An Orchestration Framework for Linguistic Task Ontologies

Catherine Chavula and C. Maria Keet

Department of Computer Science, University of Cape Town, South Africa
{cchavula,mkeet}@cs.uct.ac.za

Abstract. Ontologies provide knowledge representation formalism for expressing linguistic knowledge for computational tasks. However, natural language is complex and fluid, demanding fine-grained ontologies tailored to facilitate solving specific problems. Moreover, extant linguistic ontological resources ignore mechanisms for systematic modularisation to ensure semantic interoperability with task ontologies. We present an orchestration framework to organise and control the inheritance of ontological elements in the development of linguistic task ontologies. The framework is illustrated in the design of new task ontologies for Bantu noun classification system. Specific use is demonstrated with annotation of lexical items connected to ontology elements terms, and with the classification of nouns in the ABox into noun classes.

1 Introduction

Ontologies are increasingly being used to provide computationally ready data for Natural Language Processing (NLP) tasks in a uniform standard. Ontologies are being used to publish language resources and annotation schemes for different application scenarios on the Semantic Web. For example, natural language features are encoded in an ontology to document linguistics domain knowledge as well as to provide terminology for annotating machine readable language data in [8]. Another recent growing application for linguistic metadata frameworks or ontologies, is annotating lexicalisations of ontology elements terms with linguistic features specified in linguistic ontologies [14]. The ontology lexicalisations are used to facilitate ontology-based NLP tasks such as generating natural language descriptions of Semantic Web documents [5] and to build multilingual resources for world internationalisation (e.g. AGROVOC in many languages).

Natural language is complex and fluid, and demands modular ontologies to capture linguistic knowledge at the required level of specificity. For example, some features vary across languages and other features exist only in specific languages. However, ontological axioms are based on intensional definitions [10]; this is a problem when capturing language-specific features, which require instance level definitions. Additionally, lack of principled methodologies to link or align generic linguistic ontologies has led to isolated ontologies which can not be integrated due to conflicting representations of the domain knowledge, or not used with existing resources due to formats. Generic linguistic ontologies or frameworks attempt to address these problems by creating resources that capture all linguistic features [9, 13]. However, these linguistic resources do not capture language specific features at the desired level of granularity. Task ontologies provide a means of bridging general language knowledge with fine-grained language specific knowledge which may be tailored for specific computational tasks such as Natural Language Generation (NLG). However, the same challenges of alignment and intensional specification resurfaces.

Given the challenges for modelling language specific task oriented ontologies, the paper makes three contributions. Firstly, we present an orchestration architecture for facilitating systematic modular design and interoperability of linguistic task ontologies. Our approach merges ideas from BioTop, a domain ontology for the life sciences [2], and the DOGMA approach [11], an ontology engineering methodology. Secondly, we present Noun Class System (NCS) for Bantu languages specification in OWL ontologies based on the orchestration architecture, and thirdly, we present the classification of Bantu nouns into their noun class based on the ontology of the linguistic noun classification, therewith satisfying one of the competency questions.

The remainder of the paper is structured as follows. Section 2 describes the framework for the orchestration of linguistic task ontologies. Section 3 summarises the NCS in Bantu languages and describes the ontology development process for the Bantu noun class system ontology, and its use cases. Section 4 compares our approach with related work, and we conclude in Section 5.

2 Orchestration Framework Architecture

Human natural languages are complex and dynamic. For example, some features are universal to all languages while others exist in only specific languages. Ontologies provide an approach for specifying this complex linguistic knowledge. However, the differences in features for different languages, necessitate specialised ontology modules. Unfortunately, there is lack of principled methods for aligning fine grained conceptualisation with other high level domain conceptualisations. The orchestration framework has been developed to be used in the design of task specific linguistic ontologies to achieve semantic interoperability with the existing linguistic ontologies. The approach adopted in the architecture of the framework is inspired by ideas from BioTop, a top-domain ontology for the life sciences [2], and the DOGMA approach [11] to ontology engineering and conceptual model development. The architecture of the framework provides a systematic modular design for aligning foundational ontologies, linguistic description ontologies, and task specific linguistic ontologies.

