

## Translating Questions from Swahili to English

Arvi Hurskainen  
Institute for Asian and African Studies, Box 59  
FIN-00014 University of Helsinki, Finland  
[arvi.hurskainen@helsinki.fi](mailto:arvi.hurskainen@helsinki.fi)

### Abstract

The report discusses problems encountered in translating Swahili questions into English, using a rule-based translating system. The translation of questions involves problems that are not found in translating other types of sentences. These problems are discussed in detail and solutions to them are demonstrated using examples of each type of problems.

**Key Words:** *machine translation, questions*

### Abbreviations

**CLB	clause boundary
1/2-PL	noun class 1/2 plural
1/2-PL3-SP	noun class 1/2 third person plural subject prefix
1/2-SG	noun class 1/2 singular
1/2-SG2-SP	noun class 1/2 second person singular subject prefix
1/2-SG3-SP	noun class 1/2 third person singular subject prefix
11/10-SG	noun class 11/10 singular
3/4-SG	noun class 3/4 singular
5/6-SG	noun class 5/6 singular
9/10-PL	noun class 9/10 plural
9/10-SG	noun class 9/10 singular
<<MW	multi-word expression, head two to left
<s>	beginning of sentence
<\s>	end of sentence
<IDIOM	idiom, head one to left
<MW	multi-word expression, head one to left
@<NADJ	adjective modifying the noun on the left
@FMAINVintr	finite main verb, intransitive
@FMAINVtr-OBJ>	finite main verb, transitive but without overt object
@GCON	coordinating conjunction
@SUBJ	subject
A	adjective
ADJ	adjective
AR	word of Arabic origin
A-INFL	inflecting adjective
A-UNINFL	uninflecting adjective

CAP	word with capital initial character
CARD	cardinal number
CG	constraint grammar
COND:ngali	conditional with the marker <i>ngali</i>
COND:nge	conditional with the marker <i>nge</i>
DEM:le	demonstrative with the marker <i>le</i>
FUT:ta	future tense with the marker <i>ta</i>
INTERROG	interrogative
MT	machine translation
MW<>	part of multi-word expression, other members one on the left and one on the right
MW>	part of multi-word expression, another member one on the right
MW>>	part of multi-word expression, other members one and two on the right
MWE	multi-word expression
N	noun
NO-SP-GLOSS	no subject prefix gloss
NOART	no article
NOGLOSS	no gloss
NUM	numeral
NUM-INFL	inflecting numeral
PAST	past tense
PERF:me	perfective tense with the marker <i>me</i>
POSS	possessive pronoun
PR:na	present tense with the marker <i>na</i>
PRON	pronoun
QUESTION-MARK	question mark
SALAMA	Swahili Language Manager
SG2	second person singular
STAT	stative verb extension
SV	subject + verb, intransitive verb
SVO	subject + verb + object, transitive verb
TEMP	refers to time
TIME	refers to time
V	verb
VFIN	finite verb

## 1 Introduction

There are two major problems in translating questions from Swahili into English. First, the word order in these two languages is different, including question sentences. Second, English uses the special verb 'do' in yes/no questions in present tense and past tense, while Swahili does not. In addition, a question word in English may be represented by a bound morpheme in Swahili. Also the same morpheme in Swahili may have a different interpretation depending on its location in the sentence. All these features have the effect

that also the production of surface forms of verbs must be handled apart from the ordinary sentences.

## 2 Types of question words

Swahili has question sentences, where the question part may be composed of one or more words.

### 2.1 Single-part question words

*wapi?* where?

*nini?* what?

*lini?* when?

*gani?* what? which?

*nani?* who?

*-ngapi?* how many?

*-pi?* which?

*kwani?* why?

*mbona?* why?

### 2.2 Multi-part question words

*kwa nini?* why?

*kwa sababu gani?* why? for what reason?

*kwa njia gani?* how? in what way?

Of the single-part question words, *wapi*, *nini*, *lini*, *nani*, *kwani* and *mbona* do not allow any modifier, while the words *gani*, *-ngapi* and *-pi* expect a modifier. The modifier may be a modifying noun (*siku gani?* which day?), or in the case of inflecting question words it may be a pronominal prefix that refers to a noun. Examples include *wangapi?* how many (people)? and *yupi?* which (of people)?.

There is also the morpheme *je* that can stand as a separate word in the beginning of the question sentence, or it can be a suffix of the verb. The meaning differs depending on the location of the question morpheme.

## 3 Location of question words

A question word can be placed in Swahili in the beginning or in the end of the question sentence. Examples of words that can be either in the beginning or in the end of the sentence are in (1).

(1)

*Amekuja nani?* Who has come?

*Nani amekuja?* Who has come?

*Amekuja lini?* When has (s)he come?  
*Lini amekuja?* When has (s)he come?

*Watu wangapi wamekuja?* How many people have come?  
*Wamekuja watu wangapi?* How many people have come?

