

Evaluation of a Runyankore grammar engine for healthcare messages

Joan Byamugisha and C. Maria Keet and Brian DeRenzi

Department of Computer Science, University of Cape Town, South Africa,

{jbyamugisha,mkeet,bderenzi}@cs.uct.ac.za

Abstract

Natural Language Generation (NLG) can be used to generate personalized health information, which is especially useful when provided in one's own language. However, the NLG technique widely used in different domains and languages—templates—was shown to be inapplicable to Bantu languages, due to their characteristic agglutinative structure. We present here our use of the grammar engine NLG technique to generate text in Runyankore, a Bantu language indigenous to Uganda. Our grammar engine adds to previous work in this field with new rules for cardinality constraints, prepositions in roles, the passive, and phonological conditioning. We evaluated the generated text with linguists and non-linguists, who regarded most text as grammatically correct and understandable; and over 60% of them regarded all the text generated by our system to have been authored by a human being.

1 Introduction

The vast majority of doctor-patient interactions in a healthcare setting is through verbal communication. The provision of written information to patients, to complement and augment the face-to-face session, increases the amount of information they retain (DiMarco et al., 2005). Additionally, studies in health communication identified that patient information is likely to be more effective if it is personalized for a specific patient (Cawsey et al., 2000) and presented in an understandable form and manner (DiMarco et al., 2009). This however assumes that such information is communicated in one's first language, which,

however, may not be the case in multilingual societies. Language is important here because problems with language exacerbate literacy difficulties, which are further confounded in situations of health (DiMarco et al., 2009).

There are several Natural Language Generation (NLG) systems that generate customized patient information (Cawsey et al., 2000; DiMarco et al., 1995; de Rosis et al., 1999; Hussain et al., 2015; Lindahl, 2005; Mahamood and Reiter, 2011). These systems generate text in English, and one strategy to account for other languages could be to translate the generated English text to the target language. However, this requires correct machine translation for medical information, which does not exist for most languages, including our language of interest, Runyankore. Runyankore is a Bantu language indigenous to the south western part of Uganda (Asiimwe, 2014; Tayebwa, 2014; Turamyomwe, 2011), a country where English is the official language, whereas indigenous languages are still predominantly spoken in rural areas. There is therefore a need to investigate NLG for Runyankore.

We have limited our scope to generating patient summaries, drug prescription explanations, and treatment instructions. The kind of text generated could include: the number of pills to be taken, listing the active ingredient(s), what the medication does not contain, and the general classification of the medication (for example, that hydrocodone is an opiate). These are largely knowledge-to-text cases, for which ontologies, such as the medical terminology SNOMED-CT, can easily be used. There are several NLG systems that take ontologies as input,

mainly for English (Kaljurand and Fuchs, 2007), but also Latvian and Lithuanian (Gruzitis et al., 2010; Gruzitis and Barzdins, 2011), and Greek (Androutopoulos et al., 2013). These systems apply the template NLG technique. However, as was demonstrated in (Keet and Khumalo, 2017), templates are inapplicable to agglutinating Bantu languages, such as Runyankore.

Runyankore, like other Bantu languages has a complex verbal morphology (14 tenses), noun class system with 20 noun classes, and is highly agglutinative. A noun class (NC) determines the affixes of the nouns belonging to it, and this in turn determines the agreement markers on the associated lexical categories such as adjectives and verbs. To illustrate the agglutinative nature of Runyankore (taken from (Turamyomwe, 2011)):

Verb: *titukakimureeterahoganu*

English: ‘We have never ever brought it to him’

Decomposition: *ti-tu-ka-ki-mu-reet-er-a-ho-ga-nu*

Our previous work for Runyankore NLG (Byamugisha et al., 2016a; Byamugisha et al., 2016b) is not adequate, as it neither covers cardinality constraints (e.g., ‘take [exactly] 3 pills’); nor the passive (e.g., ‘operated by’); and ignore phonological conditioning of vowels in the agglutination process. Additionally, no evaluation on grammatical correctness of the generated text was done. All these aspects are needed for our scope of medical text generation. We undertook to address these shortcomings by: analyzing further details of the language to also process minimum, maximum, and exact cardinality, and devising algorithms for them; adding the phonological conditioning rules required for the scope; and extending the CFG of (Byamugisha et al., 2016b) with the passive. Their algorithms together with these novel ones were implemented and evaluated with 100 Runyankore speakers in rural Uganda and three Runyankore linguists. Most of the evaluated sentences were regarded as grammatically correct and understandable, and all computer-generated text was considered by a majority to have been written by a human being.