One of the challenges for aligning task ontologies with domain ontologies is to specify the alignment mechanism between task ontologies, domain ontologies and foundational ontologies. BioTop uses a 'pyramid' of one foundational ontology–Basic Formal Ontology (BFO), several top-domain ontologies (BioTop), and multiple domain ontologies (such as Cell Ontology (CL) and Gene Ontology (GO) [2]). BioTop is a top-level domain ontology that is used to create new domain ontologies which are semantically interoperable with existing ontologies as well as to improve or align existing ontologies in the life sciences domain. Our framework adopts the BioTop architecture to provide an alignment mechanism between task ontologies and domain ontologies and, within the framework, we have defined a Top-domain ontology layer that consists of generic ontologies.

DOGMA is an ontology engineering methodological framework for guiding ontology engineers to build ontological resources which are usable and reusable [11]. The DOGMA approach aims to build ontologies independent of the application requirements whilst ensuring that the specified knowledge can be reused by other applications and meet their specific requirements. DOGMA uses the *principle of double articulation* to axiomatize knowledge: domain knowledge is specified to capture the intended meaning of the vocabulary, and is reused to add application-specific constraints in order to meet application requirements or to handle highly specialised differences. Natural language is highly flexible and same concepts may vary across languages. Expressing specialised linguistic knowledge in an ontology for a single natural language is challenging because knowledge captured in ontologies is based on *intensional semantic structure* [10]. Thus, we adopted the DOGMA approach in order to accommodate the diversity of languages: an ontological conceptualization and a specific knowledge axomitization with added constraints.

The proposed approach defines four linked ontological layers: top-level, top-domain and domain ontologies,task ontologies and a fifth layer for added precision for each language:

- **Top level ontologies**, which represent high level categories of things in the world independent of a subject domain;
- **Top-domain ontologies**, which contain linguistic knowledge independent of linguistic theories and languages, and provide conceptual interlinkages with domain ontologies, task specific and domain independent knowledge; **domain ontologies** concepts can also be covered at this level, if the ontology covers sub-domain knowledge.
- **Domain ontologies**, which contain specialised knowledge of a particular sub-domain.
- **Task ontologies**, which specify language-specific scenario oriented knowledge to enhance specific computational tasks (e.g. the classification of nouns into their classes, see Section 3.6);
- **Logic-based conceptual models/axiomatizations**, which contain more precise knowledge for a specific 'application', in our case with natural language specific idiosyncrasies and additional constraints.

Fig. 1 shows the general idea of our modular architecture, which will be instantiated for linguistics knowledge and the Bantu noun class system in Section 3. The arrows in the diagram show the alignments, which can be equivalence and/or subsumption alignments between the entities in the ontologies. The purpose of the framework is to ensure that task specific ontologies can be developed in a modular and systematic fashion and that the resulting ontologies are interoperable with other ontological resources in the linguistics domain. For example, the Bantu noun classification system has different singular/plural mapping schemes across languages and it is impossible to capture this knowledge in a single conceptualisation. Modular design is suitable for this scenario but lacks mechanism for

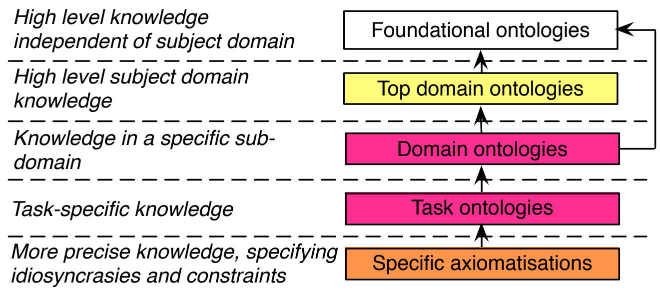

Fig. 1. Framework ontological layers

linking and aligning these modules with extant linguistic ontological resources. In the DOGMA approach, application knowledge specification uses agreed terms or vocabulary defined in the domain knowledge. Similarly, task ontologies can be defined at two levels, a task ontology and language specific task ontologies. The proposed framework has been applied in the design and implementation of Bantu NCS ontologies, which is described in the next section.

