Babel's Tower: South Africa's Wikipedias

Minor Dissertation by Michael Graaf
Submitted in Partial Requirement of M.Phil.
University of Cape Town, 2018.
Supervisors: M.Densmore & D. Johnson

Abstract: This dissertation is a comparative examination firstly of usage of, and contribution to, the Wikipedias in ten of South Africa's eleven official languages, and secondly of possible measures to address the situation discovered in the first investigative stage.

The historical context is reviewed and it is argued that the number of official languages (and therefore Wikipedias) results from decisions made in colonial and apartheid eras. Public-domain usage and contribution statistics from the Wikimedia Foundation are analysed, revealing poor growth in most cases; possible interventions via both cultural/educational strategies and technological options are reviewed.

Word count: 10 755

Note on Terminology:

(Superscript numbers refer to footnotes)

The term "South African languages" refers to the 11 official languages of the Republic, notwithstanding that English is also a world language, due to its imperial history. The other 10 are "African languages of South Africa" of which one, Afrikaans, is a creole, while the rest are "ethnic African languages" (after Alexander (2001)).

Individual ethnic African languages are named here with a term widely used in the language for itself. For example, what Wikipedia until recently called "Northern Sotho", is here "SePedi" (during the course of this study, Wikipedia began to use "Sesotho sa Leboa", which is a direct translation of "Northern Sotho", but for purposes of this study, "SePedi" was retained).

Contents:

Chapter 1: Introduction

"To Mamdani, the task of decolonising... is inextricably tied up with the need to develop African languages. 'The decolonising project has, by definition, to be multilingual,' he said 'to provide resources to nurture and develop non-Western intellectual traditions'."(Davis, 2017)

Pervasive digital networks and the enhanced gathering, storage, transport and processing of data enabled by such networks, are widely considered to have sufficiently enhanced human productivity to be labelled a "fourth industrial revolution"(Younus, 2017). The characteristic consequence of this revolution is mass collaboration across space and time, and the greatest collaborative project it has enabled is Wikipedia (together with associated projects such as Wikidata and Creative Commons), whose goal is to make all human knowledge freely available to every human[1].

However just as the adoption and benefits of digital technology are unevenly distributed among humankind, so the processing and sharing of knowledge from and to the various human languages and cultures is uneven. South Africa's 11 official languages span the entire range of such unevenness: English Wikipedia is the biggest of all, while IsiNdebele Wikipedia does not exist yet.

The intention of the present study is twofold: a diagnosis of the situations of South Africa's various Wikipedias and their associated communities, and a prognosis of possible interventions not only within Wikipedia communities but by cultural and educational authorities, practitioners and activists. The main academic lens used is that of "ICT for Development"(to be defined later); the objective of the work is to contribute to reducing the gaps between South Africa's Wikipedias, since these not only reflect social inequalities, but help perpetuate them. In this context "gaps" are not only in content, but in accessibility and prestige.

Footnote:

1) This is variously stated by assorted roleplayers, many referring to the origins of the project (see Cohen, 2008)

Dramatic differences between South African languages' Wikipedias are to be expected given the class profiles of the various populations; however, a facile dismissal of the differences as being straightforwardly determined by economics would lose insights to be gained from what Mamdani *(*in Davis*, ibid)* called "half a century of... affirmative action" which made Afrikaans "the bearer of an intellectual tradition". It should also be borne in mind that approximately half of the Afrikaans-speaking population, classified under apartheid as "coloured", is still economically disadvantaged and/or culturally marginalised, and the issue of whether and how such people benefit – or can benefit – from the relative success of the Afrikaans Wikipedia needs to be addressed along with similar questions involving other African languages.

I posit that the existence of ten such unequal Wikipedias reflects a passive response to apartheid's legacy by both South Africa's state and its civil society. The infrastructure of Wikipedia is relatively agnostic as to language policy; a user can, with the click of a button, initiate a Wikipedia in any language having a dedicated and active community, even if no longer (or not yet, as in the case of "constructed" languages) spoken natively. A different history of contestation over apartheid could have yielded different Wikipedias.

Unlike the present South African government which in a stagnant economy (both global and local) struggles to fast-track the growth of the "black middle class", the Afrikaner nationalist movement and its "apartheid" government had the luxury of decades of economic growth. While "white" Afrikaners were being lifted (largely by means of government jobs) out of poverty similar to that of many "blacks" at present, cultural tasks such as creation of curricula and textbooks up to university level were also being performed in Afrikaans, within a consciously anti-colonial discourse (O'Meara, 1983).

Clearly it would be infeasible to replicate such history for each of the nine ethnic languages; the achievements of Afrikaner nationalism (both economic and cultural) were at the expense of the majority. However, the digital age offers new possibilities of reducing the gaps relatively quickly and cheaply. These will be addressed in later chapters but first, the following chapter further examines the context in which our findings have taken place. The third chapter is a literature review; the fourth, a methodological overview; the fifth, a description and discussion of the factual findings, and the sixth and seventh, reviews of possible policy and technological interventions respectively. A concluding eighth chapter precedes the bibliography.

In sum, this dissertation addresses the following research questions in the following ways:
1) How do the various Wikipedias in South African languages compare to each other and to the hegemonic English-language version, both intrinsically and in their contributing communities?

- The factual part of the work (Chapter 5) describes and discusses the data gathered online about contributions to, and usage of, the Wikipedias.

2) What policy/methodological/technological interventions could improve the aforementioned situation going forward?
- The sixth chapter explores possible cultural and educational policies and methods which could be adopted by authorities and activists, while the seventh looks at technological interventions which can contribute.

Chapter 2: Background

Literacy reached South Africa's indigenes via Christian missionaries, who after initially prioritising translation of religious texts into indigenous languages, went on to focus strongly (in the schools which were their primary outreach strategy) on teaching converts colonial languages: in the earliest missions, Dutch (later displaced by its creole derivative, Afrikaans) but in the majority of missions, English. This tradition can be criticised as "assimilationist", but historically it is credited with producing generations of African intellectuals who were able to challenge colonial and apartheid rule confidently (Limb, 2002, 56-82) . For this reason, when the apartheid regime took power, it immediately attacked the missionary education system, replacing it with "Bantu Education" (Giliomee, 2009, 190-198).

This new paradigm, while still requiring the medium of school instruction for ethnic Africans to be either English or as time progressed, more and more Afrikaans, reduced conceptual instruction in favour of rote-learning. At the same time, ethnically-demarcated bantustans (Legassick & Wolpe, 1976, 87-107) were created, each with its "bush" university and radio station mandated to use the relevant ethnic language, and having agendas which prioritised parochial matters and avoided critical discourse. Thus these languages were relegated to secondary, stigmatised institutional status (Kamwangamalu, 2000, 50-60).

It is worth noting that there was a degree of arbitrariness and pragmatism in the apartheid regime's designation of **nine** bantustan ethnicities. On one hand, the aim was classical divide-and-rule, but on the other, this required the creation of vaguely plausible "nationalities", which could not be too many in number, or too small. Where British colonialism had designated a patchwork of hundreds of "Native Reserve" territories, apartheid sought to cluster and consolidate these into putatively separate-but-equal nation-states.

Some of the arbitration had already been done by rival missionary societies which had translated the Bible and other texts into ethnic languages, thereby establishing *de facto* standard written languages. For example, French missionaries first transcribed SeSotho, thus its orthography exaggerates its difference from SePedi which was transcribed by English missionaries. Makoni calls this process "misinvention" (Makoni, S., 2003, 132-149).

Nonetheless, the apartheid project, being not only linguistic but political, had to engage in some coercion to create bantustan nationalities. For example, communities which had not self-identified as Zulus were assigned "citizenship" of the KwaZulu bantustan: not only groups whose ancestors had successfully avoided forced assimilation into the Zulu kingdom by negotiating tributary status (AmaBhaca, AmaNgwane, AmaHlubi etc., each having their own Nguni dialect) but also mission-centred a*makholwa* (convert) communities containing families of which some within living memory had been speakers of SeSotho and other non-Zulu languages, but which happened to be located in the territory allocated to the nascent bantustan[1].

Footnote:

1) Not coincidentally, three such communities, at Ndaleni, Edendale and Mpumalanga (a township in KZN, not the contemporary province of Mpumalanga) had the highest levels of political violence in the country in the late 1980s and early 1990s as a*makholwa* resisted integration into KwaZulu (see http://natalia.org.za/Files/43/Natalia%2043-Article%20Merrett%20pp%2019%20to%2036.pdf and http://durbanhistorymuseums.org.za/tragic-legacy-in-new-light/).

Given the fraught history of social engineering and conflict in the bantustan project, Neville Alexander (1991) advanced a proposal to rationalise national language policy by delinking it from apartheid categories and reducing the number of standard versions. The essence of this proposal was to classify IsiZulu, IsiXhosa, SiSwati and IsiNdebele as dialects of an officially recognised Nguni language combining features from all of them, and likewise SeSotho, SePedi and SeTswana as dialects of a standard Sotho-Tswana language (perhaps renamed).

However, on taking power after the fall of apartheid, the ANC government, still led by alumni of the missionary school system, had deep ties to the bantustan elites; the same "royal" families which had formed the backbone of bantustan regimes, were (and still are) well-represented in the leadership of the ANC. It is therefore not surprising that Alexander's proposal gained no traction and colonial/apartheid categories remain reified to this day.

The relevance of this to South Africa's Wikipedias is that if the number of official ethnic languages was four instead of nine, the human and digital resources devoted to them could have been concentrated, in all likelihood leading to earlier "take-off" for them. It is possible in theory for Alexander's proposal still to be enacted by the creation of a Standard Nguni Wikipedia and a Sotho-Tswana Wikipedia, but whether the will to do so shall ever exist, can only be speculated upon.

