

New implementation of word order control in Swahili to English machine translation¹

Arvi Hurskainen
Department of Languages
FIN-00014 University of Helsinki, Finland
arvi.hurskainen@helsinki.fi
[DOI: 10.13140/RG.2.2.16236.44163](https://doi.org/10.13140/RG.2.2.16236.44163)

Abstract

Word order management in Swahili to English machine translation requires a set of carefully planned operations that are executed as a sequence. The re-arrangement of glosses is implemented on two levels. The first one is the word level, where the glosses are arranged in the order where they should be in the target language. For example, some verb forms have several glosses that are not necessarily in the order where they should be in target language. Another problem is the use of articles in adjective structures. Such problems should be solved on the word level, so that the article is attached to the correct word. The second level of re-arrangement of constituents is the phrase level, and in some cases the sentence level, where the word order is considered in a wider context.

Splitting the control of word order into these two phases greatly simplifies rule writing and makes it more manageable. This report illustrates the new implementation, where rules for rearranging words are written in two phases.

Key Words: *machine translation, word order control.*

1 Introduction

The main differences in word order between Swahili and English occur in noun phrases. Whereas the modifiers of the noun are located to the left of the noun in English, in Swahili they are located after the noun, although not in the same order as in English. The management of such structures is a major problem in translation.

Another major problem is the insertion of correct articles, because Swahili does not have articles, and therefore articles cannot be translated - they must be inserted. The method of doing this was described in an earlier report².

In addition to these problems, there is also a need to rearrange the order of word-internal glosses. This concerns especially some verb forms.

In my earlier implementation of the Swahili to English translation system, the reordering of words was done in one phase. In it, the sentence level reordering and word-internal reordering was done using the same rules. As a result, the rules became excessively

¹ The report is issued under licence CC BY-NC

² <http://www.njas.helsinki.fi/salama/managing-articles-in-swa-to-eng-mt.pdf>

complex. In this new implementation, I have separated the word-internal and sentence level reordering as different operations. This solution makes the rule writing simpler and makes their maintenance easier.

2 Rearranging word-internal glosses

Some forms of Swahili verbs have several morphemes, and they cannot be translated directly by converting the morphemes into glosses in target language. The morphemes may have more than one gloss, depending on the structure of the sentence. Another problem is the order of the glosses, again depending on the sentence structure.

The most complex verb structure is the one, which has a subject prefix, a tense-aspect-mood (TAM) marker, the relative prefix, the object prefix, the stem, and possibly several suffixes. The relative prefix may refer to the subject marker of the verb, or it may refer to the object marker of the verb. Both verb structures trigger a different word order, and also partly a different selection of glosses. All this makes the translation process complicated and difficult, but not impossible, to handle. I will illustrate the problem using four short example sentences. The example in (1) shows the result after analysis.

(1)

```
"<*inayonikera>"
    "kera" V CAP 4-PL-SP VFIN { they } PR:na 4-PL-REL { which }
1-SG1-OBJ OBJ { me } z [kera] SVO
    "kera" V CAP 4-PL-SP VFIN { they } PR:na 6-PL-REL { which }
1-SG1-OBJ OBJ { me } z [kera] SVO
    "kera" V CAP 4-PL-SP VFIN { they } PR:na 6-PLSG-REL { which
} 1-SG1-OBJ OBJ { me } z [kera] SVO
    "kera" V CAP 4-PL-SP VFIN { they } PR:na 9-SG-REL { which }
1-SG1-OBJ OBJ { me } z [kera] SVO
    "kera" V CAP 9-SG-SP VFIN { it } PR:na 4-PL-REL { which } 1-
SG1-OBJ OBJ { me } z [kera] SVO
    "kera" V CAP 9-SG-SP VFIN { it } PR:na 6-PL-REL { which } 1-
SG1-OBJ OBJ { me } z [kera] SVO
    "kera" V CAP 9-SG-SP VFIN { it } PR:na 6-PLSG-REL { which }
1-SG1-OBJ OBJ { me } z [kera] SVO
    "kera" V CAP 9-SG-SP VFIN { it } PR:na 9-SG-REL { which } 1-
SG1-OBJ OBJ { me } z [kera] SVO
"<ni>"
    "ni" V V-BE INIT { it , he , she } { is }
    "ni" V V-BE INIT { they } { are }
    "ni" V V-BE NOSUBJ { is , are , am }
"<uvivu>"
    "uvivu" N 11-SG NOART DER:adj (vivu)
"<.>"
    "." STOP { . } **CLB
"<*wanaonisaidia>"
    "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-SUB-REL {
who } 1-SG1-OBJ OBJ { me } z [saidia] SVO AR
    "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-OBJ-REL {
who } 1-SG1-OBJ OBJ { me } z [saidia] SVO AR
```

