

ProcedureProtocol

**Protocol: Morphological Analysis (isiZulu)Protocol: Morphological Analysis (isiXhosa)**

**X.Y.Z1.0.3**

**Name of AuthorW. Pienaar**

DOCUMENT INFORMATION

|  |  |
| --- | --- |
| **Filename** | Protocol.SADiLaR.MorphologicalAnalysisIsiXhosa.1.0.2.WP.2018-11-22.docx |
| **Type** | Protocol |
| **Title** | Protocol: Morphological Analysis (isiXhosa) |
| **Version** | 1.0.3 |
| **Source Author(s)** | W. Pienaar |
| **Commentator(s)** | M.J. Puttkammer |
| **Status** | Document Status |
| **Date** | 2018-12-04 |
| **Pages** | 21 |
| **Security Classification** | Public Domain Only Source Author(s) Only Distributors Only NWU |

|  |  |
| --- | --- |
| **Distribution** |  |
| **Key Words** |  |
| **Project** | SADiLaR |

DOCUMENT HISTORY

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Status** | **Author/Commentator** |
| Version of document when opened | Date of document as specified by previous author | Status of document when opened | Name of previous author(s) or commentator(s) |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

TABLE OF CONTENTS

[Introduction 6](#_Toc531613805)

[Aim of this protocol 6](#_Toc531613806)

[Concept and definitions 6](#_Toc531613807)

[CATEGORY 1: Nouns 8](#_Toc531613808)

[CATEGORY 2: Verbs and auxiliary verbs 8](#_Toc531613809)

[CATEGORY 3: Adjectives & Relatives 10](#_Toc531613810)

[CATEGORY 4: Adverbs ([Adv]) 11](#_Toc531613811)

[CATEGORY 5: Subject concords ([SC]) 12](#_Toc531613812)

[CATEGORY 6: Object concords ([OC]) 13](#_Toc531613813)

[CATEGORY 7: Demonstratives ([Dem]) 13](#_Toc531613814)

[CATEGORY 8: Quantitatives 14](#_Toc531613815)

[CATEGORY 9: Pronouns 15](#_Toc531613816)

[CATEGORY 10: Copulatives 16](#_Toc531613817)

[CATEGORY 11: Particles and prepositions 16](#_Toc531613818)

[CATEGORY 12: Secondary morphemes 17](#_Toc531613819)

[CATEGORY 13: Ideophones ([Ideoph]) 17](#_Toc531613820)

[CATEGORY 14: Possessive concords 17](#_Toc531613821)

[CATEGORY 15: Interjections ([Interjec]) 17](#_Toc531613822)

[CATEGORY 16: Conjunctions ([Conj]) 18](#_Toc531613823)

[CATEGORY 17: Interrogatives 18](#_Toc531613824)

[CATEGORY 18: Abbreviations ([Abbr]) 18](#_Toc531613825)

[CATEGORY 19: Demonstrative Copulatives ([DemCop]) 19](#_Toc531613826)

[CATEGORY 20: Enumeratives 19](#_Toc531613827)

[CATEGORY 21: Tense markers 20](#_Toc531613828)

[CATEGORY 22: Foreign ([Foreign]) 20](#_Toc531613829)

[CATEGORY 23: Proper name ([ProperName]) 20](#_Toc531613830)

[General principles 21](#_Toc531613831)

[Changing the protocol 21](#_Toc531613832)

[Authoritative references 21](#_Toc531613833)

[Roles and contact details 21](#_Toc531613834)

# Introduction

## Aim of this protocol

The aim of this protocol is to guide annotators in the process of morphological analysis for the development of text resources for the SADiLaR project.

This document explicates some general principles to be followed in annotating the corpus, explains the process of annotation, and provides a detailed explanation of the tag set to be used.

This is a dynamic, “living” document, and should be adapted following the procedure as explicated in the section “Changing the protocol”.

