Ontology Engineering

Lecture 9: Ontologies and natural languages

Maria Keet

email: mkeet@cs.uct.ac.za

home: http://www.meteck.org

Department of Computer Science
University of Cape Town, South Africa

Semester 2, Block I, 2019

Outline

1. Introduction
2. Multilingual ontologies
3. Ontology verbalisation

Outline

1 Introduction

2 Multilingual ontologies

3 Ontology verbalisation

Natural language and ontologies

- Using ontologies to improve NLP; e.g.:
  - To enhance precision and recall of queries
  - To enhance dialogue systems
  - To sort literature results
- Using NLP to develop ontologies (TBox)
  - Searching for candidate terms and relations
- Using NLP to populate ontologies (ABox)
  - Document retrieval enhanced by lexicalised ontologies
  - Biomedical text mining
- Natural language generation from a logic
  - Ameliorating the knowledge acquisition bottleneck
  - Other purposes; e.g., e-learning (question generation), readable medical information

Outline

1 Introduction

2 Multilingual ontologies

3 Ontology verbalisation

Multilingual ontologies

- What the previous sub-sections do not mention: they are "English ontologies" and work with natural language text in English
- How to build an ontology for, say, Spanish organic agriculture? [Organic.Lingua project] 'intelligent' eGovernment portals in the 11 official languages of South Africa?

Multilingual ontologies

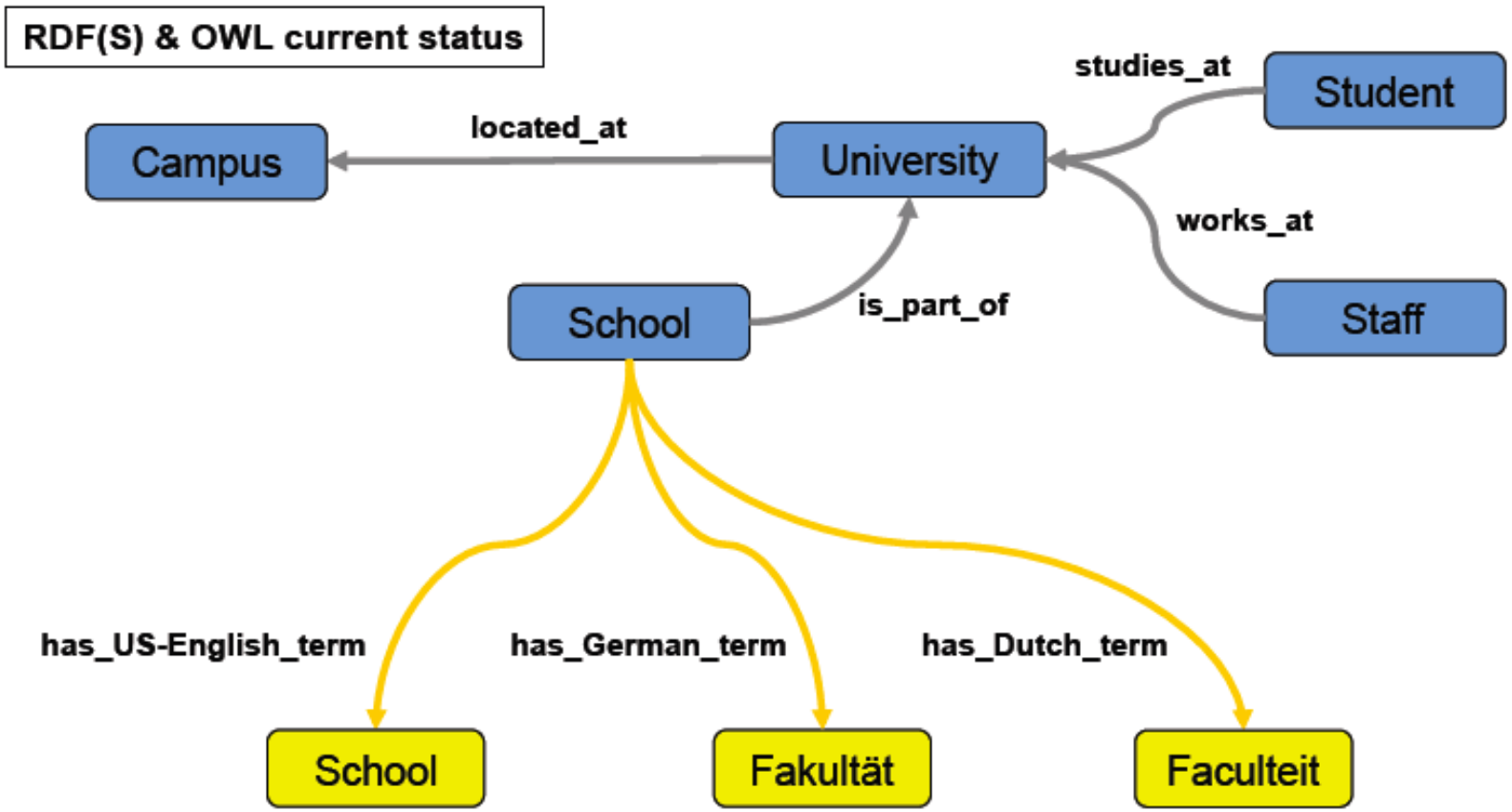
- What the previous sub-sections do not mention: they are "English ontologies" and work with natural language text in English
- How to build an ontology for, say, Spanish organic agriculture? [Organic.Lingua project] 'intelligent' eGovernment portals in the 11 official languages of South Africa?
- Multilingualism with ontologies
  - 'Ontology in different languages'?
  - NLP (NLU) for target language to learn
  - NLG for user and domain expert-friendly interface to the ontology

