



UNIVERSITY OF HELSINKI



<https://helda.helsinki.fi>

Helda

Direct and indirect questions in English to Finnish machine translation

Hurskainen, Arvi

2018

Hurskainen, A 2018, Direct and indirect questions in English to Finnish machine translation. in SALAMA - Swahili Language Manager : Technical reports on LT. Technical reports on language technology, no. 33, SALAMA - Swahili Language Manager, Helsinki.

<http://hdl.handle.net/10138/311044>

unspecified
publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Direct and indirect questions in English to Finnish machine translation

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

Abstract

English uses two methods in forming questions. In one method, special question words, such as when, what, where, which, why and who are used to indicate a question. In another method, the order of words indicates whether the sentence is a question. Also, Finnish uses question words for the first type of questions. The second type is more problematic, because, in addition to word order, also a special clitic *ko/kö* is added to the finite verb. In indirect questions Finnish behaves in the same way as in direct questions. English, on the other hand, uses the particles *if* and *whether* to indicate an indirect question in cases where special question words do not apply. This report describes how these structures are implemented in the English to Finnish MT system.

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

1 Introduction

In machine translation, it is important to identify direct and indirect questions, because they behave differently in the translation process. In the case of direct questions, the identification is easy, at least in principle. The question mark ? at the end of the sentence marks the question. This is not always a safe marker, however, because quite often the writer either omits the marker, or in other cases puts a question mark in the end of the sentence, which is not a question. In this report we omit these writer-induced inconsistencies and assume, that the writer writes correct English.

Another method of marking questions is the use of the so called question words, such as *when, what, where, which, why* and *who*. Questions of this type are quite easy to translate, because also Finnish uses corresponding question words. We concentrate here to the other type of questions, where the word order in English defines a direct question, and a particle *if* or *whether* defines an indirect question.

Finally, there are also such questions, which do not use question words or word order for defining a question. They are responses to earlier expressions and are very short, often consisting of only one word.

2 Direct questions

Some direct questions can be formed using question words. Another type of questions is to find out whether something has taken place or whether something is true or not. Since English and Finnish process these these question types differently, we discuss them separately.

2.1 Questions using question words

Let us start with questions, which use question words. Consider the example in (1).

(1)

Who says that only winners win? > Kuka sanoo, että vain voittajat voittavat?

The question word *who* is translated as *kuka*, and the translation is straightforward. If the question word is not in a subject position, it may inflect in Finnish, as in (2).

(2)

*To whom did you give it? > Kenelle sinä annoit sen?
Whom do you trust? > Keneen sinä luotat?*

The question word *who*, which refers to a human being, can be a subject or object or part of prepositional phrases. The corresponding question words referring to non-human entities are *which* and *what* (3).

(3)

*What is your opinion? > Mikä on sinun mielipiteesi?
Which is worse: foreclosure or bankruptcy? > Mikä on pahempi: etukäteen sulkeminen vai konkurssi?*

These question words may also be part of prepositional structure. They are translated using corresponding cases in Finnish (4).

(4)

*(a) To which car do you give highest credits? > Mille autolle sinä annat korkeimmat krediitit?
(b) To which car did you install a heater? > Mihin autoon sinä asensit lämmittimen?
(c) To which cars do you give highest credits? > Mille autoille sinä annat korkeimmat krediitit?
(d) To which cars did you install a heater? > Mihin autoihin sinä asensit lämmittimen?*

Some question words refer to time (when), place (where) or cause (why). These cannot be in the position of a subject, and usually they are treated as adverbs (5).

(5)

When did you come home? > **Milloin** sinä tulit kotiin?

Where is your home? > **Missä** on sinun kotisi?

Why is Obamacare so controversial? > **Miksi** Obamacare on niin ristiriitainen?

Some of the question words referring to time, place or cause can also be part of prepositional structure. The examples in (6) illuminate this.

(6)

(a) *Since when have you been here?* > **Mistä lähtien** olet ollut tässä?

(b) *Until when should I wait?* > **Mihin saakka** minun pitäisi odottaa?

(c) *From where has he come?* > **Mistä** hän on tullut?

(d) *Where will he go?* > **Minne** hän menee?

Note that in English the order of the subject and finite verb changes, if a question word is used. Finnish uses the word order of affirmative clauses.

2.2 Questions without question words

The standard method in English to form questions is to change the order of the subject and the finite verb. Examples are in (7).

(7)

(a) *Shall he go home?* > **Pitääkö** hänen mennä kotiin?

(b) *Has he come here?* > **Onko** hän tullut tänne?

(c) *Has he planned to come here?* > **Onko** hän suunnitellut tulevansa tänne?

(d) *Is the bucket full?* > **Onko** ämpäri täysi?