"saidia" V CAP 2-PL3-SP VFIN { they } PR:na 3-SG-REL { which } 1-SG1-OBJ OBJ { me } z [saidia] SVO AR
"saidia" V CAP 2-PL3-SP VFIN { they } PR:na 11-SG-REL { which } 1-SG1-OBJ OBJ { me } z [saidia] SVO AR
"<watafanikiwa>"
"fanikiwa" V 2-PL3-SP VFIN { they } FUT:ta { will } z [fana] PASS-LEX SV
"fanikiwa" V 2-PL3-SP VFIN { they } FUT:ta { will } z [fanika] SVO PREFER
"fanika" V 2-PL3-SP VFIN { they } FUT:ta { will } z [fanika] SVO APPL PASS
"fanika" V 2-PL3-SP VFIN { they } FUT:ta { will } z [fana] SV STAT APPL PASS
"<.>"
"." STOP { . } **CLB
"<*nitakaowafundisha>"
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-PL-SUB-REL { who } 2-PL2-OBJ OBJ { you } z [funda] SVO :CAUS PREFER
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-PL-SUB-REL { who } 2-PL2-OBJ OBJ { you } z [funda] SVO CAUS
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-PL-SUB-REL { who } 2-PL3-OBJ OBJ { them } z [funda] SVO :CAUS PREFER
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-PL-SUB-REL { who } 2-PL3-OBJ OBJ { them } z [funda] SVO CAUS
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-PL-OBJ-REL { who } 2-PL2-OBJ OBJ { you } z [funda] SVO :CAUS PREFER
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-PL-OBJ-REL { who } 2-PL2-OBJ OBJ { you } z [funda] SVO CAUS
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-PL-OBJ-REL { who } 2-PL3-OBJ OBJ { them } z [funda] SVO :CAUS PREFER
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-PL-OBJ-REL { who } 2-PL3-OBJ OBJ { them } z [funda] SVO CAUS
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 3-SG-REL { which } 2-PL2-OBJ OBJ { you } z [funda] SVO :CAUS PREFER
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 3-SG-REL { which } 2-PL2-OBJ OBJ { you } z [funda] SVO CAUS
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 3-SG-REL { which } 2-PL3-OBJ OBJ { them } z [funda] SVO :CAUS PREFER
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 3-SG-REL { which } 2-PL3-OBJ OBJ { them } z [funda] SVO CAUS
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 11-SG-REL { which } 2-PL2-OBJ OBJ { you } z [funda] SVO :CAUS PREFER
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 11-SG-REL { which } 2-PL2-OBJ OBJ { you } z [funda] SVO CAUS
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 11-SG-REL { which } 2-PL3-OBJ OBJ { them } z [funda] SVO :CAUS PREFER
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 11-SG-REL { which } 2-PL3-OBJ OBJ { them } z [funda] SVO CAUS
"<watafaulu>"

```
"faulu" V 2-PL3-SP VFIN { they } FUT:ta { will } z [faulu]
SVO AR
"<.>"
  "." STOP { . } **CLB
"<*nitakayemfundisha>"
  "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-SUB-REL { who } 1-SG3-OBJ OBJ { him , her } z [funda] SVO :CAUS
PREFR
  "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-SUB-REL { who } 1-SG3-OBJ OBJ { him , her } z [funda] SVO CAUS
  "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-OBJ-REL { whom } 1-SG3-OBJ OBJ { him , her } z [funda] SVO
:CAUS PREFR
  "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-OBJ-REL { whom } 1-SG3-OBJ OBJ { him , her } z [funda] SVO CAUS
"<atafaulu>"
  "faulu" V 1-SG3-SP VFIN { he } FUT:ta { will } z [faulu] SVO
AR
  "faulu" V 1-SG3-SP VFIN { she } FUT:ta { will } z [faulu]
SVO AR
"<.>"
  "." STOP { . } **CLB
```

In this phase, each morpheme except for the word stem has a gloss. From the viewpoint of translation, this result is not satisfactory, because in some contexts the subject prefix and object prefix gloss, and in some cases both, should be removed. In order to control these cases, we continue in phases. We first do the morphological disambiguation, without considering the absence of one or more glosses (2).

(2)

```
"<*inayonikera>"
  "kera" V CAP 9-SG-SP VFIN { it } PR:na 9-SG-REL { which } 1-
SG1-OBJ OBJ { me } z [kera] SVO SUB-REL
"<ni>"
  "ni" V V-BE NOSUBJ { is }
"<uvivu>"
  "uvivu" N 11-SG NOART DER:adj (vivu)
"<.>"
  "." STOP { . } **CLB
"<*wanaonisaidia>"
  "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-SUB-REL {
who } 1-SG1-OBJ OBJ { me } z [saidia] SVO AR
"<watafanikiwa>"
  "fanikiwa" V 2-PL3-SP VFIN { they } FUT:ta { will } z
[fanika] SVO PREFR
"<.>"
  "." STOP { . } **CLB
"<*nitakaowafundisha>"
```

```
"fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-
PL-OBJ-REL { who } 2-PL3-OBJ OBJ { them } z [funda] SVO :CAUS
PREFR
"<watafaulu>"
  "faulu" V 2-PL3-SP VFIN { they } FUT:ta { will } z [faulu]
SVO AR
"<.>"
  "." STOP { . } **CLB
"<*nitakayemfundisha>"
  "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-OBJ-REL { who } 1-SG3-OBJ OBJ { him } z [funda] SVO :CAUS PREFR
PROP-CAND
"<atafaulu>"
  "faulu" V 1-SG3-SP VFIN { he } FUT:ta { will } z [faulu] SVO
AR
"<.>"
  "." STOP { . } **CLB
```