Multilingual ontologies

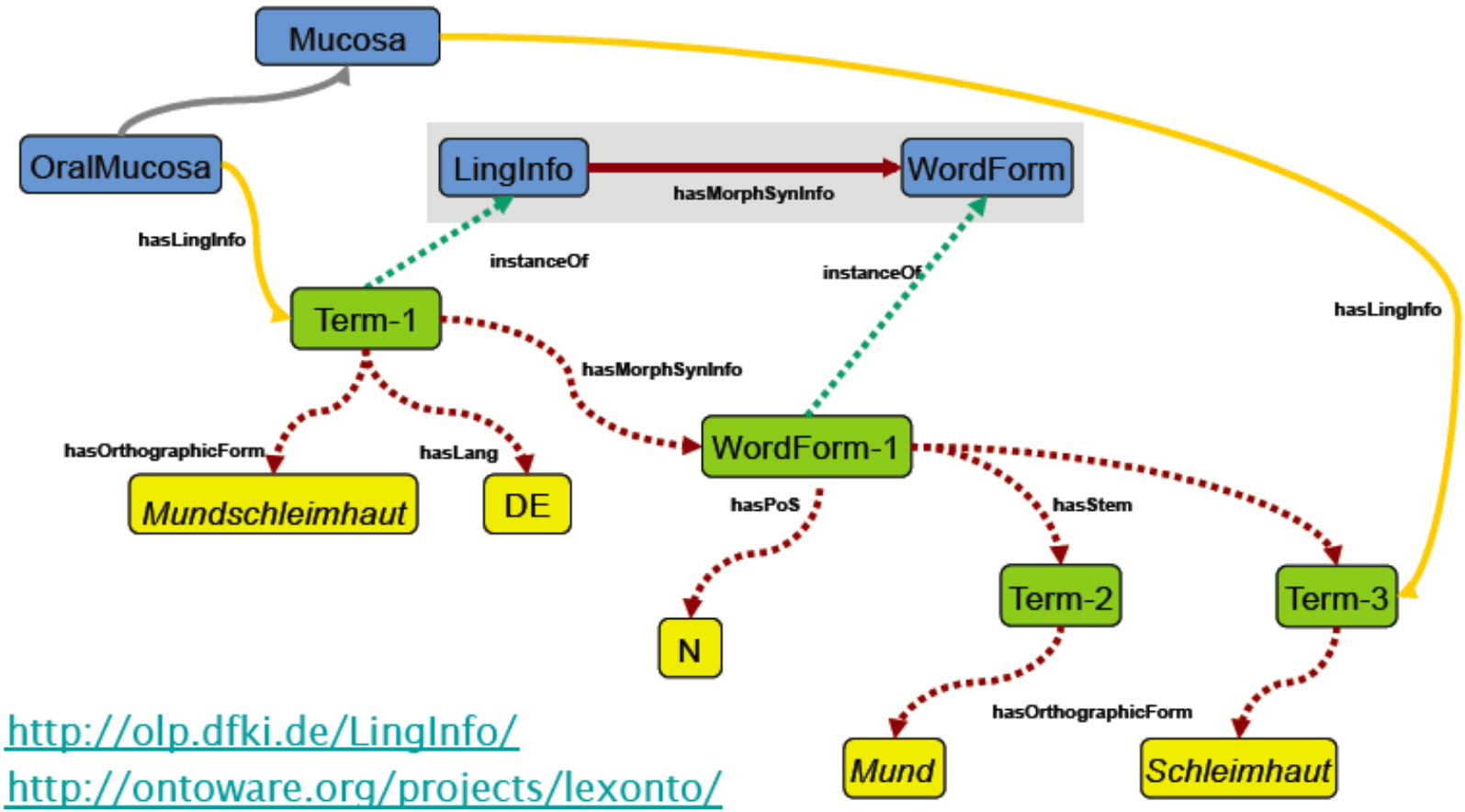
- What the previous sub-sections do not mention: they are "English ontologies" and work with natural language text in English
- How to build an ontology for, say, Spanish organic agriculture? [Organic.Lingua project] 'intelligent' eGovernment portals in the 11 official languages of South Africa?
- Multilingualism with ontologies
  - 'Ontology in different languages'?
  - NLP (NLU) for target language to learn
  - NLG for user and domain expert-friendly interface to the ontology
- Despite OWL's goal of internationalization, that has not been realised yet, and it is an active field of research

- How to create ‘ontologies in multiple languages?’
  (does that question even make sense?)
- How to manage those ontologies?
  e.g., for one subject domain, for all 11 official language of South Africa
- What to do with language peculiarities built into the current technologies?
  (can you given an example of that?)