Note that in Finnish the word order is the same as in affirmative sentences. The clitic *ko/kö* is added to the finite verb to indicate that it is a question. If the question contains negation, the situation is different (8).

(8)

Will he not go home? > **Eikö** hän mene kotiin?

Has he not come here? > **Eikö** hän ole tullut tänne?

Is he not planning to come here? > **Eikö** hän suunnittele tulla tänne?

Is the bucket not full? > **Eikö** ämpäri ole täysi?

If the question contains negation, the clitic *ko/kö* is attached to the negation word *ei* in third person singular. If the subject is something else than third person singular, the clitic is attached to the corresponding form of the negation word (9).

(9)

(a) *Shall we go home?* > **Pitääkö** meidän mennä kotiin?

(b) *Have they come here?* > **Ovatko** he tulleet tänne?

(c) *Have we planned to come here?* > **Olemmeko** me suunnitelleet tulla tänne?

- (d) *Are the buckets full?* > *Ovatko ämpärit täydet?*
- (e) *Will you not go home?* > *Etkö sinä mene kotiin?*
- (f) *Have they not come here?* > *Eivätkö he ole tulleet tänne?*
- (g) *Are they not planning to come here?* > *Eivätkö he suunnittelevat tulla tänne?*
- (h) *Are the buckets not full?* > *Eivätkö ämpärit ole täydet?*

Also here we note that the clitic *ko/kö* is attached to the finite verb and the word order is the same as in English. In negative questions, the clitic *ko/kö* is attached to the corresponding form of the negation word, and the word order is the same as it would be in affirmative clauses.

2.3 Defective questions

There are also such questions, which have no clear marking to indicate that they are questions. Consider examples in (10).

- (10)
- (a) *To everything, right?* > *Kaikkeen, niinkö?*
- (b) *Or do they?* > *Vai mitä?*
- (c) *And what then?* > *Ja mitä sitten?*
- (d) *A feminist?* > *Feministikö?*
- (e) *But how long?* > *Mutta kuinka pitkää?*
- (f) *So what?* > *Entä sitten?*
- (g) *Sounds tricky?* > *Kuulostaako pulmalliselta?*
- (h) *Thick head?* > *Paksupääkö?*
- (i) *Trump's great feat?* > *Trumpin suuri saavutusko?*
- (j) *An unpaid debt?* > *Maksamaton velkako?*
- (k) *A suicide wish?* > *Itsemurhatoiveko?*
- (l) *Vast?* > *Suunnatonko?*

These examples were extracted from a small news corpus. Some of the examples have a question word but others have not. The question mark is the only indication that it is a question. The examples are varied in that they cannot be treated in the same way. Some are idiomatic (b, f) and must be translated as idioms. Some others have a question word (c, e), which helps in translation. In the remaining examples, one (g) has a finite verb, and the clitic *ko/kö* is attached to it. If the question has no verb, the clitic is attached to the head of the noun phrase (h, j, k). In case of a single word (d, l), no choice is needed. Example (i) is problematic, because the translation can be *Trumpin suuri saavutusko?* or *Trumpinko suuri saavutus?* The choice depends on where emphasis is. If the word *feat* is emphasized, then the first alternative is correct. In the opposite case the latter alternative should be chosen.

3 Indirect questions

English uses two particles for marking an indirect question, *whether* and *if*. The first one is less problematic, because the word *whether* does not have other meanings, except for a

few cases as part of a MWE. On the other hand, the word *if* is often used as a conjunction for initiating a conditional clause.

In a survey from a corpus of 60.000 sentences of news texts it was found that the word *if* occurred 1169 times. Among those, a total of 506 sentences were such that the word *if* initiated an indirect question. In the same corpus, the word *whether* occurred in 201 sentences. Out of these the word *whether* initiated an indirect clause in 190 sentences. As a member in the construction *whether or not* it occurred 10 times, and once in the structure *as to whether*.

It was found that the word *if* was used vastly more often than *whether*, even in the sense of initiating an indirect question.

There were also a few cases, where the words *whether* and *if* both were used in the same clause. In (11) is an example with an indirect question.

