

Multi-word verbs in English to Finnish machine translation I: Single-word to multi-word verb mapping¹

Arvi Hurskainen
Department of World Cultures, Box 59
FIN-00014 University of Helsinki, Finland
arvi.hurskainen@helsinki.fi

Abstract

The report discusses such cases, where a single-word English verb maps to a multi-word Finnish verb. Translation to Finnish includes many problems, because the multi-word solution in translation affects the inflection of the modifiers of the verb. The translation system has the default rule set for single-verb to single-verb mapping. However, in the case of single-verb to multi-verb mapping, the inflection rules are different. Therefore, these cases must be kept strictly apart from the other cases. I have collected all (more than 500) such English verbs, which have at least one such Finnish gloss, which is a multi-word verb. The verbs can be subdivided into sub-groups according to the inflection patterns, which they require. The solutions for translating each type are discussed and demonstrated.

Key Words: *morphology, multi-words, machine translation.*

1 Introduction

In machine translation, multi-word expressions constitute a major problem, because they occur frequently in text, and each of them must be handled precisely in the way, which they require. Even in normal prose text, the production of correct Finnish is a challenge, mainly due the large number of inflection rules.

Many English verbs have several glosses in Finnish, and often a verb has single-word glosses and one or more multi-word glosses. Therefore, the English verb prompts different inflection patterns depending on whether the verb is translated with a single word or with a multi-word. In order to control the inflection pattern with each gloss, the instructions on inflection must be encoded, so that when disambiguation is done, the selected gloss has all required inflection instructions at hand, or at least the verb should not have any wrong instructions. A defective inflection pattern can be supplemented later in the process.

Especially the control of the Finnish case system is challenging, because the case structure is different in affirmative and passive clause structures.

¹ The report is issued under licence CC BY-NC

The Finnish multi-word verbs normally have two components. The first one is a verb, and the second one can belong to almost any of the other POS categories. They also often inflect, although with a very restricted paradigm. In order to control the rule writing setting, I have found it useful to split the multi-word verb reading into two separate readings, one for the verb itself, and the other for the modifying part of the structure.

Below we will see how various types of multi-word verbs in context are handled for producing the correct translation in each case.

2 Multi-word verbs with adverb as the second member

There are 432 such multi-word verbs, where the second member is an adverb or some other non-inflecting word, such as postposition. Because they do not inflect, inflection rules for the second member are not needed. For the verb part, separate inflection rules are needed for the verb modifier. The modifier may be an object or a modifier that is expressed with a prepositional structure. Their position in target language is normally immediately after the verb. However, there is variation, because the prosodic features, such as stress, may affect the word order. In this report, prosodology is not discussed.

Below is an example the verb *apologise*, which is translated in two different contexts. The context decides, which of the glosses is selected (1).

(1)

```
"<He>"
"he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM
SG3
"he" { hänen } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { NOGLOSS } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { itse N8 FRONT } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { PROP-CAND } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"<apologised>"
"apologise" { pahoitella V67-C O-PAR } %+FMAINV V PAST
"apologise" { pyytää V54-F FRONT O-ABL || anteeksi } %+FMAINV V
PAST
"<his>"
"he" { hän Np9 FRONT OUT } HUM MALE %A> PRON PERS GEN SG3
"he" { hänen } HUM MALE %A> PRON PERS GEN SG3
"he" { NOGLOSS } HUM MALE %A> PRON PERS GEN SG3
"he" { itse N8 FRONT } HUM MALE %A> PRON PERS GEN SG3
"<mistake>"
"mistake" { erehdys N39 FRONT } IN %OBJ DEF N SG
"mistake" { virhe N48 FRONT } IN %OBJ DEF N SG
"<.>"
"." { . }
"<<s>>"
"<s>" { <s> }
"<He>"
"he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM
SG3
"he" { hänen } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { NOGLOSS } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { itse N8 FRONT } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { PROP-CAND } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"<apologised>"
```

```

    "apologise" { pahoitella V67-C O-PAR } %+FMAINV V PAST
    "apologise" { pyytää V54-F FRONT O-ABL || anteeksi } %+FMAINV V
PAST
"<his>"
    "he" { hän Np9 FRONT OUT } HUM MALE %A> PRON PERS GEN SG3
    "he" { hänen } HUM MALE %A> PRON PERS GEN SG3
    "he" { NOGLOSS } HUM MALE %A> PRON PERS GEN SG3
    "he" { itse N8 FRONT } HUM MALE %A> PRON PERS GEN SG3
"<teacher>"
    "teacher" { opettaja N10 } HUM %OBJ DEF N SG
"<.>"
    "." { . }