*Yupi amekuja?* Who (of the numbered group) has come?  
*Amekuja yupi?* Who (of the numbered group) has come?

Some question words appear only in the beginning of the sentence (2).

(2)  
*Mbona amekwenda?* Why on earth has (s)he gone?  
*Kwani amekwenda?* Why has (s)he gone?<sup>1</sup>

The question word *gani* requires a noun modifier and it can be placed in the beginning or in the end of the sentence (3).

(3)  
*Mtu gani amekuja?* Which man has come?  
*Amekuja mtu gani?* Which man has come?

Also question words consisting of more than one element can be placed in the beginning or in the end of the sentence (4).

(4)  
*Kwa nini amekuja?* Why has (s)he come?  
*Amekuja kwa nini?* Why has (s)he come?

*Kwa sababu gani amekuja?* Why has (s)he come?  
*Amekuja kwa sababu gani?* Why has (s)he come?

Finally, the question morpheme *je* can appear in the beginning of the question sentence or as a suffix of the main verb (5).

(5)  
*Je, utakwenda?* Will you go?  
*Utakwendaje?* How will you go?

---

<sup>1</sup> *Kwani* is in fact an abbreviation of *kwa nini*.

## 4 Implementing the translation

The translation of questions from Swahili into English includes a number of problems, and their solving is not always simple. Below I shall discuss the problems using various examples and showing how each type of problems can be solved.

There are many types of questions, each with its own types of translation problems. Each type is discussed below.

### 4.1 Question word as subject

The example in (6) is a question *Amekuja nani?*. The verb has two interpretations, one with the subject marker gloss '(s)he' and the other without the gloss.

```
(6)
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
    "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
CAP @FMAINVintr
"<nani>"
    "nani" INTERROG { who } @SUBJ
    "<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

The verb is disambiguated using the following subject (@SUBJ) as key. Because there is an overt subject, the subject prefix gloss of the verb must be removed. The result is in (7).

```
(7)
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
CAP @FMAINVintr
"<nani>"
    "nani" INTERROG { who } @SUBJ
    "<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

For translation, the word order must be changed. In order to make rule writing more convenient, the presentation of the analysis is modified and the sentence is put on a single line (8).

(8)  
( <s> ) ( V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me { come } SV CAP )  
( INTERROG { who } @SUBJ ) ( "?" QUESTION-MARK { ? } \*\*CLB ) (  
<\s> )

With a rule, the word order is now changed and the correct verb form produced (9).

(9)  
( <s> ) ( INTERROG { who } @SUBJ ) ( { has } ) ( V 1/2-SG3-SP  
VFIN NO-SP-GLOSS PERF:me { come } SV CAP ) ( "?" QUESTION-MARK { ?  
} \*\*CLB ) ( <\s> )

The question *Amekuja nani?* can also be in the form *Nani amekuja?*. No word order change is needed in this example, but the subject prefix gloss must be removed (10).

(10)  
<s>  
"<\*nani>"  
    "nani" INTERROG { who } CAP @SUBJ  
"<amekuja>"  
    "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV  
@FMAINVintr  
"<?>"  
    "?" QUESTION-MARK { ? } \*\*CLB  
<\s>

Note that in (10) the verb does not yet have the correct surface form. This is implemented later in the format, where each word with its reading is on a single line (11).

(11)  
( <s> )  
( INTERROG { \*who } CAP @SUBJ )  
( V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me { has } { come } SV )  
( "?" QUESTION-MARK { ? } \*\*CLB )  
( <\s> )

## 4.2 Subject prefix of the verb as subject

In the example above, the question word is a subject. In (12) we have an example, where the subject prefix of the verb functions as a subject. Therefore, its gloss must be retained.

(12)  
<s>  
"<\*amekuja>"  
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV  
CAP @FMAINVintr  
"<lini>"  
    "lini" INTERROG { when }  
"<?>"

```
"?" QUESTION-MARK { ? } **CLB  
<\s>
```

After reordering the constituents and introducing the correct verb form we get the final form (13).

```
(13)  
( <s> ) ( INTERROG { when } TEMP ) ( { has } ) ( { (s)he } V  
1/2-SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK { ?  
} **CLB ) ( <\s> )
```

When the word order in the question is different, we get the analysis as in (14).

```
(14)  
<s>  
"<*lini>"  
  "lini" INTERROG { when } TEMP CAP  
"<amekuja>"  
  "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV  
@FMAINVintr  
"<?>"  
  "?" QUESTION-MARK { ? } **CLB  
<\s>
```

When the result is modified, we get the format as in (15).

```
(15)  
( <s> )  
( INTERROG { *when } TEMP CAP )  
( has )  
( V 1/2-SG3-SP VFIN { (s)he } PERF:me { come } SV )  
( "?" QUESTION-MARK { ? } **CLB )  
( <\s> )
```

### 4.3 Question word as a modifier to noun

There are two question words (*-ngapi* and *gani*) that function as modifiers to a noun. Such noun phrases can be before or after the verb. Of these question words, *gani* requires always the main noun, while *-ngapi* may also stand alone, in which case it has a pronominal function.