More urgently and perhaps more likely, the **non-official** languages of the Khoi/San group could also acquire Wikipedias. These were spoken by pastoral and hunter-gatherer groups which were more vulnerable to colonial genocide than the Bantu-speaking farming communities, since the latter had proto-states, military institutions, and could produce iron weapons. In South Africa, only one Khoi language, Nama, still has a community in which it is transmitted as mother-tongue (Makoni, *ibid)*. Both Nama and other Khoi/San languages have such communities in neighbouring countries, though, so the necessary investments in developing Wikipedias could be shared.

Chapter 3: Literature Review

Overview:

The differential use of, and contribution to, Wikipedias in various South African languages and issues arising from those differences can be viewed in various contexts:

● that of ICT4D, since it involves use of technology by marginalised sectors of society;
● that of Wikipedia & Education, since Wikipedia's fastest-growing area of use is in education;
● that of Wikipedia & Marginal Groups/Genders/Languages, since nine of the eleven official languages are non-hegemonic, being the mother tongues of "historically disadvantaged" communities;

I therefore review literature within those areas one by one. I have also added a section on relevant "para-academic" or "organic intellectual" literature since so much of the subject matter I address is poorly documented in formal academia.

3.1: ICT4D

Before turning to the particular topic of the current work, an overview of the concept of ICT4D is worthwhile. The concept of development itself is highly contested; hence too the application of technology to development. For example, Brewer *et al* (2005) make the case that appropriate technology (and implementation methods, and the requisite research) can greatly improve quality of life in the "developing"(less-industrialised) world. In general the ICT4D paradigm ventures beyond the market – it prioritises need over profitability, although recognising that usefulness depends on availability, which often depends at least partly on the market (Heeks, 2008).

ICT4D can thus be seen as a subset of what Trace (2016) calls "Responsible Research and Innovation (RRI)". This he suggests should focus on social and environmental objectives, such as the Sustainable Development Goals, in guiding choices. Sachs (2012) characterises the SDGs as "a historic and effective method of global mobilisation to achieve a set of important social priorities worldwide".

A context such as South Africa where "advantaged" schools already have high-speed connectivity and digital equipment to make use of such, makes the provision of similar to "disadvantaged" schools an obvious intervention. The historic legacy of apartheid continues to impose a two-tier education system, with the substitution for socio-economic criteria for racial ones. However, Gudmundsdottir, G. B., (2010) reviews four case histories of ICT deployment in Cape Town schools and concludes that provision of technology alone in disadvantaged schools can exacerbate the digital divide; teacher capacity-building is crucial.

Chetty *et al* (2012) emphasise the importance of telecommunication affordability in South Africa. Their research addresses individual users and their own devices; provision of communal infrastructure and end-use devices, as in telecentres or schools, is one way of addressing this issue. However, for the foreseeable future, low-resourced communities such as most rural ones, will struggle with access and affordability.

Early literature on the role of communal ICT facilities in development such as Rothenburg-Aalami & Pal (2005) notes the importance of ensuring that benefits of ICT are justly distributed, but takes it as given that insofar as such benefits derive from appropriate (including locally generated) content, access to the content involves the internet *per se*.

In time, lower equipment costs and wider dissemination of techniques and software have made local hosting of useful content such as educational resources and local knowledge/data more practical. This reduces the need for external traffic, thereby reducing costs as well as latency. The selection of such content should prioritise frequently-used resources, and Wikipedia is obviously one. There is a paucity of academic articles specifically about local hosting of educational material, although web searches reveal significant activity in this field. As small, cheap, low-energy devices have become available for such purposes, they have come into use not only in remote and/or low-resource locations[1], but also as an assertion of autonomy, privacy etc.[2]

Concerning methodology of investigation into "socio-technical" projects such as Wikipedia, Geiger & Ribes (2011) have presented what they call Trace Ethnography, which combines participant observation with analysis of usage logs. This would have been the logical next step if the current study had been wider in scope.

Footnotes:

1) For example, the Kenyan BRCK server and tablet array (http://education.brck.com/kiokit/)

2) Piratebox and Librarybox are Raspberry Pi projects in academic contexts (http://academhack.outsidethetext.com/home/2014/file-sharing-in-the-classroom-piratebox-and-librarybox/)

3.2: Wikipedia & Education

The literature here reveals a sweeping change under way. Whereas Wikipedia was once scorned by academia, as time goes by its usefulness is ever more recognised not as a primary source (in fact it forbids primary research data) but as a secondary and tertiary one; in a typically unplanned effect, it is becoming a canonical source of links to useful references (Baker, 2012). Its most innovative educational effect, though, is that it enables learners not just to use but to **contribute to** the public corpus of knowledge, and to achieve formal credit for doing so from their institutions (for example, Azzam *et al*, 2017 and Richardson, 2010).

The next paper referenced considers not Wikipedia itself but wiki technology *per se*, in the context of education. Chiriac (2016) reviews the use of several types of wiki (including MediaWiki, which is used for Wikipedia). She concludes that collaboration itself is a valuable feature in education, and is vastly enhanced by use of wikis.

Another writer who looks at the educational value of the "writable web" in general, Richardson (2010) devotes a book chapter to Wikipedia. In it, the particular use of assigning students to create Wikipedia articles is highlighted.

Other examinations of Wikipedia as a subject of assignments are Reimers *et al*. (2016) and Konieczny (2012). I have not yet found an academic account of the use of Wikipedia articles in assignments in schools; it is hardly surprising if such use is still minimal, given the challenging nature of the environment of adult editors. However, there is much potential – sheltered environments such as niche online versions of Wikipedia like Vikidia[3] and offline ones like WikiFundi[4] are a novelty in schools. An interesting look at the pros and cons of educational wikis being public vs. private is Guth (2009), but it deals only with tertiary education.

It may well be that in general, as use of Wikipedias becomes more encultured, both usage and contribution patterns shift. What was once marginal may become more mainstream. For example Kittur *et al*. (2007) show that in its early days, Wikipedia contributions came overwhelmingly from a relatively small group of "expert" users, but with time there was a dramatic shift in workload towards the "common" user.

Footnotes:

3) https://www.vikidia.org/

4) http://wikifundi.org/

3.3: Wikipedia & Marginal Groups/Genders/Languages

The issue here is the apparent reinforcement of unequal relative statuses of cultures, genders and languages by the emergence and growth of Wikipedias. UNESCO (2010) has argued in a policy brief that African governments should invest in African languages and multilingual education as a means of reducing social and global inequality; the question is, what can Wikipedias and their associated projects contribute to such efforts at rectification? The problem is exacerbated by the "network effect" (the increase in usefulness of a service the more participants it has)(Rask, 2008).

The most impactful reading in this regard is Ford & Wajcman (2016). This is a persuasive and well-informed critique of Wikipedia and its community (and by implication, fellow Wikimedia projects and their communities). The Wikipedian stance of "Neutral Point Of View" is deconstructed and revealed to be a pose, albeit a largely unconscious one. The authors, as women, convincingly portray the experience of many female editors – those who persist as a substantial minority, and those who drop out. The "emotional work" imposed on them by a male-dominated environment is generally invisible to male editors and community leadership. The authors clearly speak from personal experience as well as the research they cite.

The authors then extrapolate this analysis of "systemic bias" from the field of gender to that of "Western/scientific" vs. indigenous knowledge and experience, arguing that Wikipedia's lack of cultural diversity mirrors its gender deficit.

While such critiques apply to Wikipedias in general, each language's Wikipedia has a separate (if sometimes overlapping) community, so language-comparative, country-comparative and language-specific analyses come next.

Dolmaya (2017) addresses issues of translation in Wikipedia contexts, relative to linguistic justice. She identifies another network effect: translation flows more from smaller languages to larger ones, than *vice versa*. She also focuses attention on the effects of various policies on the relative extent of such phenomena.

Maseko *et al*. (2010) argue that the use of colonial/hegemonic language asserts a user's rank in a post-independence but still neocolonial milieu such as African academia. Conversely, insistence on using African languages asserts that they have prestige. However, their analysis is more focussed on the training of teachers than the (related but different) experience of pupils.

Phillipson (1996) offers an optimistic view of the then-new South African language policy, compared with the rest of Africa. In his view post-colonial African leaders have been complicit in perpetuating colonial hegemony via language. However it's unlikely that this approval would still apply in present-day South Africa, as English is the *de facto lingua franca* or at least the core of "multilingualism as *lingua franca*" (Desai,1995).

Van Dijk (2009) looks at Wikipedia in South Africa specifically, and issues related to the various languages, using statistics from the Wikimedia Foundation's Statistics website. In noting that the modest size of Afrikaans Wikipedia relative to English Wikipedia despite relatively high socio-economic status/internet access of many Afrikaans-speakers, van Dijk speculates that their high proficiency in English lowers the priority they attach to having an Afrikaans Wikipedia. Of course this possibility is less relevant to disadvantaged (generally "coloured") Afrikaans-speakers' potential benefit, but van Dijk also raises the potential of Wikipedias as advocacy/mobilising tools for disadvantaged languages.

I have yet to find any academic study specifically of the use of Wikipedia or associated projects as tools of advocacy for disadvantaged languages – however we may expect the appearance of such, bearing in mind (for example) the recent wave of attention to participatory mapping as a tool of empowerment, as in DiGessa (2008).