Simple option: Semantic Tagging

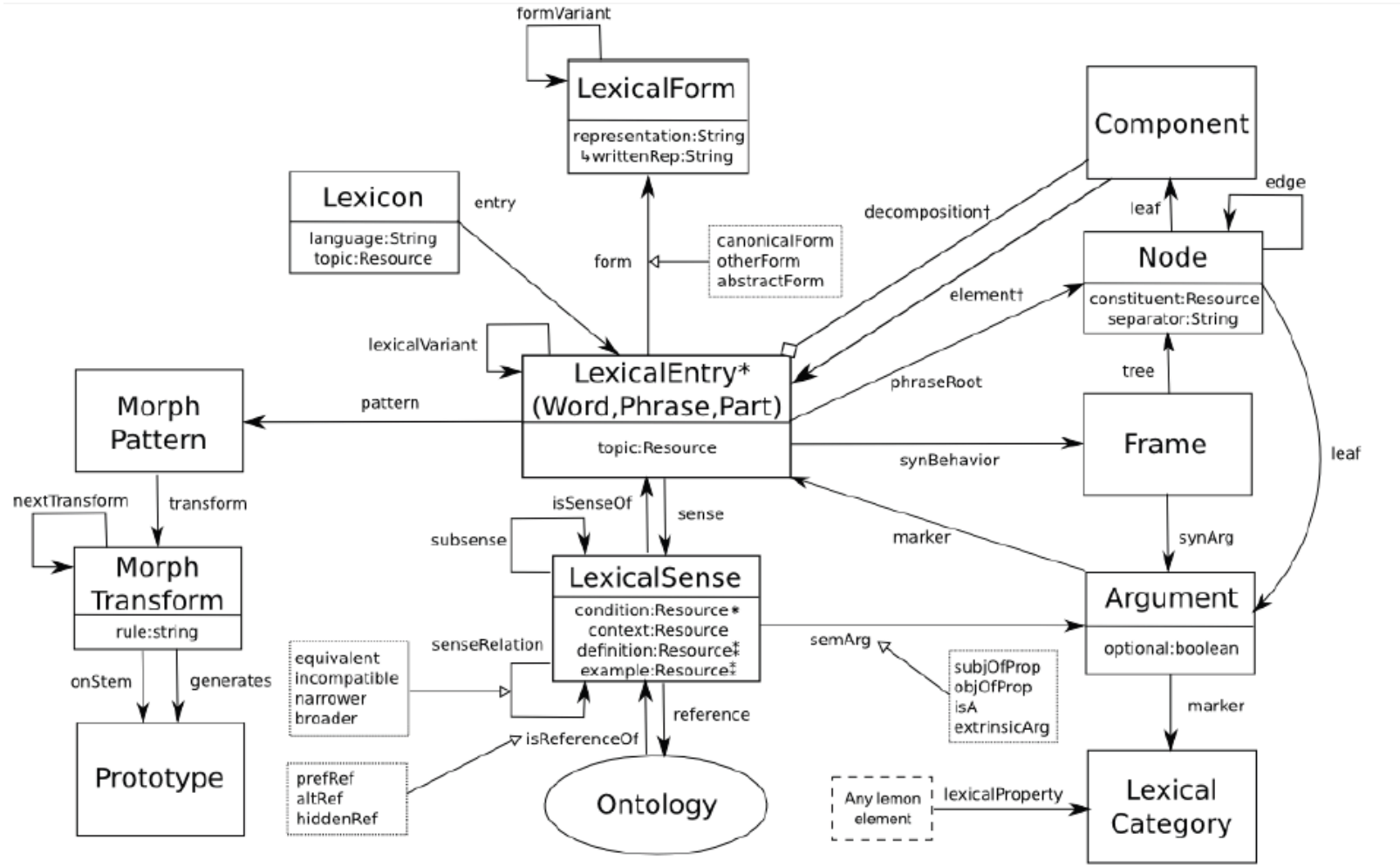


Option with some effort: Semantic Tagging with a Lexicalised Ontology



http://olp.dfki.de/LingInfo/
http://ontoware.org/projects/lexonto/

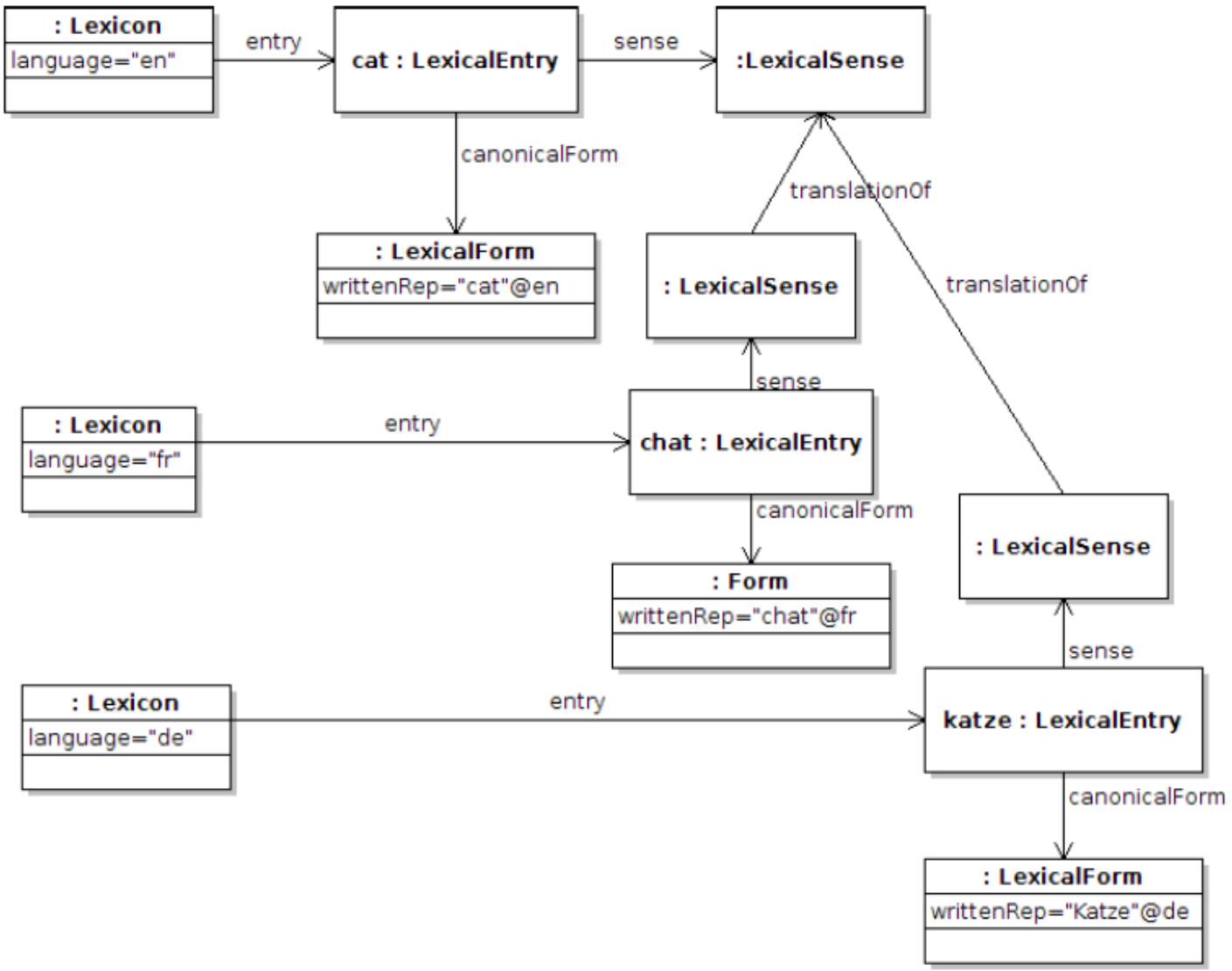
More comprehensively Lexicalised Ontologies



Lemon example

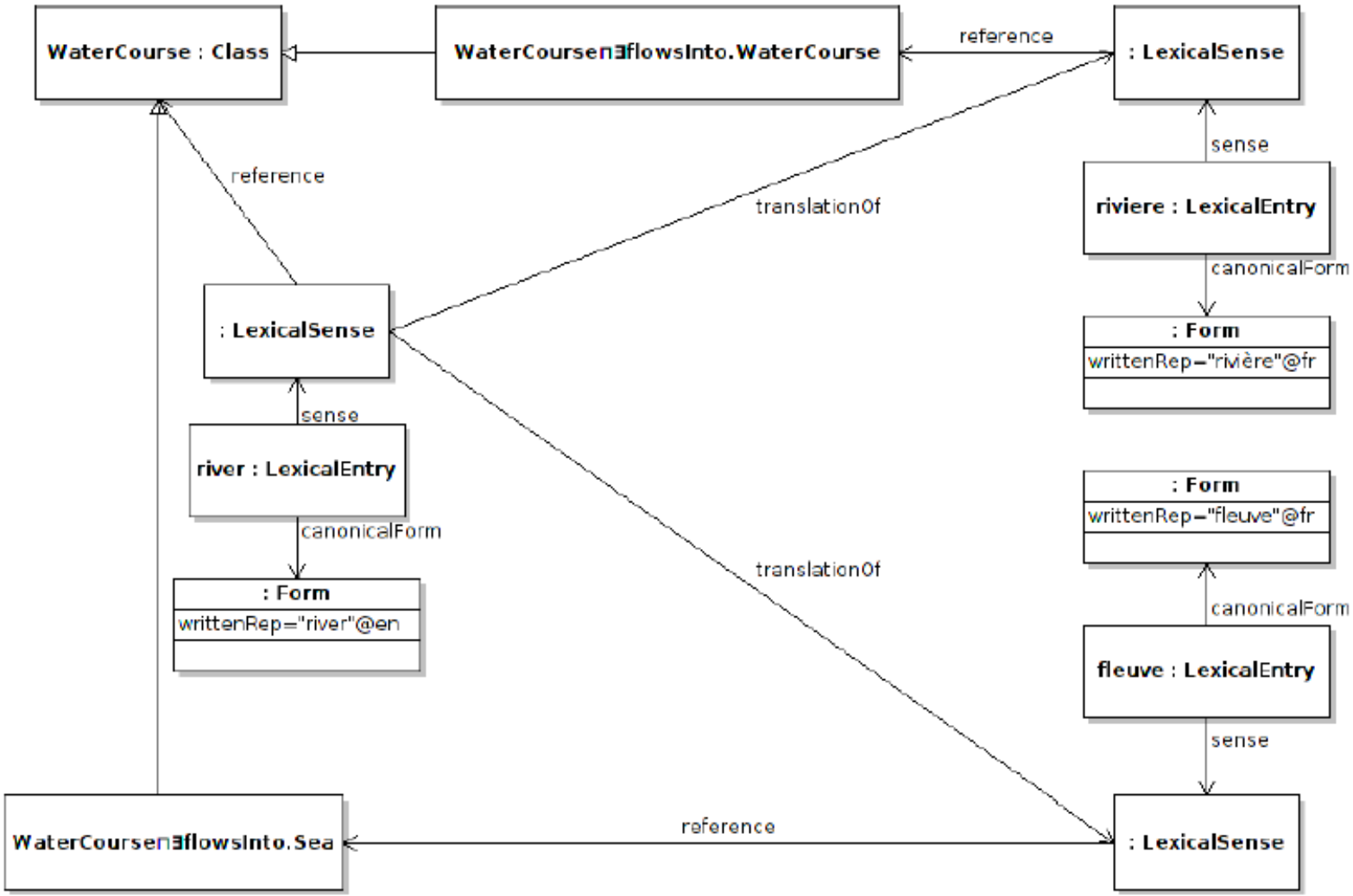
```
@base <http://www.example.org/lexicon>
@prefix ontology: <http://www.example.org/ontology#>
@prefix lemon: <http://www.monnetproject.eu/lemon#>

:myLexicon a lemon:Lexicon ;
   lemon:language "en" ;
   lemon:entry :animal .

:animal a lemon:LexicalEntry ;
   lemon:form [ lemon:writtenRep "animal"@en ] ;
   lemon:sense [ lemon:reference ontology:animal ] .
```