The sentences are now correctly disambiguated, but from the point of view of translation, the result is not satisfactory. We add also such alternatives, where one or more glosses are missing (3).

```
(3)
"<*inayonikera>"
  "kera" V CAP 9-SG-SP VFIN { it } PR:na 9-SG-REL { which } 1-
SG1-OBJ OBJ { me } z [kera] SVO SUB-REL @FMAINVtr-OBJ>
  "kera" V CAP 9-SG-SP VFIN { it } PR:na 9-SG-REL { which } 1-
SG1-OBJ OBJ NO-OBJ-GLOSS z [kera] SVO SUB-REL @FMAINVtr-OBJ>
  "kera" V CAP 9-SG-SP VFIN NO-SP-GLOSS PR:na 9-SG-REL { which
} 1-SG1-OBJ OBJ { me } z [kera] SVO SUB-REL @FMAINVtr-OBJ>
  "kera" V CAP 9-SG-SP VFIN NO-SP-GLOSS PR:na 9-SG-REL { which
} 1-SG1-OBJ OBJ NO-OBJ-GLOSS z [kera] SVO SUB-REL @FMAINVtr-OBJ>
"<ni>"
  "ni" V V-BE NOSUBJ { is } @FMAINVintr-def
"<uvivu>"
  "uvivu" N 11-SG NOART DER:adj (vivu) @<P
"<.>"
  "." STOP { . } **CLB
"<*wanaonisaidia>"
  "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-SUB-REL {
who } 1-SG1-OBJ OBJ { me } z [saidia] SVO AR @FMAINVtr-OBJ>
  "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-SUB-REL {
who } 1-SG1-OBJ OBJ NO-OBJ-GLOSS z [saidia] SVO AR @FMAINVtr-OBJ>
  "saidia" V CAP 2-PL3-SP VFIN NO-SP-GLOSS PR:na 2-PL-SUB-REL
{ who } 1-SG1-OBJ OBJ { me } z [saidia] SVO AR @FMAINVtr-OBJ>
  "saidia" V CAP 2-PL3-SP VFIN NO-SP-GLOSS PR:na 2-PL-SUB-REL
{ who } 1-SG1-OBJ OBJ NO-OBJ-GLOSS z [saidia] SVO AR @FMAINVtr-
OBJ>
"<watafanikiwa>"
  "fanikiwa" V 2-PL3-SP VFIN { they } FUT:ta { will } z
[fanika] SVO PREFR @FMAINVtr-OBJ>
```

```
"fanikiwa" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[fanika] SVO PREFER @FMAINVtr-OBJ>
"<.>"
    "." STOP { . } **CLB
"<*nitakaowafundisha>"
    "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-
PL-OBJ-REL { who } 2-PL3-OBJ OBJ { them } z [funda] SVO :CAUS
PREFER @FMAINVtr-OBJ>
    "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-
PL-OBJ-REL { who } 2-PL3-OBJ OBJ NO-OBJ-GLOSS z [funda] SVO :CAUS
PREFER @FMAINVtr-OBJ>
    "fundisha" V CAP 1-SG1-SP VFIN NO-SP-GLOSS FUT:taka { will }
2-PL-OBJ-REL { who } 2-PL3-OBJ OBJ { them } z [funda] SVO :CAUS
PREFER @FMAINVtr-OBJ>
    "fundisha" V CAP 1-SG1-SP VFIN NO-SP-GLOSS FUT:taka { will }
2-PL-OBJ-REL { who } 2-PL3-OBJ OBJ NO-OBJ-GLOSS z [funda] SVO
:CAUS PREFER @FMAINVtr-OBJ>
"<watafaulu>"
    "faulu" V 2-PL3-SP VFIN { they } FUT:ta { will } z [faulu]
SVO AR @FMAINVtr-OBJ>
    "faulu" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] SVO AR @FMAINVtr-OBJ>
"<.>"
    "." STOP { . } **CLB
"<*nitakayemfundisha>"
    "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-OBJ-REL { who } 1-SG3-OBJ OBJ { him } z [funda] SVO :CAUS PREFER
PROP-CAND @FMAINVtr-OBJ>
    "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-OBJ-REL { who } 1-SG3-OBJ OBJ NO-OBJ-GLOSS z [funda] SVO :CAUS
PREFER PROP-CAND @FMAINVtr-OBJ>
    "fundisha" V CAP 1-SG1-SP VFIN NO-SP-GLOSS FUT:taka { will }
1-SG-OBJ-REL { who } 1-SG3-OBJ OBJ { him } z [funda] SVO :CAUS
PREFER PROP-CAND @FMAINVtr-OBJ>
    "fundisha" V CAP 1-SG1-SP VFIN NO-SP-GLOSS FUT:taka { will }
1-SG-OBJ-REL { who } 1-SG3-OBJ OBJ NO-OBJ-GLOSS z [funda] SVO
:CAUS PREFER PROP-CAND @FMAINVtr-OBJ>
"<atafaulu>"
    "faulu" V 1-SG3-SP VFIN { he } FUT:ta { will } z [faulu] SVO
AR @FMAINVtr-OBJ>
    "faulu" V 1-SG3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] SVO AR @FMAINVtr-OBJ>
"<.>"
    "." STOP { . } **CLB
```

The readings of verbs have now all four alternatives to select from. After selection, using Constraint Grammar rules, the result is as in (4).

(4)
"<*inayonikera>"
 "kera" V CAP 9-SG-SP VFIN NO-SP-GLOSS PR:na 9-SG-REL { which
} 1-SG1-OBJ OBJ { me } z [kera] SVO SUB-REL @FMAINVtr-OBJ>
"<ni>"
 "ni" V V-BE NOSUBJ { is } @FMAINVintr-def
"<uvivu>"
 "uvivu" N 11-SG NOART DER:adj (vivu) @<P
"<.>"
 "." STOP { . } **CLB
"<*wanaonisaidia>"
 "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-SUB-REL {
who } 1-SG1-OBJ OBJ { me } z [saidia] SVO AR @FMAINVtr-OBJ>
"<watafanikiwa>"
 "fanikiwa" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[fanika] SVO PREFER @FMAINVtr-OBJ>
"<.>"
 "." STOP { . } **CLB
"<*nitakaowafundisha>"
 "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-
PL-OBJ-REL { who } 2-PL3-OBJ OBJ NO-OBJ-GLOSS z [funda] SVO :CAUS
PREFER @FMAINVtr-OBJ>
"<watafaulu>"
 "faulu" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] SVO AR %FMAINVtr-OBJ>
"<.>"
 "." STOP { . } **CLB
"<*nitakayemfundisha>"
 "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-OBJ-REL { who } 1-SG3-OBJ OBJ NO-OBJ-GLOSS z [funda] SVO :CAUS
PREFER PROP-CAND @FMAINVtr-OBJ>
"<atafaulu>"
 "faulu" V 1-SG3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] SVO AR @FMAINVtr-OBJ>
"<.>"
 "." STOP { . } **CLB

Now we have disambiguated readings with the correct set of word-internal glosses. We must check whether they are in correct order for transfer into target language. We see that all verbs with relative and object prefixes need reordering. We will do it in this phase, so that when we reorder the glosses of stems, we do not need to pay attention to word-internal order any longer. This simplifies greatly the reordering of stem glosses. Before reordering word-internal glosses, we need to add stem glosses (5).