The promotion of disadvantaged languages in general and in myriad specific cases and categories, is a lively academic terrain, too extensive to do justice to here. As a generalisation though, most emphasis is on **preservation**, given that 90% of languages spoken worldwide in the late 20th century were considered to be in danger of disappearing (Krauss, 1992, 4-10). While South Africa does have languages and dialects in such danger, nobody has included in that list any of the **official** languages. Rather, what is at stake is the radically unequal status among the eleven; the fact that only two are upheld as "bearer[s] of... intellectual tradition[s]" in Mamdani's phrase.

Tragically at the time of going to press, popular political discourse was focussed on whether those two should be reduced to one, rather than extended to more[5]. Academic works such as Mesthrie (2002) give a good overview of the socio-linguistic landscape in contemporary South Africa but the literature is hardly more characterised by pro-equality contestation than popular discourse is.

Footnote:

5) For example the Hoerskool Overvaal furore over retention of Afrikaans-medium instruction (https://www.iol.co.za/news/south-africa/gauteng/hoerskoolovervaal-the-long-struggle-for-a-language-of-inclusion-12820994)

3.4: Related Non-Academic Work

By its nature Wikipedia is an intellectual project; its essence is to embrace humankind in a common intellectual pursuit. Hence it itself also becomes an object of reflection within its community, with the result that informal outputs such as blogs are available. The South African chapter of the Wikimedia movement, Wikimedia-ZA, has also been party to ongoing discussions within the worldwide movement on questions of inclusivity, language equity etc[7]. Some members have also documented relative progress of the various African Wikipedias[8].

Footnotes:

7) https://blog.wikimedia.org/2018/02/05/wikimania-cape-town-ubuntu/

8) http://www.greenman.co.za/blog/

Chapter 4: Methodology

4.1: Wikipedia Comparison

This section seeks to address the research question concerning the differential status of South Africa's Wikipedias, in view of their origins in colonial and apartheid linguistic history, as reviewed in Chapter 2. User and contributor communities are profiled by means of statistics. The Wikimedia Foundation, which supports Wikipedia and similar projects, places a vast range of usage and contribution statistics in the public domain at its "Wikimedia Statistics" website[1]. This has been cited nearly a thousand times in academic literature and is being replaced by "WikiStats 2"[2].

286 Wikipedias are documented, each in a different language. All South African languages except IsiNdebele are included (perhaps IsiNdebele's similarity to the larger Nguni languages – IsiZulu, IsiXhosa and SiSwati – helps diminish demand and/or incentive for a separate work). Wikimedia Statistics also reveals the most active editors for each language and profiles their contributions, so future research building on the current work would be well supported.

From the vast array of available data it was necessary to select a sample which would usefully reveal trends. In order to address the first research question, regarding the comparative status of the various Wikipedias, one needs to assess two things in each case: its productive community and its consuming community. Hence, to construct not just a profile of current participation in and use of each project, but a series of such profiles stretching back in time to the beginning of the project. The English Wikipedia launched in 2001, so 16 years of data are available; the other South African languages' launch years range from 2002 to 2007. Thus it is possible to get an idea of shifts over at least 10, and up to 16 years.

Because statistics are given monthly, it was decided to take a "snapshot" of each language's Wikipedia's statistics for the month of September in each year of its existence. Of course a set of metrics to be analysed also had to be chosen. The particular list of metrics is less crucial than the fact that they are consistently applied across all projects, so that the comparison between Wikipedias is valid. The definition of, and justification for the use of, each metric is given in the discussion of the findings. For now, the table following shows the information gathered for September 2017.

Footnotes:

1) Wikimedia Statistics https://stats.wikimedia.org/EN/Sitemap.htm

2) https://blog.wikimedia.org/2018/01/02/wikistats-2/

Table 1: September 2017 Metrics

Language/ launch	Articles	Contributors	Active Editors	Page Views
English/ 2001	5485590	1175724	29111	7079000
Afrikaans/ 2002	47079	838	34	2500000
SePedi/ 2008	7883	28	1	235000
IsiZulu/ 2004	1078	63	3	229000
IsiXhosa/ 2002	961	41	1	186000
SeTswana/ 2004	740	12	0	171000
XiTsonga/ 2004	628	19	1	189000
SeSotho/ 2003	583	27	0	189000
SiSwati/ 2006	452	24	0	178000
TshiVenda/ 2004	325	17	0	154000

The same set of metrics was gathered for each preceding September until the launch of each project, and each metric's data points pasted into a separate LibreCalc spreadsheet, which then enabled the chronological progression of that metric to be graphed separately. The spreadsheets are appended.

Due to the vast differences in size of the various Wikipedias, in some cases a logarithmic scale is used for the vertical axis. This is taken into account in discussing the findings.

4.2: Assistive Technologies

In reporting the availability of machine-assisted translation, the relevant tool was tested by starting actual translations and observing whether the tool suggested content in the target language (where support is not offered, the tool merely suggests in the target column, the same content present in the source language article). It was later ascertained that the actual back-end translation is outsourced and thus under limited control by Wikimedia Foundation staff[3].

For the section in Chapter 7 which reviews technological issues, data usage tests were made using a Nokia 2700 feature phone (accessing Wikipedia's website via Opera Mini 5.1 browser) and a Samsung GT S5301 smartphone running Android 4.0.4 and using the stated apps & browsers.

Data usage was measured differently on each phone; in the feature phone, a balance was noted before and after the defined tasks; in the smartphone, the app "MyData" was used, which simply reports data throughput in the relevant connection since its last reset.

Footnote:

3. Amir Ahoroni, *pers. comm*.

Chapter 5: Findings & Analysis

5.1: Articles

This is the fundamental metric since an article is the basic unit of a Wikipedia, from the points of view both of users' experience of content, and of participation by editors. What should be borne in mind is that article count is a function both of the number of contributors and of the number of contributions per editor.

*Figure 1: September Article Count 2001-2017*
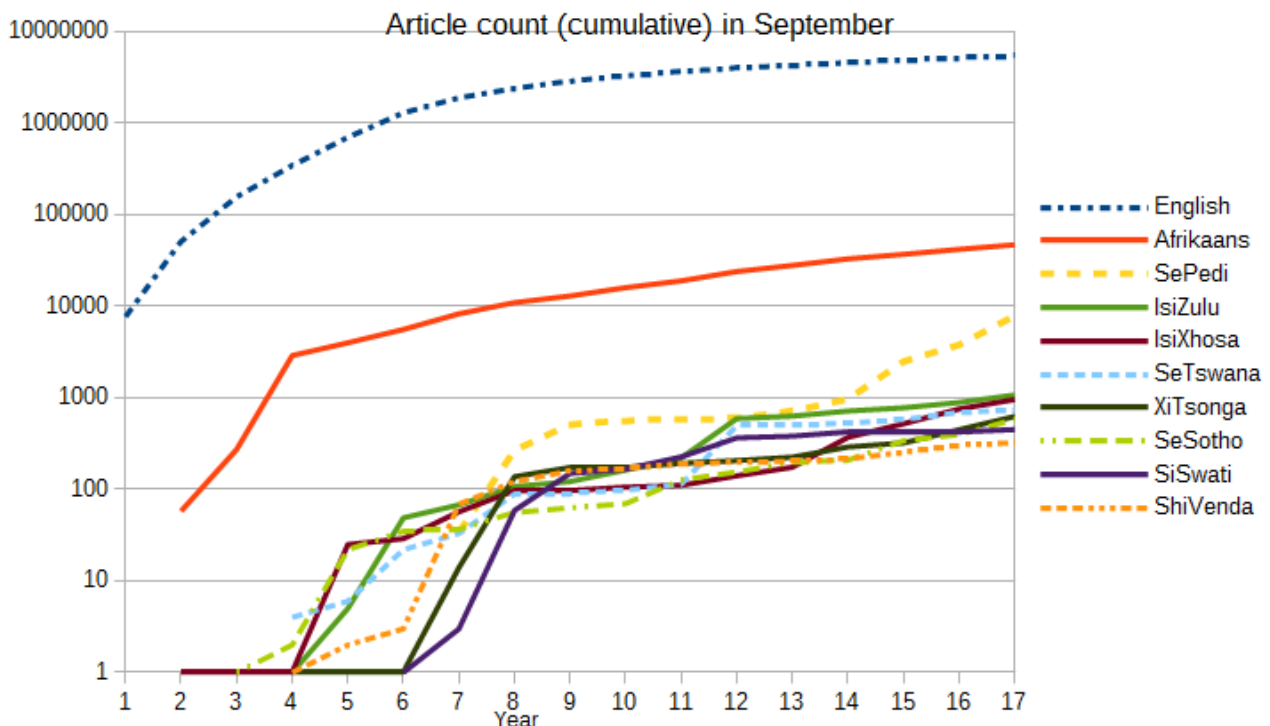
There is a superficial similarity between the various languages' curves, but this is deceptive. It may appear that each curve starts off steeply and then moves towards a gently sloping plateau, but it should be noted that several of the Wikipedias remained at a value of 1 article for a year or two after creation, which is the opposite of steep; and that the logarithmic scale used on the Y axis exaggerates small, low-level increases relative to large, higher-level ones.

Viewed absolutely, the other languages are all "flatlining" compared to English. Relatively, though, Afrikaans has achieved "take-off" while the ethnic African languages except SePedi are almost static; the differences between them are small compared to the difference between all of them and Afrikaans (not to mention English).

SePedi appears to be undergoing take-off but it may be too early to conclude that definitively. The difference between SePedi and the other ethnic African languages can be better appreciated when viewed on a non-logarithmic scale, which is practical once English and Afrikaans are removed from the picture (Figure 2).