These constructions require a number of operations, including re-ordering, deletion, insertion, and substitution. In (16), there is the case, where the question word comes after the verb.

```
(16)  
<s>  
"<*wamekuja>"
```

```
"ja" V 1/2-PL3-SP VFIN { they } PERF:me [ja] { come } SV CAP
@FMAINVintr
"ja" V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
CAP @FMAINVintr
"<watu>"
"mtu" N 1/2-PL { the } { man } @SUBJ
"<wangapi>"
"ngapi" ADJ A-INFL 1/2-PL NOART INTERROG { how many } @<NADJ
"<?>"
"?" QUESTION-MARK { ? } **CLB
<\s>
```

The verb in (16) has two interpretations, one with and another without the subject prefix gloss. The latter must be selected, because the overt verb follows the verb. The selection is performed using a disambiguation rule (17).

```
(17)
<s>
"<*wamekuja>"
"ja" V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
CAP @FMAINVintr
"<watu>"
"mtu" N 1/2-PL { the } { man } @SUBJ
"<wangapi>"
"ngapi" ADJ A-INFL 1/2-PL NOART INTERROG { how many } @<NADJ
"<?>"
"?" QUESTION-MARK { ? } **CLB
<\s>
```

The next step is to re-order the lexical items to meet the requirements of the target language (18).

```
(18)
( <s> ) ( ADJ A-INFL 1/2-PL NOART INTERROG { how many } ) ( N 1/2-
PL { man } @SUBJ ) ( V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me { come
} SV CAP ) ( "?" QUESTION-MARK { ? } **CLB ) ( <\s> )
```

Finally, the surface forms of the target language will be produced. This is done using re-writing and insertion (19). Note that the sentence-initial character has been capitalised.

```
(19)
( <s> )
( ADJ A-INFL 1/2-PL NOART INTERROG { *how many } )
( N 1/2-PL { people } @SUBJ )
( V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me { have } { come } SV CAP )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
```



If the question word is before the verb, the corresponding phases are as exemplified in (20-23).

(20)  
<s>  
"<\*watu>"  
"mtu" N 1/2-PL { the } { man } CAP @SUBJ  
"<wangapi>"  
"ngapi" ADJ A-INFL 1/2-PL NOART INTERROG { how many } @<NADJ  
"<wamekuja>"  
"ja" V 1/2-PL3-SP VFIN { they } PERF:me [ja] { come } SV  
@FMAINVintr  
"ja" V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV  
@FMAINVintr  
"<?>"  
"?" QUESTION-MARK { ? } \*\*CLB  
<\s>

(21)  
<s>  
"<\*watu>"  
"mtu" N 1/2-PL { the } { man } CAP @SUBJ  
"<wangapi>"  
"ngapi" ADJ A-INFL 1/2-PL NOART INTERROG { how many } @<NADJ  
"<wamekuja>"  
"ja" V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV  
@FMAINVintr  
"<?>"  
"?" QUESTION-MARK { ? } \*\*CLB  
<\s>

(22)  
( <s> ) ( ADJ A-INFL 1/2-PL NOART INTERROG { how many } ) ( N 1/2-  
PL { man } CAP @SUBJ ) ( V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me {  
come } SV ) ( "?" QUESTION-MARK { ? } \*\*CLB ) ( <\s> )

(23)  
( <s> )  
( ADJ A-INFL 1/2-PL NOART INTERROG { \*how many } )  
( N 1/2-PL { people } CAP @SUBJ )  
( V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me { have } { come } SV )  
( "?" QUESTION-MARK { ? } \*\*CLB )  
( <\s> )

The question word *gani* behaves much in the same way as described above. The processing phases with *gani* after the verb are in (24-27).

(24)  
<s>  
"<\*amekuja>"

```

    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
    "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
CAP @FMAINVintr
"<mtu>"
    "mtu" N 1/2-SG { the } { man } @SUBJ
"<gani>"
    "gani" ADJ A-UNINFL INTERROG { what } @<NADJ
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>

```

(25)

```

<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
CAP @FMAINVintr
"<mtu>"
    "mtu" N 1/2-SG { the } { man } @SUBJ
"<gani>"
    "gani" ADJ A-UNINFL INTERROG { what } @<NADJ
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>

```

(26)

```

( <s> ) ( ADJ A-UNINFL INTERROG { what } ) ( N 1/2-SG { man }
@SUBJ ) ( V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me { come } SV CAP )
( "?" QUESTION-MARK { ? } **CLB ) ( <\s> )

```

(27)

```

( <s> )
( ADJ A-UNINFL INTERROG { *what } )
( N 1/2-SG { man } @SUBJ )
( V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me { has } { come } SV CAP )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )

```

When the question word *gani* is before the verb, the processing phases are as exemplified in (28-31).