(5)
"<*inayonikera>"
 "kera" V CAP 9-SG-SP VFIN NO-SP-GLOSS PR:na 9-SG-REL { which
} 1-SG1-OBJ OBJ { me } z [kera] { annoy , aggravate , discomfort ,
molest , vex , trouble , irritate } SVO SUB-REL %FMAINVtr-OBJ>

```
"<ni>"
  "ni" V V-BE NOSUBJ { is } %FMAINVintr-def
"<uvivu>"
  "uvivu" N { laziness } 11-SG NOART DER:adj (vivu) %<P
"<.>"
  "." STOP { . } **CLB
"<*wanaonisaidia>"
  "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-SUB-REL {
who } 1-SG1-OBJ OBJ { me } z [saidia] { help , assist , support }
SVO AR %FMAINVtr-OBJ>
"<watafanikiwa>"
  "fanikiwa" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[fanika] { succeed , have :success } SVO PREFER %FMAINVtr-OBJ>
"<.>"
  "." STOP { . } **CLB
"<*nitakaowafundisha>"
  "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 2-
PL-OBJ-REL { who } 2-PL3-OBJ OBJ NO-OBJ-GLOSS z [funda] { teach ,
instruct , inculcate } SVO CAUS PREFER %FMAINVtr-OBJ>
"<watafaulu>"
  "faulu" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] { be successful , achieve , pass an examination } SVO AR
%FMAINVtr-OBJ>
"<.>"
  "." STOP { . } **CLB
"<*nitakayemfundisha>"
  "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT:taka { will } 1-
SG-OBJ-REL { who } 1-SG3-OBJ OBJ NO-OBJ-GLOSS z [funda] { teach ,
instruct , inculcate } SVO CAUS PREFER PROP-CAND %FMAINVtr-OBJ>
"<atafaulu>"
  "faulu" V 1-SG3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] { be successful , achieve , pass an examination } SVO AR
%FMAINVtr-OBJ>
"<.>"
  "." STOP { . } **CLB
```

Now when we have also stem glosses, we can reorder word-internal glosses. The result is in (6).

```
(6)
"<*inayonikera>"
  "kera" V CAP 9-SG-SP VFIN NO-SP-GLOSS PR:na 9-SG-REL { which
} z [kera] { annoy , aggravate , discomfort , molest , vex ,
trouble , irritate } 1-SG1-OBJ OBJ { me } SVO SUB-REL %FMAINVtr-
OBJ>
"<ni>"
  "ni" V V-BE NOSUBJ { is } %FMAINVintr-def
"<uvivu>"
  "uvivu" N { laziness } 11-SG NOART DER:adj (vivu) %<P
"<.>"
  "." STOP { . } **CLB
```



```
"<*wanaonisaidia>"
    "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-SUB-REL {
who } z [saidia] { help , assist , support } 1-SG1-OBJ OBJ { me }
SVO AR %FMAINVtr-OBJ>
"<watafanikiwa>"
    "fanikiwa" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[fanika] { succeed , have :success } SVO PREFER %FMAINVtr-OBJ>
"<.>"
    "." STOP { . } **CLB
"<*nitakaowafundisha>"
    "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT 2-PL-OBJ-REL { who
} :taka { will } 2-PL3-OBJ OBJ NO-OBJ-GLOSS z [funda] { teach ,
instruct , inculcate } SVO CAUS PREFER %FMAINVtr-OBJ>
"<watafaulu>"
    "faulu" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] { be successful , achieve , pass an examination } SVO AR
%FMAINVtr-OBJ>
"<.>"
    "." STOP { . } **CLB
"<*nitakayemfundisha>"
    "fundisha" V CAP 1-SG1-SP VFIN { *i } FUT 1-SG-OBJ-REL { who
} :taka { will } 1-SG3-OBJ OBJ NO-OBJ-GLOSS z [funda] { teach ,
instruct , inculcate } SVO CAUS PREFER PROP-CAND %FMAINVtr-OBJ>
"<atafaulu>"
    "faulu" V 1-SG3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] { be successful , achieve , pass an examination } SVO AR
%FMAINVtr-OBJ>
"<.>"
    "." STOP { . } **CLB
```

The first and second sentence have now the correct order of glosses, but the third and fourth sentence need further reordering. In them, the relative prefix refers to the object instead of subject (2-PL-OBJ-REL and 1-SG-OBJ-REL). This causes a different order of glosses in target language. We add one more rule for controlling this situation (7).

(7)

```
"<*inayonikera>"
    "kera" V CAP 9-SG-SP VFIN NO-SP-GLOSS PR:na 9-SG-REL { which
} z [kera] { annoy , aggravate , discomfort , molest , vex ,
trouble , irritate } 1-SG1-OBJ OBJ { me } SVO SUB-REL %FMAINVtr-
OBJ>
"<ni>"
    "ni" V V-BE NOSUBJ { is } %FMAINVintr-def
"<uvivu>"
    "uvivu" N { laziness } 11-SG NOART DER:adj (vivu) %<P
"<.>"
    "." STOP { . } **CLB
```

```
"<*wanaonisaidia>"
  "saidia" V CAP 2-PL3-SP VFIN { they } PR:na 2-PL-SUB-REL {
who } z [saidia] { help , assist , support } 1-SG1-OBJ OBJ { me }
SVO AR %FMAINVtr-OBJ>
"<watafanikiwa>"
  "fanikiwa" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[fanika] { succeed , have :success } SVO PREFER %FMAINVtr-OBJ>
"<.>"
  "." STOP { . } **CLB
"<*nitakaowafundisha>"
  "fundisha" V CAP 1-SG1-SP VFIN FUT 2-PL-OBJ-REL { who } {
*i } :taka { will } 2-PL3-OBJ OBJ NO-OBJ-GLOSS z [funda] { teach ,
instruct , inculcate } SVO CAUS PREFER %FMAINVtr-OBJ>
"<watafaulu>"
  "faulu" V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] { be successful , achieve , pass an examination } SVO AR
%FMAINVtr-OBJ>
"<.>"
  "." STOP { . } **CLB
"<*nitakayemfundisha>"
  "fundisha" V CAP 1-SG1-SP VFIN FUT 1-SG-OBJ-REL { whom } {
*i } :taka { will } 1-SG3-OBJ OBJ NO-OBJ-GLOSS z [funda] { teach ,
instruct , inculcate } SVO CAUS PREFER PROP-CAND %FMAINVtr-OBJ>
"<atafaulu>"
  "faulu" V 1-SG3-SP VFIN NO-SP-GLOSS FUT:ta { will } z
[faulu] { be successful , achieve , pass an examination } SVO AR
%FMAINVtr-OBJ>
"<.>"
  "." STOP { . } **CLB
```

Now the word-internal glosses are in correct order in all sentences. The verbs have many gloss alternatives. We remove the tokens and lemmas of the source language, disambiguate the glosses of target language and convert them to surface form (8).