Viewed thus, it is tempting to note that SePedi approached the current size of the other Wikipedias in this graph a few years earlier than they did, then remained for a few years in a plateau (albeit gently sloping) as they have been recently, and then to guess that they might also soon undergo dramatic growth. This would be fallacious, as article contribution is a function of a community, and each community has its own history. A more comprehensive study than the present one could investigate by means of interviews exactly what factors led to SePedi's surge. Another notable difference is that SePedi was a late starter (2007) but with a first-year count of 34 which was already ahead of some of the older Wikipedias.

Figure 2: Ethnic African-language article count (non-logarithmic)

Ethnic African Article count 2002-2017
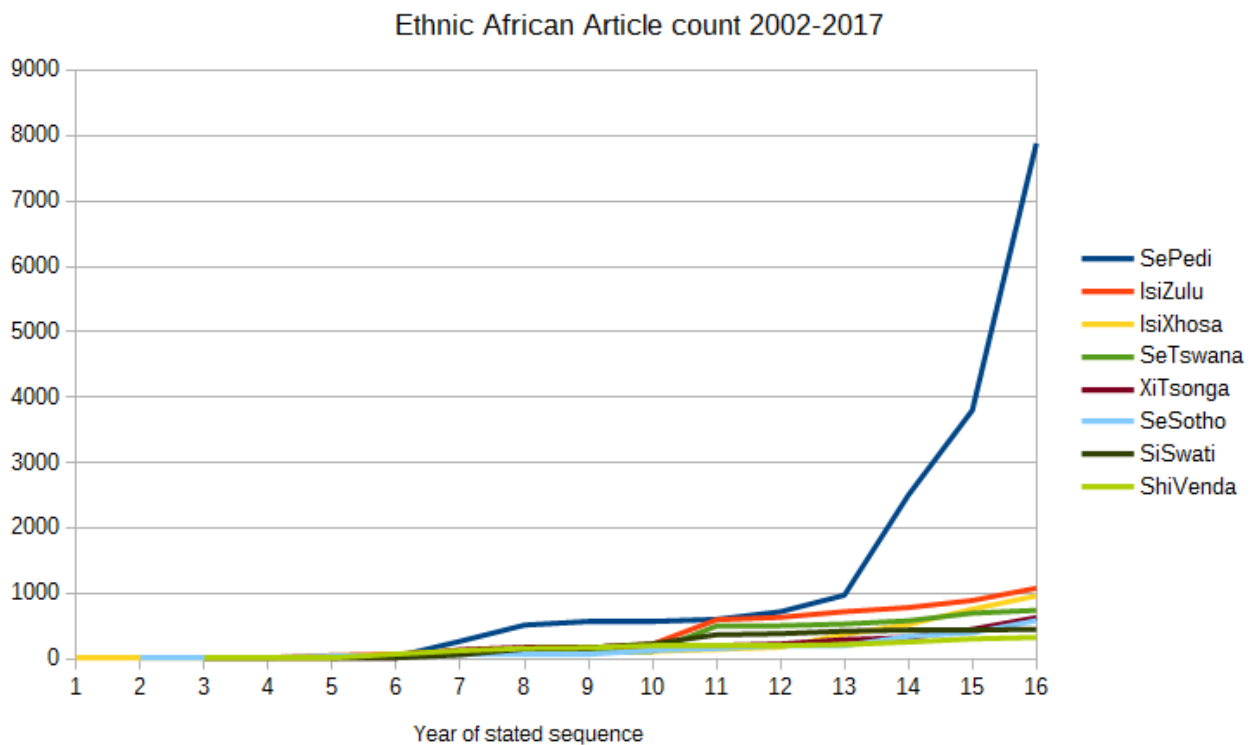
5.2: Contributors

This refers to the total number of editors who have created/edited one or more articles to date, and can be seen as a measure of the reach of a Wikipedia's appeal for contribution. A peculiarity of Wikipedias is that edits can be done either by registered and logged-in contributors, or by casual anonymous visitors. However anonymous edits are more likely to be undone ("reverted") or superseded by regular editors (and/or by "bots"(discussed later) operated by regular editors) since they are not defended in the article's Talk page, so someone feeling strongly about placing something on record has an incentive to register and log in (even if pseudonymously).

Figure 3: Contributor Counts 2001-17

Contributors (cumulative)
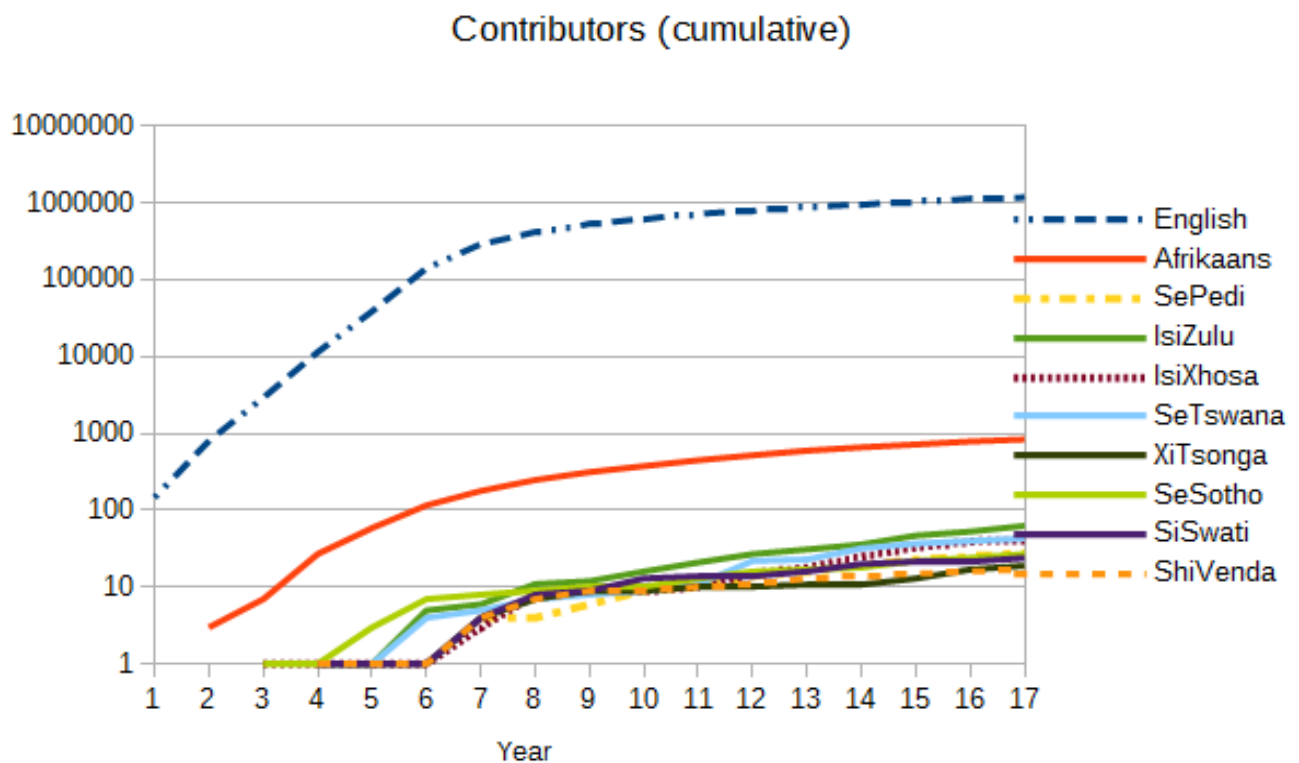
Again there is a superficial resemblance between different languages' curves; again however, some of the Wikipedias remained at a value of one for a year or two. This metric, though, describes a set of people rather than of content and must be considered in the context of the concept of community.

Because the metric is cumulative, there is no indication of what proportion of the total remains active at any given time. . The issue (noted under "Articles") of the difference between SePedi and the other ethnic African languages invites a similar examination with a non-logarithmic scale (next page).
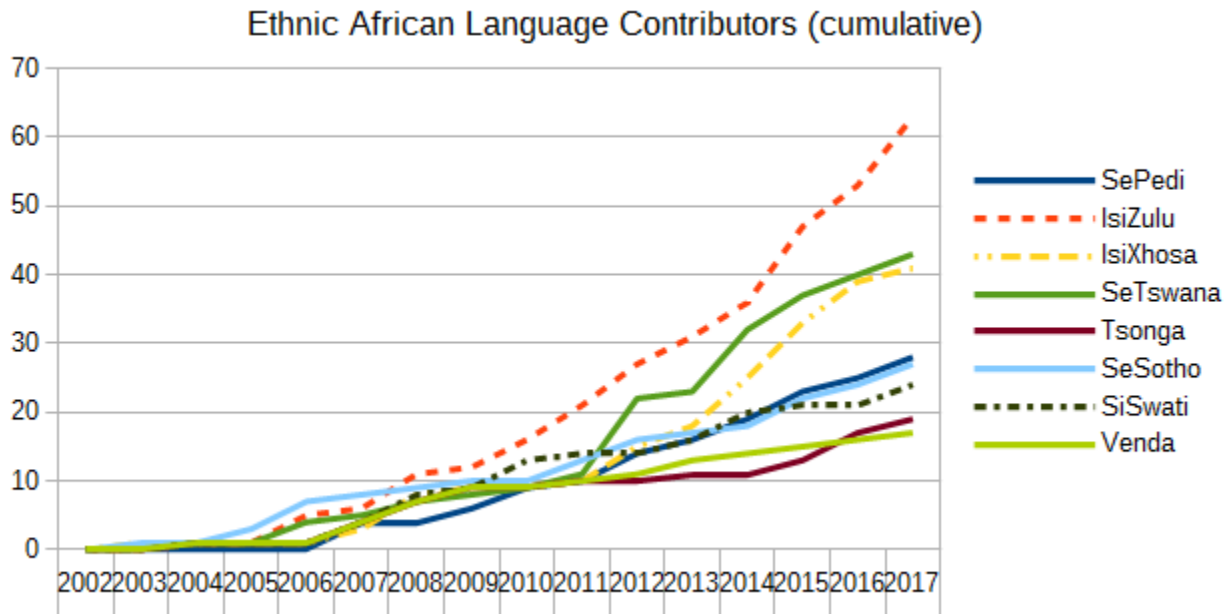
Figure 4: Non-Logarithmic View of Ethnic-African-Language Contributors

This reveals that contrary to what some might have expected, SePedi is far from having the most contributors. The surge noted previously must thus be ascribed to greater participation rather than greater recruitment. In the above dataset, rather than correlating with article count, there is a loose correlation between contributor count and (first-language) demographic representation – particularly if relative internet penetration is taken into account.