Lemon example



```
:lexicon_en lemon:entry :cat ;
lemon:language "en" .

:lexicon_de lemon:entry :katze ;
lemon:language "de".

:lexicon_fr lemon:entry :chat ;
lemon:language "fr".

:cat lemon:canonicalForm [ lemon:writtenRep "cat"@en ] ;
lemon:sense :cat_sense .

:chat lemon:canonicalForm [ lemon:writtenRep "chat"@fr ] ;
lemon:sense [ isocat:translationOf :cat_sense ] .

:katze lemon:canonicalForm [ lemon:writtenRep "katze"@de ] ;
lemon:sense [ isocat:translationOf :cat_sense ] .

isocat:translationOf rdfs:subPropertyOf lemon:senseRelation .
```

Semantic Tagging — Lemon example



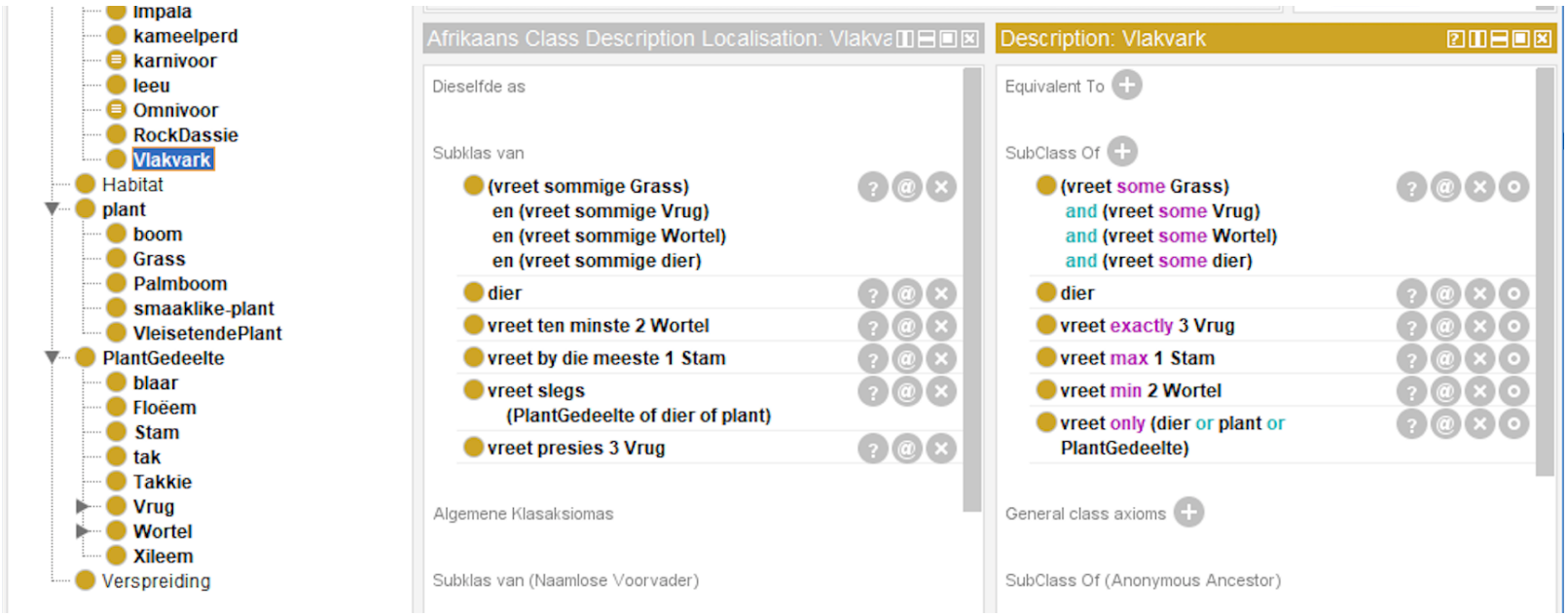
Extensions (complications) for, a.o., isiZulu

- The noun classes
- Treatment of verbs is different
  - There's no single 3rd person singular, as in English (e.g., `eats`, `teaches` vs. human eats `udla`, giraffe `idla` etc. by noun class). so no fixed string for object property name
  - The preposition (`part` *`of`* etc.) typically associates with the noun (PC or *nga*-), not verb

Extensions (complications) for, a.o., isiZulu

- The noun classes
- Treatment of verbs is different
  - There's no single 3rd person singular, as in English (e.g., `eats`, `teaches` vs. human eats `udla`, giraffe `idla` etc. by noun class). so no fixed string for object property name
  - The preposition (`part` *`of`* etc.) typically associates with the noun (PC or *nga*-), not verb
- For all languages other than English: ODE interfaces, Manchester syntax worse than useless (cognitive overload of code switching when reading an axiom)

Example of ODE issues and possible solution



Outline

1 Introduction

2 Multilingual ontologies

3 Ontology verbalisation

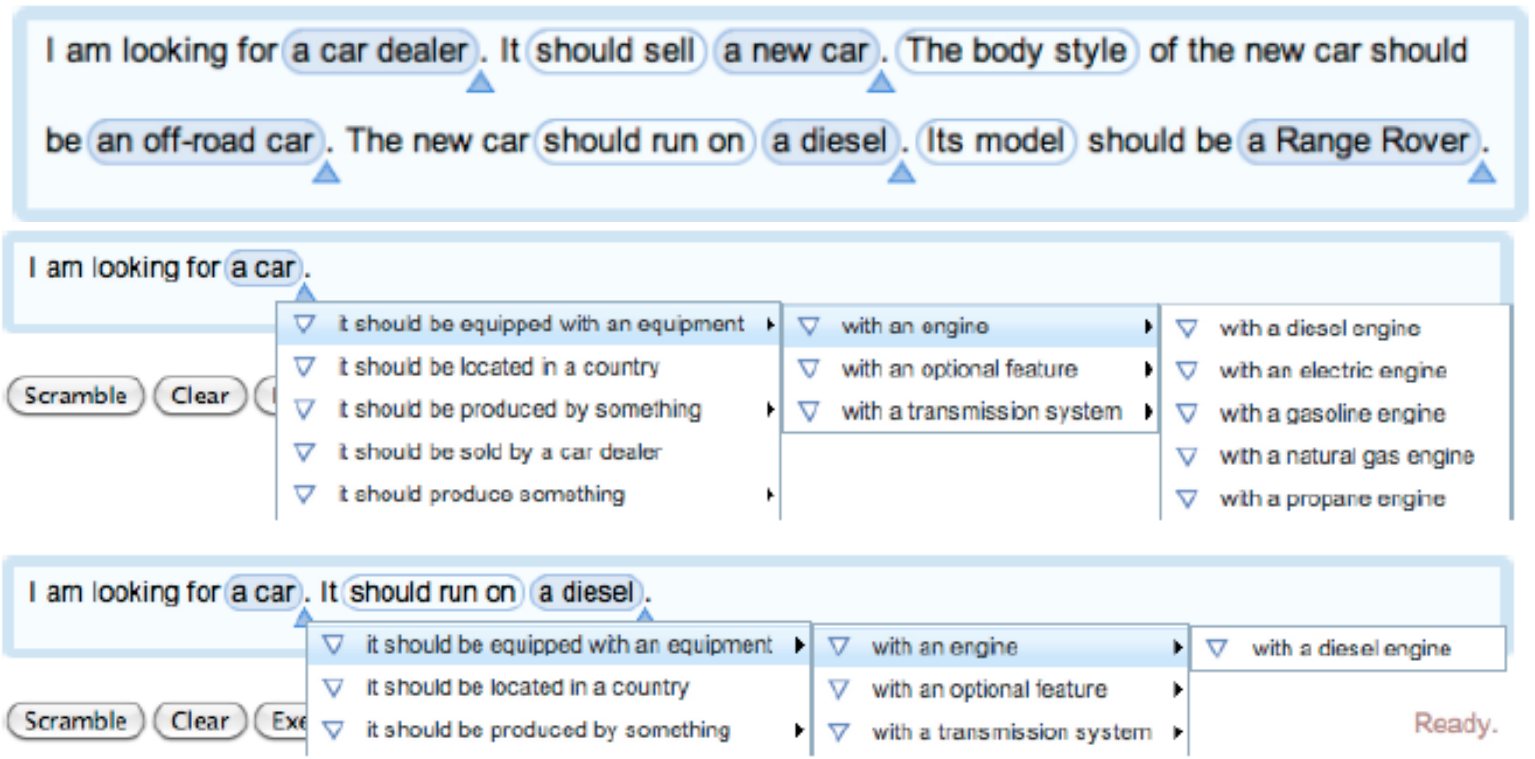
What is CNL, NLG?