## Concept and definitions

The aim of annotation on the morphological analysis layer is to identify the morphemes of each token (specifically morphologically complex words), as identified on the tokenisation layer (i.e. usually words, multi-word expressions, and punctuation). The following definitions apply:

* **Morphological analysis**: The identification and description of morphemes in complex words.
* **Morpheme**: A simplex symbolic (i.e. (grammatically) meaningful) unit in the language system. It is simplex in the sense that it does not contain smaller symbolic units as subparts; e.g. a morpheme, like the **class pre-prefix *i-*** and ***the prefix -i-*** of class 10 (‘iinkomo’).

<Protocol to be updated in terms of pre-prefixes:

**The question is how fine grained do we want to go**? The pre-prefix has a semantic value of definitive / non-definitive.

Nominal prefixes in isiXhosa are sub-divided into preprefixes and basic prefixes. This sub-division is important because in some instances the preprefix is deleted, e.g. *umfana lo* vs. *lo mfana*}

Morphemes in isiXhosa include nominal pre-prefixes and prefixes (e.g. Classes 1-17), verbal suffixes (e.g. -is-, -el-, -an-), subject and object concord, etc.>

* **Root:** The verbal stem without any affixes or extension (e.g. *-theng-)*
* **Stem**:
  + Nominal: nominal stem without affixes (e.g. *-fazi*)
  + Verbal: root with affixes and extension.(e.g*. –thenga, -thengela, -zithengela)*

Main categories for morphological analysis:

|  |
| --- |
| CATEGORY 1: Nouns |
| CATEGORY 2: Verbs and auxiliary verbs |
| CATEGORY 3: Adjectives & Relatives |
| CATEGORY 4: Adverbs |
| CATEGORY 5: Subject concords |
| CATEGORY 6: Object concords |
| CATEGORY 7: Demonstratives |
| CATEGORY 8: Quantitatives |
| CATEGORY 8: Pronouns |
| CATEGORY 9: Copulatives |
| CATEGORY 10: Particles and prepositions |
| CATEGORY 11: Secondary morphemes |
| CATEGORY 12: Ideophones |
| CATEGORY 13: Possessive concords |
| CATEGORY 14: Interjections |
| CATEGORY 15: Conjunctions |
| CATEGORY 16: Interrogatives |
| CATEGORY 17: Abbreviations |
| CATEGORY 18: Demonstrative Copulatives |
| CATEGORY 19: Enumeratives |
| CATEGORY 20: Tense markers |
| CATEGORY 23: Foreign |
| CATEGORY 24: Proper name |

# CATEGORY 1: Nouns

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Class nr.** | **Example from each class** | **Class prefix**  **([BPre] / [NPre] /**  **[NPrePre])** | **Nominal or verbal stem**  **([NStem])** | **Suffix**  **Aug**  **([AugSuf])** | **Suffix**  **Dim**  **([DimSuf])** | **Prefix e-**  **+**  **Suffix Loc**  **([LocSuf])** |
| Class 1 | umntu | u-m- | -ntu |  | umntwana | kumntwana |
| Class 1a | tata | -Ø- | -tata |  |  | kutata |
| Class 2 | abantu | ba- | -ntu |  | abantwana | kwabantwana |
| Class 2a | ootata | o-o- | -tata |  |  |  |
| Class 3 | umthi | u-m- | -thi |  |  | emthini |
| Class 4 | imithi | i-mi- | -thi |  |  | emithini |
| Class 5 | ilitye | i-li- | -tye |  | ilityana |  |
| Class 6 | amatye | a-ma- | -tye |  |  |  |
| Class 7 | isihlangu | i-si- | -hlangu |  |  | esihlangweni |
| Class 8 | izihlangu | i-zi- | -hlangu |  |  | ezihlangweni |
| Class 9 | intaba | i-Ø (N) - | -taba |  |  | entabeni |
| Class 10 | iintaba | i-i (N)- | -taba |  |  | eentabeni |
| Class 11 | usana | u-(lu)- | -sana |  |  | kusana |
| Class 14 | ubuso | u-(bu)- | -so |  |  | ebusweni |
| Class 15 | ukuhamba | u-ku- | -hamba |  |  |  |
| Class 16 |  |  |  |  |  |  |
| Class 17 | phakathi | -Ø- |  |  |  |  |