The estimate of speakers for each language given by Wikipedia[1] (below) includes non-first-language speakers, and those located in all countries (IsiZulu has a large proportion of non-native speakers, due to its history as a sub-hegemonic, subcontinental *lingua franca*):

Table 2: Approximate number of speakers (global)

English	1500 000 000
Afrikaans	13 000 000
IsiZulu	26 000 000
IsiXhosa	8 000 000
SeTswana	13 000 000
SeSotho	14 000 000
SePedi	14 000 000
SiSwati	3 000 000
XiTsonga	16 000 000
TshiVenda	3 000 000

Footnote:

1) https://stats.wikimedia.org/EN/Sitemap.htm

For example, if one notes that outside of South Africa, SeTswana is also an official language – and home language of the vast majority – in Botswana which is a relatively flourishing country, likely to have a good proportion of relatively early adopters of internet use. By comparison SiSwati is the majority (and official) language of Swaziland which has low internet penetration, SeSotho similarly in Lesotho, and IsiNdebele which lacks a Wikipedia, is the second-largest spoken language of Zimbabwe (ChiShona, Zimbabwe's most spoken language, does have a Wikipedia). XiTsonga speakers mostly live either in parts of South Africa bordering Mozambique, or in Mozambique itself – in either case, also having low internet penetration.

An interesting aside is the fact that some contribution is by "bots"—automation scripts (Halfaker & Riedl, 2012, 79-82) which users can deploy via Wikipedia's API, which perform routine tasks such as reverting an article to an earlier state (if vandalism is suspected), inserting an editor's signature after their contribution, and even generating article stubs from database items (Sauper & Barzilay, 2009, 208-216).

Here is a sample of bot activity in our various Wikipedias for May 2016 (Wikimedia Foundation, *ibid*):

Table 3: Contributions by bots

Language	All edits by bots(%)	Article creations by bots (%)
English	10	3
Afrikaans	45	6
IsiZulu	67	1
IsiXhosa	61	1
XiTsonga	73	0
SePedi	32	0
SiSwati	84	0
SeTswana	62	0
SeSotho	77	0
TshiVenda	69	0

The preponderance of human over robotic edits in English and Afrikaans with the reverse in most other languages may seem alarming, but probably reflects the thriving human editor communities in English and Afrikaans. This hypothesis is supported by the fact that SePedi, which as we have seen has a more active community than the other ethnic languages, has a lower bot edit percentage even than Afrikaans.

However the disparity in proportions of article creation is more interesting, especially considering the potential for generation of article stubs. Automatic stub creation offers a way to expand a Wikipedia with far less human effort since the more tedious parts of article creation are taken care of, allowing human editors to complete a larger number of articles. The 6% of Afrikaans articles created by bots would amount to almost 3000, which is more than the total article count of any of the ethnic languages except SePedi. The lack of such contribution in ethnic languages could be a result of bot authors not being motivated and/or able to engage with Bantu languages, which can be addressed by engaging with them and/or encouraging/enabling the creation of bots by African contributors.

Stub creation is not the only "force multiplier" - machine-assisted translation is also possible, especially between languages belonging to the same groups. The major groupings of South African languages are these: Nguni (IsiZulu, IsiXhosa, SiSwati and IsiNdebele), Sotho-Tswana (SeSotho, SePedi, SeTswana) and Germanic (English and Afrikaans). There is still a paucity of academic research documenting this within Wikipedia, but web searches reveal the level of community activity. However the Wikipedia translation-assistance tool, "Content Translation", is still in beta status and not all languages are supported yet. When translation is not supported between two languages, the tool merely suggests the source text in the target language column. Its support for South African languages in November 2017 was confined to Afrikaans and IsiXhosa - between each other, and from English (machine-assisted creation of English articles from other languages is restricted by default to highly-trusted editors (>5000 edits)).

This bleak picture belies low-hanging fruit: it could change dramatically with relatively little investment. Translation between related languages is easier than between language groups. The fact that translation support exists between one of the Nguni languages (IsiXhosa) and the Germanic group (English, Afrikaans) means that support could easily be extended to other Nguni languages and even the Sotho-Tswana group, because the Nguni and Sotho-Tswana groups are related within the larger Bantu grouping, which generally has a high degree of internal structural consistence (Maho, 1999 and Byamugisha *et al.*, 2016, 25-36).

5.3: Active Editors

This metric addresses the unanswered questions posed by the previous one, regarding the size of communities of contributors. It does so by arbitrarily defining an active editor as having made 5 or more contributions in the month prior to measurement. The terminology and criterion are those of Wikipedia Statistics.

*Figure 5: Active Editors Each September 2001-17*
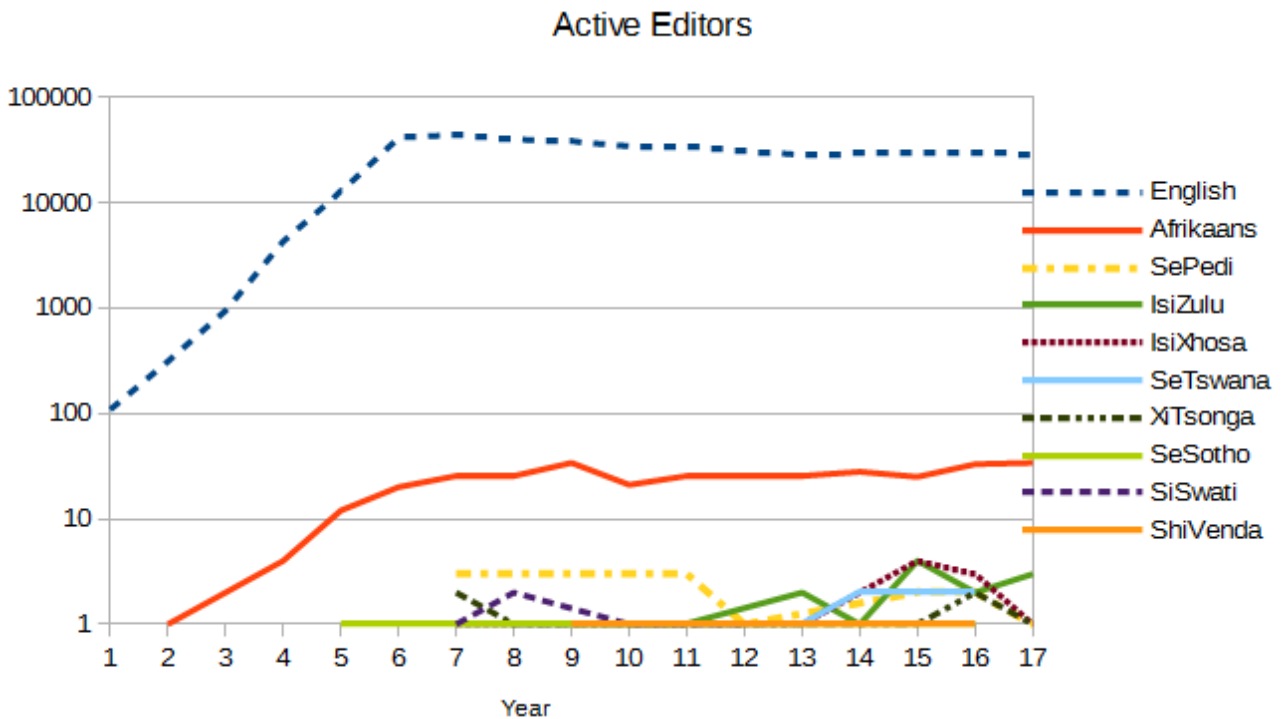
Here we see an illustration of quantitative differences so great as to become qualitative. Whereas English and Afrikaans Wikipedias clearly have stable communities of contributors, none of the other languages has. The concept of a critical mass comes to mind; a number below which contributors do not experience the reinforcement of a supportive community, and therefore tend to drop out of activity. Of course this would not be a specific number as measured by this metric, but rather a number of contributors online (even if only in commenting) as a given time, regardless of edits. Thus, degree of access to the internet becomes relevant.

However it should be borne in mind that different-sized communities can offer qualitatively different experiences. At a gathering of Afrikaans editors in 2017, some who have also participated in the English Wikipedia reported to the author that the smaller of the two is far more newcomer-friendly than the larger. It may be that there is a sweet spot for emerging Wikipedias that the South African ethnic languages have yet to reach.

5.4: Page Views

Whereas the previous two metrics concern contributors, this one reflects users. The figures given are for the month preceding measurement. Unlike the other metrics, it is only available as far back as 2008.