```

The verb apologise has the glosses *pahoitella* and *pyytää anteeksi*. When we disambiguate the sentences, we see that different alternatives will be selected depending on context (2).

```

(2)
"<He>"
    "he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM
SG3
"<apologised>"
    "apologise" { pahoitella V67-C O-PAR } %+FMAINV V PAST
"<his>"
    "he" { NOGLOSS } HUM MALE %A> PRON PERS GEN SG3
"<mistake>"
    "mistake" { erehdys N39 FRONT } IN %OBJ DEF N SG
"<.>"
    "." { . }
"<<s>>"
    "<s>" { <s> }
"<He>"
    "he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM
SG3
"<apologised>"
    "apologise" { pyytää V54-F FRONT O-ABL || anteeksi } %+FMAINV V
PAST
"<his>"
    "he" { NOGLOSS } HUM MALE %A> PRON PERS GEN SG3
"<teacher>"
    "teacher" { opettaja N10 } HUM %OBJ DEF N SG
"<.>"
    "." { . }

```

Next we add inflection tags to the readings (3),

```

(3)
"<He>"
    "he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3
NOM
"<apologised>"
    "apologise" { pahoitella V67-C } %+FMAINV O-PAR V PAST SG
"<his>"
    "he" { NOGLOSS } %A> HUM MALE PRON PERS SG3 GEN
"<mistake>"
    "mistake" { erehdys N39 FRONT } %OBJ IN DEF N SG PAR @POS-INFL

```

```
"<.>"  
  "." { . }  
"<He>"  
  "he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3  
NOM  
"<apologised>"  
  "apologise" { pyytää V54-F FRONT } %+FMAINV O-ABL V PAST SG  
"<anteeksi>"  
  "anteeksi" { +anteeksi } X SG  
"<his>"  
  "he" { NOGLOSS } %A> HUM MALE PRON PERS SG3 GEN  
"<teacher>"  
  "teacher" { opettaja N10 } %OBJ HUM DEF N SG ABL @POS-INFL  
"<.>"  
  "." { . }
```

Note that in the second sentence the verb gloss was split into two separate readings. Although the word *anteeksi* does not inflect, we may need to move it after the object. It is easier to implement with rules, when it has its own reading.

In the first sentence the object is in partitive case, but in the second sentence it is in ablative case. The verb in the first sentence has the tag O-PAR to show, that the object should be in partitive case. The verb in the second sentence has the tag O-ABL, which has the effect that the rule adds the ABL tag to the object.

The sentences are translated in (4).

(4)
Hän pahoitteli erehdystään.
Hän pyysi opettajaltaan anteeksi.

Note that the word reordering rules moved the word *anteeksi* to the place after the object. If the stress would be on the word *opettajaltaan*, the word order would be opposite.

3 Multi-word verbs with inflecting nominals as second member

A small group (six in all) are such cases, where the second member inflects in a certain case, having singular and plural forms. Two analysed sentences in (5) illustrate this.

```
(5)  
"<He>"  
  "he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM  
SG3  
  "he" { hänen } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3  
  "he" { NOGLOSS } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3  
  "he" { itse N8 FRONT } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3  
  "he" { PROP-CAND } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3  
"<treasured>"  
  "treasure" { vaalia V61 O-PAR } %+FMAINV V PAST  
  "treasure" { pitää V53-F FRONT O-PAR || arvokas N41-A ESS }  
%+FMAINV V PAST  
"<their>"  
  "they" { ne Np12 FRONT } %A> PRON PERS GEN PL3  
  "they" { he Np10 FRONT OUT HUM } %A> PRON PERS GEN PL3
```

```

    "they" { NOGLOSS } %A> PRON PERS GEN PL3
    "they" { itse N8 FRONT } %A> PRON PERS GEN PL3
    "they" { niiden } %A> PRON PERS GEN PL3
    "they" { heidän HUM } %A> PRON PERS GEN PL3
"<tree>"
    "tree" { puu N18 } %OBJ DEF N SG
"<.>"
    "." { . }
"<<s>>"
    "<s>" { <s> }
"<He>"
    "he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM
SG3
    "he" { hänen } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
    "he" { NOGLOSS } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
    "he" { itse N8 FRONT } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
    "he" { PROP-CAND } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"<treasures>"
    "treasure" { vaalia V61 O-PAR } %+FMAINV V PRES SG3
    "treasure" { pitää V53-F FRONT O-PAR || arvokas N41-A ESS }
%+FMAINV V PRES SG3
"<their>"
    "they" { ne Np12 FRONT } %A> PRON PERS GEN PL3
    "they" { he Np10 FRONT OUT HUM } %A> PRON PERS GEN PL3
    "they" { NOGLOSS } %A> PRON PERS GEN PL3
    "they" { itse N8 FRONT } %A> PRON PERS GEN PL3
    "they" { niiden } %A> PRON PERS GEN PL3
    "they" { heidän HUM } %A> PRON PERS GEN PL3
"<trees>"
    "tree" { puu N18 } %OBJ DEF N PL
"<.>"
    "." { . }

```

The verb *treasure* has two readings, and both would fit in the given context. However, for the sake of demonstration, we choose the latter one. Note that the gloss *arvokas* is in base form, and it has its inflection code N41-A, according to which it inflects. The code ESS for essive was added as part of the gloss.