The paper is structured as follows: Section 2 introduces the new rules and algorithms required to generate more expressive text. Section 3 presents the experimental evaluation. We discuss in Section 4 and conclude in Section 5.

2 New rules and algorithms

Due to the limitations in our previous work for Runyankore NLG, as well as the structure of information in our domain of interest, healthcare, we developed new rules and algorithms to account for these gaps. These rules are added to those which already take care of the basic constructors in ontology languages, being named class subsumption (‘is a’ \sqsubseteq), conjunction (‘and’ \sqcap), negation (‘not’ \neg), existential quantification (‘at least one’ \exists), and universal quantification (‘all/each’ \forall) (Byamugisha et al., 2016a).

2.1 Cardinality constraints

In terms of the knowledge-to-text input, our previous language coverage was the description logic (DL) language \mathcal{ALC} . Adding qualified cardinality constraints brings the language feature coverage to \mathcal{ALCQ} , which is an important fragment of OWL 2 DL (Motik et al., 2009). We describe the verbalization patterns for maximum (\leq), minimum (\geq), and exact cardinality ($=$) in this section.

Maximum cardinality is worded in English typically as ‘a maximum of’, ‘not more than’, or ‘at most’. We use the Runyankore equivalent of ‘not more than’, *-tarikurenga*, which is the preferred word use. However, to form the full word for ‘not more than’, the subject prefix of the concept quantified over is required. We illustrate this in the following example. Consider Axiom 1 below, where ‘symptom’ has to be pluralized to ‘symptoms’, which is in noun class (NC) 4. The plural prefix for NC4 is *emi* (making *emicucumo* ‘symptoms’) that has a subject prefix *gi* that is attached to the *-tarikurenga*, ‘not more than’. Compare this with Axiom 2, where ‘courses’ *amashomo* is in NC 6, and therewith goes with the subject prefix *ga*.

Axiom1: Diabetes $\sqsubseteq \leq 3$ has.Symptoms

Buri ndwara ya shukari eine emicucumo gitarikurenga 3.

‘Every disease of diabetes has at most 3 symptoms’

Axiom2: Student $\sqsubseteq \leq 10$ takes.Course

Buri mwegi natwaara amashomo gatarikurenga 10.

‘Every student takes not more than 10 courses’

The algorithm to generate these coordinating elements is included in Algorithm 2.1.

Algorithm 2.1 Verbalization of Maximum Cardinality (\leq)

```

1:  $A$  axiom; Variables:  $a_1, o, n_1, n_p, nc_p$ , and  $sp_p$ ;
   and functions:  $getNumber(A)$ ,  $getNext(A)$ ,
    $getNoun(o)$ ,  $getPlural(n)$ ,  $getNC(n_p)$ , and
    $getSubjectPrefix(nc_p)$ 
2:  $a_1 \leftarrow getNumber(A)$     {get the number after
    $\leq$ }
3:  $o \leftarrow getNext(A)$       {get the role after the
   number}
4:  $n_1 \leftarrow getNoun(o)$     {obtain the noun from the
   role}
5: if  $a_1 = 1$  then
6:    $nc_1 \leftarrow getNC(n_1)$     {get the noun class of
   the noun}
7:    $sp_1 \leftarrow getSubjectPrefix(nc_1)$     {use the
   noun class to obtain the subject prefix}
8:   Result  $\leftarrow$  “ $sp_1$ tarikurenga  $a_1$ ”    {verbalize
   with the appropriate subject prefix}
9: else
10:   $n_p \leftarrow getPlural(n_1)$     {pluralize the noun}
11:   $nc_p \leftarrow getNC(n_p)$       {get the plural noun
   class}
12:   $sp_p \leftarrow getSubjectPrefix(nc_p)$     {get the
   plural subject prefix}
13:  Result  $\leftarrow$  “ $sp_p$ tarikurenga  $a_1$ ”    {verbalize
   with the plural noun and subject prefix}
14: end if
15: return Result

```

Minimum cardinality (\geq) is typically rendered in English as ‘a minimum of’, ‘not less than’, or ‘at least’. We apply ‘at least’ as the Runyankore verbalization, again because this is a more directly translatable version. Like the verbalization of existential quantification (\exists) in (Byamugisha et al., 2016a), it uses *hakiri* for ‘at least’. However, unlike \exists where we always have *-mwe* for ‘one’, \geq has the number instead (unless the number is 1). Thus, a verbalization similar to \exists is used, also using the subject prefix of the concept quantified over. The examples below illustrate this case; e.g., *e* is the subject prefix of NC 9, for *diguri* ‘degree’:

Axiom1: Panado $\sqsubseteq \geq 4$ has.ActiveIngredient

Buri Panado hakiri eine ebirungo by'amaani 4
‘Every Panado has at least 4 active ingredients’

Axiom2: Student $\sqsubseteq \geq 1$ has.Degree

Buri mwegi hakiri aine diguri emwe
‘Every student has at least 1 degree’

Similar to the verbalization of \leq , there is a need to pluralize the noun whenever the number after \geq is greater than 1 (as is the case in Axiom 1, above). Due to space limitations, we omit the algorithm, as it is similar to Algorithm 2.1.