3 Applying the Framework to Bantu Noun Class System Ontologies

Bantu languages are a major language family on the African continent, with over 220 Million speakers across Sub-Saharan Africa. Bantu languages are largely agglutinative with complex structural and syntactic features [12] (as are, e.g., Finnish and Quechua). Bantu languages have several structural similarities that enable some of computational solutions to be adapted across the family. For example, noun classification is one of those pervasive features [12]. Nouns are categorized into classes to a large extent by the prefixes the nouns take. Formalising the Bantu NCS into a computational artefact is one of the requirements identified for Semantic Web NLP based applications for Bantu languages [3].

The Bantu NCS ontologies seek to provide fine-grained specification of entities and relationships for the NCS of Bantu Languages; this level of specification is necessary for deep morphological analysis of nominal phrases [7]. Further, the ontologies will serve as a computational model for the analysis of Bantu nouns and documentation of complex relationships, which may lead to further linguistic research. Also, NCS ontology can be used for annotation of nouns with their noun classes which is a necessary component in multilingual ontology-driven information systems. Clearly, the purposes of the NCS ontology require that the ontology be interoperable with existing ontological resources and the proposed framework enables Bantu NCS ontologies have been applied to achieve this. We describe basics aspects of the noun class system first, and then the ontology development methodology, design, its contents, and how the framework is applied.

3.1 Overview of Bantu Noun Class System

Nominal classification is a common feature in many languages. For instance, those in romance languages category (e.g., French and Italian), have a gender category, which classifies nouns into types such as feminine and masculine [6]. Although the Bantu noun classification has been given the treatment of gender category, Bantu classification exhibit attributes that need to be considered in its own category. The Bantu noun classification is largely based on semantics and morphological marking of nominal prefixes or word structure of a noun [12]. Early studies of Bantu nominal morphology identified individual prefixes on nouns and labelled the prefixes with Arabic numerals which were then proposed as Bantu noun classes [12]. Plural and singular forms of Bantu nouns take different prefixes. Thus, using this classification, each class can have a corresponding singular or plural form, i.e., the classes are categorised into singular and plural forms with each marked by a corresponding prefix; e.g., a pairing of noun stems and prefixes in Chichewa (in Guthrie zone (N31)) for class7/class8 are *chi-*/*zi-* and for class12/class13 they are *ka-*/*ti-*; e.g. *chi*patso ('fruit') and *zi*patso ('fruits'). These prefixes are added to other morphemes or words to create singular or plural nouns, e.g., *ka*chipatso ('small fruit'). The collection of prefixes contributes to the construction of the traditional Bantu NCS. The class of a noun determines the markers on syntactic elements in a phrase or sentence (e.g., verbs and adjectives) and contributes to their inflectional behaviour; e.g., *chi*patso *cho*koma ('tasty fruit').

The number of classes varies in different languages but the majority of the languages exhibit some similarities in the semantics of the classes, prefixing and the pairing of the classes into singular and plural forms. In the community of Bantu linguists, the Bleek-Meinhof classification is widely used [12]. The Bleek-Meinhof classification uses the prefixes as indicators of classes and the NCS is built by listing all the prefixes available in a language with Arabic Numerals prefixes. Thus, singular and plural forms of a word belong to two separate classes. In order to maintain the relationship between the singular and plural classes, linguists use the Bleek-Meinhof numbering system and may group the plural and its singular classes as one class, e.g., class1 and class2 becomes class1/2 [6].