Figure 6: Page September Page Views 2008-17
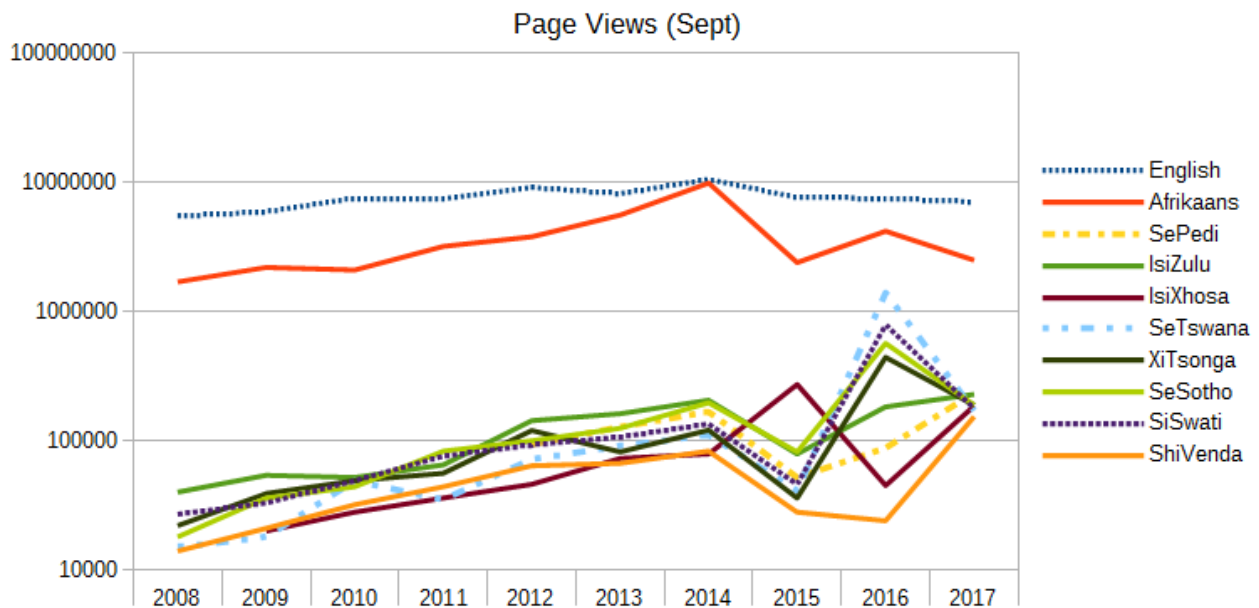
(Note: some of the perturbation visible around 2015 is due to a redefinition at that time of what constitutes a page-view. However the same change applied to all Wikipedias at the same time so the comparative value of the graph remains; note also that this is a volatile metric, as for example a news report mentioning a language's Wikipedia might lead to a spike in page-views for that month.)

Finally, we have some (modestly) good news: whereas the more mature Wikipedias have stable or even declining "viewership", the struggling ethnic African Wikipedias have a steady upwards trend over the sample period. In the 9 years, the average number of page-views per ethnic Wikipedia increased from 27 200 to 191 375, which is more than seven times. Interestingly the dramatic surge in SePedi articles is not yet accompanied by a surge in views. Of course, considering that many speakers of these languages are economically disadvantaged, this trend can be seen as a reflection of their being latecomers to the internet.

In sum, South Africa's Wikipedias, like its socio-linguistic landscape, are still deeply impacted by the legacy of apartheid. Of course, we cannot turn back time and regain lost opportunities. The question is rather, how we can make up for lost time, and address inequalities that should have been addressed long ago. Waiting for ethnic Africans' access to digital technology to increase is not only insufficient, but concedes territory to the myriad corporate distractions which vie for the attention of the newly connected. We must turn to look at proactive steps to guide future development.

Chapter 6: Policies & Methods

This chapter seeks to review possible responses to the challenging situation discovered earlier in the work. It is suggested that Wikipedias can be part of a holistic rehabilitation of disadvantaged languages in South Africa. Of course policy decisions include choosing technology, so some mention of categories of technology has to be made, although the next chapter looks at technology specifically.

Worldwide, more and more teachers in both formal and informal education, and at all levels of education are assigning learners tasks involving creation, improvement or translation of Wikipedia articles. This not only benefits the users of the Wikipedias, but also provides self-esteem to the learners (Azzam *et al.,* 2017 and Richardson, 2010).

In envisioning possible futures for South Africa's ethnic-language Wikipedias, it is worth mentioning that the English Wikipedia exists in various forms, because the canonical version is so vast, so inclusive in content, and so competitive in its community. In proposing the use of an ethnic Wikipedia in education, there are more options than just to use the existing text for that language – for example if parents or educational authorities fear exposing children or youth to "adult" content, equivalent projects in that language could be considered for the following:

- Vikidia[1] is a version for (and by) 7- to 12-year-olds;
- Wikipedia Version 1.0[2] is a project to produce variously curated versions for offline use in schools, typically having around 45000 articles and depending on whether pictures are included or not, being between 1,4 Gb and 5Gb in size;
- Wikipedia For Schools[3] is similar to Version 1.0, being a selection of articles grouped according to the UK curriculum, having 6000 articles and 50 000 images;
- among the 286 Wikipedias hosted by the Wikimedia Foundation is one in "Simple English". It has 127000 articles and occupies 170Mb.

The point here is that custom versions are relatively easy to produce. Many of the tasks involved in "forking" a Wikipedia can be automated; online collaboration by volunteers can achieve remarkable things. However, there is no reason why "real-world" resources cannot also be mobilised for such ends; even if the official education system is too overstretched and inflexible, global funding for Wikipedia activity in schools is available as demonstrated by the Wiki Education Foundation[4], and various editathons, translatathons and hackathons are supported by the Wikimedia Foundation, as is a research portal which among other things, aims to automate cross-language article suggestions[5].

Footnotes:

1) https://www.vikidia.org/

2) https://en.wikipedia.org/wiki/Wikipedia:Version_1.0_Editorial_Team

3) http://schools-wikipedia.org/

4) https://wikiedu.org/

5) https://research.wikimedia.org/knowledge-gaps.html

In envisaging future scenarios, it is worth noting that recent publications indicate a shift in "mature" internet markets whereby the digital divide moves from income-determined access to technology, to income-determined participation in content. When low-income users do eventually get affordable internet access, they often spend more time online than higher-income users, and their use of content is different. Crudely put, they are more susceptible to clickbait (Van Deursen & Van Dijk, 2014, 507-526). Given this shift towards personal-device "addiction" among youth, education planners can try to anticipate this and engage with it (most low-resource schools in South Africa currently ban personal digital devices, and school-supplied devices, if any, are usually only available to each learner for not more than a few hours per week).

The prohibitionist approach is denial of the inevitable. A better trajectory is demonstrated by Chris Hani High (Khayelitsha, Cape Town) which is a fairly typical low-resource urban public school. In 2015, after establishing a partnership with a sponsor of public WiFi, the principal persuaded the governing body to take a leap of faith by unbanning mobile phones. The principal is now an evangelist for such unbannings[6], citing online research by learners (largely Wikipedia) as the main benefit, and the cost of data as the main constraint (the sponsorship grants learners small data bundles in return for recycling cans, bottles etc.). The parents' greatest fear, that mobile phones would generate conflict and crime among the learners, did not materialise. The teachers and administrators use a capped ADSL connection funded by the governing body.

In such a situation where the cost of data is critical, locally hosted resources including Wikipedia could be of great benefit; the means for doing so are discussed in the next chapter. In many rural schools, resources are even scarcer, and the relative benefit of a digital library would be even greater. Few rural learners have WiFi-capable devices, hence the use in many such situations of shared tablet computers whose WiFi modules are configured only to access a particular wireless access point so as to have less risk of being stolen. This is a standard feature of the BRCK kit mentioned earlier.

An example of pre-emptive user engagement which could be used by educationists is the gamification of Wikipedia article translation as proposed by Moji (in press). Besides the canonical translation tool reviewed in Chapter 5, others exist in various stages of development and with various specialised intentions[7].

Footnotes:

6) The author attended a public gathering addressed by him in early 2017.

7) https://blog.wikimedia.org/2017/12/19/wikimedia-translation-tools/

In considering interventions in the context of Wikipedias and education, it can be argued that the evolution of science itself is shaped in part by the existence of Wikipedias (Thompson & Hanley, 2018). It can also be argued that making reference material available digitally is more cost-effective than the traditional provision of libraries. Institutions which lack affordable internet access can have their needs addressed by alternative networks and/or onsite hosting (Hadzic *et al.*, 2016).

Besides interventions directly related to Wikipedia and other specific content resources, attention should be given to general policy interventions to promote a conducive environment. A pertinent paper by Walton & Donner (2012) reviews the interplay in South Africa between public facilities such as libraries, and the ever greater ownership of mobile internet access devices by citizens (in the context of high mobile data costs), and makes several recommendations: not only continued roll-out of both public desktop computers and public WiFi, but better training of librarians and other facility operators.

Chapter 7: Technological Options

This chapter reviews some options for interventions in technical aspects of access to Wikipedia usage and contribution in South Africa. Devices and services are discussed and some tentative usage tests described.

The ecosystem of hardware and software involving Wikipedia is rapidly diversifying, and costs are coming down. In South Africa, the government's **SA Connect** programme aims to provide broadband to all schools, libraries and other government facilities but is behind schedule[1].