The English verbalization of exact cardinality ($=$) is ‘exactly’. However, Runyankore does not have a direct translation for ‘exactly’, but uses ‘only’ instead. The word for ‘only’, *-onka*, requires the subject prefix to form the full word. In the following examples, *bw* and *ky* are the subject prefixes of NC 14 and 7, respectively, to which the nouns *obujuma* ‘pills’ and *ekitabo* ‘book’ belong, respectively:

Axiom1: Patient $\sqsubseteq = 2$ takes.Pill

Buri murweire natwara obujuma 2 bwonka
‘Every patient takes only 2 pills’

Axiom2: Child $\sqsubseteq = 1$ has.Book

Buri mwana aine ekitabo 1 kyonka
‘Every child has only 1 book’

As is the case with \leq and \geq , the noun is pluralized whenever the number is greater than 1. The algorithm is fairly similar to the others and therefore omitted due to space limitations.

2.2 Processing of Prepositions

The presence of prepositions in roles (relations), such as ‘works for’, and passives, such as ‘operated by’ changes the pattern in which the role is verbalized. While the algorithms cannot yet deal with any arbitrary preposition, we cover those that appeared in our test ontologies. These include: ‘with’ *na*, ‘in’ *omu*, ‘of’ (depends on the NC of the noun), and ‘by’ *w*. Our implementation of ‘of’ and ‘by’ is limited to the situation where the former is present after a noun, and where the verb is in past tense for the latter. Examples of roles containing these prepositions are: works **with** \rightarrow *naakora na*; offered **in** \rightarrow *neherezibwa omu*; part **of** \rightarrow *ekicweka kya*; and driven **by** \rightarrow *naavugwa*. ‘With’ and ‘in’ are translated as *na* and *omu* respectively, except when the verb is in the past tense, then the passive is introduced (as is the case with *neherezibwa*). This case is similar to the verbalization of ‘by’ (see also Section 2.3). The verbalization of ‘of’ is different, as the NC of the noun (NC 7 in the example above) is required to obtain the genitive *ekya*, which then drops its initial vowel

to form *kya*. Algorithm 2.2 shows the verbalization process of ‘of’. The corresponding rules have been added to the ruleset.

Algorithm 2.2 Verbalization of ‘of’)

- 1: A axiom; Variables: n_1 g_1 , nc_1 , and p_1 ; and functions: $getNoun(A)$, $getNC(n_1)$, $getPreposition(n_1)$, $getGenitive(nc_1)$, and $dropIV(g_1)$
 - 2: $n_1 \leftarrow getNoun(A)$ {obtain the main noun in the axiom}
 - 3: $p_1 \leftarrow getPreposition(n_1)$ {obtain the preposition from the XML file}
 - 4: $nc_1 \leftarrow getNC(n_1)$ {obtain the noun class}
 - 5: $g_1 \leftarrow getGenitive(nc_1)$ {use the NC to obtain the genitive}
 - 6: $g'_1 \leftarrow dropIV(g_1)$ {Drop the initial vowel of the genitive}
 - 7: Result $\leftarrow n_1$ g'_1 {Verbalize with the noun and the genitive}
 - 8: **return** Result
-

2.3 The Passive in Context-Free Grammars

In Byamugisha et al., (2016b), we used a CFG for verb conjugation. In order to cater for the passive as explained in Section 2.2, we added a new non-terminal to our original CFG that had 6 non-terminals. The passive is under the ‘extensions’ grammatical slot (Turamyomwe, 2011), represented here as non-terminal *EX*, which is placed between the verb stem *VS* and the final vowel *FV*. The new CFG is now as follows:

$$\begin{aligned}
 S &\rightarrow IG\ FM\ VS\ EX\ FV \\
 IG &\rightarrow PN\ IT \\
 PN &\rightarrow ti\ | \ ni \\
 IT &\rightarrow a\ | \ o\ | \ n\ | \ tu\ | \ mu\ | \ ba\ | \ gu\ | \ gi\ | \ ri\ | \ ga\ | \ ki\ | \\
 &\quad bi\ | \ e\ | \ zi\ | \ ru\ | \ tu\ | \ ka\ | \ bu\ | \ ku\ | \ gu\ | \ ga \\
 FM &\rightarrow \emptyset \\
 VS &\rightarrow kyendez\ | \ gw\ | \ vug\ | \ gend \\
 EX &\rightarrow w\ | \ er\ | \ erer\ | \ ir\ | \ zi\ | \ is\ | \ n\ | \ ur\ | \ uur\ | \ gur\ | \\
 &\quad VS\ | \ isPN \\
 FV &\rightarrow a\ | \ e\ | \ ire
 \end{aligned}$$

This extended CFG was also added to our ruleset.