The sentences are semantically disambiguated and further inflection tags are added (6).

(6)

```

"<He>"
    "he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3
NOM
"<treasured>"
    "treasure" { pitää V53-F FRONT } %+FMAINV O-PAR V PAST SG
"<arvokas>"
    "arvokas" { +arvokas N41-A } X SG ESS
"<their>"
    "they" { he Np10 FRONT } %A> OUT HUM PRON PERS PL3 GEN
"<tree>"
    "tree" { puu N18 } %OBJ DEF N SG PAR @POS-INFL
"<.>"
    "." { . }
"<<s>>"

```

```
"<s>" { <s> }
"<He>"
  "he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3
NOM
"<treasures>"
  "treasure" { pitää V53-F FRONT } %+FMAINV O-PAR V PRES SG
"<arvokas>"
  "arvokas" { +arvokas N41-A } X PL ESS
"<their>"
  "they" { he Np10 FRONT } %A> OUT HUM PRON PERS PL3 GEN
"<trees>"
  "tree" { puu N18 } %OBJ DEF N PL PAR @POS-INFL
"<.>"
  "." { . }
```

Note that in the first sentence the object is in singular. Therefore, also the second member of the verb +*arvokas* is in singular. Correspondingly, in the second sentence the object is in plural, and also the second member of the verb is in plural.

Now the correct translation can be produced (7).

(7)

Hän piti heidän puutaan arvokkaana.

Hän pitää heidän puutaan arvokkaina.

Also here, the second member of the verb was moved to its default place, that is, after the object.

4 Multi-word verbs with adjectives in translative case as second member

There are 19 such verbs, which have the adjective in translative case as second member. Although the adjective has only one case, translative, it inflects in singular and plural. Therefore, it must be described so that inflection becomes possible. In (8) we have two almost identical analysed sentences. Only the number of the object is different.

(8)

```
"<solution>"
  "solution" { ratkaisu N2 } %SUBJ DEF N NOM SG
  "solution" { liuos N39 } %SUBJ DEF N NOM SG
"<enabled>"
  "enable" { mahdollistaa V53 TRV } %+FMAINV V PAST
  "enable" { tehdä V71 FRONT TRV || mahdollinen N38 TRA } %+FMAINV V
PAST
  "enable" { tehdä V71 FRONT TRV || kykenevä N10 FRONT TRA }
%+FMAINV V PAST
"<task>"
  "task" { tehtävä N10 FRONT } %OBJ DEF N SG
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<solution>"
  "solution" { ratkaisu N2 } %SUBJ DEF N NOM SG
```

```
"solution" { liuos N39 } %SUBJ DEF N NOM SG
"<enabled>"
  "enable" { mahdollistaa V53 TRV } %+FMAINV V PAST
  "enable" { tehdä V71 FRONT TRV || mahdollinen N38 TRA } %+FMAINV V
PAST
  "enable" { tehdä V71 FRONT TRV || kykenevä N10 FRONT TRA }
%+FMAINV V PAST
"<tasks>"
  "task" { tehtävä N10 FRONT } %OBJ DEF N PL
"<.>"
  "." { . }
```

The verb *enable* has two multi-word glosses, out of which the first one suits to this context. Also the single-word gloss would suit. For the sake of testing, we take the first multi-word gloss (9).

(9)

```
"<solution>"
  "solution" { ratkaisu N2 } %SUBJ DEF N NOM SG
"<enabled>"
  "enable" { tehdä V71 FRONT } %+FMAINV TRV V PAST
"<mahdollinen>"
  "mahdollinen" { +mahdollinen N38 } X TRA
"<task>"
  "task" { tehtävä N10 FRONT } %OBJ DEF N SG
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<solution>"
  "solution" { ratkaisu N2 } %SUBJ DEF N NOM SG
"<enabled>"
  "enable" { tehdä V71 FRONT } %+FMAINV TRV V PAST
"<mahdollinen>"
  "mahdollinen" { +mahdollinen N38 } X TRA
"<tasks>"
  "task" { tehtävä N10 FRONT } %OBJ DEF N PL
"<.>"
  "." { . }
```

The second member of the verb was converted into a separate reading for making its handling easier. Note that it still does not have a tag for singular/plural definition, although it has the case code TRA. The object has the singular/plural tag, but it does not have the case tag.

The missing tags are added using context-sensitive rules (10).