(8)

```
( V CAP 9-SG-SP VFIN NO-SP-GLOSS PR:na 9-SG-REL { *which } :z {
:annoys } 1-SG1-OBJ OBJ PRON { me } @OBJ SVO SUB-REL %FMAINVtr-
OBJ> )
( V V-BE NOSUBJ { is } %FMAINVintr-def )
( N { laziness } 11-SG NOART %<P )
( STOP { . } **CLB )
( V CAP 2-PL3-SP VFIN { *they } PR:na 2-PL-SUB-REL { who } :z {
help } 1-SG1-OBJ OBJ PRON { me } @OBJ SVO %FMAINVtr-OBJ> )
( V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } :z { succeed } SVO
PREFER %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( V CAP 1-SG1-SP VFIN FUT 2-PL-OBJ-REL { *who } PRON { *i } @SUBJ
:taka { will } 2-PL3-OBJ OBJ NO-OBJ-GLOSS z { teach } SVO CAUS
PREFER %FMAINVtr-OBJ> )
( V 2-PL3-SP VFIN NO-SP-GLOSS FUT:ta { will } :z { be successful }
SVO %FMAINVtr-OBJ> )
```

```
( STOP { . } **CLB )  
( V CAP 1-SG1-SP VFIN FUT 1-SG-OBJ-REL { *whom } PRON { *i } @SUBJ  
:taka { will } 1-SG3-OBJ OBJ NO-OBJ-GLOSS z { teach } SVO CAUS  
PREFR PROP-CAND %FMAINVtr-OBJ> )  
( V 1-SG3-SP VFIN NO-SP-GLOSS FUT:ta { will } z { } SVO %FMAINVtr-  
OBJ> )  
( STOP { . } **CLB )
```

When we remove all tags and modify the result, we get the translation (9).

(9)

Which annoys me is laziness.

They who help me will succeed.

Who I will teach will be successful.

Whom I will teach will be successful.

3 Verb forms with the relative marker of manner or of time or location

Swahili has special verb-forms, which structurally are similar with the examples above, but where the relative marker does not refer to subject or object, but rather to the manner of action, or to time or location. The prefixes are *vyo* for manner, and *po* for time or location. Such verb forms are usually translated with subordinate clauses starting with the conjunctions *as* for manner, and *when* for time or location.

For testing the translation quality of such structures, I retrieved from Helsinki Corpus of Swahili 2.0³ all such unique verb forms that have the subject prefix, the object prefix, and the relative prefix of class 8 (realised as *vyo*), or of class 16 (realised as *po*). The prefixes *vyo* and *po* are actually the relative markers of their respective noun classes.

The prefix *vyo* is actually the relative marker referring to the plural of the nouns belonging to the noun class 7/8. Yet in these examples it has another function. The corpus contains 2,694 unique verb forms of the type *vyo*, and 3,303 of the type *po*.

I have taken these examples for scrutiny, because they require a different word order in translation. For saving space, we look at the 20 most frequent verb forms and 20 least frequent verb forms of both types. In (10) we have the verbs with the relative prefix *vyo*.

(10)

Most frequent with prefix *vyo*:

Nilivyovitaja.

Alivyovitaja.

Vinavyopakana.

Vinavyojihusisha.

Vinavyojishughulisha.

Walivyozipata.

Nilivyolieleza.

³ <http://urn.fi/urn:nbn:fi:lb-201608301>

Ninavyomnukulu.
Walivyojipanga.
Ninavyokuambia.
Inavyojieleza.
Alivyozipata.
Ninavyoiona.
Vinavyojitegemea.
Vinavyoizunguka.
Nilivyokueleza.
Anavyojisikia.
Alivyoiwasilisha.
Vitakavyowawezesha.
Unavyotuona.

Least frequent with prefix vyo:

Alivyoionya.
Alivyoionesha.
Alivyoiona.
Alivyoileta.
Alivyoikuta.
Alivyoikamata.
Alivyoijua.
Alivyoibusha.
Alivyoihukumu.
Alivyoihifadhi.
Alivyoihama.
Alivyoifikia.
Alivyoifahamisha.
Alivyoindeleza.
Alivyoielezea.
Alivyoielekeza.
Alivyoibadili.
Alivyoiangalia.
Alivyoianisha.
Alivyoiacha.

The translation is in (11).

(11)

Most frequent with prefix vyo:

As I mentioned them.
As he mentioned them.
As they deny then.
As they themselves make participate.
As they are engaged in.
As they got them.

As I explained it.
As I quote him.
As they arranged themselves.
As I tell you.
As it itself explains.
As he got them.
As I see them.
As they rely on themselves.
As they go around them.
As I explained you.
As he feels.
As he communicated them.
As they will empower them.
As you see us.

Least frequent with prefix vyo:

As he warned them.
As he showed them.
As he saw them.
As he brought them.
As he found them.
As he caught them.
As he knew them.
As he roused them.
As he judged them.
As he preserved them.
As he emigrated them.
As he reached them.
As he made known them.
As he promoted them.
As he explained them.
As he directed them.
As he changed them.
As he looked at them.
As he them made.
As he left them.

As we see, the translation is reasonable, and at least the word order is correct. Also the rest of the verb forms in the list have the correct word order.

Next we take the verbs with the relative prefix po (12).

(12)

Most frequent with prefix po:

Alipomwona.
Alipojiunga.

Alipomwona.
Walipomwona.
Ilipojipatia.
Alipomuuliza.
Alipolihutubia.
Aliponiona.
Alipomuuliza.
Alipowaambia.
Alipomwambia.
Wanapowaona.
Wanapowakamata.
Ilipoivamia.
Wanapowakamata.
Walipomwona.
Tunapowaletea.
Alipowaona.
Alipoitembelea.
Yalipomkuta.

