what one of them shared:

"umsebenzi eniwenzayo muhle kakhulu."

4.   Running Projects

4.1.   Bua Lenna project

4.1.1 Bua Le Nna Sessions

The Bua Lenna sessions were held at Ansell May and John Bews Residence from April-May 2016.

In the picture are some of the students who participated in this session, and the Language Champion distributing the language exercise of the day.

The project continues to generate excitement from students to such an extent that Sesotho language is now fashionable in these residences. The sessions held in the second quarter went very well. The students attended in big numbers, which indicated that a lot of them have seen the value and importance of embracing indigenous languages. It also indicated that students have grown an interest in learning and appreciating the richness of Sesotho language and culture. This programme is important for UKZN social cohesion agenda and is also in line with the constitutional imperative of promoting multilingualism. These sessions are interactive hence encouraging student to learn through shared experiences.

4.1.2 Bua Le Nna Launch

The University Language Planning & Development Office officially launched the Bua Le Nna project on the 19th of May 2016. The launch confirmed the UKZN's commitment to adhere to the constitutional imperative of promoting multilingualism. The occasion was attended by different stakeholders, namely senior university management, academics, student leadership and students (mainly) from the two residences (Ansell May and John Bews). The presence of Prof McCraken (The Acting Dean and Head of the School of Arts), Prof S Ngubane (Head of the African Languages Cluster), Prof N

Zulu (Former Dean and Head of the School of Arts and Sesotho Language expert), and Mr. Marumola (Manager Student Housing, who represented the Executive Director of Student Affairs) was demonstrative of the importance of this program within the university.

The Director of the University Language Planning & Development Office, Dr L Khumalo with the Sesotho Language Champions and some students at the launch.

The event was celebrated through exciting performances in Sesotho, which included the poem, the Sesotho traditional dances, drama by non-Sesotho speakers demonstrating their level of conversational Sesotho, and Sesotho speeches delivered by dignitaries. The office is now left with the responsibility to transition from the pilot program to an actual language program of teaching conversational Sesotho. The planning has already started in this regard and will be finalized later in the year for the 2017 rollout.

5.   KZN Stakeholder meeting.

ULPDO also engaged in stakeholder liaison activities in the period under review. Two meetings with the Provincial and National DAC took place, on the 16th of March 2016 and 12th of April 2016 respectively, at the UTLO Boardroom.

These meetings were attended by representatives from both these departments. The purpose of the first meeting was to strengthen the synergistic relations between UKZN and DAC and to attempt to reconcile the calendar events and projects of mutual interest. A number of issues were discussed including co-hosting events, combining programmes and avoiding possible clashes. The second meeting meetings resolved to develop a Memorandum of Understanding (MOA, which is now at the finalization stages. The National DAC will give UKZN access to their term bank, and will also share literature on the development of Human Language Technologies (HLTs). Both meetings were fruitful. Another meeting will be scheduled soon in order to check the progress in relation to the decisions which were taken in these meetings.
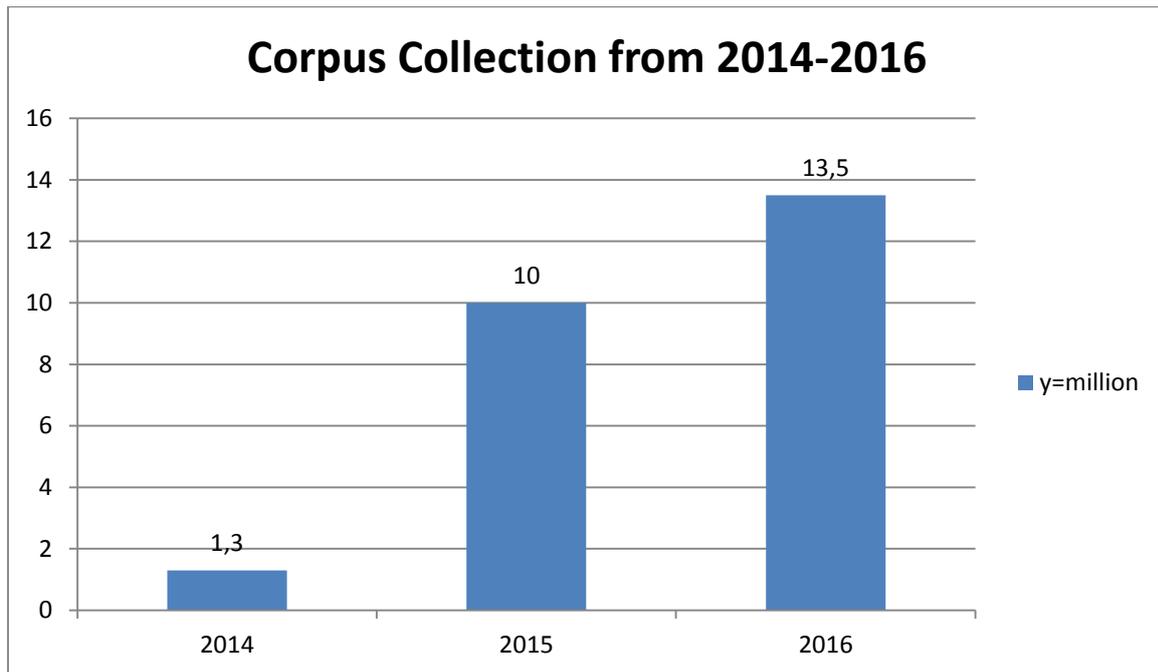
6.   ULPDO HLT's