3.2 Methodology for NCS Ontologies Development

The development of the ontologies followed a bottom-up approach [16]. In particular, this involved i) a preliminary domain analysis to establish the technical feasibility of having an NCS ontology; ii) assessment of relevant existing ontologies and non-ontological resources (databases and documentation of linguistic resources), including those described in Section 4 below; iii) identification of the concepts and relationships in Bantu noun classification, including adopting concepts from the GOLD ontology; iv) develop a first version of the ontology, based on knowledge of Chichewa and isiZulu using Meinhof's classification, for community evaluation [3].

Experiences with this NCS ontology induced a scope and structural change from the aim to lexicalise an ontology in Chichewa and isiZulu with the *lemon* model, to that it should cater for the whole Bantu language family, and more generally, be an extensible system. The bottom-up approach was followed and more resources consulted, such as [12], consulting domain experts (linguists) and Bantu language speakers and presenting (verbalised and visualised) drafts of the ontology, and competency questions formulated, including:

CQ1: Is the nominal classification feature in the ontology capturing the taxonomic structure for Bantu noun class system?

CQ2: Do the corresponding relationships capture the constraints in the relationship of nominal concepts in Bantu languages?

CQ3: Can it infer the class of a noun based on either knowing the singular or plural or noun class of a noun word?

For purposes of interoperability and extensibility, a comprehensive alignment to GOLD was carried out (GOLD was chosen, since the initial motivation for the ontology was for linguistic annotation) and a modular architecture was devised. Thereafter, the ontologies were evaluated in the tasks of noun classification into their classes and annotation of nominal lexical items (class labels in an ontology) with their noun classes.

3.3 NCS Ontologies Design and Implementation

The design of NCS ontologies captures the noun classification concepts and relationships within Bantu Languages spectrum. The current release of the ontology uses the proposed orchestration framework and has been re-engineered (cf. [3]) in the following way:

- The major improvement is the use of an orchestration framework to cater for the differences in the noun classes across Bantu languages, rather than only Chichewa and isiZulu: the use of a double articulation principle [11] to capture these differences and alignment with GOLD by applying its principles in the ontology.
- Multiple classification schemes of Bantu noun system have been used (cf. only Meinhof's).
- New concepts, relationships, and constraints to capture fine-grained linguistic domain knowledge to obtain desired inferences.

Practically, the ontologies have been represented in OWL, and are available from `http://meteck.org/files/ontologies/` in `NCS1.zip`. This contains a GOLD module (with a SUMO module), the NCS ontology, and, at the time of writing, language-specific axiomatisations for Chichewa, Xhosa, and Zulu.

3.4 Overview of the NCS Ontology

The NCS ontology design is based on the classification of Bantu nouns at fine-grained morphemic units, and the structural and lexical relationships among these units. The taxonomic structure has two main parts: the first part provide the morphological structure of Bantu nouns and the second part provides the hierarchy of concepts for properties of Bantu nouns including the NCS based on Bleek-Meinhof [12, 6]. The first part allows the labelling of Bantu nouns beyond the part-of-speech category and captures the lexical units of the nouns and how they are structured in relation to the NCS. The second component models the grammatical features of the nouns and captures the Bantu NCS concepts. The gender and grammatical number linguistic categories are included to avoid confusion with the noun classification feature. Fig. 2 shows the taxonomy of concepts in the ontology (only a subset of the noun classes are shown). Traditional Bleek-Meinhof classes and modern paired labelling schemes were used to specify the ontologies classes drawn from Bantu Languages studies [12].

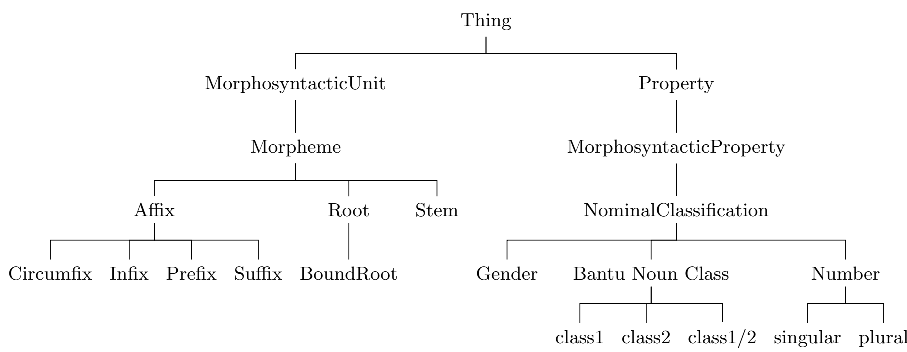

Fig. 2. Section of the class hierarchy of the Bantu NCS.