(10)

```
"<solution>"
  "solution" { ratkaisu N2 } %SUBJ DEF N SG NOM
"<enabled>"
  "enable" { tehdä V71 FRONT } %+FMAINV TRV V PAST SG
"<mahdollinen>"
  "mahdollinen" { +mahdollinen N38 } X SG TRA
"<task>"
```

```
"task" { tehtävä N10 FRONT } %OBJ DEF N SG ACC
"<.>"
"." { . }
"<<s>>"
"<s>" { <s> }
"<solution>"
"solution" { ratkaisu N2 } %SUBJ DEF N SG NOM
"<enabled>"
"enable" { tehdä V71 FRONT } %+FMAINV TRV V PAST SG
"<mahdollinen>"
"mahdollinen" { +mahdollinen N38 } X PL TRA
"<tasks>"
"task" { tehtävä N10 FRONT } %OBJ DEF N PL ACC
"<.>"
"." { . }
```

The processing can now proceed, and we get the final translation (11).

(11)
Ratkaisu teki tehtävän mahdolliseksi.
Ratkaisu teki tehtävät mahdollisiksi.

We see that the number of the second member of the verb depends on the number of the object. Also here, the second member was moved to its default place after the object.

5 Multi-word verbs with nominals in object case as second member

There are 48 such multi-word verbs, where the second member is a noun or adjective, and it inflects in object cases as well as in singular and plural. In (12) are two examples, one with noun and the other with adjective as second member.

```
(12)
"<drought>"
"drought" { kuivuus N40 } %SUBJ DEF N NOM SG
"drought" { kuivuuskausi N27 } %SUBJ DEF N NOM SG
"<dwarfs>"
"dwarf" { estää V53 FRONT O-ABL || kasvu N1 ACC } %+FMAINV V PRES
SG3
"dwarf" { kutistua V52 } %+FMAINV V PRES SG3
"<corn>"
"corn" { maissi N5 } %OBJ DEF N SG
"<.>"
"." { . }
"<<s>>"
"<s>" { <s> }
"<girls>"
"girl" { tyttö N1-C FRONT } HUM %SUBJ DEF N PL
"<excelled>"
"excel" { kunnostautua V52-F } %+FMAINV V PAST
"excel" { loistaa V53 } %+FMAINV V PAST
"excel" { olla V67b || erinomainen N38 } %+FMAINV V PAST
"<in>"
"in" { M-LOC1 } %ADVL PREP
```



```
"in" { M-LOC2 } %ADVL PREP
"in" { M-LOC3 } %ADVL PREP
"in" { NOGLOSS M-INE } %ADVL PREP
"in" { NOGLOSS M-ADE } %ADVL PREP
"in" { NOGLOSS M-ESS } %ADVL PREP
"in" { NOGLOSS M-TRA } %ADVL PREP
"in" { NOGLOSS M-ILL } %ADVL PREP
"in" { NOGLOSS M-ELA } %ADVL PREP
"in" { NOGLOSS M-ALL } %ADVL PREP
"in" { vuonna } %ADVL PREP
"in" { vuosina } %ADVL PREP
"<running>"
"running" { juokseminen N38 } %<P N NOM N-ING SG
"<.>"
"." { . }
```

The verb *dwarf* has two glosses, and only one suits with an object. The verb *excel* has three glosses, and each of them would suit to this context. We choose the last one for demonstrating its behaviour.

The disambiguated result is in (13).

```
(13)
"<drought>"
"drought" { kuivuus N40 } %SUBJ DEF N NOM SG
"<dwarfs>"
"dwarf" { estää V53 FRONT } %+FMAINV O-ABL V PRES SG3
"<kasvu>"
"kasvu" { +kasvu N1 } X ACC
"<corn>"
"corn" { maissi N5 } %OBJ DEF N SG
"<.>"
"." { . }
"<<s>>"
"<s>" { <s> }
"<girls>"
"girl" { tyttö N1-C FRONT } %SUBJ HUM DEF N PL
"<excelled>"
"excel" { olla V67b } %+FMAINV V PAST
"<erinomainen>"
"erinomainen" { +erinomainen N38 } X
"<in>"
"in" { M-LOC1 } %ADVL PREP
"<running>"
"running" { juokseminen N38 } %<P N NOM N-ING SG
"<.>"
"." { . }
```

We add missing inflection tags (14).

```
(14)
"<drought>"
"drought" { kuivuus N40 } %SUBJ DEF N SG NOM
"<dwarfs>"
"dwarf" { estää V53 FRONT } %+FMAINV O-ABL V PRES SG
"<kasvu>"
```

```
"kasvu" { +kasvu N1 } X SG ACC
"<corn>"
  "corn" { maissi N5 } %OBJ DEF N SG ABL
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<girls>"
  "girl" { tyttö N1-C FRONT } %SUBJ HUM DEF N PL NOM
"<excelled>"
  "excel" { olla V67b } %+FMAINV V PAST PL
"<erinomainen>"
  "erinomainen" { +erinomainen N38 } X PL NOM
"<in>"
  "in" { M-LOC1 } %ADVL PREP
"<running>"
  "running" { juokseminen N38 } %<P N N-ING SG INE
"<.>"
  "." { . }
```

Now the tags help in converting the text into correct translation (15).

