OWLSIZ: An isiZulu CNL for structured knowledge validation

Zola Mahlaza
Department of Computer Science
University of Cape Town
South Africa
zmahlaza@cs.uct.ac.za

C. Maria Keet
Department of Computer Science
University of Cape Town
South Africa
mkeet@cs.uct.ac.za

Abstract

In iterative knowledge elicitation, engineers are expected to be directly involved in validating the already captured knowledge and obtaining new knowledge increments, thus making the process time consuming. Languages such as English have controlled natural languages than can be repurposed to generate natural language questions from an ontology in order to allow a domain expert to independently validate the contents of an ontology without understanding a ontology authoring language such as OWL. IsiZulu, South Africa's main L1 language by number speakers, does not have such a resource, hence, it is not possible to build a verbaliser to generate such questions. Therefore, we propose an isiZulu controlled natural language, called OWL Simplified isiZulu (OWLSIZ), for producing grammatical and fluent questions from an ontology. Human evaluation of the generated questions showed that participants' judgements agree that most (83%) questions are positive for grammaticality or understandability.

1 Introduction

Ontology developers often rely on domain experts when building models or ontologies. This knowledge elicitation is often done in an iterative manner to ensure a high quality artefact. That is, the first interaction between knowledge engineers and domain experts generally focuses on only obtaining the first usable knowledge increment. Further interactions focus on obtaining further increments and also validating the previous iteration's codified knowledge. This requires the engineer to have a method of presenting the codified knowledge in an accessible manner. One method to resolve this, is to render the ontology with a controlled natural language (CNL) and provide supporting tooling for that; see (Safwat and Davis, 2017) for a recent overview. The manner in which these tools are to be used throughout the knowledge acquisition process is left to the discretion of the ontology developer, as they have not been incorporated into existing structured methods of knowledge validation. This may lead to inefficiencies, especially for inexperienced ontologists, as it requires one to be directly involved in obtaining the additional knowledge increment and also validating the already captured knowledge. For a language like English, one could repurpose a CNL such as OWL Simplified English (Power, 2012) to generate questions that domain experts could answer without involving an ontologist. This is more challenging for other languages, since fewer resources exist for them.

For our context in South Africa, isiZulu is relevant, since it is the main language by first language speakers and serves as one of the main communication languages alongside English, Afrikaans, and Setswana (out of 11 official languages and sign language). It is not possible to generate questions from ontologies in isiZulu, since such a CNL does not exist and only some statements can be generated with the CNL of (Keet and Khumalo, 2017).

We propose a solution in the form of OWL Simplified isiZulu (OWLSIZ), an isiZulu CNL for authoring questions to be used in structural questionnaires and/or Likert items to reduce the novice ontology engineer's efforts in validating a model or ontology's contents. To support its use, we developed an OWL verbaliser (in Java) that uses the CNL to generate questions. Evaluation of the quality of the generated questions shows that participants agree that most (83%) of the texts are positive since they have at most one participant who considers them to be ungrammatical and unacceptable.

The rest of the paper is structured as follows. Section 2 outlines various methods used for knowledge elicitation, Section 3 discusses existing work on generating controlled natural languages from

3rd International Workshop on Natural Language Generation from the Semantic Web (WebNLG+),
Dublin, Ireland (Virtual), 18 December 2020, pages 15–25,

ontologies, Section 4 presents our CNL, Section 5 presents the design of the verbaliser, Section 6 presents the quality of our CNL, Section 7 the discussion of results, and Section 8 concludes.

2 Knowledge elicitation and validation

Ontology engineering methodologies (OEMs) can be categorised into three groups: collaborative, non-collaborative, and custom (Kotis et al., 2020). Of interest to this work are collaborative OEMs, which have well-defined phases and involve numerous stakeholders across the different phases. The methodologies do not restrict one to specific kinds of elicitation methods as evidenced by the nine collaborative OEMs considered by (Kotis et al., 2020). Of course, many semi-automatic techniques have been proposed over the years, notably NLP-based ontology learning techniques, but the focus here is the stage where there is a domain expert in the loop, who is typically not well-versed in logic.

Some 25 years ago already, strategies like observations, interviews, or task analysis based methods were already proposed (Cooke, 1994). The chosen method impacts the overall time it takes to acquire the necessary domain knowledge, especially since most OEMs (7 out of 9) are fully iterative and the rest are either partially iterative or agile. Structured interview techniques such as the use of questionnaires, twenty-questions, and Likert scale items (Cooke, 1994) are less time consuming because they are guided. They can still be improved through the automation of question or Likert item creation. To the best of our knowledge, while there are verbalisers that can be used to create Likert items, there are currently no methods for creating questions to be used within structured techniques.

For instance, if an axiom $Human \sqsubseteq \exists hasPart.Heart$ had been added in an ontology authoring iteration, then in its validation stage, a structured questionnaire could link that to a CNL containing the template "Does each {C1} {OP1} some {C2}?" to generate the yes/no validation question 'Does each human have as part some heart?'. The domain expert would be validating the knowledge added to the ontology by answering 'yes' and indicating a mistake by answering 'no'.

3 Text generation from ontologies

There are numerous tools that take ontologies or similar models as input and produce text. For the present purpose, we will categorise them as either educational question generators and model/ontology verbalisers. Tools for the former take an ontology and generate questions to be presented to people to test their knowledge. Tools for the latter take an ontology's axioms or a model's contents and convert them into natural language text to make the input accessible to stakeholders who are not familiar with modelling languages.