(28)

```

<s>
"<*mtu>"
    "mtu" N 1/2-SG { the } { man } CAP @SUBJ
"<gani>"
    "gani" ADJ A-UNINFL INTERROG { what } @<NADJ
"<amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
@FMAINVintr

```

```
"ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
@FMAINVintr
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
```

(29)

```
<s>
"<*mtu>"
  "mtu" N 1/2-SG { the } { man } CAP @SUBJ
"<gani>"
  "gani" ADJ A-UNINFL INTERROG { what } @<NADJ
"<amekuja>"
  "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
@FMAINVintr
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
```

(30)

```
( <s> ) ( ADJ A-UNINFL INTERROG { what } ) ( N 1/2-SG { man } CAP
@SUBJ ) ( V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me { come } SV ) (
  "?" QUESTION-MARK { ? } **CLB ) ( <\s> )
```

(31)

```
( <s> )
( ADJ A-UNINFL INTERROG { *what } )
( N 1/2-SG { man } CAP @SUBJ )
( V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me { has } { come } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
```

#### 4.4 Multi-part question words

When the question part of the sentence consists of more than one word, this part is isolated as a multi-word expression (32).

(32)

```
<s>
"<*kwa_nini>"
  "kwa_nini" INTERROG { why }
"<amekuja>"
  "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
@FMAINVintr
  "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
@FMAINVintr
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
```

The verb must retain the subject prefix gloss, because there is no overt subject in the sentence (33).

```
(33)
<s>
"<*kwa_nini>"
  "kwa_nini" INTERROG { why }
"<amekuja>"
  "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
@FMAINVintr
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
```

Finally, the surface forms are produced (34).

```
(34)
( <s> )
( INTERROG { *why } )
( V 1/2-SG3-SP VFIN { (s)he } PERF:me { has } { come } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
```

Still, the order of the subject pronoun and verb modules must be corrected. This is done in the post processing phase (35).

```
(35)
<s> Why has (s)he come? <\s>
```

Processing phases of the question construction *kwa sababu gani* are shown in (36-39).

```
(36)
<s>
"<*kwa_sababu_gani>"
  "kwa_sababu_gani" INTERROG { for what reason }
"<amekuja>"
  "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
@FMAINVintr
  "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
@FMAINVintr
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
```

```
(37)
<s>
"<*kwa_sababu_gani>"
```

```
"kwa_sababu_gani" INTERROG { for what reason }
"<amekuja>"
  "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
@FMAINVintr
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
```

```
(38)
( <s> )
( INTERROG { *for what reason } )
( V 1/2-SG3-SP VFIN { (s)he } PERF:me { has } { come } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
```

```
(39)
<s> For what reason has (s)he come? <\s>
```

#### 4.5 The particle *je* as a question morpheme

There is the question type that uses *je* as a question marker. This particle can be before or after the verb. If before the verb, it is written as a separate word, often followed by a comma (40).

```
(40)
<s>
"<*je,>"
  "je," INTERROG { } NOGLOSS CAP
"<utakwenda>"
  "kwenda" V 1/2-SG2-SP VFIN { you } FUT:ta [enda] { go } SV
@FMAINVintr
  "kwenda" V 1/2-SG2-SP VFIN NO-SP-GLOSS FUT:ta [enda] { go }
SV @FMAINVintr
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
```

The interpretation of the verb with the subject prefix gloss is chosen (41).

```
(41)
<s>
"<*je,>"
  "je," INTERROG { } NOGLOSS CAP
"<utakwenda>"
  "kwenda" V 1/2-SG2-SP VFIN { you } FUT:ta [enda] { go } SV
@FMAINVintr
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
```

The word order on the lexical level does not change, as shown in (42).

```
(42)
( <s> ) ( INTERROG { } NOGLOSS CAP ) ( V 1/2-SG2-SP VFIN { you }
FUT:ta { go } SV ) ( "?" QUESTION-MARK { ? } **CLB ) ( <\s> )
```

When the surface form is produced, the word order needs modification (43).

```
(43)
( <s> )
( INTERROG { * } NOGLOSS CAP )
( V 1/2-SG2-SP VFIN { you } FUT:ta { will } { go } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
```

The final word order is achieved in the post processing phase (44).

```
(44)
<s> Will you go? <\s>
```

When the question particle *je* is after the verb, it is attached to it as a suffix. In this position it means 'how', 'in what way?'. The processing phases are in (45-48)

```
(45)
<s>
"<*utakwendaje>"
    "kwenda" V 1/2-SG2-SP VFIN { you } FUT:ta [enda] { go } SV
INTERROG { how } CAP @FMAINVintr
    "kwenda" V 1/2-SG2-SP VFIN NO-SP-GLOSS FUT:ta [enda] { go }
SV INTERROG { how } CAP @FMAINVintr
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