3.5 Application of the Framework

The design of the NCS ontologies follow the proposed architecture, and is depicted schematically in Fig. 5. At the bottom we have the logic-based conceptual models—also called 'application ontologies' or structured metadata—that capture concepts and relationships in Bantu NCS domain for a specific language. These language-specific ontologies are specialisations of the general NCS ontology as 'task ontology', following the double articulation principle. The resulting ontologies are aligned with a relevant module of GOLD, which was already aligned to the SUMO foundational ontology. The NCS ontologies are therefore linked to these resources by following these principles and the proposed framework. Conceptual models for noun classes of other Bantu Languages can easily be 'plugged in', starting from the NCS ontology as its top ontology. The framework can also be extended 'horizontally' to cater for other languages; e.g., a task ontology about verb conjugation in the Romance languages with specifics for Spanish and Italian each in its OWL file, yet remaining interoperable.

3.6 Using the Ontologies

The NCS ontologies provide language-specific linguistic properties that are useful in language studies and in language engineering tasks. One of the foreseen usage

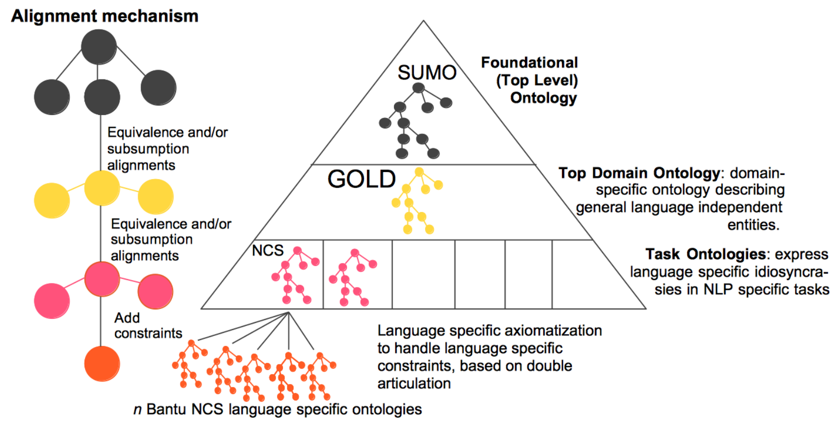

Fig. 3. GOLD and Noun Class System alignment using the proposed framework

scenarios is in the annotation of text for computational language processing such as morphological analysis as well as annotating lexical items in computational lexicons. We describe two use cases.

Use Case I: Linguistic Annotation Data on the Semantic Web consists of language independent factual knowledge which is based on formal vocabularies specified in ontologies. Unfortunately, this enormous amount of data is inaccessible to many potential human users because of the complexity of the logic-based knowledge representation model. Expressing or accessing this knowledge using natural language ensures that the knowledge is accessible to end-users. However, such interaction methods need Natural Language Processing (NLP) tasks to be incorporated into Semantic Web applications or tools. Evidently, these NLP tasks require ontologies that are grounded with rich linguistic data in multiple languages, i.e., lexical knowledge specifying how ontology elements are expressed in multiple languages and their associated linguistic properties [14].