The existing educational question generators are built for English only (Chaudhri et al., 2014; Alsubait et al., 2016; Papasalouros et al., 2008; E.V. and Kumar P., 2015; Zhang and VanLehn, 2016). They either use bare templates or SimpleNLG (Gatt and Reiter, 2009) for realisation, where 'bare' templates are a sequence of fixed words and slot, such as, Model-T's template (Puzikov and Gurevych, 2018). None of the two realisation components are usable for isiZulu, because of the language's grammatical complexity or its lack of reusable and comprehensive computational grammar rules. Contrarily, the verbalisers offer components that can be used to build an isiZulu CNL for question generation even though none of the existing verbalisers were designed for such a purpose. This is because there is a lot of variability concerning the realisation methods and languages supported by ontology verbalisers, as can be seen in Table 1.

While early verbalisers (e.g., (Wilcock, 2003)) supported only English, a number of systems have since been built to support other languages; see Table 1 for an overview. About half of these other verbalisers (52%) rely on "grammar-infused templates" (Mahlaza and Keet, 2019) for surface realisation, i.e., plain templates with some grammar rules to improve some aspects of the generated sentence. Pertinent to our considered language, isiZulu, four verbalisers (Lim and Halpin, 2016; Demey and Heath, 2014; Keet and Khumalo, 2017; Byamugisha et al., 2016) rely on what they call "patterns", which might be argued to be a form of template created to capture some linguistic dependency between items.Such patterns, and unlike a regular template, come with linguistic rules that control the value of the concords that are based on the noun class of the inserted word. An example of a pattern, for verbalising 'has part' in isiZulu, is as follows (Keet and Khumalo, 2016):

QCall$_{nc_x,pl}$ W$_{nc_x,pl}$ SC$_{nc_x,pl}$-CONJ-P$_{nc_y}$ RC$_{nc_y}$-QC$_{nc_y}$-*dwa*

For instance, the "QCall$_{nc_x,pl}$" is for the universal quantification (∀, 'for all'), which is generated

Table 1: List of verbalisers with relevant core features. Abbreviations: OWL = Web Ontology Language, ORM = Object-relational mapping, FBM = Fact Based Modeling and the template classification (classif.) abbreviations from Mahlaza and Keet (2019) are P = partial attachment, C = compulsory attachment, CE = compulsory attachment and embedding, and EP = embedding and partial attachment

Reference	**Input**	**Realisation method**	**Template classif.**	**Language(s)**
Dannélls et al. (2013b)	OWL	Grammatical framework	×	English, Bulgarian, Catalan, Danish, Dutch, Finnish, French, Hebrew, Italian, German, Norwegian, Romanian, Russian, Spanish, and Swedish
Camilleri et al. (2012)	OWL	Grammatical framework	×	English, Catalan, Dutch, Finnish, French, German, Italian, Spanish, Swedish, and Urdu
Sadoun et al. (2016)	OWL	Canned text	×	Arabic, English, French, Hindi, Japanese, Mandarin Chinese, Russian, Ukrainian and Tibetan
Jarrar et al. (2006)	ORM	Template	×	English, Dutch, German, Italian, Spanish, Catalan, French, Lithuanian, Russian, and Arabic
Lim and Halpin (2016)	ORM	"Pattern"	P	English, Malay, and Mandarin
Dannélls et al. (2013a)	OWL	Grammatical framework	C	English, French, Italian, Finnish and Swedish
Dannélls (2012)	OWL	Grammatical framework	C	English, Swedish, and Hebrew
Androutsopoulos et al. (2013)	OWL	"Sentence plan"	CE	English and Greek
Davis et al. (2012)	?	Grammatical framework	C	English and Dutch
Bouayad-Agha et al. (2012)	OWL	Meaning-text theory	×	English and Finnish
Liang et al. (2011)	OWL	Template	×	English and Mandarin
Gruūzitis (2011)	OWL	Grammatical Framework	CE	English and Latvian
Dannélls (2010)	OWL	Grammatical Framework	C	English, Swedish, and Hebrew
Dannélls (2008)	OWL	Grammatical Framework	CE	English and Swedish
Demey and Heath (2014)	FBM	"Pattern"	E	English and Chinese
Keet and Khumalo (2017)	OWL	"Pattern"	EP	IsiZulu
Sanby et al. (2016)	OWL	Template	×	Afrikaans
Byamugisha et al. (2016)	OWL	"Pattern"	EP	Runyankore
Aguado et al. (1998)	-	KPML	×	Spanish
Halpin and Curland (2006)	ORM	Template	×	English
Bouayad-Agha et al. (2011)	OWL	Template	×	Spanish

based on what the plural of the noun class is of the noun that denotes the entity that plays the whole in the has-part relation (the "W$_{nc_x,pl}$"), combining the appropriate quantitative concord (e.g., *ba-* for noun class 2) with *-onke* to generate the appropriate surface realisation (e.g., *bonke*, for nouns in noun class 2). An axiom with ∀Isibhedlela ('hospital', in noun class 7) then verbalises as *zonke izibhedlela* ('all hospitals') after applying the appropriate algorithms associated with the pattern. Compare this with straightforward templates for English for the same scenario:

Each [Whole] has as part at least one [Part].
All [Whole]$_{pl}$ have as part at least one [Part].

where the text remains the same and the ontology vocabulary can simply be plugged into the variable slots, or, at most, generate a plural version of a variable in the singular.

Demey and Heath (2014)'s patterns are syntax templates, which are not suitable for isiZulu due to lack of existing computational syntax rules. The other patterns are potentially suitable because they are created to introduce rules for noun-related dependencies, which are essential to all the languages in the Niger-Congo B family. For instance, Lim and Halpin (2016) use them to encode Malay and Mandarin noun classifiers and Keet and Khumalo (2017); Byamugisha et al. (2016) use theirs to capture agreement between nouns and a number of parts-of-speech in isiZulu and Runyankore.