```
(46)
<s>
"<*utakwendaje>"
    "kwenda" V 1/2-SG2-SP VFIN { you } FUT:ta [enda] { go } SV
INTERROG { how } CAP @FMAINVintr
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

```
(47)
( <s> ) ( INTERROG { how } CAP ) ( { will } ) ( { you } V 1/2-
SG2-SP VFIN FUT:ta { go } SV ) ( "?" QUESTION-MARK { ? } **CLB ) (
<\s> )
```

```
(48)
( <s> )
( INTERROG { *how } CAP )
( { will } )
( { you } V 1/2-SG2-SP VFIN FUT:ta { go } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
```

#### 4.6 Other ways of expressing 'how?' in Swahili

There are also other ways to express 'how?' in Swahili, such as *kwa njia gani? kwa namna gani?* These are treated as multi-word expressions (MWE). These MWEs are isolated using the Constraint Grammar parser (49).

```
(49)
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
"<kwa>"
    "kwa" MW>>
"<njia>"
    "njia" MW<>
"<gani>"
    "gani" INTERROG <<MW { how }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
"<kwa>"
    "kwa" MW>>
"<namna>"
    "namna" MW<>
"<gani>"
    "gani" INTERROG <<MW { how }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

Then various modifications are done to meet the requirements of the target language (50). Note that in this phase both strings are identical.

(50)  
( <s> ) ( INTERROG { how } <<MW ) ( { has } ) ( { (s)he } V 1/2-  
SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK { ? }  
\*\*CLB ) ( <\s> )  
( <s> ) ( INTERROG { how } <<MW ) ( { has } ) ( { (s)he } V 1/2-  
SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK { ? }  
\*\*CLB ) ( <\s> )

When the verb follows the question in these structures, only the glosses of the verb need to be re-ordered (51).

(51)  
<s>  
"<\*kwa>"  
    "kwa" MW>>  
"<namna>"  
    "namna" MW<<  
"<gani>"  
    "gani" INTERROG <<MW { how }  
"<amekuja>"  
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV  
@FMAINVintr  
"<?>"  
    "?" QUESTION-MARK { ? } \*\*CLB  
<\s>

The surface form of the verb is first produced (52).

(52)  
( <s> )  
( INTERROG { \*how } <<MW )  
( V 1/2-SG3-SP VFIN { (s)he } PERF:me { has } { come } SV )  
( "?" QUESTION-MARK { ? } \*\*CLB )  
( <\s> )

Then the order of the glosses in the verb is re-arranged (53).

(53)  
( <s> )  
( INTERROG { \*how } <<MW )  
( { has } )  
( V 1/2-SG3-SP VFIN { (s)he } PERF:me { come } SV )  
( "?" QUESTION-MARK { ? } \*\*CLB )  
( <\s> )



#### 4.7 Questions referring to time

There are frequently occurring questions referring to time, such as *saa ngapi? siku gani? juma gani? mwezi gani? mwaka gani? wakati gani?* etc. These structures are first analysed as in (54).

```
(54)
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP
"<siku>"
    "siku" N 9/10-SG { the } { day } TIME
"<gani>"
    "gani" ADJ A-UNINFL INTERROG { what }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP
"<juma>"
    "juma" N 5/6-SG { the } { week } TIME AR
"<gani>"
    "gani" ADJ A-UNINFL INTERROG { what }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP
"<mwezi>"
    "mwezi" N 3/4-SG { the } { month } TIME
    "mwezi" N 3/4-SG { the } { moon }
"<gani>"
    "gani" ADJ A-UNINFL INTERROG { what }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP
"<mwaka>"
    "mwaka" N 3/4-SG { the } { year } TIME
"<gani>"
    "gani" ADJ A-UNINFL INTERROG { what }
```

```
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP
"<wakati>"
    "wakati" N 11/10-SG { the } { time } TIME AR
"<gani>"
    "gani" ADJ A-UNINFL INTERROG { what }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP
"<saa>"
    "saa" N 9/10-PL { the } { hour } TIME AR
"<ngapi>"
    "ngapi" ADJ A-INFL 9/10-PL NOART INTERROG { how many }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

Then the question structures of the sentence are isolated as MWEs (55).

```
(55)
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
"<siku>"
    "siku" MW>
"<gani>"
    "gani" INTERROG { what day } <MW
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
"<juma>"
    "juma" MW>
"<gani>"
    "gani" INTERROG { what week } <MW
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

```
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
"<mwezi>"
    "mwezi" MW>
"<gani>"
    "gani" INTERROG { what month } <MW
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
"<mwaka>"
    "mwaka" MW>
"<gani>"
    "gani" INTERROG { what year } <MW
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
"<wakati>"
    "wakati" MW>
"<gani>"
    "gani" INTERROG { what time } <MW
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*amekuja>"
    "ja" V 1/2-SG3-SP VFIN { (s)he } PERF:me [ja] { come } SV
CAP @FMAINVintr
"<saa>"
    "saa" MW>
"<ngapi>"
    "ngapi" INTERROG { what time } <MW
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

This having been done, the constituents will be re-ordered and the surface forms produced (56).