Before examining technology itself, we should consider what supportive organisational infrastructure is available – equipment and connectivity alone could, at best, exacerbate inequality and at worst, be a waste of money. Besides the formal education authorities' technical efforts, numerous (largely international) outreach projects are available, for example those listed at UCT's ConnectED website[2]. Such institutional support is easily overlooked in the quest for the content it is associated with; sometimes these organisation even assist in sponsoring equipment or channelling donations thereof.

Although schools and libraries are supposed to have desktop computers, these will probably never be sufficient in number to meet the needs of all users. Thus WiFi hotspots will be important for the remainder. Hence, for the foreseeable future, the most affordable access to Wikipedias by the disadvantaged masses in South Africa will be hand-held: mobile phones and tablets (see the discussion of school device policy in the previous chapter). Hence a brief survey of how these can be used to support Wikipedia activity is in order, along with suggestions for interventions.

The cheapest (from R200/USD14 in November 2017) are feature phones using embedded operating systems such as Java Mobile Edition, Windows CE and (in older models) Nokia's System 40, which access specific online services including **.mobi** sites (featuring modified XHTML, which succeeded the modified XML used by Wireless Application Protocol (WAP) devices/services)(Saha *et al.*, 2001, 54-58) . Not only are these phones cheaper, but they have longer battery life than smartphones – an important feature for rural or informal-settlement dwellers who are often far from the electricity grid.

Footnotes:

1) https://www.itweb.co.za/content/xo1Jr5qxbDo7KdWL

2) http://connectedafrica.org.za/?page_id=162

However, feature phones lack WiFi capability and so the only way they are likely to be used by the poor to access Wikipedia is via zero-rating (the provision of a service free of charge, typically resulting from deals between content providers and mobile operators). For Wikipedia, this has so far only been offered by one mobile operator in South Africa and only in conjunction with the Opera Mini Browser[4]. Zero-rating has been criticised as manipulative when offered by social media corporations (Sen *et al.*, 2017, 12-24), but Wikipedia cannot be classified thus.

A test edit was attempted on the feature phone described in chapter 4; while it was possible both to log in to Wikipedia and to navigate within an article, editing was problematic. Firstly, on pressing the Edit button, Opera Mini presented an ominous notification about data security (presumably because of its use of intermediary servers which relay page requests and compress the content before serving it to the user device). Secondly, when selecting text to edit, the text area enlarged and while "space" and "backspace/delete" functions worked, character entry did not (both with and without predictive text enabled). The device in question has a typical 4 by 3 keyboard; no test was made with a QWERTY-enabled feature phone, since this test sought to establish accessibility via the cheapest devices.

After the conclusion of the current investigation (i.e. during the examination and revision stage) the Wikimedia Foundation announced a policy decision, to discontinue support for its zero-rated programs. This was at the height of the net-neutrality debate, although the announcement made no reference to that. The historical importance free access to Wikipedia was highlighted by a keynote speaker at Wikimania 2018 a few months later[6], and at the time of going to press there is still discussion in the Wikimedia-ZA community about revisiting the policy change.

Besides resumption of official support, the Foundation should consider extension of zero-rating to other networks and other platforms available on feature phones, e.g. as an app. In so doing, a useful exercise would be to revisit the history of WAPedia which was an unofficial, advertising-funded, low-data-use, read-only implementation of Wikipedia accessible by low-end feature phones (discontinued in 2013). It supported "several" languages and offered to extend to others on request. Its decline accompanied the decline in use of feature phones in the first-world market/s at which its sponsors targeted their adverts[7]. Hence a revival is conceivable given the long tail of feature phones in Africa, either commercially as in the original, or publicly/charitably funded. While the read-only nature of the service is regrettable, reaching new audiences not only provides immediate use-value, but can ultimately create more incentive for further participation across the relevant communities.

Footnotes:

4) https://blog.wikimedia.org/2014/03/18/mtn-south-africa-responds-to-sinenjongo-high-school-open-letter-and-launches-wikipedia-zero/

6) https://twitter.com/Wikimedia_ZA/status/1020208906063286273

7) https://en.wikipedia.org/wiki/Wapedia

Another initiative worth revisiting is the Praekelt Foundation's Vumi project, which renders partial access to Wikipedia even to "dumb" phones, using the USSD protocol[8]. In future, attention should be maintained upon the feature phone market since the "long tail" might become even more of an "on-ramp" into the world of smartphones if Google's recent investment in KaiOS[9] bears fruit.

The cheapest smartphones (priced from R350/USD25 in November 2017) feature the Android operating system. These, besides full web access via browsers, also enable apps to be downloaded and installed, typically from the Google Play Store. Dozens of these centre on Wikipedia; here follows a short overview of the "official" options – those approved by the Wikimedia Foundation.

The Kiwix app is an open-source offline reader which, once installed, offers downloads not only of various forms of Wikipedia and other Wikimedia products, but even packaged compilations of TED talks. A search (as opposed to the suggested downloads when the app is loaded) yields 3 versions of Wikipedia (text-only, text and pictures, and text, pictures and video) as well as assorted other WMF products for the following South African languages: Afrikaans, IsiZulu, IsiXhosa, SiSwati, SePedi, SeTswana, XiTsonga, TshiVenda and two versions of SeSotho,

For a typical cheap Android phone which has limited storage, downloading and installing Kiwix and one or more of the existing South African language Wikipedias in the phone's inbuilt storage and/or on a memory card should be no problem because Kiwix uses the compressed ZIM file format, which is indexed to enable search without decompression. The app occupies 35,7 Mb and the Wikipedias range from 298 Kb upwards. The app automatically splits files larger than 4Gb to prevent problems with the Android file-system.

The official (online) Wikipedia app for Android occupies 11,4 Mb and version 2.4 160-r-2016-10-14 was tested. A small edit consumed 1,43 Mb. Of note was that the editing interface was near-WYSIWYG (as opposed to Wiki Markup Language (WML) which until recently was the default in the desktop, and considered a major constraint on participation).

All these modest data figures must however, for present purposes, be seen through the eyes of the "less-connected" for whom a 10Mb bundle is overpriced and a serious investment (de Lanerolle *et al.*, 2017). Besides the existing model of zero-rating at the whim of mobile operators and social media corporations, the "freemium" model has been proposed (Esselaar et al., 2017). This extends the usage of the term (coined for software) to connectivity, and differentiates between legacy 2G infrastructure and newer networking, proposing that unrestricted internet access via the former be free of charge.

Footnotes:

8) http://blog.praekeltfoundation.org/post/65981723628/wikipedia-zero-over-text-with-praekelt-foundation

9) https://techcrunch.com/2018/06/27/google-kaios/

Offline versions of Wikipedia and similar content are, as previously mentioned, often hosted by means of ever-more-affordable nano-servers such as the Raspberry Pi. In such contexts, a small investment can have a high impact. Typically the content is made available via a pre-configured dashboard which also includes offers course management and/or learner management systems, and sometimes server admin functions. These platforms first appeared in low-resource universities but are now spreading to low-resource schools[12]. A modest but totally off-the-shelf solution is offered by the Mesh Potato, a node router popular in community telecoms networks; it can be ordered with a pre-installed locally hosted library which includes a version of Wikipedia.

Whereas an isolated nano-server can only serve a few clients in its immediate network, emerging techniques such as the "community CDN" or cloudlet paradigm discussed by Hadzic *et al.* (*ibid)* offer new possibilities of resource-sharing over wide areas, without the cost of internet *per se*, and often with less latency.

Ironically the best resourced current Wikipedia outreach in Africa, **Afripedia**, promotes use of the French version, mostly offline via Kiwix; however, it is not a project of the Foundation. Although *Francophonie Africaine* is large, most of it consists of non-first-language and/or creole speakers. The metropolitan language of the French Wikipedia is not African any more than English is[13]. Similar projects should be launched for African languages, including South African ones. Funding could be sought from public and private sectors, including the Wikimedia Foundation.

The communities of (mostly) volunteers who develop and maintain these hardware and software tools are evolving and interacting, and the Wikimedia Foundation plays an important part in fostering this. Furthermore, in 2018 the Foundation is convening its largest global event in South Africa, and has chosen to focus on bridging gaps between languages, cultures and Wikipedias[14].

Footnotes:

12) For example eGranary https://www.widernet.org/eGranary/

13) https://en.wikipedia.org/wiki/Afripedia_Project

14) https://blog.wikimedia.org/2018/02/05/wikimania-cape-town-ubuntu/

Chapter 8: Conclusion

We have seen that the inequality among South Africa's Wikipedias mirrors those in its society and in the world as a whole. We have also seen that the exact configuration of that inequality has been shaped first by colonial and apartheid agendas, then by post-apartheid default or passivity. Likewise, the technical means of access to Wikipedias, namely internet infrastructure and connectivity, and end-user equipment, are unequally distributed. The situation is exacerbated by the unpreparedness of public institutions and the limited capacity of volunteer communities.

However, the Wikimedia Foundation is a dynamic entity and has plenty of work cut out for it. It serves as a figurehead for a social movement which is still weak in South Africa but stronger, and growing in strength, in many countries. During the revision of the current document Wikimania 2018 was successfully hosted in South Africa[1], with its theme touching on the concerns mentioned here. Furthermore the occasion served to dramatically boost membership of the local chapter, Wikimedia-ZA, both numerically and in terms of diversity. This momentum should not be allowed to dissipate.

The existing range of hardware and software tools and resources is fast-evolving and still under-utilised. A lot can be achieved with modest investments of time and funds, particularly if coupled with the collaboration which is the hallmark of the Wikiverse.\

Of particular interest is the nascent community telecommunications movement. Also during the revision of this text, the 3rd African Community Network Summit[2] was hosted in South Africa, and locally hosted content including Wikipedias, were on the agenda. Logistical and advocacy support for the movement from the Internet Society[3], which is a global player, is significant.