(11)

```
"<I>"
  "i" { minä Np5 } HUM OUT %SUBJ PRON PERS NOM SG1
  "i" { NOGLOSS } HUM OUT %SUBJ PRON PERS NOM SG1
  "i" { itse N8 FRONT } HUM OUT %SUBJ PRON PERS NOM SG1
"<do>"
  "do" { tehdä V71 FRONT } %+FAUXV V PRES
  "do" { tehdä V71 FRONT TRV } %+FAUXV V PRES
  "do" { NOGLOSS } %+FAUXV V PRES
  "do" { harjoittaa V53-C O-PAR } %+FAUXV V PRES
  "do" { samoin } %+FAUXV V PRES
"<n't>"
  "not" { ei } FRONT %ADVL NEG-PART
  "not" { en } FRONT %ADVL NEG-PART
  "not" { et } FRONT %ADVL NEG-PART
  "not" { emme } FRONT %ADVL NEG-PART
  "not" { ette } FRONT %ADVL NEG-PART
  "not" { eivät :2 } FRONT %ADVL NEG-PART
  "not" { NOGLOSS } FRONT %ADVL NEG-PART
"<know>"
  "know" { tietää V76-F FRONT TRV } %-FMAINV V INF
  "know" { tuntee V59-J TRV } %-FMAINV V INF
"<if>"
  "if" { jos } %CS CS
  "if" { KO } %CS CS
  "if" { PA } %CS CS
"<I>"
  "i" { minä Np5 } HUM OUT %SUBJ PRON PERS NOM SG1
  "i" { NOGLOSS } HUM OUT %SUBJ PRON PERS NOM SG1
  "i" { itse N8 FRONT } HUM OUT %SUBJ PRON PERS NOM SG1
"<regret>"
  "regret" { pahoitella V67-C O-PAR } %+FMAINV V PRES
  "regret" { valittaa V53-C O-PAR } %+FMAINV V PRES
  "regret" { katua V52-F O-PAR } %+FMAINV V PRES
"<those>"
  "that" { tuo Np3 } %SUBJ PRON DEM PL
  "that" { nuo Np4 } %SUBJ PRON DEM PL
```

```
"that" { NOGLOSS } %SUBJ PRON DEM PL
"that" { , että :2 } %SUBJ PRON DEM PL
"<roles>"
"role" { rooli N5 } IN %SUBJ N PL NOM
"role" { asema N10 } IN %SUBJ N PL NOM
"role" { tehtävä N10 FRONT } IN %SUBJ N PL NOM
"role" { rooli-- COMP } IN %SUBJ N PL NOM
"<or>"
"or" { tai } %CC CC
"or" { vai } %CC CC
"<not>"
"not" { ei } FRONT %ADVL NEG-PART
"not" { en } FRONT %ADVL NEG-PART
"not" { et } FRONT %ADVL NEG-PART
"not" { emme } FRONT %ADVL NEG-PART
"not" { ette } FRONT %ADVL NEG-PART
"not" { eivät :2 } FRONT %ADVL NEG-PART
"not" { NOGLOSS } FRONT %ADVL NEG-PART
"<,>"
", " { , }
", " { NOGLOSS }
"<but>"
"but" { mutta } %CC CC
"but" { vaan } %CC CC
"but" { kuin } %CC CC
"<result>"
"result" { tulos N39 } %SUBJ DEF N NOM SG
"result" { lopputulos N39 } %SUBJ DEF N NOM SG
"result" { seuraus N39 } %SUBJ DEF N NOM SG
"<does>"
"do" { tehdä V71 FRONT } %+FAUXV V PRES SG3
"do" { tehdä V71 FRONT TRV } %+FAUXV V PRES SG3
"do" { NOGLOSS } %+FAUXV V PRES SG3
"do" { harjoittaa V53-C O-PAR } %+FAUXV V PRES SG3
"do" { samoin } %+FAUXV V PRES SG3
"<n't>"
"not" { ei } FRONT %ADVL NEG-PART
"not" { en } FRONT %ADVL NEG-PART
"not" { et } FRONT %ADVL NEG-PART
"not" { emme } FRONT %ADVL NEG-PART
"not" { ette } FRONT %ADVL NEG-PART
"not" { eivät :2 } FRONT %ADVL NEG-PART
"not" { NOGLOSS } FRONT %ADVL NEG-PART
"<satisfy>"
"satisfy" { tyydyttää V53-C FRONT O-PAR } %-FMAINV V INF
"<neither>"
"neither" { ei } %AD-A> ADV
"neither" { eivät :2 } %AD-A> ADV
"<me>"
"i" { minä Np5 } HUM OUT %OBJ PRON PERS SG1
"i" { NOGLOSS } HUM OUT %OBJ PRON PERS SG1
```

```
"i" { itse N8 FRONT } HUM OUT %OBJ PRON PERS SG1
"<nor>"
  "nor" { eikä } %CC CC
  "nor" { eivätkä :2 } %CC CC
"<audience>"
  "audience" { kuulijakunta N10-J } HUM %OBJ DEF N SG
  "audience" { audienssi N5 } HUM %OBJ DEF N SG
  "audience" { kohdeyleisö N2 FRONT } HUM %OBJ DEF N SG
"<.>"
  "." { . }
```

We see that before semantic disambiguation, words have several interpretations each. Especially we pay attention to the word *if*. It has three interpretations, one of them KO. This is a special tag to indicate that, if constraint conditions are fulfilled, it can initiate an indirect question. The disambiguated result is in (12).