(15)

*Kuivuus estää maissilta kasvun.
Tytöt olivat erinomaiset juoksemisessa.*

The second member of the verb was moved to the place after the object.

6 Multi-word verbs with nominals in object case as second member

There are five such multi-word verbs, where the second member is a nominal in essive case. Although the case is fixed, it inflects according to singular and plural, and must therefore be described accordingly. The verbs in this category are intransitive. Therefore, the object case is not relevant.

Two sentences in (16) illustrate the process.

(16)

```
"<He>"
  "he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM
SG3
  "he" { hänen } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
  "he" { NOGLOSS } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
  "he" { itse N8 FRONT } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
  "he" { PROP-CAND } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"<volunteered>"
  "volunteer" { toimia V61 || vapaaehtoinen N38 ESS } %+FMAINV V
PAST
"<in>"
  "in" { M-LOC1 } %ADVL PREP
  "in" { M-LOC2 } %ADVL PREP
  "in" { M-LOC3 } %ADVL PREP
  "in" { NOGLOSS M-INE } %ADVL PREP
  "in" { NOGLOSS M-ADE } %ADVL PREP
```

```
"in" { NOGLOSS M-ESS } %ADVL PREP
"in" { NOGLOSS M-TRA } %ADVL PREP
"in" { NOGLOSS M-ILL } %ADVL PREP
"in" { NOGLOSS M-ELA } %ADVL PREP
"in" { NOGLOSS M-ALL } %ADVL PREP
"in" { vuonna } %ADVL PREP
"in" { vuosina } %ADVL PREP
"<group>"
  "group" { ryhmä N10 FRONT } %<P DEF N NOM
  "group" { ryhmittymä N10 FRONT } %<P DEF N NOM
  "group" { ryhmä-- COMP } %<P DEF N NOM
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<They>"
  "they" { ne Np12 FRONT } %SUBJ PRON NOM PL3 CAPINIT
  "they" { he Np10 FRONT OUT HUM } %SUBJ PRON NOM PL3 CAPINIT
  "they" { NOGLOSS } %SUBJ PRON NOM PL3 CAPINIT
  "they" { itse N8 FRONT } %SUBJ PRON NOM PL3 CAPINIT
  "they" { niiden } %SUBJ PRON NOM PL3 CAPINIT
  "they" { heidän HUM } %SUBJ PRON NOM PL3 CAPINIT
"<volunteered>"
  "volunteer" { toimia V61 || vapaaehtoinen N38 ESS } %+FMAINV V
PAST
"<in>"
  "in" { M-LOC1 } %ADVL PREP
  "in" { M-LOC2 } %ADVL PREP
  "in" { M-LOC3 } %ADVL PREP
  "in" { NOGLOSS M-INE } %ADVL PREP
  "in" { NOGLOSS M-ADE } %ADVL PREP
  "in" { NOGLOSS M-ESS } %ADVL PREP
  "in" { NOGLOSS M-TRA } %ADVL PREP
  "in" { NOGLOSS M-ILL } %ADVL PREP
  "in" { NOGLOSS M-ELA } %ADVL PREP
  "in" { NOGLOSS M-ALL } %ADVL PREP
  "in" { vuonna } %ADVL PREP
  "in" { vuosina } %ADVL PREP
"<group>"
  "group" { ryhmä N10 FRONT } %<P DEF N NOM
  "group" { ryhmittymä N10 FRONT } %<P DEF N NOM
  "group" { ryhmä-- COMP } %<P DEF N NOM
"<.>"
  "." { . }
```

The verb *volunteer* has only one gloss and does not need disambiguation. When other words are disambiguated and the verb structure is modified, we get a reading as in (17).

(17)

```
"<He>"
  "he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS NOM
SG3
"<volunteered>"
  "volunteer" { toimia V61 } %+FMAINV V PAST
"<vapaaehtoinen>"
  "vapaaehtoinen" { +vapaaehtoinen N38 } X ESS
```

```
"<in>"
  "in" { M-LOC1 } %ADVL PREP
"<group>"
  "group" { ryhmä N10 FRONT } %<P DEF N NOM
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<They>"
  "they" { he Np12 FRONT } %SUBJ PRON NOM PL3 CAPINIT
"<volunteered>"
  "volunteer" { toimia V61 } %+FMAINV V PAST
"<vapaaehtoinen>"
  "vapaaehtoinen" { +vapaaehtoinen N38 } X ESS
"<in>"
  "in" { M-LOC1 } %ADVL PREP
"<group>"
  "group" { ryhmä N10 FRONT } %<P DEF N NOM
"<.>"
  "." { . }
```

Some inflection codes need to be added (18).