- **C**ontrolled **N**aural **L**anguage: constrain the grammar/vocabulary of a natural language
- **N**atural **L**anguage **G**eneration: generate natural language text from structured data, information, or knowledge

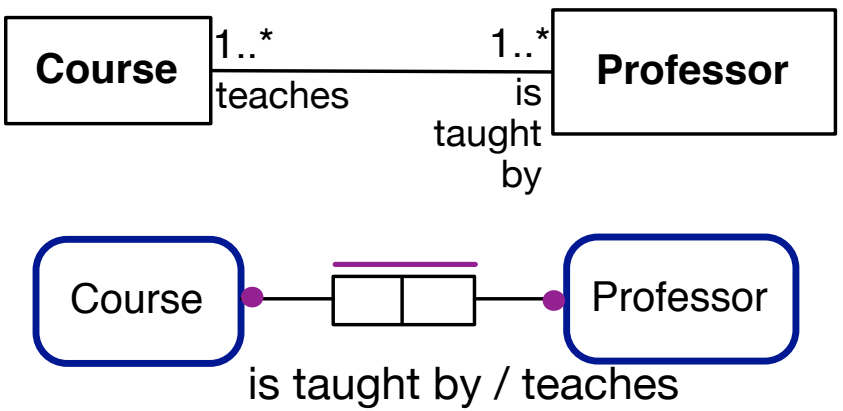
Natural language interfaces with some CNL or NLG

- Many tools, webpages, etc. with some natural language component
- Querying of information in natural language (cf. a query language SQL, SPARQL)
- Business rules typically specified in a natural language
- etc.

Example: Query formulation with Quelo [Franconi et al.(2010)]



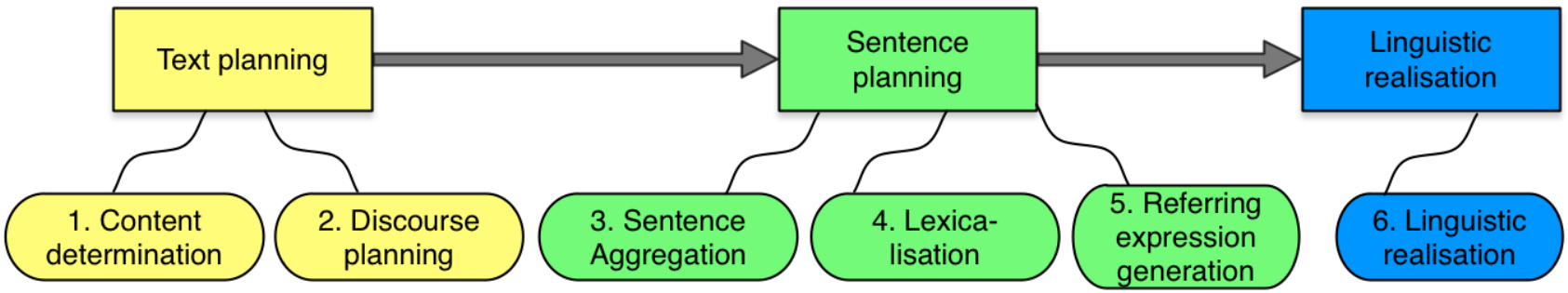
Example: Business rules and conceptual data models



Each Course is taught by **at least one** Professor

Each Professor teaches **at least one** Course

The 'NLG pipeline'



1. What structured data/info/ knowledge do you want to put into NL sentences?
2. In what order should it be presented?

3. Which messages to put together into a sentence?
4. Which words and phrases will it use for each domain concept and relation?
5. Which words or phrases to select to identify domain entities?

6. Use grammar rules to produce syntactically, morphologically, and orthographically correct (and is also meaningful)

NLG, principal approaches to generate the text

- Canned text
- Templates
  - Notably for English [Fuchs et al.(2010), Schwitter et al.(2008), Third et al.(2011), Curland and Halpin(2007)],
  - but also other languages [Jarrar et al.(2006)] (see list)
- Grammar engines, such as [Kuhn(2013)], Grammatical Framework (`http://www.grammaticalframework.org/`), SimpleNLG

NLG, principal approaches to generate the text

- Canned text
- Templates
  - Notably for English [Fuchs et al.(2010), Schwitter et al.(2008), Third et al.(2011), Curland and Halpin(2007)],
  - but also other languages [Jarrar et al.(2006)] (see list)
- Grammar engines, such as [Kuhn(2013)], Grammatical Framework (`http://www.grammaticalframework.org/`), SimpleNLG

$\Rightarrow$ CNL, NLG

Business rules/conceptual data models and logic reconstruction

BR: **Each** Course is taught by **at least one** Professor

FOL: $\forall x\ (\text{Course}(x) \rightarrow \exists y\ (\text{is\_taught\_by}(x, y) \wedge \text{Professor}(y)))$

DL: $\text{Course} \sqsubseteq \exists\ \text{is\_taught\_by}.\text{Professor}$

Example of templates

```
<Constraint xsi:type="Mandatory">
 <Text> -[Mandatory] Cada</Text>
 <Object index="0"/>
 <Text>debe</Text>
 <Role index="0"/>
 <Text>al menos un(a)</Text>
 <Object index="1"/>
</Constraint>
```

```
<Constraint xsi:type="Mandatory">
 <Text> -[Mandatory]  Each</Text>
 <Object index="0"/>
 <Text>must</Text>
 <Role index="0"/>
 <Text>at least one</Text>
 <Object index="1"/>
</Constraint>
```

for a large fragment of ORM, and 11 languages [Jarrar et al.(2006)]

Example of templates

```
<Constraint xsi:type="Mandatory">
 <Text> -[Mandatory] Cada</Text>
 <Object index="0"/>
 <Text>debe</Text>
 <Role index="0"/>
 <Text>al menos un(a)</Text>
 <Object index="1"/>
</Constraint>
```

```
<Constraint xsi:type="Mandatory">
 <Text> -[Mandatory]  Each</Text>
 <Object index="0"/>
 <Text>must</Text>
 <Role index="0"/>
 <Text>at least one</Text>
 <Object index="1"/>
</Constraint>
```

for a large fragment of ORM, and 11 languages [Jarrar et al.(2006)]

Example of templates

```
<Constraint xsi:type="Mandatory">
 <Text> -[Mandatory] Cada</Text>
 <Object index="0"/>
 <Text>debe</Text>
 <Role index="0"/>
 <Text>al menos un(a)</Text>
 <Object index="1"/>
</Constraint>
```

```
<Constraint xsi:type="Mandatory">
 <Text> -[Mandatory]  Each</Text>
 <Object index="0"/>
 <Text>must</Text>
 <Role index="0"/>
 <Text>at least one</Text>
 <Object index="1"/>
</Constraint>
```

for a large fragment of ORM, and 11 languages [Jarrar et al.(2006)]