These two kinds of noun-centered patterns were not used to generate questions from OWL in isiZulu, but statements, rather. Moreover, they are either limited capturing a single POS that is controlled by noun (Lim and Halpin, 2016) or tightly couple the linearization algorithm with the template (Keet and Khumalo, 2017; Byamugisha et al., 2016). In this work, in contrast, we design a CNL that is able to generate questions, whose grammar-infused templates are not tightly coupled with the linearisation algorithm, and are able to capture dependencies between the noun and numerous other parts-of-speech.

4 OWL Simplified isiZulu

The approach used to design the CNL to generate validation questions in isiZulu from ontologies, is incremental and bottom-up. We begin by selecting the first set of OWL constructors to support by selecting ones already supported by an existing English CNL and use its templates as inspiration for our isiZulu templates. The chosen CNL for that is OWL Simplified English (Power, 2012) and the resulting OWLSIZ templates are provided in Table 2, together with an approximate English translation for indicative purpose.

The isiZulu templates use several concords[1], copulas, and locative affixes. They are polymorphic affixes whose values change depending on the noun in which they are found in the case of the locative prefix and copula or due to the noun class of another noun in the sentence in the case of concords. In Table 2, we use a box around a sequence of affixes to illustrate a decomposed word whose underlying morphemes are provided in a sequential manner. We also use an arrow from a concord to the noun that controls its value. For instance, in template 10, the third word is made up of two affixes: the subject concord (SC) and *-odwa*. The subject concord's value depends on the noun class of the value inserted into the class slot (i.e., {C}) and there are 17 such classes in isiZulu. When the subject concord's value is inserted, then phonological conditioning may be applied when combining the two affixes, since isiZulu does not permit consecutive vowels. For instance, when noun class 15's *ku-* is appended to *-odwa*, we obtain *kodwa*.

To illustrate the use of the isiZulu templates, let us consider how to verbalise as a question an axiom of the type $A \sqsubseteq B$, e.g., $ihebhu \sqsubseteq umuthi$ and in OWL functional syntax style, SubClassOf(ihebhu umuthi). Template 1 would be chosen, since the B in the position of C2 is a named class. One would then obtain the following text

(i) *Ingabe **lo**nke ihebhu **ling**umuthi?*

Is SC-every herb SC-COP-plant?

'Is every herb a plant?'

In this example, the concords and copula are indicated in bold: the concords are the *lo-* (from *li-* + *o-*) and *li-* components of the highlighted text, which are governed by the noun class of the first noun, *ihebhu* 'herb', and the copula is the *-ng-* component, which is determined by the first character of the second noun, *u-*.

[1] consult Meeussen (1967) for a detailed categorisation of concords

5 Verbaliser implementation

We designed the verbaliser for OWLSIZ in the way shown with the architecture diagram in Figure 1. It was implemented in Java and it uses the OWL API[2] to parse the ontology. Unlike the predominant realisation method as listed in Table 1, GF, we have chosen to rely on Java, because that enables us to have a verbaliser that can readily generate text from any ontology (provided it has isiZulu labels). GF-based verbalisers, on the other hand, require two additional time-consuming steps, being converting OWL ontologies into GF (e.g., (Angelov and Enache, 2010)), and the high start-up costs in developing a resource grammar that faces the usual difficulty for under-resourced languages in that there is scant documentation of the grammar.

We will discuss the planning and realisation components in the remainder of this section.

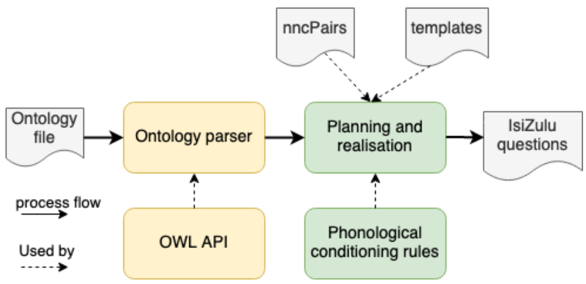

Figure 1: Verbaliser architecture

There are minimal rules for selecting a template for each supported axiom type. Only one logical axiom (SubClassOf) and one class expression (ObjectSomeValuesFrom) have multiple template forms. Template 1 verbalises SubClassOf axioms where the range is an OWL class and Template 1.1 handles cases where the range is a class expressions. A single rule is used by the verbaliser to choose between the two forms of templates.

There are four different templates for the ObjectSomeValuesFrom class expressions (number 6).

[2] https://github.com/owlcs/owlapi

Table 2: List of OWL Simplified IsiZulu templates. The arrows indicate dependencies, pointing from dependent to the determiner, and a box around elements means that those elements inside it will result in one word in the sentence generated. Abbreviations: I = individual, C = class, DP = data property, OP = Object property, CE = Class expression, OP_noun = Object property that is noun, DP = Data property, L = Literal, N = Cardinality, *SC* = subject concord, *OC* = object concord, *COP* = copula prefix, *RelC* = relative concord, *LocPre* = locative prefix.