Somewhat of a wild card at this stage is the arrival of open cellular technology[4]. Although this holds great promise for community networks, its origins in corporate philanthropy and the fact that cellular communications are more tightly regulated than those using unlicensed spectrum, make its future opaque, and likely to differ from country to country.

Footnotes:

1) https://wikimania2018.wikimedia.org/w/index.php?title=Wikimania&oldid=21561

2) https://www.internetsociety.org/events/summit-community-networks-africa/2018/

3) https://www.internetsociety.org/about-internet-society/

4) https://telecominfraproject.com/opencellular-wireless-access-platform-design/

More generally, various factors are contributing to the rise of "edge computing" as a paradigm. Shortly before the time of going to press the first large-scale consumer-level rollout of local hosting was announced: the installation of Nextcloud instances into millions of routers in Japan[5]. These are capable of federation with each other but equally, of independent operation - at the discretion of the owner. This shift of control is nothing less than a blow to the power of the corporate cloud operators and to the surveillance industry. Time will show that autonomy (which includes privacy) is as important to justice as is access to communications, although more attention is still (understandably) focussed on access in Africa.

Besides social and institutional action, there is urgent potential for further research. This dissertation has touched on some areas where worldly events are far ahead of their documentation by academia, hence its heavy dependence on non-academic sources. The ICT4D discipline has an important part to play in the future of South Africa, Wikipedia and the world as a whole.

Footnote:

1) https://nextcloud.com/blog/japan-to-add-millions-of-new-nodes-to-federated-nextcloud-network/

Bibliography

Alexander, N. (1991). *Language policy and national unity in South Africa/Azania*. Buchu Books.

Alexander, N. (2001). *Language politics in South Africa*. In Bekker, S., Dodds, M. and K. Meshack (eds) *Shifting African Identities.* Pretoria: Human Research Science Council, 141-152.

Ames, M. G. (2016). *Learning consumption: Media, literacy, and the legacy of One Laptop per Child.* The Information Society*, 32*(2), 85-97.

Azzam, A., Bresler, D., Leon, A., Maggio, L., Whitaker, E., Heilman, J., ... & Trotter, F. (2017). *Why medical schools should embrace Wikipedia: final-year medical student contributions to Wikipedia articles for academic credit at one School.* Academic Medicine, 92(2), 194.

Baker, S. (2012). *A place for Wikipedia or putting Wikipedia in its place*. New Horizons for Learning, 10(1).

Brewer, E., Demmer, M., Du, B., Ho, M., Kam, M., Nedevschi, S., ... & Fall, K. (2005). *The case for technology in developing regions.* Computer, 38(6), 25-38.

Byamugisha, J., Keet, C. M., & DeRenzi, B. (2016,). *Bootstrapping a Runyankore CNL from an isiZulu CNL*. In *International Workshop on Controlled Natural Language*. Springer International Publishing.

Chetty, M., Banks, R., Brush, A. J., Donner, J., & Grinter, R. (2012, May). Y*ou're capped: understanding the effects of bandwidth caps on broadband use in the home*. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 3021-3030). ACM

Chiriac, T. (2106) *Wiki as collaborative learning: Moldova experience* in Beseda,J.(ed) DISCO 2016 proceedings.

Cohen, N. (March 17, 2008). <u>*Open-Source Troubles in the Wiki World*</u> The New York Times.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. Teachers College Press.

Davis, R. (23 August 2017) *Mahmood Mamdani: Sixteen years on, UCT's old nemesis returns to talk decolonisation* Daily Maverick

de Lanerolle, I., Walton, M., & Schoon, A. (2017). *Izolo: mobile diaries of the less connected.*

Desai, Z. (1995). *The evolution of a post-apartheid language policy in South Africa: an ongoing site of struggle*. European Journal of Intercultural Studies, *5*(3), 18-25.

Di Gessa, S., Poole, P., & Bending, T. (2008). *Participatory mapping as a tool for empowerment: Experiences and lessons learned from the ILC network*. Rome: ILC/IFAD, *45*.

Dolmaya, J. M. (2017): *Expanding the sum of all human knowledge: Wikipedia, translation and linguistic justice*, The Translator, DOI: 10.1080/13556509.2017.1321519

Esselaar, S., Song, S., & Stork, C. (2017) *Freemium Internet: Next Generation Business Model to connect next billion*, 28th European Regional Conference of the International Telecommunications Society (ITS).

Heeks, R. ICT4D 2.0: *The Next Phase of Applying ICT for International Development*. Computer, 41(6).

Ford, Heather, and Judy Wajcman. *'Anyone can edit', not everyone does: Wikipedia and the gender gap.* Social Studies of Science (2016).

Geiger, R. Stuart, and David Ribes. *Trace ethnography: Following coordination through documentary practices.* System Sciences (HICSS), 2011 44th Hawaii International Conference on. IEEE, 2011.

Giliomee, H. (2009). *A note on Bantu Education, 1953 to 1970*. South African Journal of Economics, 77(1)

Gudmundsdottir, G. B. (2010). *From digital divide to digital equity: Learners' ICT competence in four primary schools in Cape Town, South Africa*. International Journal of Education and Development using Information and Communication Technology, 6(2), 1H.

Guth, S. (2007). *Wikis in education: is public better?*. In Proceedings of the 2007 international

symposium on Wikis (pp. 61-68). ACM.

Hadzic, S., Phokeer, A., & Johnson, D. (2016). T*ownshipNet: A localized hybrid TVWS-WiFi and cloud services network*. In *IEEE International Symposium on Technology and Society (ISTAS)* (Vol. 20, p. 22).

Halfaker, A., & Riedl, J. (2012). *Bots and cyborgs: Wikipedia's immune system*. Computer, 45(3).

Kamwangamalu, N. M. (2000). *A new language policy, old language practices: Status planning for African languages in a multilingual South Africa*. South African Journal of African languages, 20(1).

Kittur, A., Chi, E., Pendleton, B. A., Suh, B., & Mytkowicz, T. (2007). *Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie*. World wide web, 1(2), 19.

Konieczny, P. (2012). *Wikis and Wikipedia as a teaching tool: Five years later*. First Monday.

Legassick, M., & Wolpe, H. (1976). *The Bantustans and capital accumulation in South Africa*. Review of African Political Economy, 3(7).

Limb, P. (2002). *Early ANC leaders and the British world: ambiguities and identities*. Historia.

Maho, J. (1999). *A Comparative Study of Bantu Noun Classes*

Maseko, P., Nosilela, B., Sam, M., Terzoli, A., & Dalvit, L. (2010). *The role of the web in the promotion of African languages*. Alternation, 17(1), 312-327.

Mesthrie, R. (Ed.). (2002). *Language in South Africa*. Cambridge University Press.

Moji, J. *The Temporal Effects of Gamification in The Context of African Language Content Generation*. In press, Univ. of Cape Town.

O'Meara, D. (1983). *Volkskapitalisme: class, capital and ideology in the development of Afrikaner nationalism, 1934-1948*. Cambridge University Press.

Ouane, Adama, and Christine Glanz. *How and why Africa should invest in African languages and multilingual education*. UNESCO Institute for Lifelong Learning, 2010.

Phillipson, R. (1996). *Linguistic imperialism: African perspectives*. ELT journal, 50(2), 160-167.

Rask, M. (2008). *The reach and richness of Wikipedia: Is Wikinomics only for rich countries?*. First Monday, 13(6).

Richardson, W. (2010). *Blogs, wikis, podcasts, and other powerful web tools for classrooms*. Corwin Press.

Reimers, G., & Neovesky, A. (2016). *Contributing to Wikipedia as an assignment for undergraduate students. PDF).* Proceedings of DisCo.

Rothenberg-Aalami, J., & Pal, J. (2005). *Rural telecenter impact assessments and the political economy of ICT of Development* ICT4D.

Sachs, J. D. (2012). *From millennium development goals to sustainable development goals.* The Lancet, 379(9832), 2206-2211.

Saha, S., Jamtgaard, M., & Villasenor, J. (2001). *Bringing the wireless Internet to mobile devices.* Computer, 34(6).

Sanders, L., & Stappers, P. J. (2014). *From designing to co-designing to collective dreaming: three slices in time.* **interactions**, *21*(6), 24-33.

Sauper, C., & Barzilay, R. (2009). *Automatically generating Wikipedia articles: A structure-aware approach.* In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1* Association for Computational Linguistics.

Sen, R., Ahmad, S., Phokeer, A., Farooq, Z. A., Qazi, I. A., Choffnes, D., & Gummadi, K. P. (2017). *Inside the walled garden: Deconstructing facebook's free basics program.* ACM SIGCOMM Computer Communication Review, 47(5).

Thompson, N, & Hanley, D. *Science is shaped by Wikipedia: evidence from a randomized control trial* (2018).

Trace, C. (2016). *Rethink, Retool, Reboot: Technology as if people and planet mattered.* Rugby, UK: Practical Action Publishing.

Van Deursen, A. J., & Van Dijk, J. A. (2014). *The digital divide shifts to differences in usage.* In *New media & society, 16*(3).

Van Dijk, Z. (2009). *Wikipedia and lesser-resourced languages.* In *Language Problems and Language Planning,* 33(3).

Walton, M., & Donner, J. (2012). *Public access, private mobile: The interplay of shared access and the mobile Internet for teenagers in Cape Town.* Global Impact Study research report series.

Younus, U.M. (2017) *Fourth industrial revolution.*