2.4 Phonological Conditioning

Due to the agglutinative nature of Runyankore, the text resulting from our algorithms sometimes con-

tains letter combinations that do not exist in Runyankore phonology. When this happens, phonological rules are used to make the required changes that reflect the sound change in language. This is referred to as phonological conditioning, which is considered as a last step of text generation. Table 1 shows a sample of the inputs into this step, under which process it occurs, the output after phonological conditioning, and the reason why this is the case. Note that while some cases that require phonological conditioning appear similar, and would thus be assumed to have the same solution, this is actually not the case. Take the example of *baona* and *baonka*, which could be assumed to both be solved by a double vowel, leading to *boona* and *boonka* respectively. Instead, each case is assessed individually, due to the presence of the nasal compound *nk*, which therefore results in *boona* and *bonka*.

All algorithms and rules have been implemented and first verified with four ontologies: SNOMED-CT, university¹, people², and family³. We were able to generate text for all axioms which contained our selected constructors.

3 Evaluation of Generated Text

The typical method of evaluating the performance of NLG systems is to ask subjects to read and judge the generated text, as compared to human-authored text (Bouayad-Agha et al., 2012). Another form of evaluating NLG systems is to present people with a text composed of both human-authored and computer generated text, and ask them to identify which is which (de Rosis et al., 1999; Hussain et al., 2015). We used both methods in our evaluation. Similar to Hussain et al., (2015), we both rated the generated text for grammatical correctness and understandability, as well as distinguished between human-authored and computer generated text.

3.1 Materials and Methods

A questionnaire survey was used to evaluate the generated text. The questionnaire had three main sec-

¹<http://www.mindswap.org/ontologies/debugging/university.owl>

²<http://owl.man.ac.uk/2005/07/sssw/people>

³<http://www.mindswap.org/ontologies/family.owl>

Table 1: Examples of when phonological conditioning is required.

Before	Process	After	Reason
<i>abaegi</i>	Pluralization	<i>abeegi</i>	Vowel coalescence
<i>kiona</i>	Agglutination	<i>kyona</i>	Whenever <i>i</i> is followed by a vowel, it is converted to <i>y</i>
<i>baona</i>	Agglutination	<i>boona</i>	Vowel coalescence
<i>baonka</i>	Agglutination	<i>boonka</i> , then <i>bonka</i>	When a double vowel (<i>oo</i>) precedes a nasal compound (<i>-nk-</i>), vowel elision occurs
<i>niavuga</i>	Conjugation	<i>navuga</i> , then <i>naavuga</i>	First, vowel elision occurs with <i>ni</i> , then the vowel is doubled due to vowel harmony

tions: (1) age, highest level of education, occupation, and first language; (2) 10 generated sentences, which were varied based on the DL constructor being verbalized, as well as the presence of special conditions like prepositions, the passive, and definitions; and (3) 10 sentences, of which 5 human-authored and 5 computer generated. The study was conducted in Mbarara, a district in Uganda, where Runyankore is ethnically and predominantly spoken. We used purposive sampling by only selecting participants who could read, write, and speak Runyankore. We evaluated with both linguists and non-linguists. We obtained 100 non-linguists from a single village, Mirama. In order to inform our target population that we were looking for study participants, an announcement was made at the local catholic church after the Sunday service. This information was related to the headmaster of a nearby school, who agreed to let his students and staff take part in our study. All our study participants were at least 18 years old. We also contacted 3 linguists from the Department of African Languages, College of Humanities and Social Sciences, Makerere University in Uganda.

We used a modified version of the questionnaire for linguists, which had 4 sections. The first 3 were similar to the questionnaire given to non-linguists; but we added a forth section to evaluate the output from the CFG. This section had 99 conjugated verbs, testing both the standard CFG and deviations from the standard CFG, negation, several verb stems from the ontologies, and phonological conditioning.