Axiom type (in OWL functional syntax style)	Approximate English template translation	IsiZulu template
1. SubClassOf(C1 C2)	Is every {C1} a(n) {C2}?	Ingabe {*SC*} onke {C1} {*SC*} {*COP*} {C2}?
1.1. SubClassOf(C1 CE)	Does/Is every {C1} {CE}?	Ingabe {*SC*} onke {C1} {*SC*} {CE}?
2. ClassAssertion(C I)	Is {I} a(n) {C}?	Ingabe {I} {*SC*} {*COP*} {C}?
3. ObjectPropertyAssertion(OP I1 I2)	{I1} {OP} {I2}?	{I1} {*SC*} {OP} {I2} ?
4. EquivalentClasses(C1 C2)	Is every {C2} a(n) {C1}?	Ingabe noma yi {*OC*} phi {C2} {*SC*} {*COP*} {C1}?
5. DisjointClasses(C1 C2)	Is there no {C1} that is a(n) {C2}?	A {*SC*} kho yini {C1} {*RelC*} {*COP*} {C2}?
6. ObjectSomeValuesFrom(OP C)	{OP} a(n) {C}?	{OP} {C}?
6.1. ObjectSomeValuesFrom(OP CE)	{OP} a {CE}?	{OP} {CE}?
6.2. ObjectSomeValuesFrom(OP_{noun} C)	{OP_{noun}}s a(n) {C}?	yi {OP} {*LocPre*} {C} ini?
6.3. ObjectSomeValuesFrom	—	{OP} {*LocPre*} {C} ini?
6.4. ObjectSomeValuesFrom(OP_{noun} C)	{OP_{noun}}s a(n) {C}?	yi {OP} {*LocPre*} {C}?
7. ObjectHasValue(OP I)	{OP} {I}?	{OP} {I}?
8. DataPropertyAssertion(DP I L)	Does/Will/Did/Is {I} {DP} {L}?	Ingabe {I} {DP} {L}?
9. DataHasValue(DP L)	{DP} {L}?	{DP} {L}?
10. ObjectAllValuesFrom(OP C)	{OP} only {C}?	{OP} {C} {*SC*} odwa
11. ObjectExactCardinality(N OP C)	{OP} exactly {N} {C}?	{OP} {C} {*RelC*} yi- {N} ncamashi/ngqo ?
12. ObjectMinCardinality(N OP C)	{OP} at least {N} {C}?	{OP} {C} {*RelC*} ngaphezulu kuka-{N} ?
13. ObjectMaxCardinality(N OP C)	{OP} at most {N} {C}?	{OP} {C} {*RelC*} mbalwa ku-{N} ?

Template 6.4 is applicable when the object property is a noun and the OWL class belongs to noun classes 1, 1a, 2, or 2a, whereas template 6.2 is used when that OWL class does *not* belong to classes 1a, 2, or 2a. Template 6.3 is applicable when the object property is the containment part-whole relation (Keet and Khumalo, 2016). Template 6 and 6.1 are applicable when the range is an OWL class and class expression respectively. The verbaliser has a planning rule that chooses an appropriate template when given a class expression.

The nncPairs of Figure 1 is a file `nncPairs.txt` that is used for determining whether an object property is or contains a noun and to retrieve the noun class of the noun, which is reused from Keet et al.'s isiZulu verbaliser (Keet et al., 2017). We extended the file by adding 8 words to add more variation (*umakhalekhukhwini*, *uZola*, *iNokia 3310*, *ifoni*, *umfundi*, *ukubhukuda*, *ihebhu*, and *umdlalo*) and changed the noun class annotations of 6 nouns (*ufulawa* 3a, *amanzi* 6, *irhaba* 9, *ivazi*-9, *ithiyetha yokuhlinzela* 9, *indoda* 5) due to two typos, mass noun marker deletion, and differences of opinion on whether noun class 9a is indeed separate from noun class 9.

The realisation module takes a template and slot fillers, inserts the fillers into their respective slots, resolves the values of concords, locatives, and copulas in polymorphic words, and then forms words by appending the affixes together while relying on isiZulu phonological conditioning rules taken from (Naidoo, 2002; Sibanda, 2007; Pretorius and Bosch, 2010; Van der Spuy, 2014; Posthumus, 2016).

6 Evaluation

In order to evaluate the quality of questions generated by the verbaliser, we make use of the test ontology presented in (Keet et al., 2017). We extended the ontology's 82 logical axioms with 12 axioms to ensure coverage for OWLSIZ: ClassAssertion (2), ObjectProperty (2), ObjectPropertyAssertion (1), EquivalentClasses (1), DataPropertyAssertion (1), DataProperty (1), ObjectAllValuesFrom (1), ObjectExactCardinality (1), ObjectMinCardinality (1), and ObjectMaxCardinality (1). This update resulted in an ontology with 91 axioms (it is not 94 because 3 axioms were used to constrain the existing axioms).

For internal evaluation, the ontology was verbalised and we categorized the resulting texts into *verbalisable* and *unverbalisable*. These two classes capture our verbaliser's (un)supported axioms. For each of the *verbalisable* axioms, we analysed their corresponding text to determine whether there are phonological conditioning errors, morphological agreement errors, and any other grammatical error.

For external evaluation, the ontology was verbalised and we packaged the resulting questions into a survey. Participants were recruited via snowball sampling using the first author's Twitter and WhatsApp accounts. The participants were asked to judge the quality of each question by choosing either "grammatical and acceptable", "grammatical and ambiguous", "ungrammatical and understandable", or "ungrammatical and unacceptable". In order to obtain high quality judgements, we ensured that each participants did not judge more than 40 sentences by randomly diving the 76 texts into two surveys each containing 38 texts (henceforth, survey A and B). Survey A includes texts generated by templates 3, 10, 12, and 13 while survey B does not and survey B includes texts generated by templates 4, 8, and 11 while survey A does not. Participants were randomly assigned to a survey.

The verbaliser's input, chosen templates, and numbered output texts are given as supplementary material at `https://github.com/AdeebNqo/grammarinfusedtemplates`.