# CATEGORY 2: Verbs and auxiliary verbs

|  |  |  |
| --- | --- | --- |
| Verbal root | [VRoot] | -hamb-, -bon- |
|  |  |  |
| Intransitive verbal root | [Intrans] | -bel-, -mis- |
|  |  |  |
| Verb terminative | [VerbTerm] | -a, -e, -i |

Prefixes:

|  |  |  |
| --- | --- | --- |
| Reflexive prefix | [ReflPre] | -zi- |
| Imperative prefix | [ImpPre] | -yi- |
| Locative prefix | [LocPre] | -ku- |
| Progressive prefix | [ProgPre] | -sa- |
| Pre-locative “s” | [PreLoc-s] | -s- |
| Bound Object concord 1st person Singular | n/a | n/a |
| Bound Object concord Class 1 | n/a | n/a |
| Mono-syllabic marker *e-* | n/a | n/a |

Verbal suffixes:

The following tags are present in the data, use the specific tag in the table below:

|  |  |  |
| --- | --- | --- |
| Neuter | [NeutExt] | -ek- / -akal- |
| Intensive | [IntensExt] | -isis- |
| Reciprocal | [RecipExt] | -an- |
| Applicative | [ApplExt] | -el- |
| Causative | [CausExt] | -is- |
| Passive | [PassExt] | -iw- / -w- |

The following tags exist in isiXhosa, use the specific tag in the table below if possible, otherwise use [VerbExt].

|  |  |  |
| --- | --- | --- |
| Associative | [Ass] | (see Reciprocal) |
| Iterative | [Ite] | -ezel- |
| Imperative suffix | [ImpSuf] | -ni |
| Dispersive | [Dis] | -alal- |
| Positional | [Pos] | -am- |
| Alternative causative | [ACau] | -z- /-s- |
| Contactive | [Con] | -ath- |
| Neutro passive | [NPas] | -ek- |
| Perfect | [Per] | -il- |
| Reversive intransitive | [RIntrans] | -uk- |
| Reversive transitive | [RTrans] | -ul- /-ulul- |

Auxiliary verbs ([AuxVStem])

|  |
| --- |
| Verb stem |
| -ba |
| -hle |
| -andula |
| -buya |
| -fana |
| -mana |
| -phinda |

Copulative verbs ([Cop])

|  |
| --- |
| Particles |
| ng- |
| y- |
| aku- |
| u- (See subject concords) |
| zi - (See subject concords) |
|  |
| Verbs |
| -ba |

# CATEGORY 3: Adjectives & Relatives

Adjective stems ([AdjStem])

|  |
| --- |
| -bi |
| -bini |
| -dala |
| -de |
| -fuphi |
| -futshane |
| -hlanu |
| -hle |
| -khulu |
| -ncinci |
| -ne |
| -ngaphi |
| -ninzi |
| -nye |
| -thandathu |
| -thathu |
| -tsha |

Adjective prefix ([AdjPref])

|  |  |
| --- | --- |
| **Persons & classes** | **Adjective prefix** |
| Class 1 | om- |
| Class 1a | om- |
| Class 2 | aba- |
| Class 2a | aba- |
| Class 3 | om- |
| Class 4 | emi- |
| Class 5 | eli- |
| Class 6 | ama- |
| Class 7 | esi- |
| Class 8 | ezi- |
| Class 9 | e- |
| Class 10 | ezi- |
| Class 11 | olu- |
| Class 14 | obu- |
| Class 15 | oku- |
| Class 16 |  |
| Class 17 | oku- |

Relative stems ([RelStem])

|  |
| --- |
| -banzi |
| -dikidiki |
| -bomvu |
| -lula |
| -krakra |
| -shushu |
| -mnyama |
| -luhlaza |
| -manzi |

Relative concord ([RelConc])

|  |  |
| --- | --- |
| Class 1 | o- / u- |
| Class 1a | o- / u- |
| Class 2 | ba- |
| Class 2a | ba- |
| Class 3 | o- / u- |
| Class 4 | e- / i- |
| Class 5 | li- |
| Class 6 | a- |
| Class 7 | (e)si- |
| Class 8 | (e)zi- |
| Class 9 | e- / i- |
| Class 10 | (e)zi- |
| Class 11 | lu- |
| Class 14 | bu- |
| Class 15 | ku- |
| Class 16 |  |
| Class 17 | ku- |