Grammatical Correctness and/or understandability The 10 sentences were required to be graded each according to four criteria: **grammati-**

cally correct and understandable, incorrect grammar but understandable, grammatically correct but not understandable, and incorrect grammar and not understandable. Table 2 shows all the sentences in the questionnaire, as well as the DL axioms they verbalize, and the specific constructor whose verbalization we were testing for.

Sentence G originally had ‘ProfessorInHCIorAI’ and ‘AIStudent’ as concepts in the axiom. However, ‘AI’ and ‘HCI’ were replaced with ‘Science’ before the evaluation, because they were unfamiliar to the study participants, and this could have negatively affected how the sentence was graded.

Computer Generated versus Human-Authored 10 sentences, with 5 authored by a Runyankore Linguist (H) and 5 computer generated (C), were presented to study participants. They were then required to grade each sentence either as human-authored or computer generated, based on its construction. The sentences used in this part of the questionnaire are presented in Table 3.1, along with the DL axioms verbalized.

All study participants received a questionnaire containing the questions for evaluating grammatical correctness and understandability, and computer generated vs. human-authored.

3.2 Results and Analysis

99% of the 100 study participants were aged between 18 and 30 years of age. 94% were students, 5% were teachers, and 1% were retired nurse. 54% were female, and 94% had high school as their highest level of education. All study participants spoke Runyankore as their first language.

Table 2: Sentences evaluated for grammatical correctness and understandability.

	DL Axiom	Constructor Investigated	Runyankore Sentence
A	Chlordiazepoxide_hydrochloride $\sqsubseteq \exists$ Has_active_ingredient(Chlordiazepoxide)	\exists in medical domain	<i>Buri mubazi gwa hydrochloride ya Chlordiazepoxide hakiri gwine ekirungo ekyamaani kimwe kiri omubazi gwa Chlordiazepoxide.</i>
B	Giraffe $\sqsubseteq \forall$ eats.Leaf	\forall	<i>Buri ntwiga nerya amapapa goona.</i>
C	Leaf $\sqsubseteq \exists$ part_of.Leaf	\exists with preposition	<i>Buri eipapa hakiri n'ekicweka kya omuti gumwe.</i>
D	Newspaper \sqsubseteq Publication	\sqsubseteq	<i>Buri rupapura rwamakuru n'ekihandiiko ekishohozibwe.</i>
E	Person $\sqsubseteq = 2$ has.Parent	$=$	<i>Buri muntu aine abazaire 2 boonka.</i>
F	Student $\sqsubseteq \geq 1$ hasDegree	\geq	<i>Buri mwegi hakiri aine diguri emwe.</i>
G	ScienceProfessor $\sqsubseteq \forall$ advisorOf.ScienceStudent	\forall with preposition	<i>Buri purofeesa wa sayansi n'omuhabuzi wa boona abari abeege ba sayansi.</i>
H	Man $\sqsubseteq \neg$ Woman	disjointness ($\sqsubseteq \neg$)	<i>Omukazi ti mushaija.</i>
I	Old.Lady $\sqsubseteq \exists$ has_pet.Animal $\sqcap \forall$ has_pet.Cat	\sqcap	<i>Buri mukaikuru hakiri aine enyamaishwa erikutungwa abantu kuzaanisa emwe eri enyamishwa, kandi aine enyamaishwa erikutungwa abantu kuzaanisa zoona eziri enjangu.</i>
J	LecturerTaking4Courses $\sqsubseteq = 4$ takes.Course	$=$ with noun phrase	<i>Buri mushomesa orikushomesa amashomo ana natwara amashomo 4 goonka.</i>

Grammatical Correctness and/or Understandability Our preferred outcome during this evaluation was to have all sentences graded as ‘grammatically correct and understandable’ by more than 50% of study participants. The results are summarised in Figure 1. Sentences A, D, E, G, and H, which evaluated the verbalization of: medical \exists , \sqsubseteq , $=$, \forall with preposition, and \neg , were regarded as ‘grammatically correct and understandable’ by over 50% of the study participants (66%, 80%, 86%, 71%, and 92% respectively), which was a very positive outcome. Sentences B and J received the highest scores of ‘grammatically correct and understandable’ among its scores (47% and 38% respectively), but this was followed closely by ‘incorrect grammar but understandable’ (with 41% for B and 35% for J).

Sentences C, F, and I, which evaluated the verbalization of \exists -with preposition, \geq , and \sqcap , were regarded as ‘grammatically correct and understandable’ by only 8%, 34%, and 26% of the study participants, respectively. They were graded for the worst outcome (incorrect grammar and not understandable) by 18%, 13%, and 35% of study participants respectively.

