

Translating compound words from Finnish to English¹

Arvi Hurskainen
Department of Languages
FIN-00014 University of Helsinki, Finland
arvi.hurskainen@helsinki.fi

Abstract

Finnish and English differ from each other in how they form concepts that require more than one word. Finnish uses compounding, that is, segments are written together as single strings. English typically separates the segments as separate words. Also the genitive structure is used for expressing such concepts.

In relation to translation, compounds can be grouped into two categories. Some compounds are translated using a single English word. Other compounds can be translated using the translation of the glosses of individual segments. In the latter case, the translation is a sequence of processes. In this report, I describe the translation method and discuss the problems involved in it.

Key Words: *compound words, morphological description, translation.*

1 Introduction

In Technical Reports No 75² and 77³ I described the analysis of compound words in Finnish. Because the number of compounds in Finnish is almost limitless, the listing of compounds into the lexicon is not a feasible solution. Therefore, I suggested a method, where each compound part, a segment, is listed separately into their own sub-lexicons, and the word recognition is assembled from these segments.

It was found out that this method reduces the size of the lexicon considerably, and at the same time number of compound possibilities increases dramatically.

It was also found out that it was possible to describe all compounds in this way, including those compounds that were lexicalized as single words, and which were only translatable using a single word.

However, in translation we must take into account the fact that those compounds, which are translated using a single word, must be treated as single strings, although they were identified and analysed using the split methodology.

The translation processes of various types of compounds will be described below.

¹ The report is issued under licence CC BY-NC

² <http://www.njas.helsinki.fi/salama/describing-compound-words-in-finnish.pdf>

³ <http://www.njas.helsinki.fi/salama/enhanced-method-for-describing-compound-words.pdf>

2 Segment translation method: two segments

First, we take a look at the method, where we translate each segment separately, and then modify the result as needed. We take such compounds, which have two segments. We take examples of such compounds, which have the string *ydin* as the first segment (1).

(1)
ydinsota
ydinase
ydinfysiikka
ydinpommi
ydinenergia
ydinkoe
ydintehtävä

For the sake of simplicity, only base forms are handled here. In (2) is the analysis result. In order to prune away unnecessary readings, also such disambiguation was carried out which does not need context-sensitive rules.

(2)
"<ydinsota>"
 "ydin|sota" C2 N SG NOM
 "ydin|sota" C2 N SG ACC-N
"<ydinase>"
 "ydin|ase" C2 N SG NOM
 "ydin|ase" C2 N SG ACC-N
"<ydinfysiikka>"
 "ydin|fysiikka" C2 N SG NOM
 "ydin|fysiikka" C2 N SG ACC-N
"<ydinpommi>"
 "ydin|pommi" C2 N SG NOM
 "ydin|pommi" C2 N SG ACC-N
"<ydinenergia>"
 "ydin|energia" C2 N SG NOM
 "ydin|energia" C2 N SG ACC-N
"<ydinkoe>"
 "ydin|koe" C2 N SG INE
"<ydintehtävä>"
 "ydintehtävä N SG NOM
 "ydintehtävä N SG ACC-N

Note that all compounds except for the last one, *ydintehtävä*, were constructed using the split method. The reason is that the appropriate gloss of its first segment is not the default gloss. This will be seen below.

The translation of the base forms starts with the translation of the first segment (3).

(3)
"<ydinsota>"
" { nuclear|sota } C2 N SG NOM
" { nuclear|sota } C2 N SG ACC-N
"<ydinase>"
" { nuclear|ase } C2 N SG NOM
" { nuclear|ase } C2 N SG ACC-N
"<ydinfysiikka>"
" { nuclear|fysiikka } C2 N SG NOM
" { nuclear|fysiikka } C2 N SG ACC-N
"<ydinpommi>"
" { nuclear|pommi } C2 N SG NOM
" { nuclear|pommi } C2 N SG ACC-N
"<ydinenergia>"
" { nuclear|energia } C2 N SG NOM
" { nuclear|energia } C2 N SG ACC-N
"<ydinkoe>"
" { nuclear|koe } C2 N SG NOM
" { nuclear|koe } C2 N SG ACC-N
"<ydintehtävä>"
"ydintehtävä N SG NOM
"ydintehtävä N SG ACC-N

We see that the segment *ydin* was replaced by the gloss *nuclear* in those cases, where the analysis was done using the split method. The last example was left untouched. In the next phase we translate the last segment (4).