7 Results and discussion

Out of the total 91 axioms, 76 were verbalisable axioms and of their corresponding texts, 74 texts are free of morphological agreement errors, phonological conditioning errors, or any other grammatical errors as assessed by the first author, an L2 isiZulu speaker with experience with various aspects of isiZulu grammar (e.g., (Mahlaza, 2018)). Template 5 was selected when verbalising DisjointClasses(isidlanyama isidlazitshalo), and the system generated the following output:

(i) *asikho yini isidlanyama esiyisidlazitshalo?*
NEG-SC-exist carnivore$_{[NC7]}$ RelC-COP-herbivore$_{[NC7]}$?
'Is there no carnivore that is a herbivore?'

In the output, *isidlazitshalo* is prefixed with *esi-* and *-y-* where *esi-* is a relative concord that would have a different value if *isidlanyama* was not used (e.g, it would be *eli-* if a noun that belongs to class 5 was used instead). Similarly, *-y-* is the copulative prefix value; if the value inserted into {C2} had *u*, *o*, or *a* as preceding vowel, then *-ng-* would be used instead.

One out of the two questions determined to have errors had an morphological agreement error (question 42 in the supplementary material). The second was generated by Template 8 and, when verbalising DataPropertyAssertion(neminyaka uZola 50), the system generated the following output:

(ii) *Ingabe uZola neminyaka 50?*
Is Zola$_{[NC1a]}$ CONJ-years 50?
'Is Zola aged 50?'

This is lacking agreement markers; hence it is grammatically incorrect: the correct one should have been *Ingabe uZola [SC]neminyaka [RelC]ngu-50?* where the subject concord (SC) depends on the individual (*uZola*) and the relative concord (RelC) depends on the noun found in the object property (i.e., *iminyaka*). We cannot correct template 8 by introducing the subject and relative concords as done for this particular example because different categories of data properties may require different solutions.

The 15 that were classified as unverbalisable had several causes. 10 had nouns whose classes were unresolved; e.g., one where the compound noun's noun class was unresolved (*isampula igazi* ought to have been *isampula egazi* 'blood sample' in noun class 5). The other five included unsupported axiom types, such as ObjectComplementOf.

Six participants filled in one of the surveys, of

Table 3: Number of participants' judgements. Abbreviations: Gramm. + ambig. = grammatical and ambiguous, Gramm. + accept. = grammatical and acceptable, Ungramm. + understand. = ungrammatical and understandable, Ungramm. + unaccept. = ungrammatical and unacceptable, and Pct. = percent

Survey	Gramm. + ambig.	Gramm. + accept.	Ungramm. + understand.	Ungramm. + unaccept.
A	17	41	6	12
B	23	78	19	32
A+B	40	119	25	44
A+B Pct.	18%	52%	11%	19%

which five L1 isiZulu speakers and one L2. Survey A had two participants (one L1 and one L2 isiZulu speakers) who made a total of 76 judgements. Survey B had four participants and they made a total of 152 judgements (38 per participant). The separation of the participants' judgements into their respective categories is listed in Table 3. A majority of the judgements were that the texts are grammatical and acceptable; there were 41/76 in Survey A and 78/152 in Survey B. Overall, i.e., combining Survey A and B, most of the judgements were positive, since only 19% of the judgements were ungrammatical + unacceptable.

The participants' judgements of each question are given in Figure 2. Observe that there are only two questions (labelled 25 and 42 in the figure) for which participants agree that they are ungrammatical and unacceptable, which are:

25 : *iNokia 3310 lifundisa uZola?*
'The Nokia 3310 teaches Zola'
42 : *Ingabe noma yiyiphi indlu eyinyama?*
'Is every house (the same as) meat?

These two questions were generated by template 3 and 4, respectively. This does not mean that they are of low quality, for the following reasons. Question 25 is ungrammatical and unacceptable due to the presence of the word *eyinyama* as opposed to *iyinyama* in that specific context. Analysis of the reason why an 'e' was used as opposed to 'i' shows that the serialised template used by Java verbaliser is slightly different from the one listed in Table 2 as we mistakenly used the relative concord in place of the subject concord. Question 42 is likely judged ungrammatical and unacceptable by participants due to unfamiliarity of the noun, as they may be unsure to which noun class 'Nokia 3310' belongs: by default, foreign objects are allocated to noun class 5 (hence *i*Nokia), but 'mobile phone', that the Nokia is, is *umakhalekhukhwini*, which is in noun class 3.

Figure 2 also shows that 83% of the texts are judged positively as they have at most one participant who considers them to be ungrammatical and unacceptable. Furthermore, since Survey A only had two participants, even if we adjust how we determine the number texts judged positively and define them as the texts where is no participant who considers the text to be ungrammatical and unacceptable, we still find that most (71%) of Survey A's questions were judged positively by the participants.

Participants agreed in their judgements (i.e., they chose the same options out of the ones listed in Section 6) for 25/76 texts (more specifically, 16/38 for Survey A and 9/38 for survey B). Furthermore, out of the 25, they agreed positively to 24 of the texts and agreed that remaining question is ungrammatical and unacceptable. While in (Keet and Khumalo, 2014) it seemed that there was more disagreement in the human evaluation the longer the sentences were, here, the participants' disagreement is not due to differences in text length. This since the texts for which the participants agree and disagree are of similar length (agree = average of 4 words and disagree = 5 average words), save for a single outlier with 7 words in the texts for which they disagree. The disagreement may be due to a misunderstanding that the participants have regarding how to evaluate the texts. For instance, when given text that reflects an unacceptable conceptualisation of the real world, then participants may be selecting "ungrammatical and unacceptable" to reflect the unacceptability of the conceptualisation, as opposed to evaluating the quality of the text. This is suspected in the evaluation of the question *Ingabe lonke ibhotela lenza ifoni eliyi-1 ncamashi?* ('Does every butter make exactly 1 phone?') where three participants selected "Ungrammatical and unacceptable" and one selected "Grammatical and ambiguous", even though the text is grammatically correct as judged by the first author. Furthermore, regional differences in the dialects spoken by the participants may result in one participant judging a particular text acceptable while others judge it ambiguous.