The grading of sentences C, F,

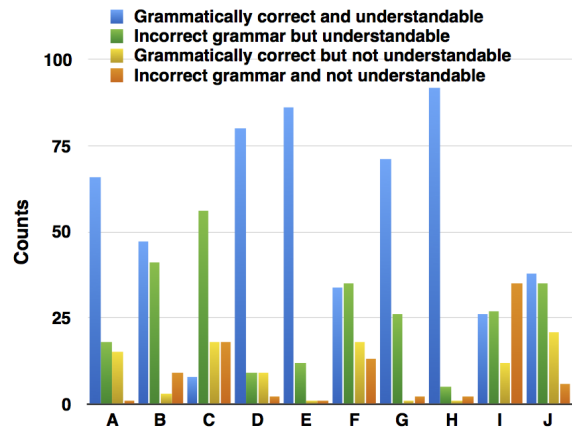


Figure 1: Evaluation results: A, D, E, G, and H performed very well (> 65%); C, F, and I performed poorly (< 35%); B and J performed marginally

and I was due to a lack of vowel assimilation, issues of syntax versus semantics, and wrong pluralization of definitions, respectively. The algorithms were updated to perform vowel assimilation and pluralize definitions accordingly.

Table 3: Sentences evaluated as either computer generated (C) or human-authored (H)

Label	DL Axiom	Runyankore Sentence
C1	Dog $\sqsubseteq \exists$ eats.Bone	<i>Buri mbwa hakiri nerya eigufa rimwe.</i>
C2	Hydrocodone \sqsubseteq Morphine_derivative	<i>Buri mubazi gwa Hydrocodone n'omubazi gwokukyendeeza obusaasi.</i>
H1	Lecturer $\sqsubseteq \neg$ Professor	<i>Omushomesa wa yunivasite ti purofeesa.</i>
C3	Sheep $\sqsubseteq \forall$ eats.Grass	<i>Buri ntaama nerya ebinyaansi byoona.</i>
H2	Student $\sqsubseteq \leq 7$ studies.Course	<i>Buri mwegi naashoma amashomo gatarikurenga 7.</i>
C4	TeachingFaculty $\sqsubseteq \leq 3$ takes.Course	<i>Abashomesa omu kitongore boona nibatwara amashomo gatarikurenga 3.</i>
H3	Hydrocodone \sqsubseteq Opiate	<i>Buri mubazi gwa Hydrocodone gurimu ebirungo ebirikukyendeeza obusaasi.</i>
H4	Parent $\sqsubseteq \exists$ hasChild	<i>Buri muzaire nomuntu oine haakiri omwana omwe.</i>
C5	Cat $\sqsubseteq \neg$ Dog	<i>Enjangu ti mbwa.</i>
H5	Van \sqsubseteq Vehicle	<i>Buri vaani nemotoka.</i>

Computer Generated Versus Human-Authored

The desired outcome for this part of the evaluation was to have all computer generated sentences graded as human-authored by at least 66% of study participants. Hussain et al., (2015) evaluated for the same, but with 3 professionals, and their best result was that 64% of the overall text were regarded as human-authored. Our results are summarised in Figure 2. All computer generated sentences (C) except for C4 (with 64%) were above this bar. C1, C2, C3, and C5 were regarded as human-authored by 78%, 71%, 90%, and 97% of study participants respectively.

On the other hand, some human-authored text performed under the desired threshold. H2 just made it with 66%, while H3 and H5 were below with 64% and 56% of study participants respectively, regarding them as human-authored. They actually performed worse than computer generated text.

The implication that most study participants (> 60%) regarded all generated text as having been written by a human being is positive outcome.

Results from Linguist We have so far received feedback from one of the 3 linguists we contacted, and we present that feedback here. From the same sentences in Table 2 (except I) for evaluating grammatical correctness and/or understandability, A, D, E, and G were graded as ‘incorrect grammar but understandable’, due to: the issue of translating medical terminologies, the nature of the text from the axiom, a lack of proper phonological conditioning,

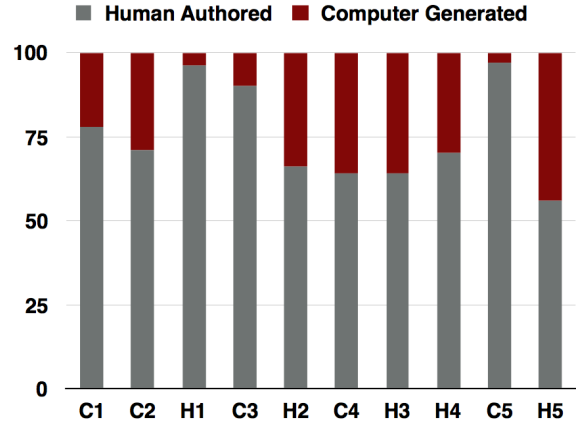


Figure 2: Computer vs Human: C1, C2, C3, C5 considered human authored by more than 66%; H2, H3, H5 performed worse than C