(4)
"<ydinsota>"
" { nuclear|war } C2 N SG NOM
" { nuclear|war } C2 N SG ACC-N
"<ydinase>"
" { nuclear|veapon , gun } C2 N SG NOM
" { nuclear|veapon , gun } C2 N SG ACC-N
"<ydinfysiikka>"
" { nuclear|physics } C2 N SG NOM
" { nuclear|physics } C2 N SG ACC-N
"<ydinpommi>"
" { nuclear|bomb , bombshell } C2 N SG NOM
" { nuclear|bomb , bombshell } C2 N SG ACC-N
"<ydinenergia>"
" { nuclear|energy } C2 N SG NOM
" { nuclear|energy } C2 N SG ACC-N
"<ydinkoe>"
" { nuclear|experiment , trial } C2 N SG NOM

```
"{ nuclear|experiment , trial }" C2 N SG ACC-N
"<ydintehtävä>"
  "{ core task } N SG NOM
  "{ core task } N SG ACC-N
```

Also the last segment is now translated. There are also cases, where the segment has more than one gloss. These should be disambiguated using the context-sensitive rules. However, because we have here only a list of separate words, the disambiguation is omitted, and we take the first gloss, which is interpreted as a default.

The last example was translated separately as a single string. Note that in it, the segment *ydin* was translated as *core*.

The result is in (5).

```
(5)
"<ydinsota>"
  "{ nuclear|war }" C2 N SG NOM
  "{ nuclear|war }" C2 N SG ACC-N
"<ydinase>"
  "{ nuclear|veapon }" C2 N SG NOM
  "{ nuclear|veapon }" C2 N SG ACC-N
"<ydinfysiikka>"
  "{ nuclear|physics }" C2 N SG NOM
  "{ nuclear|physics }" C2 N SG ACC-N
"<ydinpommi>"
  "{ nuclear|bomb }" C2 N SG NOM
  "{ nuclear|bomb }" C2 N SG ACC-N
"<ydinenergia>"
  "{ nuclear|energy }" C2 N SG NOM
  "{ nuclear|energy }" C2 N SG ACC-N
"<ydinkoe>"
  "{ nuclear|experiment }" C2 N SG NOM
  "{ nuclear|experiment }" C2 N SG ACC-N
"<ydintehtävä>"
  "{ core task } N SG NOM
  "{ core task } N SG ACC-N
```

The translated segments can now be separated as separate words (6).

```
(6)
"<ydinsota>"
  "{ nuclear war }" C2 N SG NOM
  "{ nuclear war }" C2 N SG ACC-N
"<ydinase>"
  "{ nuclear veapon }" C2 N SG NOM
  "{ nuclear veapon }" C2 N SG ACC-N
"<ydinfysiikka>"
```

"{ nuclear physics }" C2 N SG NOM
"{ nuclear physics }" C2 N SG ACC-N
"<ydinpommi>"
"{ nuclear bomb }" C2 N SG NOM
"{ nuclear bomb }" C2 N SG ACC-N
"<ydinenergia>"
"{ nuclear energy }" C2 N SG NOM
"{ nuclear energy }" C2 N SG ACC-N
"<ydinkoe>"
"{ nuclear experiment }" C2 N SG NOM
"{ nuclear experiment }" C2 N SG ACC-N
"<ydintehtävä>"
"{ core task } N SG NOM
"{ core task } N SG ACC-N

3 Segment translation method: more than two segments

The analysis system was constructed in a way that allows for up to four segments in a compound. The last one of these segments is the inflecting word. The segments of the three first slots must be translated separately from the last slot, because the needed gloss depends on its position in the string. An example of this is above. The noun *ydin* as a word of its own is not translated as *nuclear*, which in fact is an adjective.

In (7) there are four words, two of them having three segments and two having four segments.

(7)
"<öljynpaineanturiin>"
"öljyn|paine|anturi" C2 C3 N SG ILL
"<öljynpaineanturijärjestelmä>"
"öljyn|paine|anturijärjestelmä C1 C2 C3 N SG NOM
"öljyn|paine|anturijärjestelmä C1 C2 C3 N SG ACC-N
"<öljyvärimaalaus>"
"öljy|väri|maalaus" C2 C3 N SG NOM
"öljy|väri|maalaus" C2 C3 N SG ACC-N
"<öljyvärimaalausalusta>"
"öljy|väri|maalaus|alusta" C1 C2 C3 N SG NOM
"öljy|väri|maalaus|alusta" C1 C2 C3 N SG ACC-N
"öljy|väri|maalaus|alunen" C1 C2 C3 N SG PAR
"öljy|väri|maalaus|alus" C1 C2 C3 N SG PAR
"öljy|väri|maalaus|alku" C1 C2 C3 N SG ELA

The last word is semantically ambiguous. However, we leave it as such for demonstrating how each variety is translated.