Relative suffix ([RelSuf])

|  |
| --- |
| -yo |

# CATEGORY 4: Adverbs ([Adv])

Basic adverbs

|  |  |
| --- | --- |
| njalo | ‘always’ |
| nje | 'thus/in this manner' |
| impela | ‘truly’ |

Derived adverbs

|  |  |
| --- | --- |
| kabini | ‘twice’ |
| kakhulu | ‘greatly’ |

Nouns used as adverbs

|  |
| --- |
| kumuntu |
| ngabantu |
| njengomuthi |

Adverb prefixes ([AdvPre])

|  |
| --- |
| ka- |
| na- |
| nga- |

# CATEGORY 5: Subject concords ([SC])

|  |  |  |  |
| --- | --- | --- | --- |
| **Persons and classes** |  | **Std** | **Consecutive** |
| 1st Sing. |  | ndi- | nda- |
| 1st Pl. |  | si- | sa- |
| 2nd Sing. |  | u- | wa- |
| 2nd Pl. |  | ni- | na- |
| Class 1 | umntu | u- | wa- |
| Class 1a | tata | u- | wa- |
| Class 2 | abantu | ba- | ba- |
| Class 2a | ootata | ba- | ba- |
| Class 3 | umthi | u- | wa- |
| Class 4 | imithi | i- | ya- |
| Class 5 | ilitye | li- | la- |
| Class 6 | amatye | a- | a- |
| Class 7 | isihlangu | si- | sa- |
| Class 8 | izihlangu | zi- | za- |
| Class 9 | intaba | i- | ya- |
| Class 10 | iintaba | zi- | za- |
| Class 11 | usana | lu- | lwa- |
| Class 14 | ubuso | bu- | ba- |
| Class 15 | ukuhamba | ku- | kwa- |
| Class 16 |  |  |  |
| Class 17 | phakathi | ku- | kwa- |

Indefinite and classless

|  |  |
| --- | --- |
| Indefinite | ku- (e.g. *kuyashisa, kuyabanda* |
| Classless | <occurrence in isiXhosa to be discussed> |

# CATEGORY 6: Object concords ([OC])

|  |  |
| --- | --- |
| **Persons & classes** | **Object**  **concord** |
| 1st p.s. | -ndi- |
| 1st p.p. | -si- |
| 2nd p.s. | -ku- |
| 2nd p.p. | -ni- |
| Class 1 | -m- |
| Class 1a | -m- |
| Class 2 | -ba- |
| Class 2a | -ba- |
| Class 3 | -wu- |
| Class 4 | -yi- |
| Class 5 | -li- |
| Class 6 | -wa- |
| Class 7 | -si- |
| Class 8 | -zi- |
| Class 9 | -yi- |
| Class 10 | -zi- |
| Class 11 | -lu- |
| Class 14 | -bu- |
| Class 15 | -ku- |
| Class 16 |  |
| Class 17 | -ku- |

# CATEGORY 7: Demonstratives ([Dem])

|  |  |  |  |
| --- | --- | --- | --- |
| **Persons & classes** | **Demonstratives** | | |
|  | **Pos 1 ‘*this’*** | **Pos 2 ‘*that’*** | **Pos 3 ‘*yonder’*** |
| Class 1 | lo | lowo (loo) | lowa (laa) |
| Class 1a | lo | lowo (loo) | lowa (laa) |
| Class 2 | aba | abo | abaya (abaa) |
| Class 2a | aba | abo | abaya (abaa) |
| Class 3 | lo | lowo (loo) | lowa (laa) |
| Class 4 | le | leyo (loo) | leya (laa) |
| Class 5 | eli | elo | eliya (elaa) |
| Class 6 | la | lawo (loo) | lawa (laa) |
| Class 7 | esi | eso | esiya (esaa) |
| Class 8 | ezi | ezo | eziya (ezaa) |
| Class 9 | le | leyo (loo) | leya (laa) |
| Class 10 | ezi | ezo | eziya (ezaa) |
| Class 11 | olu | olo | oluya (olwaa) |
| Class 14 | obu | obo | obuya (obaa) |
| Class 15 | oku | oko | okuya (okwaa) |
| Class 16 |  |  |  |
| Class 17 | apha | apho | phaya |