Example of templates

```
<Constraint xsi:type="Mandatory">
 <Text> -[Mandatory] Cada</Text>
 <Object index="0"/>
 <Text>debe</Text>
 <Role index="0"/>
 <Text>al menos un(a)</Text>
 <Object index="1"/>
</Constraint>
```

```
<Constraint xsi:type="Mandatory">
 <Text> -[Mandatory]  Each</Text>
 <Object index="0"/>
 <Text>must</Text>
 <Role index="0"/>
 <Text>at least one</Text>
 <Object index="1"/>
</Constraint>
```

for a large fragment of ORM, and 11 languages [Jarrar et al.(2006)]

NL Grammars, illustration

$$
\begin{aligned}
\textit{Sentence} &\longrightarrow \textit{NounPhrase} \mid \textit{VerbPhrase} \\
\textit{NounPhrase} &\longrightarrow \textit{Adjective} \mid \textit{NounPhrase} \\
\textit{NounPhrase} &\longrightarrow \textit{Noun} \\
&\ldots \\
\\
\textit{Noun} &\longrightarrow \textit{car} \mid \textit{train} \\
\textit{Adjective} &\longrightarrow \textit{big} \mid \textit{broken} \\
&\ldots
\end{aligned}
$$

(and complexity of the grammar)

Question

- Can the template-based approach be used also for isiZulu?

Question

- Can the template-based approach be used also for isiZulu?
  - If so, create those templates
  - If not, start with basics for a grammar engine

Question

- Can the template-based approach be used also for isiZulu?
  - If so, create those templates
  - If not, start with basics for a grammar engine
- Use a practically useful language to benefit both ICT and linguists and, possibly, some subject domain (e.g., medicine)
- Details in [Keet and Khumalo(2014b), Keet and Khumalo(2014a), Keet and Khumalo(2017)]

A logic foundation for isiZulu knowledge-to-text

- Roughly OWL 2 EL
- OWL 2 EL is a W3C-standardised profile of OWL 2
- Tools, ontologies in OWL 2 (notably SNOMED CT)

Universal Quantification

- Consider here only the universal quantification at the start of the concept inclusion axiom ('nominal head')
- 'all'/'each' uses *-onke*, prefixed with the oral prefix of the noun class of that first noun (OWL class/DL concept) on lhs of $\sqsubseteq$

(U1) Boy $\sqsubseteq$ ...

wonke umfana ...	('each boy...'; *u-* + *-onke*)
bonke abafana ...	('all boys...'; *ba-* + *-onke*)

(U2) Phone $\sqsubseteq$ ...

lonke ifoni ...	('each phone...'; *li-* + *-onke*)
onke amafoni ...	('all phones...'; *a-* + *-onke*)

NC	QC (all)		NEG SC	PRON	RC	QC_{dwa}	EC
	$QC_{oral+onke}$	QC_{nke}					
1	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke → wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke → wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke → yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke → lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke → sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke → zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke → lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke → bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke → konke	zo-	aku-	khona	oku-	zo-	ku-

NC	QC (all)		NEG SC	PRON	RC	QC_{dwa}	EC
	$QC_{oral+onke}$	QC_{nke}					
1	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke → wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke → wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke → yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke → lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke → sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke → zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke → lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke → bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke → konke	zo-	aku-	khona	oku-	zo-	ku-

NC	QC_{oral}	QC (all) -onke	QC_{nke}	NEG SC	PRON	RC	QC_{dwa}	EC
1	u-onke →	wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke →	bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke →	wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke →	bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke →	wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke →	bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke →	wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke →	yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke →	lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke →	onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke →	sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke →	zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke →	yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke →	onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke →	yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke →	zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke →	lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke →	zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke →	bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke →	konke	zo-	aku-	khona	oku-	zo-	ku-

Subsumption

- Two different ways of carving up the nouns to determine which rules apply: semantic and syntactic
- Need to choose between
  - singular and plural
  - with or without the universal quantification voiced
  - generic or determinate

(S1) `MedicinalHerb` $\sqsubseteq$ `Plant`

ikhambi ngumuthi ('medicinal herb is a plant')

amakhambi yimithi ('medicinal herbs are plants')

wonke amakhambi ngumuthi ('all medicinal herbs are a plant')

(S2) `Giraffes` $\sqsubseteq$ `Animals`

izindlulamithi yizilwane ('giraffes are animals'; generic)

(S3) `Cellphone` $\sqsubseteq$ `Phone`

Umakhalekhukhwini uyifoni ('cellphone is a phone'; determ.)

Possible subsumption patterns

a. N_1 <copulative *ng*/*y* depending on first letter of N_2>N_2.

b. <plural of N_1> <copulative *ng*/*y* depending on first letter of plural of N_2><plural of N_2>.

c. <All-concord for NC_x>onke <plural of N_1, being of NC_x> <copulative *ng*/*y* depending on first letter of N_2>N_2.

Subsumption: adding negation

- Need to choose between
  - singular and plural, and with or without the universal quantification voiced
- Copulative is omitted
- Combines the negative subject concord (NEG SC) of the noun class of the first noun (*aku-*) with the pronomial (PRON) of the noun class of second noun (*-yona*)

(SN1) $\texttt{Cup} \sqsubseteq \neg\texttt{Glass}$

indebe akuyona ingilazi ('cup not a glass')

zonke izindebe aziyona ingilazi ('all cups not a glass')

NC	QC (all)		NEG SC	PRON	RC	QC_{dwa}	EC
	$QC_{oral+onke}$	QC_{nke}					
1	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke → wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke → wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke → yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke → lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke → sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke → zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke → lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke → bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke → konke	zo-	aku-	khona	oku-	zo-	ku-

NC	QC (all)		NEG SC	PRON	RC	QC_{dwa}	EC
	$QC_{oral+onke}$	QC_{nke}					
1	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke → wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke → wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke → yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke → lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke → sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke → zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke → lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke → bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke → konke	zo-	aku-	khona	oku-	zo-	ku-

NC	QC (all)		NEG SC	PRON	RC	QC_{dwa}	EC
	$QC_{oral+onke}$	QC_{nke}					
1	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke → wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke → wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke → yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke → lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke → sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke → zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke → lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke → bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke → konke	zo-	aku-	khona	oku-	zo-	ku-

Possible negation (disjointness) patterns

a. <N_1 of NC_x> <NEG SC of NC_x><PRON of NC_y> <N_2 of NC_y>.

b. <All-concord for NC_x>onke <plural N_1, being of NC_x> <NEG SC of NC_x><PRON of NC_y> <N_2 with NC_y>.

Existential Quantification

(E1) Giraffe $\sqsubseteq \exists$eats.Twig

yonke indlulamithi idla ihlamvana elilodwa	(‘each giraffe eats at least one twig’)
zonke izindlulamithi zidla ihlamvana elilodwa	(‘all giraffes eat at least one twig’)

a. <All-concord for NC_x>onke <pl. N_1, is in NC_x> <conjugated verb> <N_2 of NC_y> <RC for NC_y><QC for NC_y>dwa.