We also determined the number of texts that all participants judge as being grammatically correct irrespective of whether they are acceptable or ambiguous. This was calculated by counting the

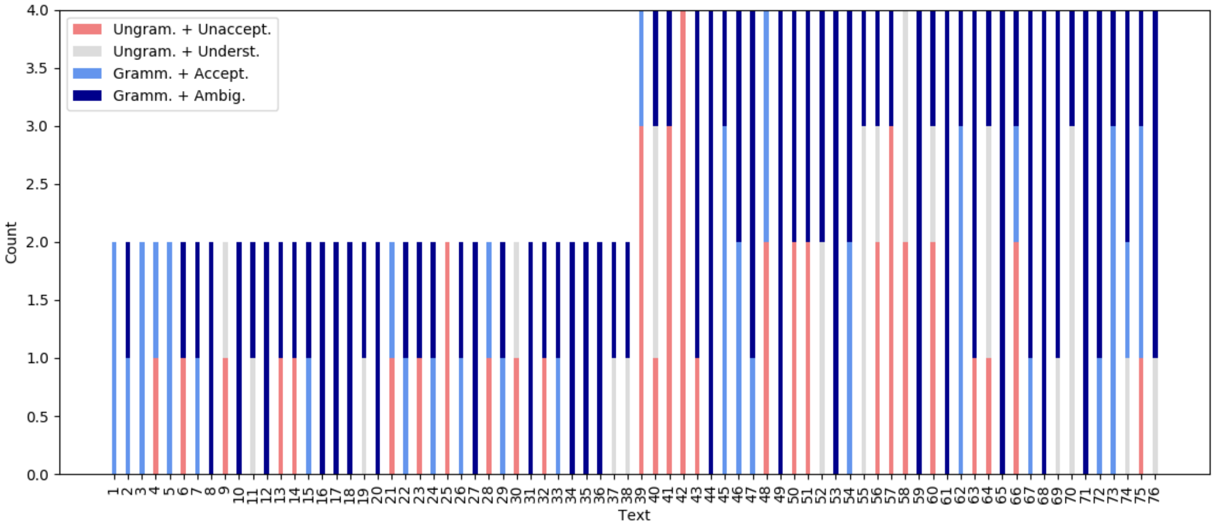

Figure 2: Ratings of the sentences, aggregated by sentence. Survey A texts are numbered 1-38 and Survey B texts are numbered 39-76. Abbreviations: Gramm. + Ambig. = grammatical and ambiguous, Gramm. + Accept. = grammatical and acceptable, Ungramm. + Underst. = ungrammatical and understandable, Ungramm. + unaccept. = ungrammatical and unacceptable.

number of texts where participants judged them as any combination of 'grammatical and ambiguous' and 'grammatical and acceptable'. We found that only 51% were judged as being grammatically correct hence 37/76 are ungrammatical. Moreover, the participants all agree that a majority (20/37) of the ungrammatical texts are still acceptable.

The evaluation shows that the templates produce texts, whose majority is found to be understandable and grammatically correct by isiZulu speakers. Moreover, even when generated texts are ungrammatical, most of them are still acceptable. Most of the texts generated by the templates are free of morphological agreement errors. In the only case where there was such an error (question 42), it was due to human error when creating the serialised template and not a problem with the templates listed in Table 2. The above observations suggest that the verbaliser can be used successfully (i.e., with isiZulu domain experts able to understand the meaning of the questions) when validating an ontology that has the axioms listed in Table 2, with the exception of axioms involving the EquivalentClassess type.

8 Conclusions and future work

We have created the first isiZulu CNL and verbaliser capable to generating questions from an ontology for the purpose of knowledge elicitation and validation. Evaluation of the quality of text generated by the implemented verbaliser shows that most (81%) of the participants' judgements, overall, are positive. Moreover, when we analysed judgements aggregated for each question, we find that most of the texts (83%) are judged positively because they have at most one participant who considers them to be ungrammatical and unacceptable. Adjusting how we determine the number of texts judged positively for Survey A since there were only two participants, we found that most (71%) of Survey A's questions were also judged positively because they had no participant who considers them to be ungrammatical and unacceptable. Overall, while there is a sizeable number (37/76) of questions for which the participants agree that they are ungrammatical, a majority of them (20/37) are still judged as being acceptable.

Future work includes determining which templates are generating unacceptable texts, soliciting an isiZulu grammarian's feedback regarding its grammaticality, and correcting the identified problems.

Acknowledgments

This work was financially supported by Hasso Plattner Institute for Digital Engineering through the HPI Research School at UCT and the National Research Foundation (NRF) of South Africa (Grant Number 120852).

References

G Aguado, A Bañón, J Bateman, Socorro Bernardos, M Fernández, Asunción Gómez-Pérez, Elena Nieto, A Olalla, R Plaza, and Antonio Sánchez. 1998. Ontogeneration: Reusing domain and linguistic ontologies for Spanish text generation. In *Workshop on Applications of Ontologies and Problem Solving Methods, ECAI*, volume 98.

Tahani Alsubait, Bijan Parsia, and Ulrike Sattler. 2016. Ontology-based multiple choice question generation. *Künstliche Intell.*, 30(2):183–188.