(18)

```
"<He>"
  "he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3
NOM
"<volunteered>"
  "volunteer" { toimia V61 } %+FMAINV V PAST SG
"<vapaaehtoinen>"
  "vapaaehtoinen" { +vapaaehtoinen N38 } X SG ESS
"<in>"
  "in" { M-LOC1 } %ADVL PREP
"<group>"
  "group" { ryhmä N10 FRONT } %<P DEF N SG INE
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<They>"
  "they" { he Np12 FRONT } %SUBJ PRON PL3 CAPINIT NOM
"<volunteered>"
  "volunteer" { toimia V61 } %+FMAINV V PAST PL
"<vapaaehtoinen>"
  "vapaaehtoinen" { +vapaaehtoinen N38 } X PL ESS
"<in>"
  "in" { M-LOC1 } %ADVL PREP
"<group>"
  "group" { ryhmä N10 FRONT } %<P DEF N SG INE
"<.>"
  "." { . }
```

Below (19) are the sentences in the intermediate phase, where the word formation process can be seen. In all earlier examples this phase was bypassed.

(19)

```
"<He>"
    "he" { h:än :Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3
NOM
"<volunteered>"
    "volunteer" { toim:ia+i :V61 } %+FMAINV V PAST SG
"<vapaaehtoinen>"
    "vapaaehtoinen" { +vapaaehtoi:nen+sena :N38 } X SG ESS
"<in>"
    "in" { M-LOC1 } %ADVL PREP
"<group>"
    "group" { ryhm:ä+assa :N10 FRONT } %<P DEF N SG INE
"<.>"
    "." { . }
"<<s>>"
    "<s>" { <s> }
"<They>"
    "they" { h:e :Np12 FRONT } %SUBJ PRON PL3 CAPINIT NOM
"<volunteered>"
    "volunteer" { toim:ia+ivat :V61 } %+FMAINV V PAST PL
"<vapaaehtoinen>"
    "vapaaehtoinen" { +vapaaehtoi:nen+sina :N38 } X PL ESS
"<in>"
    "in" { M-LOC1 } %ADVL PREP
"<group>"
    "group" { ryhm:ä+assa :N10 FRONT } %<P DEF N SG INE
"<.>"
    "." { . }
```

The surface form is further pruned (20).

(20)

```
"<He>"
    "he" { hän } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3 NOM
"<volunteered>"
    "volunteer" { toimi } %+FMAINV V PAST SG
"<vapaaehtoinen>"
    "vapaaehtoinen" { +vapaaehtoisena } X SG ESS
"<in>"
    "in" { M-LOC1 } %ADVL PREP
"<group>"
    "group" { ryhmässä } %<P DEF N SG INE
"<.>"
    "." { . }
"<<s>>"
    "<s>" { <s> }
"<They>"
    "they" { he } %SUBJ PRON PL3 CAPINIT NOM
"<volunteered>"
    "volunteer" { toimivat } %+FMAINV V PAST PL
"<vapaaehtoinen>"
    "vapaaehtoinen" { +vapaaehtoisina } X PL ESS
"<in>"
    "in" { M-LOC1 } %ADVL PREP
"<group>"
    "group" { ryhmässä } %<P DEF N SG INE
```

"<.>"
"." { . }

Now the final translation can be produced (21).

(21)
Hän toimi vapaaehtoisena ryhmässä.
He toimivat vapaaehtoisina ryhmässä.

Note that when the sentence has no object, the second member of the verbs was not moved to the right.

7 Multi-word verbs in passive constructions

We saw above that if the verb has a multi-word gloss, its inflection structure is often different than in single-word structures. The same applies also to passive constructions. A further factor that has effect on the case of the object is whether the sentence is a true passive sentence, or whether it is a passive form of a sentence, which has a subject, that is, a sentence formed using the agent particle *of*. Examples of these sentences are in (22).

(22)
The task was formalised.
He was impressed.
He was impressed by their wisdom.

After adding the Finnish glosses, the result is as in (23)