NC	QC (all)		NEG SC	PRON	RC	QC_{dwa}	EC
	$QC_{oral+onke}$	QC_{nke}					
1	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke → wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke → wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke → yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke → lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke → sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke → zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke → lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke → bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke → konke	zo-	aku-	khona	oku-	zo-	ku-

NC	QC (all)		NEG SC	PRON	RC	QC_{dwa}	EC
	$QC_{oral+onke}$	QC_{nke}					
1	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke → wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke → wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke → yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke → lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke → sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke → zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke → lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke → bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke → konke	zo-	aku-	khona	oku-	zo-	ku-

NC	QC (all)		NEG SC	PRON	RC	QC_{dwa}	EC
	$QC_{oral+onke}$	QC_{nke}					
1	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
1a	u-onke → wonke	wo-	aka-	yena	o-	ye-	mu-
2a	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3a	u-onke → wonke	wo-	aka-	wona	o-	ye-	mu-
(2a)	ba-onke → bonke	bo-	aba-	bona	aba-	bo-	ba-
3	u-onke → wonke	wo-	awu-	wona	o-	wo-	mu-
4	i-onke → yonke	yo-	ayi-	yona	e-	yo-	mi-
5	li-onke → lonke	lo-	ali-	lona	eli-	lo-	li-
6	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
7	si-onke → sonke	so-	asi-	sona	esi-	so-	si-
8	zi-onke → zonke	zo-	azi-	zona	ezi	zo-	zi-
9a	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
(6)	a-onke → onke	o-	awa-	wona	a-	wo-	ma-
9	i-onke → yonke	yo-	ayi-	yona	e-	yo-	yi-
10	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
11	lu-onke → lonke	lo-	alu-	lona	olu-	lo-	lu-
(10)	zi-onke → zonke	zo-	azi-	zona	ezi-	zo-	zi-
14	ba-onke → bonke	bo-	abu-	bona	obu-	bo-	bu-
15	ku-onke → konke	zo-	aku-	khona	oku-	zo-	ku-

Example

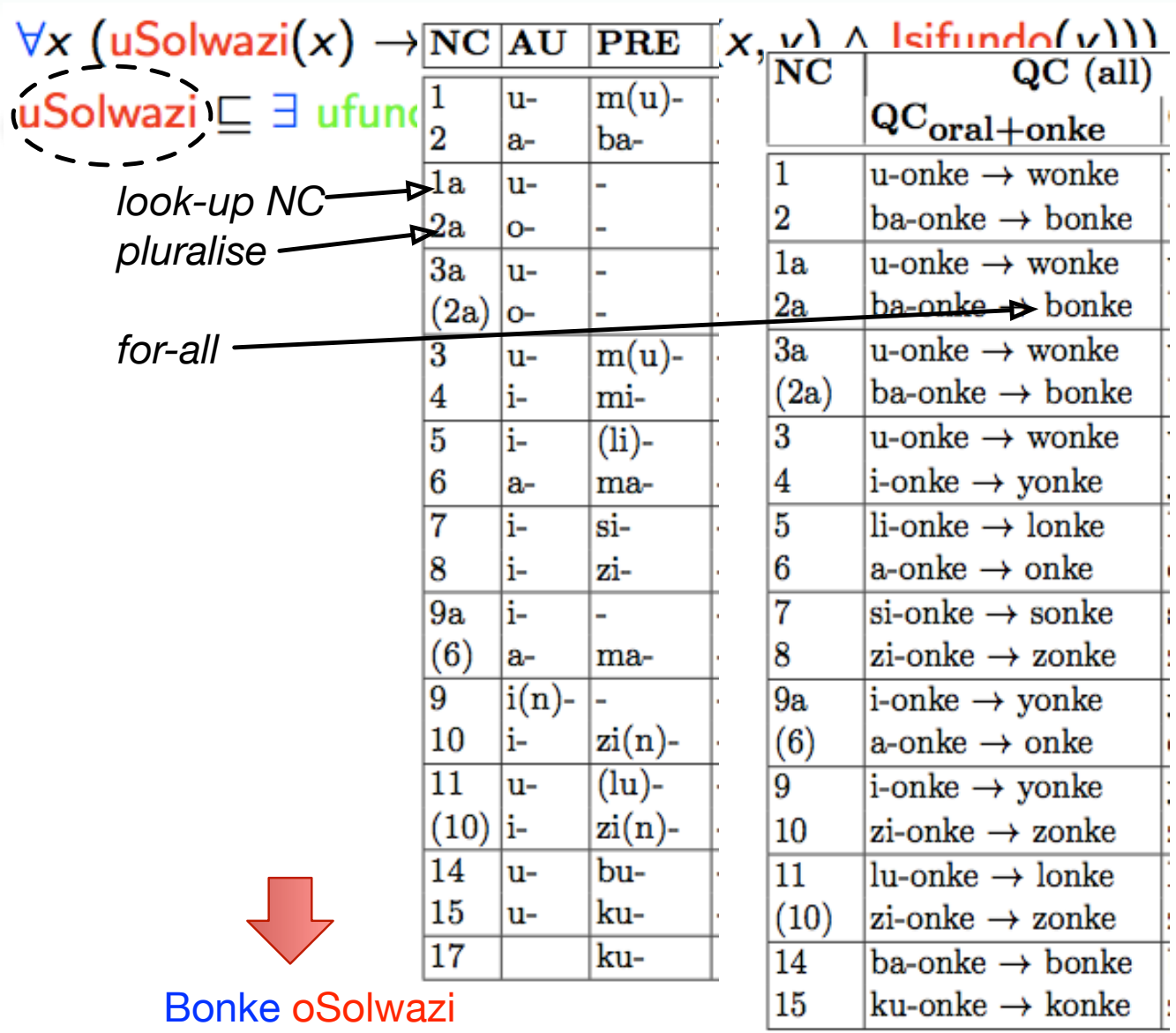
- $\forall x\ (\text{Professor}(x) \rightarrow \exists y\ (\text{teaches}(x, y) \wedge \text{Course}(y)))$
- $\text{Professor} \sqsubseteq \exists\ \text{teaches}.\text{Course}$
- **Each** Professor teaches **at least one** Course

Example

- $\forall x\ (\text{uSolwazi}(x) \rightarrow \exists y\ (\text{ufundisa}(x, y) \wedge \text{Isifundo}(y)))$
- $\text{uSolwazi} \sqsubseteq \exists\ \text{ufundisa}.\text{Isifundo}$
- ?

$$\forall x\ (\mathsf{uSolwazi}(x) \rightarrow \exists y\ (\mathsf{ufundisa}(x, y) \wedge \mathsf{Isifundo}(y)))$$

$$\mathsf{uSolwazi} \sqsubseteq \exists\ \mathsf{ufundisa}.\mathsf{Isifundo}$$



NC	AU	PRE
1	u-	m(u)-
2	a-	ba-
1a	u-	-
2a	o-	-
3a	u-	-
(2a)	o-	-
3	u-	m(u)-
4	i-	mi-
5	i-	(li)-
6	a-	ma-
7	i-	si-
8	i-	zi-
9a	i-	-
(6)	a-	ma-
9	i(n)-	-
10	i-	zi(n)-
11	u-	(lu)-
(10)	i-	zi(n)-
14	u-	bu-
15	u-	ku-
17		ku-

NC	QC (all) $QC_{oral+onke}$
1	u-onke → wonke
2	ba-onke → bonke
1a	u-onke → wonke
2a	ba-onke → bonke
3a	u-onke → wonke
(2a)	ba-onke → bonke
3	u-onke → wonke
4	i-onke → yonke
5	li-onke → lonke
6	a-onke → onke
7	si-onke → sonke
8	zi-onke → zonke
9a	i-onke → yonke
(6)	a-onke → onke
9	i-onke → yonke
10	zi-onke → zonke
11	lu-onke → lonke
(10)	zi-onke → zonke
14	ba-onke → bonke
15	ku-onke → konke