6.1   Corpus Project

N	text file	file size	tokens (running words) in text	tokens used for word list	sum of entries	types (distinct words)	type/token ratio	type/token standardised TTR	STTR std.dev.	STTR basis	mean word length	word length std.dev.	entences mean (in words)	std.dev.	ragraphs mean (in words)	std.dev.	headings mean (in words)
1	Overall		13 556 784	13 089		1 592	12.17	69.44	32.41	1 000	7.05	3.14	730 5	18.52	444.0(	32 66(400.7'	3 573.
2	zuluamanothi00001.txt	163 695	19 674	19 635		7 684	39.13	64.24	33.95	1 000	7.11	2.57	1 597	12.29	31.38	7 2 805.	2 784.
3	zuluamanothi00002.txt	2 377	356	293		172	58.70			1 000	4.88	3.34	16	18.31	19.80	6 48.83	79.78
4	zuluamanothi00003.txt	13 383	1 525	1 514		873	57.66	63.30		1 000	7.52	2.61	150	10.09	5.22	1 1 514.	
5	zuluamanothi00004.txt	53 823	5 892	5 652		2 749	48.64	58.58	34.25	1 000	7.19	3.74	213	26.54	149.3:	150 37.68	106.95
6	zulubhayibheli00001.txt	3 886 2(	435 481	404 291		71 37:	17.65	58.29	44.43	1 000	7.44	3.13	29 96	20.82	1 269.	1 404 29	
7	zuluezasegagasini00001.txt	43	4	4		4	100.0(			1 000	7.50	2.08	1	4.00		1 4.00	
8	zuluezasegagasini00002.txt	82 642	8 948	8 505		4 432	52.11	70.40	24.24	1 000	7.35	3.17	459	18.53	14.73	1 8 505.	
9	zuluezasegagasini00003.txt	97 109	10 966	9 781		3 589	36.69	52.39	34.24	1 000	7.01	3.24	486	20.13	17.07	1 9 781.	
10	zuluezasegagasini00004.txt	95 277	10 457	9 533		4 135	43.38	58.63	28.31	1 000	7.23	3.28	448	21.28	23.80	1 9 533.	
11	zuluezasegagasini00005.txt	1 862	42	41		33	80.49			1 000	4.98	3.63	1	41.00		1 41.00	
12	zuluezasegagasini00006.txt	22 326	2 390	2 313		1 536	66.41	71.50	20.15	1 000	7.45	3.01	133	17.39	13.10	1 2 313.	
13	zuluezasegagasini00007.txt	89 640	10 006	9 356		4 092	43.74	59.17	29.66	1 000	7.11	3.23	473	19.78	17.68	1 9 356.	
14	zuluezasegagasini00008.txt	69 906	7 697	7 167		3 563	49.71	64.06	26.11	1 000	7.19	3.14	384	18.66	16.32	1 7 167.	
15	zuluezasegagasini00009.txt	77 117	8 374	7 944		4 012	50.50	66.88	25.79	1 000	7.34	3.12	413	19.23	15.92	1 7 944.	
16	zuluezasegagasini00010.txt	86 934	9 204	8 853		4 431	50.05	67.72	25.45	1 000	7.56	3.09	485	18.25	14.75	1 8 853.	
17	zuluezasegagasini00011.txt	20 432	2 135	2 023		1 280	63.27	65.70	24.25	1 000	7.60	3.15	115	17.60	17.28	1 2 023.	
18	zuluezasegagasini00012.txt	89 768	9 619	8 980		4 150	46.21	63.32	27.65	1 000	7.42	3.20	480	18.71	15.91	1 8 980.	
19	zuluezasegagasini00013.txt	59 143	6 655	6 036		2 724	45.13	55.68	30.56	1 000	7.04	3.21	298	20.26	18.67	1 6 036.	
20	zuluezasegagasini00014.txt	75 887	8 340	7 789		3 822	49.07	63.63	26.96	1 000	7.22	3.18	431	18.07	17.62	1 7 789.	
21	zuluezasegagasini00015.txt	51 849	5 692	5 388		2 816	52.26	66.92	25.13	1 000	7.23	3.12	280	19.25	16.95	1 5 388.	
22	zuluezasegagasini00016.txt	6 760	691	687		538	78.31			1 000	7.84	3.18	38	18.08	8.89	1 687.0(	
23	zuluezasegagasini00017.txt	31 947	3 443	3 286		2 060	62.69	70.20	22.89	1 000	7.44	3.17	190	17.29	12.63	1 3 286.	
24	zuluezasegagasini00018.txt	120 387	13 086	11 803		4 504	38.16	52.74	35.03	1 000	7.24	3.24	560	21.08	20.59	1 11 80:	
25	zuluezasegagasini00019.txt	81 556	9 220	8 417		3 677	43.69	58.64	29.58	1 000	6.97	3.23	404	20.83	21.24	1 8 417.	
26	zuluezasegagasini00020.txt	77 759	8 791	8 275		3 834	46.33	65.34	28.73	1 000	7.00	3.19	426	19.43	15.46	1 8 275.	