(23)
"<He>"
"he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM
SG3
"he" { hänen } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { NOGLOSS } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { itse N8 FRONT } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"he" { PROP-CAND } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"<was>"
"be" { olla V67b BE TRV-N V-4INF-TRA } O-LOC1 %+FAUXV V PAST SG
"be" { olla V67b V-3INF-ILL } O-LOC1 %+FAUXV V PAST SG
"be" { olla V67b V-3INF-INE } O-LOC1 %+FAUXV V PAST SG
"be" { olla V67b BE O-PAR } O-LOC1 %+FAUXV V PAST SG
"be" { eivät ole O-PAR V-4INF-TRA :2 } O-LOC1 %+FAUXV V PAST SG
"be" { eivät olleet O-PAR V-4INF-TRA :3 } O-LOC1 %+FAUXV V PAST SG
"be" { emme :6 } O-LOC1 %+FAUXV V PAST SG
"be" { emme ole V-3INF-INE } O-LOC1 %+FAUXV V PAST SG
"be" { emme olleet V-3INF-INE } O-LOC1 %+FAUXV V PAST SG
"be" { ei ollut V-4INF-TRA } O-LOC1 %+FAUXV V PAST SG
"be" { ei ole O-PAR V-4INF-TRA } O-LOC1 %+FAUXV V PAST SG
"be" { NOGLOSS } O-LOC1 %+FAUXV V PAST SG
"be" { joka Np13 } O-LOC1 %+FAUXV V PAST SG
"be" { jotka Np14 } O-LOC1 %+FAUXV V PAST SG
"be" { tulla V67 V-3INF-ILL } O-LOC1 %+FAUXV V PAST SG

```
"<impressed>"
  "impress" { vaikuttaa V53-C O-ILL } %-FMAINV V EN
  "impress" { tehdä V71 FRONT O-ILL || vaikutus N39 ACC } %-FMAINV V
EN
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<He>"
  "he" { hän Np9 FRONT OUT } HUM MALE %SUBJ CAPINIT PRON PERS NOM
SG3
  "he" { hänen } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
  "he" { NOGLOSS } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
  "he" { itse N8 FRONT } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
  "he" { PROP-CAND } HUM MALE %SUBJ CAPINIT PRON PERS NOM SG3
"<was>"
  "be" { olla V67b BE TRV-N V-4INF-TRA } O-LOC1 %+FAUXV V PAST SG
  "be" { olla V67b V-3INF-ILL } O-LOC1 %+FAUXV V PAST SG
  "be" { olla V67b V-3INF-INE } O-LOC1 %+FAUXV V PAST SG
  "be" { olla V67b BE O-PAR } O-LOC1 %+FAUXV V PAST SG
  "be" { eivät ole O-PAR V-4INF-TRA :2 } O-LOC1 %+FAUXV V PAST SG
  "be" { eivät olleet O-PAR V-4INF-TRA :3 } O-LOC1 %+FAUXV V PAST SG
  "be" { emme :6 } O-LOC1 %+FAUXV V PAST SG
  "be" { emme ole V-3INF-INE } O-LOC1 %+FAUXV V PAST SG
  "be" { emme olleet V-3INF-INE } O-LOC1 %+FAUXV V PAST SG
  "be" { ei ollut V-4INF-TRA } O-LOC1 %+FAUXV V PAST SG
  "be" { ei ole O-PAR V-4INF-TRA } O-LOC1 %+FAUXV V PAST SG
  "be" { NOGLOSS } O-LOC1 %+FAUXV V PAST SG
  "be" { joka Np13 } O-LOC1 %+FAUXV V PAST SG
  "be" { jotka Np14 } O-LOC1 %+FAUXV V PAST SG
  "be" { tulla V67 V-3INF-ILL } O-LOC1 %+FAUXV V PAST SG
"<impressed>"
  "impress" { vaikuttaa V53-C O-ILL } %-FMAINV V EN-AG
  "impress" { tehdä V71 FRONT O-ILL || vaikutus N39 ACC } %-FMAINV V
EN-AG
"<by>"
  "by" { kanssa M-GEN POST } %ADVL AG-PART
  "by" { mennessä M-ILL POST :1 } %ADVL AG-PART
  "by" { NOGLOSS M-ADE } %ADVL AG-PART
  "by" { NOGLOSS AG-PART } %ADVL AG-PART
  "by" { NOGLOSS M-INS } %ADVL AG-PART
  "by" { avulla M-GEN POST } %ADVL AG-PART
  "by" { vuoksi M-GEN POST } %ADVL AG-PART
  "by" { taholta M-GEN POST } %ADVL AG-PART
"<their>"
  "they" { ne Np12 FRONT } %A> PRON PERS GEN PL3
  "they" { he Np10 FRONT OUT HUM } %A> PRON PERS GEN PL3
  "they" { NOGLOSS } %A> PRON PERS GEN PL3
  "they" { itse N8 FRONT } %A> PRON PERS GEN PL3
  "they" { niiden } %A> PRON PERS GEN PL3
  "they" { heidän HUM } %A> PRON PERS GEN PL3
"<wisdom>"
  "wisdom" { viisaus N40 } %<P DEF N NOM SG
"<.>"
  "." { . }
```

The disambiguated sentences are in (24). Also, the second member of the multi-word verb is converted into a separate reading.