and the incorrect arrangement of the prepositional phrase respectively. C was graded as both incorrect and not understandable because of an error in the algorithm. This led to a modification of the grammar rules and algorithms. After the poor performance of sentence I in the non-linguists’ questionnaire, arising from the translation of ‘pet’, we used a different axiom for the linguists, in order to focus their evaluation on the verbalization of conjunction. We used: Old_Lady $\sqsubseteq \exists$ reads.Publication $\sqcap \forall$ reads.Tabloid. This was verbalized as: *Buri mukaikuru hakiri nashoma ekihandiiko ekishohozibwe kimwe, kandi naashoma taburoyidi zoona.* It was regarded as ‘grammatically correct but not un-

derstandable’ because ‘tabloid’ was naturalized as *taburoyidi*. This is however the conventional way of translating loan words, and its negative effect on the grading of the sentence was unexpected. Sentences B, F, H, and J (except for the issue with the nasal compound) were graded as ‘grammatically correct and understandable’. Except for H, the rest differ from the evaluation by non-linguists, where F was regarded as ‘incorrect grammar but understandable’ by 35%, 1% more than those who regarded it as ‘grammatically correct and understandable’. Sentences B and J were only marginally regarded as being grammatically correct and understandable. It will be interesting to see whether all linguists will have a similar evaluation. From the same sentences in Table 3.1, C1, C2, H2, C4, and H3 were graded as human-authored; while H1, C3, H4, C5, and H5 were graded as computer generated. The reasons for differences in assessing this between linguists and non-linguists are unclear, as neither group explained their choice. It is however encouraging that more computer generated text was considered human-authored by a linguist.

When evaluating the performance of the CFG, 11 of the 99 conjugated verbs presented in the questionnaire were considered as incorrect: 1 due to a wrong subject prefix, 2 due to the need for an extra suffix, and 8 because vowel harmony was not implemented. The subject prefix and vowel harmony errors have since been fixed. The need for an extra suffix was only identified for the verb stem for ‘eat’ (*ry*). We are still investigating this error.

4 Discussion

Our work here adds to the growing efforts to verbalize ontologies in multiple languages. It has also provided a basis to consider that the underlying theories could be generalizable to other Bantu languages. In our previous work (Byamugisha et al., 2016a; Byamugisha et al., 2016c), we showed that the factors affecting verbalization in isiZulu and Runyankore are the same, and both Runyankore and isiZulu had similar exceptions to pluralizing nouns according to the standard NC table. Further, the passive as a grammatical slot is present in the verbal morphologies of both languages.

Perhaps most interestingly for our domain of in-

terest, is that the use of a small sample of SNOMED-CT, a very large healthcare ontology, during testing, helped to investigate how to translate medical jargon to Runyankore. In some cases, the term is maintained, which is the case for common terms like ‘Panado’; in others, it is given context, e.g., hydrocodone translated as *omubazi gwa hydrocodone* ‘medicine of hydrocodone’; or a mixed translation and context approach, e.g., *endwara ya shukari* ‘disease of sugar’, where ‘sugar’ is a common translation for ‘diabetes’ in a healthcare context; while in extreme cases, the term is defined, e.g., opiate translated as *omubazi ogukusinza ogukwejunisibwa kukyendeza obusaasi* ‘medicine which intoxicates as treatment to reduce pain’. There are several terms for anatomy, diseases, and drugs which are not directly translatable to Runyankore, but this offers a starting point to show alternative ways to handle it.

Further, the results from the evaluation of the generated text are very encouraging, both from a linguist and non-linguists. Additionally, important feedback was obtained, which enabled us to make modifications to the initial rules and algorithms.

5 Conclusion

New algorithms for knowledge-to-text verbalisation into Runyankore have been developed. These include: i) rules and algorithms for cardinality constraints (maximum, minimum, and exact), therewith extending the language coverage, ii) handling some prepositions in complex roles; iii) processing the passive with a CFG, and iv) phonological conditioning. We have evaluated the text generated by the algorithms and rules with a group from the general population on grammatical correctness, understandability, and whether they were computer generated and human-authored text. The data demonstrated that most sentences were evaluated as grammatically correct and understandable, and human authored. We are currently investigating the architecture required to implement this as a NLG system.