Least frequent with prefix *po*:

Alipoifungua.
Alipoificha.
Alipoifahamu.
Alipoielezea.
Alipoielekea.
Alipoiegamiza.
Alipoichezea.
Alipoichambua.
Alipoichachamalia.
Alipoithiri.
Alipoianza.
Alipoiangalia.
Alipoiangalia.
Alipoandikia.
Alipoandika.
Alipoandaa.
Alipoiambia.
Alipoahirisha.
Alipoiaga.
Alipoiacha.

The translation of the verb structures with the relative marker *po* is in (13).

(13)

Most frequent with prefix *po*:

When he saw him.

When he himself joined.
When he saw him.
When they saw him.
When it itself got.
When he asked him.
When he gave speech to it.
When he saw me.
When he asked him.
When he told them.
When he told him.
When they see them.
When they catch them.
When it invaded them.
When they catch them.
When they saw him.
When we bring them.
When he saw them.
When he visited them.
When they found him.

Least frequent with prefix *po*:

When he opened them.
When he hid them.
When he knew them.
When he explained them.
When he was inclined to them.
When he leaned them.
When he played them.
When he analysed them.
When he stood firm to them.
When he affected them.
When he began them.
When he looked at them.
When he looked at them.
When he wrote to them.
When he wrote them.
When he prepared them.
When he told them.
When he postponed them.
When he took leave from them.
When he left them.

In order to get a fuller picture of these structures, we should retrieve them in a wider sentence context. However, this would not give more light to the problem itself, that is, the order of words deriving from Swahili verbs.

4 Managing articles

Swahili language does not use articles, which is why we cannot convert articles of the source language into target language. We must add them, using context sensitive rules. Because I have described the method of doing this in another report⁴, I will not discuss it here. I just state that this is a problem that must be dealt with on word level.

When all word-level problems have been resolved, we can move to the problems encountered on sentence level.

5 Reordering glosses on sentence level

When word-internal glosses are in control, that is, they are in the order, where they should be in target language, we need to control the order of separate words. Reordering is needed especially in noun phrases, where word order is drastically different in Swahili and English. The noun phrase types of Swahili are listed in (14).⁵

(14)

1. *Vitabu vyangu vipya vile vitatu vimepotea.*
2. *Vitabu vyangu vipya vile vyote vimepotea.*
3. *Vitabu vyangu vipya vile vimepotea.*
4. *Vitabu vyangu vipya vimepotea.*
5. *Vitabu vyangu vile vimepotea.*
6. *Vitabu vyangu vitatu vimepotea.*
7. *Vitabu vyangu vyote vimepotea.*
8. *Vitabu vipya vitatu vimepotea.*
9. *Vitabu vipya vyote vimepotea.*
10. *Vitabu vitatu vyote vimepotea.*
11. *Vitabu vyote vitatu vimepotea.*
12. *Vitabu vyangu vimepotea.*
13. *Vitabu vipya vimepotea.*
14. *Vitabu vitatu vimepotea.*
15. *Vitabu vyote vimepotea.*
16. *Vitabu vile vimepotea.*

As I said earlier, the word-internal reordering as well as the control of articles was already discussed earlier. In this phase, the processing result of the above sentences is as in (15).

(15)

(N { [book](#) } CAP 7/8-PL %SUBJ) (PRON POSS 8-PL SG1 { [my](#) } %GCON) (ADJ { [new](#) } A-INFL 8-PL %<NADJ) (PRON DEM :le 8-PL { [those](#) })

⁴ <http://www.njas.helsinki.fi/salama/managing-articles-in-swa-to-eng-mt.pdf>

⁵ The list contains only the basic types of noun phrases. The demonstrative pronoun may also precede the noun, and such cases need their own reordering rules. They are not included into this report.

%SUBJ) (NUM 8-PL NUM CARD INFL { three } %SUBJ) (V 8-PL-SP
VFIN NO-SP-GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP
{ . } **CLB)
(N { book } CAP 7/8-PL %SUBJ) (PRON POSS 8-PL SG1 { my } %GCON
) (ADJ { new } A-INFL 8-PL %<NADJ) (PRON DEM :le 8-PL { those }
%SUBJ) (PRON :OTE 8-PL { all } %SUBJ) (V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { . }
**CLB)
(N { book } CAP 7/8-PL %SUBJ) (PRON POSS 8-PL SG1 { my } %GCON
) (ADJ { new } A-INFL 8-PL %<NADJ) (PRON DEM :le 8-PL { those }
%SUBJ) (V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ>) (STOP { . } **CLB)
(N { book } CAP 7/8-PL %SUBJ) (PRON POSS 8-PL SG1 { my } %GCON
) (ADJ { new } A-INFL 8-PL %<NADJ) (V 8-PL-SP VFIN NO-SP-GLOSS
PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { . } **CLB)
(N { book } CAP 7/8-PL %SUBJ) (PRON POSS 8-PL SG1 { my } %GCON
) (PRON DEM :le 8-PL { those } %SUBJ) (V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { . }
**CLB)
(N { book } CAP 7/8-PL %SUBJ) (PRON POSS 8-PL SG1 { my } %GCON
) (NUM 8-PL NUM CARD INFL { three } ORD %SUBJ) (V 8-PL-SP VFIN
NO-SP-GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { .
} **CLB)
(N { book } CAP 7/8-PL %SUBJ) (PRON POSS 8-PL SG1 { my } %GCON
) (PRON :OTE 8-PL { all } %SUBJ) (V 8-PL-SP VFIN NO-SP-GLOSS
PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { . } **CLB)
(N { book } CAP 7/8-PL %SUBJ) (ADJ { new } A-INFL 8-PL %<NADJ)
(NUM 8-PL NUM CARD INFL { three } %SUBJ) (V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { . }
**CLB)
(N { book } CAP 7/8-PL %SUBJ) (ADJ { new } A-INFL 8-PL %<NADJ)
(PRON :OTE 8-PL { all } %SUBJ) (V 8-PL-SP VFIN NO-SP-GLOSS
PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { . } **CLB)
(N { book } CAP 7/8-PL %SUBJ) (NUM 8-PL NUM CARD INFL { three }
%SUBJ) (PRON :OTE 8-PL { all } %SUBJ) (V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { . }
**CLB)
(N { book } CAP 7/8-PL %SUBJ) (PRON :OTE 8-PL { all } %<DN) (
NUM 8-PL NUM CARD INFL { three } %SUBJ) (V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ>) (STOP { . }
**CLB)
(N { book } CAP 7/8-PL %SUBJ) (PRON POSS 8-PL SG1 { my } %GCON
) (V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ>) (STOP { . } **CLB)
(N { book } CAP 7/8-PL %SUBJ) (ADJ { the } { new } A-INFL 8-PL
%<NADJ) (V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ>) (STOP { . } **CLB)
(N { book } CAP 7/8-PL %SUBJ) (NUM 8-PL NUM CARD INFL { three }
%SUBJ) (V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ>) (STOP { . } **CLB)

```
( N { book } CAP 7/8-PL %SUBJ ) ( PRON :OTE 8-PL { all } %<DN ) (
V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO %FMAINVtr-
OBJ> ) ( STOP { . } **CLB )
( N { book } CAP 7/8-PL %SUBJ ) ( PRON DEM :le 8-PL { those }
%<NDEM ) ( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ> ) ( STOP { . } **CLB )
```