Ion Androutsopoulos, Gerasimos Lampouras, and Dimitrios Galanis. 2013. Generating natural language descriptions from OWL ontologies: the NaturalOWL system. *J. Artif. Intell. Res.*, 48:671–715.

Krasimir Angelov and Ramona Enache. 2010. Typeful ontologies with direct multilingual verbalization. In *Controlled Natural Language - Second International Workshop, CNL 2010, Revised Papers*, volume 7175 of *Lecture Notes in Computer Science*, pages 1–20. Springer. September 13-15, Marettimo Island, Italy.

Nadjet Bouayad-Agha, Gerard Casamayor, Simon Mille, Marco Rospocher, Horacio Saggion, Luciano Serafini, and Leo Wanner. 2012. From ontology to NL: Generation of multilingual user-oriented environmental reports. In *Natural Language Processing and Information Systems*, pages 216–221, Berlin, Heidelberg. Springer Berlin Heidelberg.

Nadjet Bouayad-Agha, Gerard Casamayor, Leo Wanner, Fernando Díez, and Sergio López Hernández. 2011. FootbOWL: Using a generic ontology of football competition for planning match summaries. In *The Semantic Web: Research and Applications - 8th Extended Semantic Web Conference, ESWC 2011, Heraklion, Crete, Greece, May 29-June 2, 2011, Proceedings, Part I*, volume 6643 of *Lecture Notes in Computer Science*, pages 230–244. Springer.

Joan Byamugisha, C. Maria Keet, and Brian DeRenzi. 2016. Bootstrapping a Runyankore CNL from an isiZulu CNL. In *Controlled Natural Language - 5th International Workshop, CNL 2016, Aberdeen, UK, July 25-27, 2016, Proceedings*, volume 9767 of *Lecture Notes in Computer Science*, pages 25–36. Springer.

John J. Camilleri, Norbert E. Fuchs, and Kaarel Kaljurand. 2012. Deliverable d11.1. ACE grammar library. Technical report, MOLTO project.

Vinay K. Chaudhri, Peter E. Clark, Adam Overholtzer, and Aaron Spaulding. 2014. Question generation from a knowledge base. In *Knowledge Engineering and Knowledge Management - 19th International Conference, EKAW 2014, Linköping, Sweden, November 24-28, 2014. Proceedings*, volume 8876 of *Lecture Notes in Computer Science*, pages 54–65. Springer.

Nancy J. Cooke. 1994. Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, 41(6):801849.

Dana Dannélls. 2008. Generating tailored texts for museum exhibitss. In *Proceedings of the Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008).*, pages 17–20, Marrakech, Morocco. Association for Computational Linguistics.

Dana Dannélls. 2010. Discourse generation from formal specificationsusing the grammatical framework, gf. *Research in Computing Science.*, 46:167–178. Special issue: Natural Language Processing and its Applications.

Dana Dannélls. 2012. On generating coherent multilingual descriptions of museum objects from semantic web ontologies. In *INLG 2012 - Proceedings of the Seventh International Natural Language Generation Conference, 30 May 2012 - 1 June 2012, Starved Rock State Park, Utica, IL, USA*, pages 76–84. The Association for Computer Linguistics.

Dana Dannélls, Aarne Ranta, and Ramona Enache. 2013a. Multilingual grammar for museum object descriptions. In Ramona Enache, editor, *Frontiers of Multilingual Grammar Development*, chapter 3, pages 99–107. Chalmers University of Technology and Gteborg University, Sweden.

Dana Dannélls, Aarne Ranta, Ramona Enache, Mariana Damova, and Maria Mateva. 2013b. Multilingual access to cultural heritage content on the semantic web. In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities, LaTeCH@ACL 2013, August 8, 2013, Sofia, Bulgaria*, pages 107–115. The Association for Computer Linguistics.

Brian Davis, Ramona Enache, Jeroen van Grondelle, and Laurette Pretorius. 2012. Multilingual verbalisation of modular ontologies using GF and lemon. In *Controlled Natural Language*, pages 167–184, Berlin, Heidelberg. Springer Berlin Heidelberg.

Yan Tang Demey and Clifford Heath. 2014. Towards verbalizing multilingual n-ary relations. In Paul Buitelaar and Philipp Cimiano, editors, *Towards the Multilingual Semantic Web, Principles, Methods and Applications*, pages 83–98. Springer.

Vinu E.V. and Sreenivasa Kumar P. 2015. A novel approach to generate MCQs from domain ontology: Considering DL semantics and open-world assumption. *Journal of Web Semantics*, 34:40 – 54.

Claire Gardent and Aldo Gangemi, editors. 2016. *Proceedings of the 2nd International Workshop on Natural Language Generation and the Semantic Web, WebNLG 2016, Edinburgh, UK, September 6, 2016*. Association for Computational Linguistics.

Albert Gatt and Ehud Reiter. 2009. Simplenlg: A realisation engine for practical applications. In *ENLG*

2009 - Proceedings of the 12th European Workshop on Natural Language Generation, March 30-31, 2009, Athens, Greece, pages 90–93. The Association for Computer Linguistics.

Normunds Gruūzītis. 2011. *Formal Grammar and Semantics of Controlled Latvian Language*. Ph.D. thesis, University of Latvia.

Terry Halpin and Matthew Curland. 2006. Automated verbalization for ORM 2. In *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, pages 1181–1190, Berlin, Heidelberg. Springer Berlin Heidelberg.

Mustafa Jarrar, C Maria Keet, and Paolo Dongilli. 2006. Multilingual verbalization of ORM conceptual models and axiomatized ontologies. Technical report, Starlab, Vrije Universiteit Brussel, Belgium.