# CATEGORY 8: Quantitatives

Quantitative Stem ([QuantStem])

E.g. -dwa, -nke

Quantitative Concord ([QuantConc])

|  |  |  |
| --- | --- | --- |
| **Persons & classes** | **Quantifiers** | |
| **Inclusive** | **Exclusive** |
| 1st p.s. | ndo- | ndo- / nde- |
| 1st p.p. | so- | so- / se- |
| 2nd p.s. | wo- | we- |
| 2nd p.p. | no- | no- / ne- |
| Class 1 | wo- | ye- |
| Class 1a | wo- | ye- |
| Class 2 | bo- | bo- |
| Class 2a | bo- | bo- |
| Class 3 | wo- | wo- |
| Class 4 | yo- | yo- |
| Class 5 | lo- | lo- |
| Class 6 | o- | o- |
| Class 7 | so- | so- |
| Class 8 | zo- | zo- |
| Class 9 | yo- | yo- |
| Class 10 | zo- | zo- |
| Class 11 | lo- | lo- |
| Class 14 | bo- | bo- |
| Class 15 | ko- | ko- |
| Class 16 |  |  |
| Class 17 | ko- | ko- |

# CATEGORY 9: Pronouns

Emphatic

Absolute Pronouns ([AbsPron])

|  |  |
| --- | --- |
| **Persons & classes** | **Absolute Pronoun** |
| 1st p.s. | mina |
| 1st p.p. | thina |
| 2nd p.s. | wena |
| 2nd p.p. | nina |
| Class 1 | yena |
| Class 1a | yena |
| Class 2 | bona |
| Class 2a | bona |
| Class 3 | wona |
| Class 4 | yona |
| Class 5 | lona |
| Class 6 | wona |
| Class 7 | sona |
| Class 8 | zona |
| Class 9 | yona |
| Class 10 | zona |
| Class 11 | lona |
| Class 14 | bona |
| Class 15 | khona |
| Class 16 |  |
| Class 17 | khona |

Possessive Pronouns ([PossPron])

|  |  |
| --- | --- |
| **Persons & classes** | **Possessive Pronoun** |
| Class 1 | wakho |
| Class 1a | wakho |
| Class 2 | bakho |
| Class 2a | bakho |
| Class 3 | wakho |
| Class 4 | yakho |
| Class 5 | lakho |
| Class 6 | akho |
| Class 7 | sakho |
| Class 8 | zakho |
| Class 9 | yakho |
| Class 10 | zakho |
| Class 11 | lwakho |
| Class 14 | bakho |
| Class 15 | kwakho |
| Class 17 | kwakho |

Pronoun Stem ([PronStem])

|  |  |
| --- | --- |
| **Persons & classes** | **Pronoun Stem** |
| 1st p.s. | mi- |
| 1st p.p. | thi- |
| 2nd p.s. | we- |
| 2nd p.p. | ni- |
| Class 1 | ye- |
| Class 1a | ye- |
| Class 2 | bo- |
| Class 2a | bo- |
| Class 3 | wo- |
| Class 4 | yo- |
| Class 5 | lo- |
| Class 6 | wo- |
| Class 7 | so- |
| Class 8 | zo- |
| Class 9 | yo- |
| Class 10 | zo- |
| Class 11 | lo- |
| Class 14 | bo- |
| Class 15 | kho- |
| Class 16 |  |
| Class 17 | kho- |

Pronoun Suffix ([PronSuf])

|  |
| --- |
| -na |

# CATEGORY 10: Copulatives

Copulative Particles ([CopPre]): see Category 11

|  |
| --- |
| ng- |
| -y- |
| aku- |
| u- (See subject concords) |
| zi - (See subject concords) |