(56)

```
( <s> ) ( INTERROG { what day } <IDIOM ) ( { has } ) ( { (s)he }
V 1/2-SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK {
? } **CLB ) ( <\s> )
( <s> ) ( INTERROG { what week } <IDIOM ) ( { has } ) ( { (s)he
} V 1/2-SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK
{ ? } **CLB ) ( <\s> )
( <s> ) ( INTERROG { what month } <IDIOM ) ( { has } ) ( { (s)he
} V 1/2-SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK
{ ? } **CLB ) ( <\s> )
( <s> ) ( INTERROG { what year } <IDIOM ) ( { has } ) ( { (s)he
} V 1/2-SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK
{ ? } **CLB ) ( <\s> )
( <s> ) ( INTERROG { what time } <IDIOM ) ( { has } ) ( { (s)he
} V 1/2-SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK
{ ? } **CLB ) ( <\s> )
( <s> ) ( INTERROG { what time } <IDIOM ) ( { has } ) ( { (s)he
} V 1/2-SG3-SP VFIN PERF:me { come } SV CAP ) ( "?" QUESTION-MARK
{ ? } **CLB ) ( <\s> )
```

#### 4.8 Question sentences with overt subject

So far the examples have been without an overt subject, except for the cases, where the question word itself is a subject. Let us see what happens, when we have an overt subject, perhaps with modifications, such as adjectives, pronouns and numbers. In (57) are examples of such structures, analyzed and disambiguated.

(57)

```
<s>
"<*mwalimu>"
  "mwalimu" N 1/2-SG { the } { teacher } CAP @SUBJ
"<yule>"
  "yule" PRON DEM:le 1/2-SG { that }
"<atakuja>"
  "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS FUT:ta [ja] { come } SV
@FMAINVintr
"<lini>"
  "lini" INTERROG { when } TEMP
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*mwalimu>"
  "mwalimu" N 1/2-SG { the } { teacher } CAP @SUBJ
"<wako>"
  "ako" PRON POSS 1/2-SG SG2 { your } @GCON
"<ameondoka>"
  "ondoka" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ondoka] {
leave } PFR SVO STAT @FMAINVtr-OBJ>
"<lini>"
```

```

    "lini" INTERROG { when } TEMP
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*mwalimu>"
    "mwalimu" N 1/2-SG { the } { teacher } CAP @SUBJ
"<hodari>"
    "hodari" ADJ A-UNINFL { skilful } AR @<NADJ
"<amekuja>"
    "ja" V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me [ja] { come } SV
@FMAINVintr
"<lini>"
    "lini" INTERROG { when } TEMP
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*walimu>"
    "mwalimu" N 1/2-PL { the } { teacher } CAP @SUBJ
"<wawili>"
    "wili" NUM 1/2-PL NUM-INFL CARD { two } @SUBJ
"<wamekwenda>"
    "kwenda" V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me [enda] { go }
SV @FMAINVintr
"<wapi>"
    "wapi" INTERROG { where }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>

```

In (57), we see that the subject prefix gloss has already been deleted, because the overt subject replaces it. Also when the modifier of the noun is an adjective, the article is retained, while in the case of pronouns and numerals it is deleted.

```

(57)
( <s> ) ( INTERROG { when } TEMP ) ( { will } ) ( PRON DEM:le
1/2-SG { that } ) ( N 1/2-SG { teacher } CAP @SUBJ ) ( V 1/2-SG3-
SP VFIN NO-SP-GLOSS FUT:ta { come } SV ) ( "?" QUESTION-MARK { ? }
**CLB ) ( <\s> )
( <s> ) ( INTERROG { when } TEMP ) ( { has } ) ( PRON POSS 1/2-SG
SG2 { your } ) ( N 1/2-SG { teacher } CAP @SUBJ ) ( V 1/2-SG3-SP
VFIN NO-SP-GLOSS PERF:me { leave } PREFR SVO STAT ) ( "?"
QUESTION-MARK { ? } **CLB ) ( <\s> )
( <s> ) ( INTERROG { when } TEMP ) ( { has } ) ( { the } ) ( ADJ
A-UNINFL { skilful } ) ( N 1/2-SG { teacher } CAP @SUBJ ) ( V 1/2-
SG3-SP VFIN NO-SP-GLOSS PERF:me { come } SV ) ( "?" QUESTION-MARK
{ ? } **CLB ) ( <\s> )
( <s> ) ( INTERROG { where } ) ( { have } ) ( NUM 1/2-PL NUM-INFL
CARD { two } @SUBJ ) ( N 1/2-PL { teacher } CAP @SUBJ ) ( V 1/2-

```

PL3-SP VFIN NO-SP-GLOSS PERF:me { go } SV ) ( "?" QUESTION-MARK {  
 ? } \*\*CLB ) ( <\s> )

The surface forms in the target language are still defective in (57). In (58) they are completed, and also the word order is correct.