(12)

```
"<I>"
  "i" { minä Np5 } HUM OUT %SUBJ PRON PERS NOM SG1
"<do>"
  "do" { NOGLOSS } %+FAUXV V PRES
"<n't>"
  "not" { en } FRONT %ADVL NEG-PART
"<know>"
  "know" { tietää V76-F FRONT TRV } %-FMAINV V INF
"<if>"
  "if" { KO } %CS CS
"<I>"
  "i" { minä Np5 } HUM OUT %SUBJ PRON PERS NOM SG1
"<regret>"
  "regret" { pahoitella V67-C O-PAR } %+FMAINV V PRES
"<those>"
  "that" { nuo Np4 } %SUBJ PRON DEM PL
"<roles>"
  "role" { rooli N5 } IN %SUBJ N PL NOM
"<or>"
  "or" { vai } %CC CC
"<not>"
  "not" { en } FRONT %ADVL NEG-PART
"<,>"
  "," { , }
"<but>"
  "but" { mutta } %CC CC
"<result>"
  "result" { tulos N39 } %SUBJ DEF N NOM SG
"<does>"
  "do" { NOGLOSS } %+FAUXV V PRES SG3
"<n't>"
  "not" { ei } FRONT %ADVL NEG-PART
"<satisfy>"
```

```
"satisfy" { tyydyttää V53-C FRONT O-PAR } %-FMAINV V INF
"<neither>"
  "neither" { ei } %AD-A> ADV
"<me>"
  "i" { minä Np5 } HUM OUT %OBJ PRON PERS SG1
"<nor>"
  "nor" { eikä } %CC CC
"<audience>"
  "audience" { kuulijakunta N10-J } HUM %OBJ DEF N SG
"<.>"
  "." { . }
```

In the result we see that the rule system has selected the interpretation KO for the word *if*. The selection rule according to the Constraint Grammar formalism has this form (13).

```
(13)
SELECT ("if" KO) (-1 ("know") OR ("sure") OR ("clear") OR ("see")
OR ("wonder") OR ("unclear") OR ("or") OR ("ask") OR ("say") OR
("investigate") OR ("find_out")) (*2 ("or") BARRIER CLB LINK 1
("not"));
```

The rule reads: Select the interpretation KO for the word *if*, if in the first cohort to the left there is one of the words (*know, sure, clear, see, wonder, unclear, or, ask, say, investigate, or find out*), and after the second cohort to the right there is the word *or* and immediately after it the word *not*. Do not scan beyond the clause boundary mark.

The rule in (13) could also be written so that instead of listing individual words we define a set and use the set name instead in the rule. The set definition could be as in (14).

```
(14)
LIST KO-CONS = "know" "sure" "clear" "see" "wonder" "unclear" "or"
"ask" "say" "investigate" "find_out";
```

Now the rule in (13) could be written in a much shorter form (15).

```
(15)
SELECT ("if" KO) (-1 KO-KONS) (*2 ("or") BARRIER CLB LINK 1
("not"));
```

In the next phase we add the inflection tag to the appropriate word (16).

```
(16)
"<I>"
  "i" { minä Np5 } %SUBJ HUM OUT PRON PERS SG1 NOM
"<do>"
  "do" { NOGLOSS } %+FAUXV V PRES SG1
"<n't>"
  "not" { en FRONT } %ADVL NEG-PART
"<know>"
  "know" { tietää V76-F FRONT } %-FMAINV TRV V INF SG NEG-PRES
```

```
"<if>"
  "if" { KO } %CS CS
"<I>"
  "i" { minä Np5 } %SUBJ HUM OUT PRON PERS SG1 NOM
"<regret>"
  "regret" { pahoitella V67-C } %+FMAINV O-PAR V PRES SG1 KO
"<those>"
  "that" { nuo Np4 } %SUBJ PRON DEM PL PAR
"<roles>"
  "role" { rooli N5 } %SUBJ IN N PL PAR
"<or>"
  "or" { vai } %CC CC
"<not>"
  "not" { en FRONT } %ADVL NEG-PART
"<,>"
  "," { , }
"<but>"
  "but" { mutta } %CC CC
"<result>"
  "result" { tulos N39 } %SUBJ DEF N SG NOM
"<does>"
  "do" { NOGLOSS } %+FAUXV V PRES SG3
"<n't>"
  "not" { ei FRONT } %ADVL NEG-PART
"<satisfy>"
  "satisfy" { tyydyttää V53-C FRONT } %-FMAINV O-PAR V INF SG
NEG-PRES
"<neither>"
  "neither" { ei } %AD-A> ADV
"<me>"
  "i" { minä Np5 } %OBJ HUM OUT PRON PERS SG1 PAR
"<nor>"
  "nor" { eikä } %CC CC
"<audience>"
  "audience" { kuulijakunta N10-J } %OBJ HUM DEF N SG PAR
"<.>"
  "." { . }
```