The development and growth of the isiZulu National Corpus (INC) continues at an impressive pace. Since the launch of the INC in 2014 at Terminology Development and Dissemination Indaba at an impressive size of 1.3 million tokens, the INC reached the 10 million threshold at the end of 2015. This is significant because it means the INC is growing at a fast rate, and has surpassed other African languages corpora within a short space of time. At a strategic level the office has created crucial synergies with important stakeholders such as Shutter and Shooter Publishers and the Killie Campbell Library. Access to texts produced and archived at these two institutions, and other institutional partners has helped in accelerating the growth of the INC. Currently (as shown above) the INC stands at just over 13.5 million tokens. Fig. 1. below illustrates the pattern of growth since the launch of the INC in November 2014.

Figure 1

6.2   Term Bank, Spellchecker and ZuMoLex

The development of the HLTs is articulated in one of the objectives of the office as follows: to 'develop an isiZulu National Corpus and isiZulu Term Bank as important reservoirs for the development of robust Human Language Technologies (HLTs) and for posterity'. Human Language Technology as enablers in the cultivation of isiZulu include the isiZulu Terminology Bank, the isiZulu Spellchecker and the Mobile compatible lexicon called ZuMoLex. The development of both the isiZulu Term Bank and ZuMoLex is progressing very well. The development of the application (ZuMoLex) is envisaged to be completed in June 2016. In the same vein, Dr Khumalo (ULPDO director) and other computer engineers have produced a prototype isiZulu spellchecker using n-grams and training the spellchecker on the INC.

7. Workshops and Conference attendance

7.1 CICLing and IST Africa Conferences

The ULPDO director (Dr Khumalo) had two papers presented at two prestigious international conferences. The first paper scooped First Place in Verifiability, Reproducibility, and Working Description Award, at the 17th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2016), which was held in Konya, Turkey from the 3rd to 9th of April 2016. The award is for the paper that

he co-authored with Joan Byamugisha and Maria Keet both from the University of Cape Town. The paper "Pluralising Nouns in isiZulu and Related Languages" describes the design and implementation of an algorithm to automatically pluralise isiZulu nouns.

The second paper was presented at the IST Africa conference which took place on the 11-13 May 2016 at the eLangeni Hotel in Durban. The title of the paper, 'The Effects of a Corpus on IsiZulu Spellchecker Based on N-grams' has already been published in the peer-reviewed IST conference proceedings. The paper was well received by the conference delegates. It was particularly interesting for participants to learn that the highest accuracy rate for this spellchecker is now 89%.

7.2 Digital Humanities Association of Southern Africa (DHASA) Workshop

The workshop to formally create the DHASA took place from the 4th – 8th of April 2016 at the North West University. The ULPDO director (Dr. Khumalo) attended this workshop. The objectives of DHASA amongst others are to integrate digital technology in research across all disciplines (from humanities to sciences). DHASA was formally created after the adoption of its constitution and the election of its inaugural executive management committee. Dr L Khumalo was elected as a member of the inaugural Executive Management Committee of the Digital Humanities in Southern Africa (DHASA), after serving for a year as its steering committee member.

7.3 DHASA Meeting in Cape Town
The meeting took place from the 17th – 18th May 2016 in Cape Town

DHASA Executive Management Committee Members: (L-R) Prof Roux, Dr Khumalo, Ms Thomson, Dr Calteaux, Prof de Lange.

The ULPDO director (Dr. Khumalo) attended the first executive management committee meeting, which was held at the University of Cape Town. The purpose of this meeting was to craft the three year program of action for DHASA and the hosting of its first international conference in 2017. A detailed plan on the development of South Africa's eleven official languages was also discussed. The business of the meeting was done successfully.

Looking ahead

The office is growing in its reach and stature. The language programme at UKZN is producing tangible results. Progress towards developing an isiZulu Term Bank, and isiZulu App are at an advanced stage and both tools will be launched in November 2016. The building of the isiZulu National Corpus is breath-taking owing to the synergies that the office has put in place with strategic partner institutions.