In the first phase, we translate the first segment (8).

(8)
"<öljyinpaineanturiin>"
" { oil|paine|anturi" C2 C3 N SG ILL
"<öljyinpaineanturijärjestelmä>"
" { oil|paine|anturijärjestelmä C1 C2 C3 N SG NOM
" { oil|paine|anturijärjestelmä C1 C2 C3 N SG ACC-N
"<öljyvärimaalaus>"
" { oil|väri|maalau" C2 C3 N SG NOM
" { oil|väri|maalau" C2 C3 N SG ACC-N
"<öljyvärimaalausalusta>"
" { oil|väri|maalau|alusta" C1 C2 C3 N SG NOM
" { oil|väri|maalau|alusta" C1 C2 C3 N SG ACC-N
" { oil|väri|maalau|alunen" C1 C2 C3 N SG PAR
" { oil|väri|maalau|alus" C1 C2 C3 N SG PAR
" { oil|väri|maalau|alku" C1 C2 C3 N SG ELA

The first segments *öljy* and *öljyn* were replaced with the gloss *oil*.
In the next phase we translate the second segments, *paine* and *väri* (9).

(9)
"<öljyinpaineanturiin>"
" { oil|pressure|anturi" C2 C3 N SG ILL
"<öljyinpaineanturijärjestelmä>"
" { oil|pressure|anturijärjestelmä C1 C2 C3 N SG NOM
" { oil|pressure|anturijärjestelmä C1 C2 C3 N SG ACC-N
"<öljyvärimaalaus>"
" { oil|color|maalau" C2 C3 N SG NOM
" { oil|color|maalau" C2 C3 N SG ACC-N
"<öljyvärimaalausalusta>"
" { oil|color|maalau|alusta" C1 C2 C3 N SG NOM
" { oil|color|maalau|alusta" C1 C2 C3 N SG ACC-N
" { oil|color|maalau|alunen" C1 C2 C3 N SG PAR
" { oil|color|maalau|alus" C1 C2 C3 N SG PAR
" { oil|color|maalau|alku" C1 C2 C3 N SG ELA

The second slot was filled with the glosses *pressure* and *color*.

The environment of the third slot is identical with the environment of the second slot - both are surrounded with the vertical bar '|'. Therefore, the third slot words can be translated using the same environment constraints as in the second slot. The result is in (10).

(10)
"<öljyinpaineanturiin>"
" { oil|pressure|anturi" C2 C3 N SG ILL
"<öljyinpaineanturijärjestelmä>"
" { oil|pressure|sensor|järjestelmä C1 C2 C3 N SG NOM

```
"{ oil|pressure|sensor|järjestelmä C1 C2 C3 N SG ACC-N
"<öljyväriemaalaus>"
  "{ oil|color|maalaus" C2 C3 N SG NOM
  "{ oil|color|maalaus" C2 C3 N SG ACC-N
"<öljyväriemaalausalusta>"
  "{ oil|color|painting|alusta" C1 C2 C3 N SG NOM
  "{ oil|color|painting|alusta" C1 C2 C3 N SG ACC-N
  "{ oil|color|painting|alunen" C1 C2 C3 N SG PAR
  "{ oil|color|painting|alus" C1 C2 C3 N SG PAR
  "{ oil|color|painting|alku" C1 C2 C3 N SG ELA
```

Note that the word *anturi* in the third slot was translated as *sensor*, but the same word in the last slot was not translated. The same is the situation with the word *maalaus*, which was translated in the third slot, but not when it was in the last slot.

Now we translate the last segment (11).

```
(11)
"<öljynpaineanturiin>"
  "{ oil|pressure|anturi { sensor }" C2 C3 N SG ILL
"<öljynpaineanturijärjestelmä>"
  "{ oil|pressure|sensor|järjestelmä { system , regime } C1 C2 C3 N SG NOM
  "{ oil|pressure|sensor|järjestelmä { system , regime } C1 C2 C3 N SG ACC-N
"<öljyväriemaalaus>"
  "{ oil|color|maalaus { painting , portrayal }" C2 C3 N SG NOM
  "{ oil|color|maalaus { painting , portrayal }" C2 C3 N SG ACC-N
"<öljyväriemaalausalusta>"
  "{ oil|color|painting|alusta { platform }" C1 C2 C3 N SG NOM
  "{ oil|color|painting|alusta { platform }" C1 C2 C3 N SG ACC-N
  "{ oil|color|painting|alunen { pad , underlay }" C1 C2 C3 N SG PAR
  "{ oil|color|painting|alus { boat , ship , vessel }" C1 C2 C3 N SG PAR
  "{ oil|color|painting|alku { onset , start }" C1 C2 C3 N SG ELA
```