We see that the tag **KO** was added to the verb *regret*. The rule for adding the tag is in (17).

```
(17)
MAP (@KO) TARGET VFIN (*-1 ("if") + (KO) BARRIER CLB) (*1 ("or")
BARRIER SNTB);
```

The rule reads: Map the tag **@KO** to the finite verb, if somewhere on the left there is the word *if* with the tag **KO**. Do not scan beyond clause boundary or verb. In addition, somewhere on the right in the same sentence there should be the word *or*.

Tis is then converted to the surface form and added to the inflected verb (18).

(18)

Minä en tiedä pahoittelenko minä noita rooleja vai en, mutta tulos ei tyydytä ei minua eikä kuulijakuntaa.

Sometimes there is more than one finite verb, to which the question clitic should be added. Consider the example in (19)

(19)

```
"<Of>"
  "of" { M-LOC2 } %ADVL CAPINIT PREP
  "of" { NOGLOSS M-GEN } %ADVL CAPINIT PREP
  "of" { NOGLOSS M-ELA } %ADVL CAPINIT PREP
  "of" { NOGLOSS M-ABL } %ADVL CAPINIT PREP
  "of" { NOGLOSS M-ACC-N } %ADVL CAPINIT PREP
  "of" { NOGLOSS M-PAR } %ADVL CAPINIT PREP
  "of" { NOGLOSS :2 } %ADVL CAPINIT PREP
  "of" { PROP-CAND } %ADVL CAPINIT PREP
"<far>"
  "far" { kaukana } %AD-A> ADV
  "far" { kauemmin ADV-CMP } %AD-A> ADV
  "far" { paljon } %AD-A> ADV
  "far" { kauaksi } %AD-A> ADV
  "far" { kauempana } %AD-A> ADV
  "far" { kauemmaksi } %AD-A> ADV
"<greater>"
  "great" { suuri N26 } %A> A CMP
  "great" { suurempi N16-H } %A> A CMP
  "great" { suurin N51 } %A> A CMP
"<importance>"
  "importance" { merkitys N39 FRONT } %<P N NOM SG
  "importance" { tärkeys N40 FRONT } %<P N NOM SG
"<is>"
  "be" { olla V67b BE TRV-N V-4INF-TRA } O-LOC1 %+FMAINV V
PRES SG3
  "be" { olla V67b V-3INF-ILL } O-LOC1 %+FMAINV V PRES SG3
  "be" { olla V67b V-3INF-INE } O-LOC1 %+FMAINV V PRES SG3
  "be" { olla V67b BE O-PAR } O-LOC1 %+FMAINV V PRES SG3
  "be" { eivät ole O-PAR V-4INF-TRA :2 } O-LOC1 %+FMAINV V
PRES SG3
  "be" { eivät olleet O-PAR V-4INF-TRA :3 } O-LOC1 %+FMAINV V
PRES SG3
  "be" { emme :6 } O-LOC1 %+FMAINV V PRES SG3
  "be" { emme ole V-3INF-INE } O-LOC1 %+FMAINV V PRES SG3
  "be" { emme olleet V-3INF-INE } O-LOC1 %+FMAINV V PRES SG3
  "be" { ei ollut V-4INF-TRA } O-LOC1 %+FMAINV V PRES SG3
  "be" { ei ole O-PAR V-4INF-TRA } O-LOC1 %+FMAINV V PRES SG3
  "be" { NOGLOSS } O-LOC1 %+FMAINV V PRES SG3
  "be" { joka Np13 } O-LOC1 %+FMAINV V PRES SG3
  "be" { jotka Np14 } O-LOC1 %+FMAINV V PRES SG3
  "be" { tulla V67 V-3INF-ILL } O-LOC1 %+FMAINV V PRES SG3
```

```
"<whether>"
  "whether" { KO } %CS CS
  "whether" { PA } %CS CS
  "whether" { tokko } %CS CS
  "whether" { josko } %CS CS
  "whether" { joko } %CS CS
"<Dacre's>"
  "dacre" { dacre } %A> CAP N GEN SG
  "dacre" { PROP-CAND } %A> CAP N GEN SG
"<Mail>"
  "mail" { posti N5 } %OBJ DEF N SG
  "mail" { postilähetys N39 FRONT } %OBJ DEF N SG
"<represents>"
  "represent" { edustaa V53 O-PAR } %+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
"<opinions>"
  "opinion" { mielipide N48-F FRONT } %OBJ DEF N PL
  "opinion" { lausunto N1-J } %OBJ DEF N PL
  "opinion" { mieli N26 FRONT } %OBJ DEF N PL
"<,>"
  "," { , }
  "," { NOGLOSS }
"<or>"
  "or" { tai } %CC CC
  "or" { vai } %CC CC
"<provokes>"
  "provoke" { provosoida V62 O-PAR } %+FMAINV V PRES SG3
"<them>"
  "they" { ne Np12 FRONT } %OBJ PRON PERS PL3
  "they" { he Np10 FRONT OUT HUM } %OBJ PRON PERS PL3
  "they" { NOGLOSS } %OBJ PRON PERS PL3
  "they" { itse N8 FRONT } %OBJ PRON PERS PL3
  "they" { niiden } %OBJ PRON PERS PL3
  "they" { heidän HUM } %OBJ PRON PERS PL3
"<.>"
  "." { . }
```