(58)  
 ( <s> )  
 ( INTERROG { \*when } TEMP )  
 ( { will } )  
 ( PRON DEM:le 1/2-SG { that } )  
 ( N 1/2-SG { teacher } CAP @SUBJ )  
 ( V 1/2-SG3-SP VFIN NO-SP-GLOSS FUT:ta { come } SV )  
 ( "?" QUESTION-MARK { ? } \*\*CLB )  
 ( <\s> )  
 ( <s> )  
 ( INTERROG { \*when } TEMP )  
 ( { has } )  
 ( PRON POSS 1/2-SG SG2 { your } )  
 ( N 1/2-SG { teacher } CAP @SUBJ )  
 ( V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me { left } PREFER SVO STAT )  
 ( "?" QUESTION-MARK { ? } \*\*CLB )  
 ( <\s> )  
 ( <s> )  
 ( INTERROG { \*when } TEMP )  
 ( { has } )  
 ( { the } )  
 ( ADJ A-UNINFL { skilful } )  
 ( N 1/2-SG { teacher } CAP @SUBJ )  
 ( V 1/2-SG3-SP VFIN NO-SP-GLOSS PERF:me { come } SV )  
 ( "?" QUESTION-MARK { ? } \*\*CLB )  
 ( <\s> )  
 ( <s> )  
 ( INTERROG { \*where } )  
 ( { have } )  
 ( NUM 1/2-PL NUM-INFL CARD { two } @SUBJ )  
 ( N 1/2-PL { teachers } CAP @SUBJ )  
 ( V 1/2-PL3-SP VFIN NO-SP-GLOSS PERF:me { gone } SV )  
 ( "?" QUESTION-MARK { ? } \*\*CLB )  
 ( <\s> )

#### 4.9 Questions in various tenses

The examples above have been mostly in perfective tense. Various tenses require separate treatment for producing the correct word order and surface forms in target language. Consider the conditional present and conditional perfective tenses in (59).

(59)  
 <s>  
 "<\*angekwenda>"

```

      "kwenda" V 1/2-SG3-SP VFIN { (s)he } COND:nge [enda] { go }
SV CAP @FMAINVintr
"<wapi>"
  "wapi" INTERROG { where }
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*angalikwenda>"
  "kwenda" V 1/2-SG3-SP VFIN { (s)he } COND:ngali [enda] { go
} SV CAP @FMAINVintr
"<wapi>"
  "wapi" INTERROG { where }
"<?>"
  "?" QUESTION-MARK { ? } **CLB
<\s>

```

The final word order and correct surface forms are produced with corresponding rules (60).

```

(60)
( <s> )
( INTERROG { *where } )
( { would } )
( { (s)he } V 1/2-SG3-SP VFIN COND:nge { go } SV CAP )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
( <s> )
( INTERROG { *where } )
( { would } )
( { (s)he } V 1/2-SG3-SP VFIN ( { have } )
COND:ngali { gone } SV CAP )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )

```

When we add an overt subject and a modifier to it, we get the results as in (61-62).

```

(61)
<s>
"<*mwalimu>"
  "mwalimu" N 1/2-SG { the } { teacher } CAP @SUBJ
"<yule>"
  "yule" PRON DEM:le 1/2-SG { that }
"<angekwenda>"
  "kwenda" V 1/2-SG3-SP VFIN NO-SP-GLOSS COND:nge [enda] { go
} SV @FMAINVintr
"<wapi>"
  "wapi" INTERROG { where }
"<?>"
  "?" QUESTION-MARK { ? } **CLB

```

```
<\s>
<s>
"<*mwalimu>"
    "mwalimu" N 1/2-SG { the } { teacher } CAP @SUBJ
"<yule>"
    "yule" PRON DEM:le 1/2-SG { that }
"<angalikwenda>"
    "kwenda" V 1/2-SG3-SP VFIN NO-SP-GLOSS COND:ngali [enda] {
go } SV @FMAINVintr
"<wapi>"
    "wapi" INTERROG { where }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
```

```
(62)
( <s> )
( INTERROG { *where } )
( { would } )
( PRON DEM:le 1/2-SG { that } )
( N 1/2-SG { teacher } CAP @SUBJ )
( V 1/2-SG3-SP VFIN NO-SP-GLOSS COND:ngali { go } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
( <s> )
( INTERROG { *where } )
( { would } )
( PRON DEM:le 1/2-SG { that } )
( N 1/2-SG { teacher } CAP @SUBJ )
( { have } )
( V 1/2-SG3-SP VFIN NO-SP-GLOSS COND:ngali { gone } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
```