Ontology lexicalisation provides a means of enriching ontologies with linguistic knowledge [14]. Several models have been proposed to express how ontology elements are linguistically realised. For example, the lexicon model for ontologies, *lemon*, is a descriptive model for structuring and publishing ontology lexicalisations on the Semantic Web [14]. *lemon* defines the structure for lexical entries and how the entries interface with ontology elements terms. Externally defined linguistic properties, e.g., linguistic annotation ontologies, are used to describe the entries in the lexicons. In the context of NCS ontology, Bantu noun entries in *lemon* format can be annotated with their noun classes. Linguistic properties defined in the upper layers of the orchestration framework can be used with properties defined in the NCS ontology consistently. For instance, Fig. 4 shows the lexical entry for the property dcterms:language (`http://purl.org/dc/terms/language`) from the Dublin Core Metadata Initiative (DCMI), and the entry uses NCS ontology elements to specify a noun class of a Chichewa lexicon.

```
@prefix dcterms: <http://purl.org/dc/terms/>.
@prefix rdfs: <http://www.w3.org/2001/02/rdf-schema#>.
@prefix ncsNY: <http://www.meteck.org/files/ontologies/ncsNY/>.
@prefix lemon: <http://www.lemon-model.net/>.
@prefix gold: <http://purl.org/linguistics/gold/>.
@prefix :  <http://www.mteck.org/id/dcterms/lexiconNY>.
:myDCLexicon a lemon:Lexicon ;
lemon:language "ny" ;
lemon:entry :chiyankhulo.
:chiyankhulo a lemon:LexicalEntry ;
ncsNY:BantuNounClass ncsNY:class7;
gold:PartOfSpeechProperty gold:noun;
lemon:canonicalForm [lemon:writtenRep "chiyankhulo"@ny];
lemon:sense [lemon:reference dcterms:language] .
ncs:BantuNounClass rdfs:subPropertyOf lemon:property.
gold:PartOfSpeechProperty rdfs:subPropertyOf lemon:property.
```

Fig. 4. Chichewa dcterms:language entry.

Use Case II : ABox Classification Modelling linguistic properties in ontologies provide more expressiveness to specify the complex relationships that exist among concepts. Using the proposed orchestration framework, language specific idiosyncrasies can be captured and formalised in a generic paradigm without interfering language universals. Positively, the combined knowledge from all the framework layers can be used to infer new relationships not explicitly specified; this is useful in language processing because automatic individual classification may compensate incomplete linguistic annotation especially for under-resourced languages. One of the requirements (CQ3) of the NCS ontologies is to be able to infer the class of a noun (ABox individual) with a singular or plural relationship to another annotated noun (see Section 3.2). A task-based evaluation of this requirement requires that the reasoner returns correct ABox classification or accurate responses to DL queries concerning ABox classification. For example, the NCS ontology of Chichewa specifies the relationship between classes 7 and 8 using ncs:hasPlural and ncs:hasSingular, so that with the singular asserted, it can deduce the plural (where the plural relationship has been specified), or vv, which is illustrated in Fig. 5 for *chiyankhulo*.

4 Related Work

Ontologies have been widely used by researchers to formalise linguistic knowledge for use in ontology driven information systems and the Semantic Web. For example, GOLD is a linguistic ontology that documents expert linguistic knowledge in an ontology. GOLD is aligned with Suggested Upper Merged Ontology

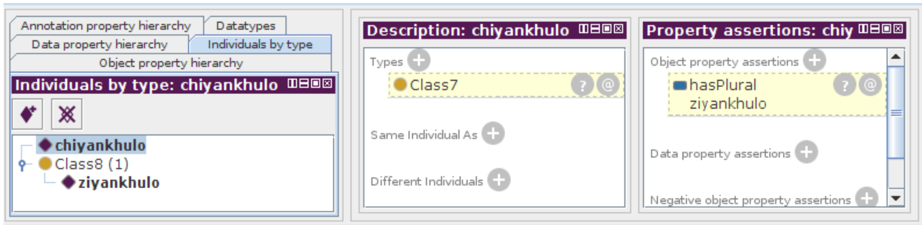

Fig. 5. Example of deductions for *chiyankhulo*, given *ziyankhulo* is in class 8.