Copulative Verbs ([Cop]): see Category 2

|  |
| --- |
| -be |
| -se |
| -ba |

# CATEGORY 11: Particles and prepositions

|  |
| --- |
| Locative / temporal {{example of what is meant by temporal, please?}} |
| Copulative: ng-, y-, w-, aku-, akusi- |
| Copulative particles resembling subject concords: see Category 5  In most cases copulative “particles” used for adjectives, resemble the subject concord, but not always:  Ngingumuntu > ngi-[SC]-ng-[CopPre]-umuntu[NStem]  Umuntu muhle > mu-[AdjConc]-hle[AdjStem]  Izitsha zinhle > zin-[AdjConc]-hle[AdjStem] |

# CATEGORY 12: Secondary morphemes

|  |  |  |
| --- | --- | --- |
| Future | [Fut] | -za-, -zo-, -zoku-, -ya-, -yo-, -yoku- |
| Progressive | [Pro] | -sa- |
| Potential | [Pot] | -nga- |
| Exclusive | [ExclNeg] | -ka- |
| Negative | [NegPre] | -a-, -ka-, -nga-, -nge-, -anga-, -i- |
| Prefix | [Pref] | ka-, a-, |
| Comparative | [Comp] | ubu-, beku |

# CATEGORY 13: Ideophones ([Ideoph])

|  |  |  |  |
| --- | --- | --- | --- |
| bhungubhungu | bofu | chaphu-chaphu | chu |
| chuku | cobo-cobo | cwaka | cwe |
| dewe | duke | fixi | gubungu |
| gwa | ja | jwi | manga |
| mbo | nca | nku | nqa |
| nqo | qikili | qukulu | qxuphu |
| roxoxoxoxo | tshe | vunga-vunga |  |

# CATEGORY 14: Possessive concords

Possessive concords ([PossConc])

|  |  |
| --- | --- |
| **Persons & classes** | **Possessive Concord** |
| Class 1 | wa- |
| Class 1a | wa- |
| Class 2 | ba- |
| Class 2a | ba- |
| Class 3 | wa- |
| Class 4 | ya- |
| Class 5 | la- |
| Class 6 | a- |
| Class 7 | sa- |
| Class 8 | za- |
| Class 9 | ya- |
| Class 10 | za- |
| Class 11 | lwa- |
| Class 14 | ba- |
| Class 15 | kwa- |
| Class 16 |  |
| Class 17 | kwa- |

Possessive prefix “ka” ([PossKA])

# CATEGORY 15: Interjections ([Interjec])

E.g. *ai*!, arha!, awu!, hayi!, he!, heke!, inene!, kuhle!, nantso ke!, nene-nene!, ngxe!, nyani!, safa!, shu!, sii!, tyhini!, tyhu!, yoo!

# CATEGORY 16: Conjunctions ([Conj])

|  |
| --- |
| futhi |
| hleze |
| kaloku |
| kambe |
| kodwa |
| koko |
| nje |
| njenga |
| nokuba |
| noxa |
| ukuba |
| ukuze |
| uma |
| zama |

# CATEGORY 17: Interrogatives

Question words ([Interrog]), e.g.

|  |  |
| --- | --- |
| ubani? | ‘who’ |
| njani? | ‘how’ |
| oobani? | ‘who plural’ |
| kangaphi? | ‘how many times’ |
| nini | ‘when’ |

Question particles ([InterrogSuf]), e.g.

|  |
| --- |
| -phi? |
| -ni? |
| -ngaphi? |
| -ntoni ? |

# CATEGORY 18: Abbreviations ([Abbr])

E.g.:

Mnu. (mnumzana)

Nkskz. (nkosikazi)

Nkszn. (nkosazana)

Njing. (njingalwazi)