Acknowledgements This work is based on the research supported by the Hasso Plattner Institute (HPI) Research School in CS4A at UCT and the National Research Foundation of South Africa (Grant Number 93397).

References

- Ion Androutsopoulos, Gerasimos Lampouras, and Dimitrios Galanis. 2013. Generating natural language descriptions from owl ontologies: The naturalowl system. *Journal of Artificial Intelligence Research*, 48:671–715.
- Allen Asiimwe. 2014. *Definiteness and Specificity in Runyankore-Rukiga*. Ph.D. thesis, Stellenbosch University, Cape Town, South Africa.
- Nadjet Bouayad-Agha, Gerard Casamayor, and Leo Wanner. 2012. Natural language generation in the context of the semantic web. *Semantic Web Journal*.
- Joan Byamugisha, C. Maria Keet, and Brian DeRenzi. 2016a. Bootstrapping a runyankore cnl from an isizulu cnl. In *CNL 2016*, Aberdeen, Scotland. Springer LNCS.
- Joan Byamugisha, C. Maria Keet, and Brian DeRenzi. 2016b. Tense and aspect in runyankore using a context-free grammar. In *INLG 2016*, Edinburgh, Scotland.
- Joan Byamugisha, C. Maria Keet, and Langa Khumalo. 2016c. Pluralizing nouns in isizulu and related languages. In *CICLing 2016*, Konya, Turkey.
- J. Alison Cawsey, B. Ray Jones, and Janne Pearson. 2000. The evaluation of a personalized health information system for patients with cancer. *User Modeling and User-Adapted Interaction*, 10(1):47–72.
- Fiorella de Rosis, Floriana Grasso, and C. Dianne Berry. 1999. Refining instructional text generation after evaluation. *Artificial Intelligence in Medicine*, 17(1):1–36.
- Chrysanne DiMarco, Graeme Hirst, Leo Wanner, and John Wilkinson. 1995. Healthdoc: Customizing patient information and health education by medical condition and personal characteristics. In *Workshop on Artificial Intelligence in Patient Education*, Glasgow, Scotland.
- Chrysanne DiMarco, Peter Bray, Dominic Covey, Don Cowan, Vic DiCiccio, Eduard Hovy, Joan Lipa, and Cathy Yang. 2005. Authoring and generation of tailored preoperative patient education materials. In *Workshop on Personalization in e-Health User Modeling, Conference*, Edinburgh, Scotland.
- Chrysanne DiMarco, David Wiljer, and Eduard Hovy. 2009. Self-managed access to personalized healthcare through automated generation of tailored health educational materials from electronic health records. In *AAAI Fall Symposium on Virtual Health Interaction*, Washington D. C.
- Normunds Gruzitis and Guntis Barzdins. 2011. Towards a more natural multilingual controlled language interface to owl. In *IWCS*, pages 335–339.
- Normunds Gruzitis, Gunta Nespore, and Baiba Saulite. 2010. Verbalizing ontologies in controlled baltic languages. In *4th International Conference on Human Language Technologies-The Baltic Perspective*, Riga, Latvia.
- Sazzad Mohammed Hussain, A. Rafael Calvo, Louise Ellis, Juchen Li, Laura Ospina-Pinillos, Tracey Davenport, and Ian Hickie. 2015. Nlg-based moderator response generator to support mental health. In *CHI’2015*, Seoul, South Korea. ACM.
- Kaarel Kaljurand and E. Norbert Fuchs. 2007. Verbalizing owl in attempto controlled english. In *Third International Workshop on OWL: Experiences and Directions*, Innsbruck, Austria.
- C. Maria Keet and Langa Khumalo. 2017. Towards a knowledge-to-text controlled natural language of isizulu. *Language Resources and Evaluation*, 51:131–157.
- Fredrik Lindahl. 2005. Practical text generation in clinical medicine. Chalmers Computer Science Department Winter Meeting.
- Saad Mahamood and Ehud Reiter. 2011. Generating affective natural language for parents of neonatal infants. In *ENLG 2011*, pages 12–21, Stroudsburg, Pa, USA. ACM.
- Boris Motik, Peter F. Patel-Schneider, and Bijan Parsia. 2009. OWL 2 web ontology language structural specification and functional-style syntax. W3c recommendation, W3C, 27 Oct. <http://www.w3.org/TR/owl2-syntax/>.
- Doreen Daphine Tayebwa. 2014. Demonstrative determiners in runyankore-rukiga. Master’s thesis, Norwegian University of Science and Technology, Norway.
- Justus Turamyomwe. 2011. Tense and aspect in runyankore-rukiga: Linguistic resources and analysis. Master’s thesis, Norwegian University of Science and Technology, Norway.