(SUMO), a foundational ontology, to ensure semantic interoperability with other ontologies. GOLD captures linguistic properties independent of any linguistic theory and the ontology contains general and language specific linguistic properties. Due to these attributes, we used GOLD as a Top Domain Ontology for our instantiated framework. However, GOLD encodes the Bantu noun classes as a type of gender by defining a Roman numeral based gender concept. We capture the classification in a different way as noun classes are mostly based on the underlying meaning, e.g., humans are in classes 1 and 2, and other morphological aspects (recall Section 3.1).

Ontologies have also been used to mediate between domain ontologies and natural language realisations of ontological entities. For example, General Upper Model (GUM) ontology implements an interface for the interaction of domain-specific knowledge and general linguistic resources [1]. Thus, GUM provides an abstraction between surface realization and domain specific knowledge. Although, GUM can be categorised as a task ontology targeting NLP tasks such as NLG, the ontology does not provide any means for linking with other linguistic ontologies. Our work is different as the proposed orchestration framework provide a method for linking linguistic ontologies to task linguistic ontologies to ensure interoperability.

Due to the heterogeneity of terminology for annotating linguistic properties, different data models have been proposed to make language data and metadata interoperable. The ISO TC37/SC4 Data Category Registry (DCR) is a community maintained repository for linguistics concepts and metadata categories [13]. The terminologies or data categories can be imported for use in applications on the Semantic Web; the categories have been used to create LexInfo ontology, which is used in annotating ontology-lexicons in *lemon* format [5]. Still, the available categories are limited, lacking complete noun class information. For example, only Zulu noun classes have been proposed for DCR categories [15] and that consists of a subset of the noun classes identified for all Bantu languages. We have demonstrated how our framework can be used to accommodate Bantu noun classes for all languages in its family. Additionally, we have proposed a framework for linguistic task ontologies but DCR only focuses on terminologies for linguistic annotation.

Similar to DCR, Ontologies of Linguistic Annotation (OLiA) is a repository of linguistic data categories. OLiA formalises the semantics of linguistic annotation terminologies as OWL2/DL ontologies to achieve both structural and conceptual interoperability between different annotation schemes in the extant repositories such as GOLD and ISOcat [4]. However, Bantu languages being under-resourced, are not covered to an adequate level of detail. The NCS ontology focuses on language specific attributes of nouns which can be applied to NLP applications within Bantu languages and this sets it apart from resources such as OLiA, which attempt to align general linguistic ontologies to ensure interoperability. Furthermore, the orchestration framework adds a modular design architecture at a lower level, allowing language-specific idiosyncrasies to be accommodated.

A repository for PartOfSpeech features for tagging two South African Languages is proposed in [7]. The repository is designed to have a taxonomic representation of linguistic categories for Bantu languages and the design of the repository is to be implemented in a relational database. This work is similar to our NCS ontologies, but the representation of the NCS is not considered as part of the ontology. Additionally, the repository does not consider the formalisation of the linguistic properties into a formal ontology.

5 Discussion and Conclusions

The representation issue of the tension between genericity and specificity of representing domain knowledge, has been solved by merging into a single framework, a pyramidal modular architecture with the *double articulation principle*. The proposed framework can be applied in developing task-oriented ontologies whose conceptualisation does not match any of the existing (linguistic) ontologies but has to be used with the existing resources and refine existing ones. This framework was applied to linguistic ontologies so as to control the development of task specific linguistic ontologies to ensure that concepts are aligned with extant domain and foundational ontologies, with as finer-grained instantiation the design of noun class ontologies. Multiple noun class ontologies (conceptual models/structured metadata) have been developed for different Bantu language using the proposed framework. Thanks to alignment with GOLD and SUMO, the NCS ontologies can be used with other linguistic ontologies to annotate text and other structured linguistic resources. In addition, the NCS ontologies can be used to classify nouns of a specific Bantu language using a specific Bantu NCS ontology. This can be used as pre-processing stage of language resources and can reduce the cost of developing such resources and improve the performance of NLP tasks such as morphological processing. We have also illustrated how the ontology can be used in the classification of nouns where the nouns are individuals in the ontology and annotation of lexical entries with linguistic properties. The ontologies may be used with other community-maintained terminology repositories that capture other linguistic properties. Our future direction of this work includes using the framework to further align other task ontologies and building a library of ontologies which have been aligned using this approach, and use this repository to conduct an empirical evaluation of the framework. We are currently adding the NCS ontologies to the linguistic Linked Open Data

(LOD) cloud so that it also can be used for ontology-driven multilingual information systems.