We see that this sentence has the word *whether* to indicate the beginning of the indirect question. It has several meanings, but the most frequent is the role as the marker of the indirect question clause. Therefore KO is the default interpretation, and no disambiguation rule is needed. The disambiguated version is in (20). Also clitic tags are added to appropriate verbs.

(20)
"<Of>"
 "of" { NOGLOSS M-ACC-N } %ADVL CAPINIT PREP
"<far>"
 "far" { paljon } %AD-A> ADV
"<greater>"
 "great" { suurempi N16-H } %A> A CMP SG NOM
"<importance>"
 "importance" { merkitys N39 FRONT } %<P N SG NOM
"<is>"
 "be" { olla V67b } %+FMAINV BE TRV-N V-4INF-TRA O-LOC1 V
PRES SG3
"<whether>"
 "whether" { **KO** } %CS CS
"<Dacre's>"
 "dacre" { dacre N7 } %A> CAP N SG GEN
"<Mail>"
 "mail" { posti N5 } %OBJ DEF N SG NOM
"<represents>"
 "represent" { edustaa V53 } %+FMAINV O-PAR V PRES SG3 **KO**
"<their>"
 "they" { he Np10 FRONT } %A> OUT HUM PRON PERS PL3 GEN
"<opinions>"
 "opinion" { mielipide N48-F FRONT } %OBJ DEF N PL PAR POS-
INFL
"<,>"
 "," { , }
"<or>"
 "or" { vai } %CC CC
"<provokes>"
 "provoke" { provosoida V62 } %+FMAINV O-PAR V PRES SG **KO**
"<them>"
 "they" { ne Np12 FRONT } %OBJ PRON PERS PL3 PAR
"<.>"
 "." { . }

Note that the finite verbs *represent* and *provoke* have now the added tag KO. These were added using two separate rules. The translation is in (21).

(21)
Paljon suurempi merkitys on, edustaako Dacren posti heidän mielipiteitään, vai provosoiko niitä.

Finally, there are cases, where the words *whether* and *if* both occur in the same sentence (22).