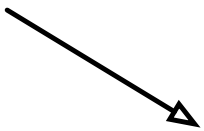
$\forall x\ (\text{uSolwazi}(x) \rightarrow \exists y\ (\text{ufundisa}(x, y) \wedge \text{lsifundo}(y)))$

$\text{uSolwazi} \sqsubseteq \exists(\text{ufundisa}.\text{l}$

... for relevant NC. Here:

ngi-
u-
u-
si-
ni-
ba-

AlgoConjugate



Bonke oSolwazi bafundisa

$$\forall x\ (\text{uSolwazi}(x) \rightarrow \exists y\ (\text{ufundisa}(x, y) \wedge \text{Isifundo}(y)))$$

$$\text{uSolwazi} \sqsubseteq \exists\ \text{ufundisa}.\text{Isifundo}$$

Bonke oSolwazi bafundisa Isifundo

$\forall x\ (\text{uSolwazi}(x) \rightarrow \exists y\ (\text{ufundisa}(\ldots) \wedge \ldots))$

$\text{uSolwazi} \sqsubseteq \exists\ \text{ufundisa}.\text{Is}\ldots$

look-up NC

get RC

get QC

add -dwa

NC	AU	PRE
1	u-	m(u)-
2	a-	ba-
1a	u-	-
2a	o-	-
3a	u-	-
(2a)	o-	-
3	u-	m(u)-
4	i-	mi-
5	i-	(li)-
6	a-	ma-
7	i-	si-
8	i-	zi-
9a	i-	-
(6)	a-	ma-
9	i(n)-	-
10	i-	zi(n)-
11	u-	(lu)-
(10)	i-	zi(n)-
14	u-	bu-
15	u-	ku-
17		ku-

RC	QC_{dwa}
o-	ye-
aba-	bo-
o-	ye-
aba-	bo-
o-	ye-
aba-	bo-
o-	wo-
e-	yo-
eli-	lo-
a-	wo-
esi-	so-
ezi	zo-
e-	yo-
a-	wo-
e-	yo-
ezi-	zo-
olu-	lo-
ezi-	zo-
obu-	bo-
oku-	zo-

Bonke oSolwazi bafundisa Isifundo esisodwa

example

(1) $\texttt{Grandmother} \sqsubseteq \exists \texttt{eats.Apple}$

bonke ogogo badla i-aphula elilodwa

Each grandmother eats at least one apple

(2) $\texttt{Human} \sqsubseteq \exists \texttt{hasPart.Hearth}$

bonke abantu banenhliziyo eyodwa

Each human has part some heart

(3) $\texttt{Herbivore} \sqsubseteq \neg \texttt{Carnivore}$

Onke amahebhivo awalona ikhanivo

Each herbivore is not a carnivore

How to evaluate?

- Typical way of evaluating: ask linguists and/or intended target group
- Questions depend on what you want to know; e.g.,
  - Does the text capture the semantics adequately?
  - Must it really be grammatically correct or is understandable also acceptable?
  - Compared against alternate representation (figures, tables) or human-authored text?

How to evaluate?

- Typical way of evaluating: ask linguists and/or intended target group
- Questions depend on what you want to know; e.g.,
  - Does the text capture the semantics adequately?
  - Must it really be grammatically correct or is understandable also acceptable?
  - Compared against alternate representation (figures, tables) or human-authored text?
- Survey, asked linguists and non-linguists for their preferences
- 10 questions pitting the patterns against each other
- Online, with isiZulu-localised version of Limesurvey

Summary

1. Introduction
2. Multilingual ontologies
3. Ontology verbalisation

References I

Sonja E. Bosch and Roald Eisele.

The effectiveness of morphological rules for an isiZulu spelling checker.
South African Journal of African Languages, 25(1):25–36, 2005.

M. Curland and T. Halpin.

Model driven development with NORMA.
In *Proceedings of the 40th International Conference on System Sciences (HICSS-40)*, pages 286a–286a.
IEEE Computer Society, 2007.
Los Alamitos, Hawaii.

Enrico Franconi, Paolo Guagliardo, and Marco Trevisan.

An intelligent query interface based on ontology navigation.
In *Workshop on Visual Interfaces to the Social and Semantic Web (VISSW'10)*, 2010.
Hong Kong, February 2010.

Norbert E. Fuchs, Kaarel Kaljurand, and Tobias Kuhn.

Discourse Representation Structures for ACE 6.6.
Technical Report ifi-2010.0010, Department of Informatics, University of Zurich, Zurich, Switzerland, 2010.

Mustafa Jarrar, C. Maria Keet, and Paolo Dongilli.

Multilingual verbalization of ORM conceptual models and axiomatized ontologies.
Starlab technical report, Vrije Universiteit Brussel, Belgium, February 2006.
URL `http://www.meteck.org/files/ORMmultiverb_JKD.pdf`.

C. M. Keet and L. Khumalo.

Toward a knowledge-to-text controlled natural language of isiZulu.
Language Resources and Evaluation, 51(1):131–157, 2017.
doi: 10.1007/s10579-016-9340-0.

References II

C. Maria Keet and Langa Khumalo.

Toward verbalizing logical theories in isiZulu.
In B. Davis, T. Kuhn, and K. Kaljurand, editors, *Proceedings of the 4th Workshop on Controlled Natural Language (CNL'14)*, volume 8625 of *LNAI*, pages 78–89. Springer, 2014a.
20-22 August 2014, Galway, Ireland.

C. Maria Keet and Langa Khumalo.

Basics for a grammar engine to verbalize logical theories in isiZulu.
In A. Bikakis et al., editors, *Proceedings of the 8th International Web Rule Symposium (RuleML'14)*, volume 8620 of *LNCS*, pages 216–225. Springer, 2014b.
August 18-20, 2014, Prague, Czech Republic.

Langa Khumalo.

Advances in developing corpora in African languages.
Kuwala, 1(2):21–30, 2015.

Tobias Kuhn.

A principled approach to grammars for controlled natural languages and predictive editors.
Journal of Logic, Language and Information, 22(1):33–70, 2013.

B. Ndaba, H. Suleman, C. M. Keet, and L. Khumalo.

The effects of a corpus on isizulu spellcheckers based on n-grams.
In Paul Cunningham and Miriam Cunningham, editors, *IST-Africa 2016*. IIMC International Information Management Corporation, 2016.
11-13 May, 2016, Durban, South Africa.

References III

D. J. Prinsloo and G.-M. de Schryver.

Spellcheckers for the south african languages, part 2: the utilisation of clusters of circumfixes.
South African Journal of African Languages, §:83–94, 2004.

R. Schwitter, K. Kaljurand, A. Cregan, C. Dolbear, and G. Hart.

A comparison of three controlled natural languages for OWL 1.1.
In *Proc. of OWLED 2008 DC*, 2008.
Washington, DC, USA metropolitan area, on 1-2 April 2008.

Sebastian Spiegler, Andrew van der Spuy, and Peter A. Flach.

Ukwabelana – an open-source morphological zulu corpus.
In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING'10)*, pages 1020–1028. Association for Computational Linguistics, 2010.
Beijing.

Allan Third, Sandra Williams, and Richard Power.

OWL to English: a tool for generating organised easily-navigated hypertexts from ontologies.
poster/demo paper, Open Unversity UK, 2011.
10th International Semantic Web Conference (ISWC'11), 23-27 Oct 2011, Bonn, Germany.