COSATU

ANC

# CATEGORY 19: Demonstrative Copulatives ([DemCop])

|  |  |  |  |
| --- | --- | --- | --- |
| **Persons & classes** | **1** | **2** | **3** |
| Class 1 | nanku | nanko | nankuya |
| Class 1a | nanku | nanko | nankuya |
| Class 2 | naba | nabo | nabaya |
| Class 2b | naba | nabo | nabaya |
| Class 3 | nangu | nango | nanguya |
| Class 4 | nantsi | nantso | nantsiya |
| Class 5 | nali | nalo | naliya |
| Class 6 | nanga | nango | nangaya |
| Class 7 | nasi | naso | nasiya |
| Class 8 | nazi | nazo | naziya |
| Class 9 | nantsi | nantso | nantsiya |
| Class 10 | nanzi/nazi | nanzo | nanziya |
| Class 11 | nalu | nalo | naluyo |
| Class 14 | nabu | nabo | nabuya |
| Class 15 | naku | nako | nakuya |
| Class 16 |  |  |  |
| Class 17 |  |  |  |

# CATEGORY 20: Enumeratives

Enumerative stem ([EnumStem]):

|  |
| --- |
| -nye |
| -ni |
| -elowo |

Enumerative Concord ([EnumConc]):

|  |  |
| --- | --- |
| **Persons & classes** | **Enumerative Concord** |
| 1st p.s. | -ndi- |
| 1st p.p. | -si- |
| 2nd p.s. | -ku- |
| 2nd p.p. | -ni- |
| Class 1 | -m- |
| Class 1a | -m- |
| Class 2 | -ba- |
| Class 2a | -ba- |
| Class 3 | -wu- |
| Class 4 | -yi- |
| Class 5 | -li- |
| Class 6 | -wa- |
| Class 7 | -si- |
| Class 8 | -zi- |
| Class 9 | -yi- |
| Class 10 | -zi- |
| Class 11 | -lu- |
| Class 14 | -bu- |
| Class 15 | -ku- |
| Class 16 |  |
| Class 17 | -ku- |

# CATEGORY 21: Tense markers

|  |  |  |
| --- | --- | --- |
| Past tense | [Past] | -uku- |
| Present tense | [Pres] | -ya- |
| Perfect | [Perf] | -ile, -e |
| Future tense | [Fut] | -za-/-zo-/-zoku-, -ya-/-yo-/-yoku- |
| Negative present | [NegPre] | -ka-, -nga-, -anga, -i |
| Negative future | [NegFut] | -aba- |

# CATEGORY 22: Foreign ([Foreign])

Tag used for any foreign words e.g. finance, management, service.

# CATEGORY 23: Proper name ([ProperName])

Tag used for proper names.

# General principles

The following general principles apply:

* Annotation on the morphological analysis layer should be in excess of 98%
* The general principles of the Text Encoding Initiative P5 Guidelines (TEI P5; <http://www.tei-c.org/Guidelines/>) should be followed when in doubt.

# Changing the protocol

Since this protocol is considered to be a dynamic document, it is subject to change from time to time. For purposes of good governance, the following principles and procedures should be followed:

* The language expert (in consultation with the project leader and project manager) has the final say in changing this procedure document, and is the only authoritative person who might make final changes to the protocol.
* If an annotator (or other project member) wishes change a specification, s/he should inform the linguistic manager by email: Wikus.Pienaar@nwu.ac.za. The linguistic manager in liaison with the project leader considers the request and adds the change to the protocol document. The protocol document is saved with a new version number. All other annotators are then informed of the change, and requested to change previous annotations, where applicable.

# Authoritative references

The following language-specific works of reference should be considered authoritative when in doubt:

* Du Plessis, JA. 1978. isiXhosa 4. Oudiovista-Produksies, Goodwood.
* Louw, JA & Jubase, JB. 1963. Handboek van Xhosa. Bona Pers. Johannesburg.

# Roles and contact details

The following roles in this project can be identified:

**Linguistic manager**: Responsible for overall linguistic quality in liaison with the language experts and the project leader.

*Wikus Pienaar*

*Wikus.Pienaar@nwu.ac.za*

**Project leader**: Responsible for successful technical delivery of project.

*Martin Puttkammer*

*Martin.Puttkammer@nwu.ac.za*

**Project manager**: Responsible for timeous delivery of project, within budget. If you suspect that you will not be able to deliver on a set milestone, contact the project manager.

*Sunny Gent*

*Sunny.Gent@nwu.ac.za*

**Language expert**: Responsible for language correctness.

**Annotator**: Responsible for annotating the corpus, according to the above procedures and principles.