(22)
"<Investigators>"
 "investigator" { tutkija N12 } HUM %SUBJ CAPINIT N PL NOM
 "investigator" { PROP-CAND } HUM %SUBJ CAPINIT N PL NOM

```
"<said>"
  "say" { sanoa V52 TRV O-ALL SVOO } V-INF-PART HUM %+FMAINV V
PAST
  "say" { todeta V74-F TRV } V-INF-PART HUM %+FMAINV V PAST
"<it>"
  "it" { se Np11 FRONT OUT } %SUBJ PRON NOM SG3
  "it" { sen } %SUBJ PRON NOM SG3
  "it" { NOGLOSS } %SUBJ PRON NOM SG3
  "it" { itse N8 FRONT } %SUBJ PRON NOM SG3
"<was>"
  "be" { olla V67b BE TRV-N V-4INF-TRA } O-LOC1 %+FMAINV V
PAST
  "be" { olla V67b V-3INF-ILL } O-LOC1 %+FMAINV V PAST
  "be" { olla V67b V-3INF-INE } O-LOC1 %+FMAINV V PAST
  "be" { olla V67b BE O-PAR } O-LOC1 %+FMAINV V PAST
  "be" { eivät ole O-PAR V-4INF-TRA :2 } O-LOC1 %+FMAINV V
PAST
  "be" { eivät olleet O-PAR V-4INF-TRA :3 } O-LOC1 %+FMAINV V
PAST
  "be" { emme :6 } O-LOC1 %+FMAINV V PAST
  "be" { emme ole V-3INF-INE } O-LOC1 %+FMAINV V PAST
  "be" { emme olleet V-3INF-INE } O-LOC1 %+FMAINV V PAST
  "be" { ei ollut V-4INF-TRA } O-LOC1 %+FMAINV V PAST
  "be" { ei ole O-PAR V-4INF-TRA } O-LOC1 %+FMAINV V PAST
  "be" { NOGLOSS } O-LOC1 %+FMAINV V PAST
  "be" { joka Np13 } O-LOC1 %+FMAINV V PAST
  "be" { jotka Np14 } O-LOC1 %+FMAINV V PAST
  "be" { tulla V67 V-3INF-ILL } O-LOC1 %+FMAINV V PAST
"<too soon>"
  "too soon" { liian aikaisin } %ADVL ADV
  "too soon" { liian aikaista :2 } %ADVL ADV
  "too soon" { liian pian } %ADVL ADV
"<to>"
  "to" { NOGLOSS } %INFMARK> INFMARK>
"<say>"
  "say" { sanoa V52 TRV O-ALL SVOO } V-INF-PART HUM %-FMAINV V
INF
  "say" { todeta V74-F TRV } V-INF-PART HUM %-FMAINV V INF
"<whether>"
  "whether" { KO } %CS CS
  "whether" { PA } %CS CS
  "whether" { tokko } %CS CS
  "whether" { josko } %CS CS
  "whether" { joko } %CS CS
"<Wong>"
  "wong" { wong } %SUBJ CAP N NOM SG
  "wong" { PROP-CAND } %SUBJ CAP N NOM SG
"<specifically>"
  "specifically" { erityisesti STI } %ADVL ADV
  "specifically" { erityisen } %ADVL ADV
"<targeted>"
```

```
"target" { tähdätä V73-F FRONT O-PAR O-ILL } %+FMAINV V PAST
"target" { kohdentaa V54-J TRV } %+FMAINV V PAST
"target" { kohdistaa V53 TRV } %+FMAINV V PAST
"<trio>"
  "trio" { trio N1 } HUM %OBJ DEF N SG
  "trio" { kolmikko N4-A } HUM %OBJ DEF N SG
"<or>"
  "or" { tai } %CC CC
  "or" { vai } %CC CC
"<if>"
  "if" { jos } %CS CS
  "if" { KO } %CS CS
  "if" { PA } %CS CS
"<they>"
  "they" { ne Np12 FRONT } %SUBJ PRON PERS NOM PL3
  "they" { he Np10 FRONT OUT HUM } %SUBJ PRON PERS NOM PL3
  "they" { NOGLOSS } %SUBJ PRON PERS NOM PL3
  "they" { itse N8 FRONT } %SUBJ PRON PERS NOM PL3
  "they" { niiden } %SUBJ PRON PERS NOM PL3
  "they" { heidän HUM } %SUBJ PRON PERS NOM PL3
"<were>"
  "be" { olla V67b BE TRV-N V-4INF-TRA } O-LOC1 %+FAUXV V PAST
PL COND-CAND
  "be" { olla V67b V-3INF-ILL } O-LOC1 %+FAUXV V PAST PL COND-
CAND
  "be" { olla V67b V-3INF-INE } O-LOC1 %+FAUXV V PAST PL COND-
CAND
  "be" { olla V67b BE O-PAR } O-LOC1 %+FAUXV V PAST PL COND-
CAND
  "be" { eivät ole O-PAR V-4INF-TRA :2 } O-LOC1 %+FAUXV V PAST
PL COND-CAND
  "be" { eivät olleet O-PAR V-4INF-TRA :3 } O-LOC1 %+FAUXV V
PAST PL COND-CAND
  "be" { emme :6 } O-LOC1 %+FAUXV V PAST PL COND-CAND
  "be" { emme ole V-3INF-INE } O-LOC1 %+FAUXV V PAST PL COND-
CAND
  "be" { emme olleet V-3INF-INE } O-LOC1 %+FAUXV V PAST PL
COND-CAND
  "be" { ei ollut V-4INF-TRA } O-LOC1 %+FAUXV V PAST PL COND-
CAND
  "be" { ei ole O-PAR V-4INF-TRA } O-LOC1 %+FAUXV V PAST PL
COND-CAND
  "be" { NOGLOSS } O-LOC1 %+FAUXV V PAST PL COND-CAND
  "be" { joka Np13 } O-LOC1 %+FAUXV V PAST PL COND-CAND
  "be" { jotka Np14 } O-LOC1 %+FAUXV V PAST PL COND-CAND
  "be" { tulla V67 V-3INF-ILL } O-LOC1 %+FAUXV V PAST PL COND-
CAND
"<randomly>"
  "randomly" { satunnaisesti STI } %ADVL ADV
  "randomly" { satunnaisen } %ADVL ADV
"<gunned>"
```

```
"gun" { ampua V52-H TRV } %-FMAINV V EN
"<down>"
"down" { alas } %ADVL ADV
"down" { alhaalla } %ADVL ADV
"<.>"
"." { . }
```