Most of the phrase types need a separate rule for reordering the constituents. However, some short phrases can be handled with a single rule, as we see below in (17). When we apply the reordering rules, the result will be as in (16).

(16)

```
( PRON DEM :le 8-PL { those } %SUBJ ) ( PRON POSS 8-PL SG1 { my }
%GCON ) ( NUM 8-PL NUM CARD INFL { three } %SUBJ ) ( ADJ { new }
A-INFL 8-PL %<NADJ ) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP
VFIN NO-SP-GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP
{ . } **CLB )
( PRON DEM :le 8-PL { those } %SUBJ ) ( PRON POSS 8-PL SG1 { my }
%GCON ) ( PRON :OTE 8-PL { all } %SUBJ ) ( ADJ { new } A-INFL 8-PL
%<NADJ ) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP { . }
**CLB )
( PRON DEM :le 8-PL { those } %SUBJ ) ( PRON POSS 8-PL SG1 { my }
%GCON ) ( ADJ { new } A-INFL 8-PL %<NADJ ) ( N { book } CAP 7/8-PL
%SUBJ ) ( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ> ) ( STOP { . } **CLB )
( PRON POSS 8-PL SG1 { my } %GCON ) ( ADJ { new } A-INFL 8-PL
%<NADJ ) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP { . }
**CLB )
( PRON POSS 8-PL SG1 { my } %GCON ) ( PRON DEM :le 8-PL { those }
%SUBJ ) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP { . }
**CLB )
( PRON POSS 8-PL SG1 { my } %GCON ) ( NUM 8-PL NUM CARD INFL {
three } ORD %SUBJ ) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP
VFIN NO-SP-GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP
{ . } **CLB )
( PRON POSS 8-PL SG1 { my } %GCON ) ( PRON :OTE 8-PL { all } %SUBJ
) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP VFIN NO-SP-GLOSS
PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP { . } **CLB )
( NUM 8-PL NUM CARD INFL { three } %SUBJ ) ( ADJ { new } A-INFL 8-
PL %<NADJ ) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP { . }
**CLB )
( PRON :OTE 8-PL { all } %SUBJ ) ( ADJ { new } A-INFL 8-PL %<NADJ
) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP VFIN NO-SP-GLOSS
PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP { . } **CLB )
( PRON :OTE 8-PL { all } %SUBJ ) ( NUM 8-PL NUM CARD INFL { three
} %SUBJ ) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP VFIN NO-SP-
```

```
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP { . }
**CLB )
( PRON :OTE 8-PL { all } %<DN ) ( NUM 8-PL NUM CARD INFL { three }
%SUBJ ) ( N { book } CAP 7/8-PL %SUBJ ) ( V 8-PL-SP VFIN NO-SP-
GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ> ) ( STOP { . }
**CLB )
( PRON POSS 8-PL SG1 { my } %GCON ) ( N { book } CAP 7/8-PL %SUBJ
) ( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO %FMAINVtr-
OBJ> ) ( STOP { . } **CLB )
( ADJ { the } { new } A-INFL 8-PL %<NADJ ) ( N { book } CAP 7/8-PL
%SUBJ ) ( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ> ) ( STOP { . } **CLB )
( NUM 8-PL NUM CARD INFL { three } %SUBJ ) ( N { book } CAP 7/8-PL
%SUBJ ) ( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ> ) ( STOP { . } **CLB )
( PRON :OTE 8-PL { all } %<DN ) ( N { book } CAP 7/8-PL %SUBJ ) ( V
8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO %FMAINVtr-OBJ>
) ( STOP { . } **CLB )
( PRON DEM :le 8-PL { those } %<NDEM ) ( N { book } CAP 7/8-PL
%SUBJ ) ( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me z { be lost } SVO
%FMAINVtr-OBJ> ) ( STOP { . } **CLB )
```

All glosses are now in correct order, and the use of articles is also controlled. The rules for reordering words are in (17).