Acknowledgements We thank Dr Langa Khumalo and Atikonda Mtenje for their input and feedback. This work is based upon research supported by the National Research Foundation of South Africa (Project UID: 93397).

References

1. Bateman, J.A.: Upper modeling: organizing knowledge for natural language processing. In: Proc. of the 5th INLG Workshop. pp. 54–60. Pittsburgh, PA. (June 1990)
2. Beisswanger, E., Schulz, S., Stenzhorn, H., Hahn, U.: Biotop: An upper domain ontology for the life sciences: A description of its current structure, contents and interfaces to obo ontologies. Applied Ontology 3(4), 205–212 (Dec 2008)
3. Chavula, C., Keet, C.: Is Lemon Sufficient for Building Multilingual Ontologies for Bantu Languages? In: Proc. of OWLED'14. vol. CEUR-WS vol. 1265, pp. 61–72. Riva del Garda, Italy (Oct 17-18 2014)
4. Chiarcos, C.: Ontologies of Linguistic Annotation: Survey and perspectives. In: Proc. of LREC'12. ELRA proceedings, Istanbul, Turkey (May 2012)
5. Cimiano, P., Nagel, D., Unger, C.: Exploiting ontology lexica for generating natural language texts from RDF data. In: Proc. of 14th European Workshop on NLG. pp. 10–19 (2013), august, Sofia, Bulgaria
6. Corbett, G.G.: Gender. Cambridge University Press., Cambridge (1991)
7. Faab, G., Bosch, S., Taljard, E.: Towards a Part-of-Speech Ontology; Encoding Morphemic Units of Two South African Languages. Nordic Journal of African Studies 21(3), 118–140 (May 2012)
8. Farrar, S., Langendoen, D.: Markup and the GOLD Ontology. In: Proce. of Ws. on Digitizing and Annotating Text and Field Recordings (2003)
9. Farrar, S., Langendoen, D.T.: A linguistic ontology for the semantic web. In: GLOT International. 3, vol. 7, pp. 97–100 (2003)
10. Guarino, N.: Formal ontology and information systems. In: Proc. of FOIS'98. pp. 3–15. IOS Press (1998), trento, Italy
11. Jarrar, M., Meersman, R.: Ontology Engineering The DOGMA Approach. In: Advances in Web Semantics I, LNCS, vol. 4891, pp. 7–34. Springer (2009)
12. Katamba, F.: Bantu Nominal Morphology. In: Nurse, D., Phillippson, G. (eds.) The Bantu Languages. Curzon Press, London (2003)
13. Kemps-Snijders, M., Windhouwer, M., Wittenburg, P., Wright, S.: ISOcat: Remodelling metadata for language resources. International Journal of Metadata, Semantics and Ontologies 4(4), 261–276 (2009)
14. McCrae, J., Aguado-de Cea, G., Buitelaar, P., Cimiano, P., Declerck, T., Gómez-Pérez, A., Gracia, J., Hollink, L., Montiel-Ponsoda, E., Spohr, D., Wunner, T.: Interchanging lexical resources on the Semantic Web. LRE 46(4), 701–719 (2012)
15. Pretorius, L., Bosch, S.E.: Towards Extending the ISOcat Data Category Registry with Zulu Morphosyntax. In: Proceedings of the 10th Joint ACL - ISO Workshop on Interoperable Semantic Annotation (2014)
16. Suárez-Figueroa, M., Gómez-Pérez, A., Fernández-López, M.: The NeOn methodology for ontology engineering. In: Ontology Engineering in a Networked World, pp. 9–34. Springer (2012)