We see that both words, *whether* and *if*, have several meanings. The default meaning of *whether* is KO, but the default meaning of *if* is *jos*, because it is more frequent than KO. The rule for selecting the appropriate meaning for the word *if* is as follows:

(23)
SELECT ("if" KO) (-1 KO-CONS) (*2 VFIN BARRIER CLB LINK NOT *1
SUBJ BARRIER CLB OR ("when") OR (CONJ) OR COMMA);

The rule reads: Select for the word *if* the reading KO, if the word immediately to the left belongs to the set KO-CONS (see (14) above). After the second cohort to the right there should be a finite verb, but do not scan beyond clause boundary. In addition, after that there should not be a subject or the word *when* or conjunction or comma. Do not scan beyond clause boundary.

The disambiguated sentence is in (24).

(24)
"<Investigators>"
"investigator" { tutkija N12 } HUM %SUBJ CAPINIT N PL NOM
"<said>"
"say" { sanoa V52 TRV O-ALL SVOO } V-INF-PART HUM %+FMAINV V
PAST
"<it>"
"it" { NOGLOSS } %SUBJ PRON NOM SG3
"<was>"
"be" { olla V67b V-3INF-INE } O-LOC1 %+FMAINV V PAST
"<too soon>"
"too soon" { liian aikaista :2 } %ADVL ADV
"<to>"
"to" { NOGLOSS } %INFMARK> INFMARK>
"<say>"
"say" { sanoa V52 TRV O-ALL SVOO } V-INF-PART HUM %-FMAINV V
INF
"<whether>"
"whether" { KO } %CS CS
"<Wong>"
"wong" { wong } %SUBJ CAP N NOM SG
"<specifically>"
"specifically" { erityisesti STI } %ADVL ADV
"<targeted>"
"target" { tähdätä V73-F FRONT O-PAR O-ILL } %+FMAINV V PAST
"<trio>"
"trio" { trio N1 } HUM %OBJ DEF N SG
"<or>"

```
"or" { vai } %CC CC
"<if>"
  "if" { KO } %CS CS
"<they>"
  "they" { ne Np12 FRONT } %SUBJ PRON PERS NOM PL3
"<were>"
  "be" { NOGLOSS } O-LOC1 %+FAUXV V PAST PL COND-CAND
"<randomly>"
  "randomly" { satunnaisesti STI } %ADVL ADV
"<gunned>"
  "gun" { ampua V52-H TRV } %-FMAINV V EN
"<down>"
  "down" { alas } %ADVL ADV
"<.>"
  "." { . }
```

The rule for selecting the correct word, to which the clitic *ko/kö* will be attached, is a bit complicated. For the verb *target* it is as in (25).

(25)

```
MAP (@KO) TARGET V (*-2 (KO) BARRIER CLB OR COMMA OR (&\() LINK
NOT 2 VFIN) (NOT 1 BE OR AG-PART) (NOT -1 ("should") OR ("could")
OR REL OR HAVE OR BE + EN OR COMMA OR ("when") OR ("no one") OR
INF OR ("get")) (NOT -2 ("should")) (NOT 0 ("should"));
```

The rule reads: Map the tag @KO to the verb, if after the second cohort to the left there is the tag KO, but do not scan beyond clause boundary or &(. Also the second word after the verb should not be a finite verb. The first word to the right should not be the verb *be* or agentive particle. The first word to the left should not be the word *should* or *could* or relativ pronoun or the verb *have* or the verb *be* in participial form or the word *when* or the MWE *no one* or infinitive or the verb *get*. The second word to the left should not be *should*. Also the target word should not be *should*.

The rule is an example of how the rule grows when new constraints are needed. It is not in ideal form and it could be made more straightforward through thorough testing.

For the verb *gun* the rule is not quite so complex (26).

(26)

```
MAP (@KO) TARGET VFIN OR EN (*-1 (KO) BARRIER CLB LINK -1 ("or"));
```

The rule reads: Map the tag @KO to the finite verb or participial if to the left there is the tag KO, but do not scan beyond clause boundary. Also next to the left from KO there should be the word *or*.

The translation is in (27).

(27)

Tutkijat sanoivat, että oli liian aikaista sanoa tähtäsikö Wong erityisesti trioa vai ammuttiinko ne satunnaisesti alas.

4 Conclusion

In this report we have discussed problems in translating direct and indirect questions from English to Finnish. We have seen that English has its unique ways in formulating various types of questions, and Finnish has almost idiomatic ways of formulating the same questions. Especially the questions without question words require special attention for converting the questions correctly into Finnish. Nevertheless, no nonsolvable problems occur.