(24)

```
"<task>"
  "task" { tehtävä N10 FRONT } %SUBJ DEF N NOM SG
"<was>"
  "be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<formalised>"
  "formalise" { antaa V56-J } %-FMAINV O-ALL TRV V EN
"<varma>"
  "varma" { +varma N9 muoto } X N1-F ACC
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<He>"
  "he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS NOM
SG3
"<was>"
  "be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<impressed>"
  "impress" { tehdä V71 FRONT } %-FMAINV O-ILL V EN
"<vaikutus>"
  "vaikutus" { +vaikutus N39 } X ACC
"<.>"
  "." { . }
"<<s>>"
  "<s>" { <s> }
"<He>"
  "he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS NOM
SG3
"<was>"
  "be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<impressed>"
  "impress" { tehdä V71 FRONT } %-FMAINV O-ILL V EN-AG
"<vaikutus>"
  "vaikutus" { +vaikutus N39 } X ACC
"<by>"
  "by" { NOGLOSS M-ADE } %ADVL AG-PART
"<their>"
  "they" { he Np10 FRONT } %A> OUT HUM PRON PERS GEN PL3
"<wisdom>"
  "wisdom" { viisaus N40 } %<P DEF N NOM SG
"<.>"
  "." { . }
```

Note that in each sentence, the second member of the multi-word verb is in accusative by default. However, the final case depends on the sentence structure. We add further inflection tags (25).

(25)

```
"<task>"
  "task" { tehtävä N10 FRONT } %SUBJ DEF N SG ALL
"<was>"
```



```
"be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<formalised>"
"formalise" { antaa V56-J } %-FMAINV O-ALL TRV V EN SG PASS-PAST
"<varma>"
"varma" { +varma N9 muoto } X N1-F SG ACC-N
"<.>"
"." { . }
"<<s>>"
"<s>" { <s> }
"<He>"
"he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3
ILL
"<was>"
"be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<impressed>"
"impress" { tehdä V71 FRONT } %-FMAINV O-ILL V EN SG PASS-PAST
"<vaikutus>"
"vaikutus" { +vaikutus N39 } X SG ACC-N
"<.>"
"." { . }
"<<s>>"
"<s>" { <s> }
"<He>"
"he" { hän Np9 FRONT } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3
ILL
"<was>"
"be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<impressed>"
"impress" { tehdä V71 FRONT } %-FMAINV O-ILL V EN-AG SG PAST
"<vaikutus>"
"vaikutus" { +vaikutus N39 } X SG ACC
"<by>"
"by" { NOGLOSS M-ADE } %ADVL AG-PART
"<their>"
"they" { he Np10 FRONT } %A> OUT HUM PRON PERS PL3 GEN
"<wisdom>"
"wisdom" { viisaus N40 } %<P DEF N SG NOM @POS-SGPL
"<.>"
"." { . }
```

In the two first sentences, the plain accusative case (ACC), which most often takes the genitive form, was changed into the nominative accusative case (ACC-N). This is, because the object in passive structures gets this case. In the third sentence, the object remains in the basic accusative case (ACC), because the sentence has an agent, and the sentence is not a true passive sentence.

Note also, that the subject case is not the normal nominative. The subject in the passive sentence gets its case definition according to the object inflection tag of the main verb. The formal subject of the passive sentence is in fact object and follows the rules of object in inflection.

The translation phase in (26) shows, how the Finnish words are formed from morphemes.

(26)

```
"<task>"
    "task" { tehtäv+älle } %SUBJ DEF N SG ALL
"<was>"
    "be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<formalised>"
    "formalise" { ann%+ettiin } %-FMAINV O-ALL TRV V SG PASS-PAST
"<varma>"
    "varma" { +varm+a muoto } X SG ACC-N +o
"<.>"
    "." { . }
"<<s>>"
    "<s>" { <s> }
"<He>"
    "he" { h+äneen } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3 ILL
"<was>"
    "be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<impressed>"
    "impress" { te+hitiin } %-FMAINV O-ILL V SG PASS-PAST
"<vaikutus>"
    "vaikutus" { +vaikutu+s } X SG ACC-N
"<.>"
    "." { . }
"<<s>>"
    "<s>" { <s> }
"<He>"
    "he" { h+äneen } %SUBJ OUT HUM MALE CAPINIT PRON PERS SG3 ILL
"<was>"
    "be" { NOGLOSS } %+FAUXV O-LOC1 V PAST SG
"<impressed>"
    "impress" { te+ki } %-FMAINV O-ILL V -AG SG PAST
"<vaikutus>"
    "vaikutus" { +vaikutu+ksen } X SG ACC
"<by>"
    "by" { NOGLOSS M-ADE } %ADV L AG-PART
"<their>"
    "they" { h+eidän } %A> OUT HUM PRON PERS PL3 GEN
"<wisdom>"
    "wisdom" { viisau:te=nsa } %<P DEF N SG NOM
"<.>"
    "." { . }
```

Now finally we can produce the translation (27).

(27)

*Tehtävälle annettiin varma muoto.
Häneen tehtiin vaikutus.
Häneen teki vaikutuksen heidän viisautensa.*

8 Conclusion

In this report, we have looked into types of multi-word verbs in English to Finnish machine translation. Only such verbs were included, where a single-word English verb maps to a multi-word Finnish gloss. Although it has been possible to discuss the problem

only on the basis of a few examples, much more testing has already been done. Tests show that it is possible to get correct translations of all types of structures, although several problems related to cases must be solved simultaneously. The solution proposed here works, but no claim is done that this is the only possible solution.