(15)

```
perl -pe 's/(\( N [^\)]+ \)) (\( PRON POSS [^\)]+ \)) (\( ADJ
[^\)]+ \)) (\( PRON DEM [^\)]+ \)) (\( NUM [^\)]+ \)) /$4 $2 $5 $3
$1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( PRON POSS [^\)]+ \)) (\( ADJ
[^\)]+ \)) (\( PRON DEM [^\)]+ \)) (\( PRON \:OTE [^\)]+ \)) /$4
$2 $5 $3 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( PRON POSS [^\)]+ \)) (\( ADJ
[^\)]+ \)) (\( PRON DEM [^\)]+ \)) /$4 $2 $3 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( PRON POSS [^\)]+ \)) (\( ADJ
[^\)]+ \)) /$2 $3 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( PRON POSS [^\)]+ \)) (\( PRON DEM
[^\)]+ \)) /$2 $3 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( PRON POSS [^\)]+ \)) (\( NUM
[^\)]+ \)) /$2 $3 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( PRON POSS [^\)]+ \)) (\( PRON
\:OTE [^\)]+ \)) /$2 $3 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( ADJ [^\)]+ \)) (\( NUM [^\)]+ \))
/$3 $2 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( ADJ [^\)]+ \)) (\( PRON \:OTE  
[^\)]+ \)) /$3 $2 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( NUM [^\)]+ \)) (\( PRON \:OTE  
[^\)]+ \)) /$3 $2 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( PRON \:OTE [^\)]+ \)) (\( NUM  
[^\)]+ \)) /$2 $3 $1/gm' | \
```

```
perl -pe 's/(\( N [^\)]+ \)) (\( (PRON POSS|ADJ|NUM|PRON DEM|PRON  
\:OTE) [^\)]+ \)) /$2 $1/gm'
```

Note that the last rule handles five types of short reordering tasks.

Words of the target language are still in the lemma form, and we must convert them to surface form (18).

(18)

```
( PRON DEM :le 8-PL { *those } %SUBJ )  
( PRON POSS 8-PL SG1 { my } %GCON )  
( NUM 8-PL NUM CARD INFL { three } %SUBJ )  
( ADJ { new } A-INFL 8-PL %<NADJ )  
( N { books } CAP 7/8-PL %SUBJ )  
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }  
SVO %FMAINVtr-OBJ> )  
( STOP { . } **CLB )  
( PRON DEM :le 8-PL { *those } %SUBJ )  
( PRON POSS 8-PL SG1 { my } %GCON )  
( PRON :OTE 8-PL { all } %SUBJ )  
( ADJ { new } A-INFL 8-PL %<NADJ )  
( N { books } CAP 7/8-PL %SUBJ )  
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }  
SVO %FMAINVtr-OBJ> )  
( STOP { . } **CLB )  
( PRON DEM :le 8-PL { *those } %SUBJ )  
( PRON POSS 8-PL SG1 { my } %GCON )  
( ADJ { new } A-INFL 8-PL %<NADJ )  
( N { books } CAP 7/8-PL %SUBJ )  
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }  
SVO %FMAINVtr-OBJ> )  
( STOP { . } **CLB )  
( PRON POSS 8-PL SG1 { *my } %GCON )  
( ADJ { new } A-INFL 8-PL %<NADJ )  
( N { books } CAP 7/8-PL %SUBJ )  
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }  
SVO %FMAINVtr-OBJ> )  
( STOP { . } **CLB )  
( PRON POSS 8-PL SG1 { *my } %GCON )  
( PRON DEM :le 8-PL { those } %SUBJ )  
( N { books } CAP 7/8-PL %SUBJ )
```

```
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( PRON POSS 8-PL SG1 { *my } %GCON )
( NUM 8-PL NUM CARD INFL { three } ORD %SUBJ )
( N { books } CAP 7/8-PL %SUBJ )
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( PRON POSS 8-PL SG1 { *my } %GCON )
( PRON :OTE 8-PL { all } %SUBJ )
( N { books } CAP 7/8-PL %SUBJ )
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( NUM 8-PL NUM CARD INFL { *three } %SUBJ )
( ADJ { new } A-INFL 8-PL %<NADJ )
( N { books } CAP 7/8-PL %SUBJ )
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( PRON :OTE 8-PL { *all } %SUBJ )
( ADJ { new } A-INFL 8-PL %<NADJ )
( N { books } CAP 7/8-PL %SUBJ )
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( PRON :OTE 8-PL { *all } %SUBJ )
( NUM 8-PL NUM CARD INFL { three } %SUBJ )
( N { books } CAP 7/8-PL %SUBJ )
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( PRON :OTE 8-PL { *all } %<DN )
( NUM 8-PL NUM CARD INFL { three } %SUBJ )
( N { books } CAP 7/8-PL %SUBJ )
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( PRON POSS 8-PL SG1 { *my } %GCON )
( N { books } CAP 7/8-PL %SUBJ )
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( ADJ { *the } { new } A-INFL 8-PL %<NADJ )
( N { books } CAP 7/8-PL %SUBJ )
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }
SVO %FMAINVtr-OBJ> )
( STOP { . } **CLB )
( NUM 8-PL NUM CARD INFL { *three } %SUBJ )
( N { books } CAP 7/8-PL %SUBJ )
```

```
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }  
SVO %FMAINVtr-OBJ> )  
( STOP { . } **CLB )  
( PRON :OTE 8-PL { *all } %<DN )  
( N { books } CAP 7/8-PL %SUBJ )  
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }  
SVO %FMAINVtr-OBJ> )  
( STOP { . } **CLB )  
( PRON DEM :le 8-PL { *those } %<NDEM )  
( N { books } CAP 7/8-PL %SUBJ )  
( V 8-PL-SP VFIN NO-SP-GLOSS PERF:me :z { :have } { :been lost }  
SVO %FMAINVtr-OBJ> )  
( STOP { . } **CLB )
```

We remove all tags and do the final pruning. The translation is in (19).

(19)

1. *Those my three new books have been lost.*
2. *Those my all new books have been lost.*
3. *Those my new books have been lost.*
4. *My new books have been lost.*
5. *Those my books have been lost.*
6. *My three books have been lost.*
7. *My all books have been lost.*
8. *Three new books have been lost.*
9. *All new books have been lost.*
10. *All three books have been lost.*
11. *All three books have been lost.*
12. *My books have been lost.*
13. *The new books have been lost.*
14. *Three books have been lost.*
15. *All books have been lost.*
16. *Those books have been lost.*

6 Conclusion

In this report I have demonstrated how reordering of words in target language can be implemented in two phases, whereby the rules become clearer and easier to maintain. The presence and absence of word-internal glosses of verbs are controlled using CG rules. The word-internal glosses are then rearranged to meet the requirements of the target language. And finally, the phrase-level and sentence-level rearrangement of words is performed. By splitting the rearrangement task into separate phases, it is easier to control the process and make corrections to rules when needed.