C. Maria Keet and Langa Khumalo. 2014. Toward verbalizing ontologies in isiZulu. In *Controlled Natural Language - 4th International Workshop, CNL 2014, Galway, Ireland, August 20-22, 2014. Proceedings*, volume 8625 of *Lecture Notes in Computer Science*, pages 78–89. Springer.

C. Maria Keet and Langa Khumalo. 2016. On the verbalization patterns of part-whole relations in isiZulu. In *INLG 2016 - Proceedings of the Ninth International Natural Language Generation Conference, September 5-8, 2016, Edinburgh, UK*, pages 174–183. The Association for Computer Linguistics.

C. Maria Keet and Langa Khumalo. 2017. Toward a knowledge-to-text controlled natural language of isiZulu. *Language Resources and Evaluation*, 51(1):131–157.

C. Maria Keet, Musa Xakaza, and Langa Khumalo. 2017. Verbalising OWL ontologies in IsiZulu with python. In *The Semantic Web: ESWC 2017 Satellite Events - ESWC 2017 Satellite Events, Portorož, Slovenia, May 28 - June 1, 2017, Revised Selected Papers*, volume 10577 of *Lecture Notes in Computer Science*, pages 59–64. Springer.

Konstantinos I. Kotis, George A. Vouros, and Dimitris Spiliotopoulos. 2020. Ontology engineering methodologies for the evolution of living and reused ontologies: status, trends, findings and recommendations. *The Knowledge Engineering Review*, 35:e4.

Fennie Liang, Robert Stevens, and Alan L. Rector. 2011. OntoVerbal-M: a multilingual verbaliser for SNOMED CT. In *Proceedings of the 2nd International Workshop on the Multilingual Semantic Web, Bonn, Germany, October 23, 2011*, volume 775 of *CEUR Workshop Proceedings*, pages 13–24. CEUR-WS.org.

Shin Huei Lim and Terry Halpin. 2016. Automated verbalization of ORM models in Malay and Mandarin. *International Journal of Information System Modeling and Design*, 7(4):1–16.

Zola Mahlaza. 2018. Grammars for generating isiXhosa and isiZulu weather bulletin verbs. Msc thesis, Department of Computer Science, University of Cape Town, South Africa.

Zola Mahlaza and C. Maria Keet. 2019. A classification of grammar-infused templates for ontology and model verbalisation. In *Metadata and Semantic Research - 13th International Conference, MTSR 2019, Rome, Italy, October 28-31, 2019, Revised Selected Papers*, volume 1057 of *Communications in Computer and Information Science*, pages 64–76. Springer.

Achille Emile Meeussen. 1967. Bantu grammatical reconstructions. *Africana linguistica*, 3(1):79–121.

Shamila Naidoo. 2002. The palatalisation process in isiZulu revisited. *South African Journal of African Languages*, 22(1):59–69.

Andreas Papasalouros, Konstantinos Kanaris, and Konstantinos Kotis. 2008. Automatic generation of multiple choice questions from domain ontologies. In *IADIS International Conference e-Learning 2008, Amsterdam, The Netherlands, July 22-25, 2008. Proceedings*, pages 427–434. IADIS.

Lionel Posthumus. 2016. A systemized explanation for vowel phoneme change in the inadmissible phonological structure /vv/ in Zulu.

Richard Power. 2012. OWL simplified english: A finite-state language for ontology editing. In *Controlled Natural Language - Third International Workshop, CNL 2012, Zurich, Switzerland, August 29-31, 2012. Proceedings*, volume 7427 of *Lecture Notes in Computer Science*, pages 44–60. Springer.

Laurette Pretorius and Sonja Bosch. 2010. Finite state morphology of the Nguni language cluster: Modelling and implementation issues. In *Finite-State Methods and Natural Language Processing*, pages 123–130, Berlin, Heidelberg. Springer Berlin Heidelberg.

Yevgeniy Puzikov and Iryna Gurevych. 2018. E2E NLG challenge: Neural models vs. templates. In *Proceedings of the 11th International Conference on Natural Language Generation, Tilburg University, The Netherlands, November 5-8, 2018*, pages 463–471. Association for Computational Linguistics.

Driss Sadoun, Satenik Mkhitaryan, Damien Nouvel, and Mathieu Valette. 2016. ReadME generation from an OWL ontology describing NLP tools. In (Gardent and Gangemi, 2016), pages 46–49.

Hazem Safwat and Brian Davis. 2017. CNLs for the semantic web: a state of the art. *Language Resources & Evaluation*, 51(1):191–220.

Lauren Sanby, Ion Todd, and C. Maria Keet. 2016. Comparing the template-based approach to GF: the case of Afrikaans. In (Gardent and Gangemi, 2016), pages 50–53.

Galen Sibanda. 2007. Vowel processes in Nguni: Resolving the problem of unacceptable vv sequences. In *Selected Proceedings of the 38th Annual Conference on African Linguistics, Gainesville, Florida, March 22-25, 2007*, pages 38–55. Cascadilla Proceedings Project, Somerville, Massachusetts, USA.

Andrew Van der Spuy. 2014. Bilabial palatalisation in Zulu: A morphologically conditioned phenomenon. *Stellenbosch Papers in Linguistics Plus*, 44:71–87.

Graham Wilcock. 2003. Talking OWLs: Towards an ontology verbalizer. In *Proceedings of the Human Language Technology for the Semantic Web and Web Services Workshop, ISWC 2003, Second International Semantic Web Conference, Sanibel Island, FL, USA, October 20-23, 2003*, pages 109–112.

Lishan Zhang and Kurt VanLehn. 2016. How do machine-generated questions compare to human-generated questions? *Research and Practice in Technology Enhanced Learning*, 11(1):7.