#### 4.10 Translating questions with 'do' as the compulsory element

English is a peculiar language in that in present and past tenses it uses the special verb 'do' in questions. This again causes problems in translation. Consider the disambiguated examples in (63).

```
(63)
<s>
"<*mwalimu>"
    "mwalimu" N 1/2-SG { the } { teacher } CAP @SUBJ
"<yule>"
    "yule" PRON DEM:le 1/2-SG { that }
"<anakwenda>"
    "kwenda" V 1/2-SG3-SP VFIN NO-SP-GLOSS PR:na [enda] { go }
SV @FMAINVintr
"<wapi>"
```



```

      "wapi" INTERROG { where }
"<?>"
      "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*mwalimu>"
      "mwalimu" N 1/2-SG { the } { teacher } CAP @SUBJ
"<yule>"
      "yule" PRON DEM:le 1/2-SG { that }
"<alikhwenda>"
      "kwenda" V 1/2-SG3-SP VFIN NO-SP-GLOSS PAST [enda] { go } SV
@FMAINVintr
"<wapi>"
      "wapi" INTERROG { where }
"<?>"
      "?" QUESTION-MARK { ? } **CLB
<\s>

```

The correct forms in the target language are in (64).

```

(64)
( <s> )
( INTERROG { *where } )
( { does } )
( PRON DEM:le 1/2-SG { that } )
( N 1/2-SG { teacher } CAP @SUBJ )
( V 1/2-SG3-SP VFIN NO-SP-GLOSS PR:na { go } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
( <s> )
( INTERROG { *where } )
( { did } )
( PRON DEM:le 1/2-SG { that } )
( N 1/2-SG { teacher } CAP @SUBJ )
( V 1/2-SG3-SP VFIN NO-SP-GLOSS PAST { go } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )

```

When we change the subject into plural, we get the forms as in (65-66).

```

(65)
<s>
"<*walimu>"
      "mwalimu" N 1/2-PL { the } { teacher } CAP @SUBJ
"<wale>"
      "wale" PRON DEM:le 1/2-PL { those }
"<wanakwenda>"
      "kwenda" V 1/2-PL3-SP VFIN NO-SP-GLOSS PR:na [enda] { go }
SV @FMAINVintr
"<wapi>"

```

```

    "wapi" INTERROG { where }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>
<s>
"<*walimu>"
    "mwaliimu" N 1/2-PL { the } { teacher } CAP @SUBJ
"<wale>"
    "wale" PRON DEM:le 1/2-PL { those }
"<walikwenda>"
    "kwenda" V 1/2-PL3-SP VFIN NO-SP-GLOSS PAST [enda] { go } SV
@FMAINVintr
"<wapi>"
    "wapi" INTERROG { where }
"<?>"
    "?" QUESTION-MARK { ? } **CLB
<\s>

```

```

(66)
( <s> )
( INTERROG { *where } )
( { do } )
( PRON DEM:le 1/2-PL { those } )
( N 1/2-PL { teachers } CAP @SUBJ )
( V 1/2-PL3-SP VFIN NO-SP-GLOSS PR:na { go } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )
( <s> )
( INTERROG { *where } )
( { did } )
( PRON DEM:le 1/2-PL { those } )
( N 1/2-PL { teachers } CAP @SUBJ )
( V 1/2-PL3-SP VFIN NO-SP-GLOSS PAST { go } SV )
( "?" QUESTION-MARK { ? } **CLB )
( <\s> )

```

## 5 Discussion and conclusion

Due to the very different structures of Swahili and English languages, the MT from Swahili to English requires careful planning. This is particularly true of questions, where the differences are especially complex. The word order in English questions is rather fixed, but in Swahili most question sentences may have the question part before or after the verb. English changes the order of constituents in questions, so that the subject comes between the auxiliary verb and main verb. In Swahili this change does not happen. English uses the special verb 'do' in present and past tense questions, which means that this has to be added in translating from Swahili to English.

In the current implementation of SALAMA, the standard processing order is such that first the lexical glosses are put into the order required by the target language, and then the surface forms are produced. Questions need a different approach, because the

subject comes between the auxiliary verb and main verb. Therefore, the final word order can be fixed only after the surface forms have been completed.

This problem can be solved in two ways. One method is to re-order the words in questions after the surface forms have been completed. This requires three phases. First, the lexical glosses are re-ordered, Second, the surface forms are produced. Third, the surface words in question sentences are re-ordered. Another method is to do all these three phases with a single rule. What cannot be done with this method is the production of the surface forms for the main verb, because each of them has to be treated individually. However, the production of the surface forms of the main verb can be done with the same rules that handle the verbs in other types of questions. In the current implementation, the question sentences were implemented using the latter solution.