The translation result of the last segment is different from the previous ones. The original Finnish segment is retained, and the English gloss is added after it and surrounded with curly braces. This is the standard way of translating words, and there is no need to change it because of compound words. We can prune the result as in (12). Also the extra glosses are removed and the default gloss retained.

```
(12)
"<öljynpaineanturiin>"
  "{ oil|pressure|anturi { sensor }" C2 C3 N SG ILL
"<öljynpaineanturijärjestelmä>"
  "{ oil|pressure|sensor|system } C1 C2 C3 N SG NOM
  "{ oil|pressure|sensor|system } C1 C2 C3 N SG ACC-N
"<öljyväriemaalaus>"
  "{ oil|color|painting }" C2 C3 N SG NOM
```

"{ oil|color|painting }" C2 C3 N SG ACC-N
"<öljyvärimaalaus>"
"{ oil|color|painting|platform }" C1 C2 C3 N SG NOM
"{ oil|color|painting|platform }" C1 C2 C3 N SG ACC-N
"{ oil|color|painting|pad }" C1 C2 C3 N SG PAR
"{ oil|color|painting|boat }" C1 C2 C3 N SG PAR
"{ oil|color|painting|onset }" C1 C2 C3 N SG ELA

The translated segments can now be separated from each other (13).

(13)
"<öljynpaineanturiin>"
"{ oil pressure anturi { sensor }" C2 C3 N SG ILL
"<öljynpaineanturijärjestelmä>"
"{ oil pressure sensor system } C1 C2 C3 N SG NOM
"{ oil pressure sensor system } C1 C2 C3 N SG ACC-N
"<öljyvärimaalaus>"
"{ oil color painting }" C2 C3 N SG NOM
"{ oil color painting }" C2 C3 N SG ACC-N
"<öljyvärimaalaus>"
"{ oil color painting platform }" C1 C2 C3 N SG NOM
"{ oil color painting platform }" C1 C2 C3 N SG ACC-N
"{ oil color painting pad }" C1 C2 C3 N SG PAR
"{ oil color painting boat }" C1 C2 C3 N SG PAR
"{ oil color painting onset }" C1 C2 C3 N SG ELA

The compound words now have translations. The three first compounds have non-ambiguous interpretations, but the last one has four semantically different alternatives. It depends on the inflection form of the compound, how many alternative readings the word has. Examples are in (14).

(14)
"<öljyvärimaalaus>"
"{ oil|color|painting|platform }" C1 C2 C3 N PL NOM
"{ oil|color|painting|platform }" C1 C2 C3 N PL ACC
"{ oil|color|painting|platform }" C1 C2 C3 N PL ACC-N
"<öljyvärimaalaus>"
"{ oil|color|painting|platform }" C1 C2 C3 N SG GEN
"{ oil|color|painting|platform }" C1 C2 C3 N SG ACC
"<öljyvärimaalaus>"
"{ oil|color|painting|platform }" C1 C2 C3 N SG ADE
"<öljyvärimaalaus>"
"{ oil|color|painting|platform }" C1 C2 C3 N PL ADE CLIT

We see that when putting the nominative form of the compound into plural, the reading is semantically non-ambiguous, and only the morphological ambiguity remains. Also other forms are semantically non-ambiguous, and the morphological ambiguity varies.

4 Problematic cases

There are several cases, where the segment-based translation is not feasible. We take some examples of compounds, which have the segment *yksityis* as its first element (15).

(15)
"<yksityiskodit>"
 "yksityis|koti" C2 N PL NOM
 "yksityis|koti" C2 N PL ACC
 "yksityis|koti" C2 N PL ACC-N
"<yksityiskohdan>"
 "yksityiskohta" N SG GEN
"<yksityiskohtaisuus>"
 "yksityis|kohtaisuus" C2 N SG NOM
 "yksityis|kohtaisuus" C2 N SG ACC-N
"<yksityiskäyttöön>"
 "yksityis|käyttö" C2 N SG ILL
"<yksityiskäytöstä>"
 "yksityis|käytös" C2 N SG PAR
 "yksityis|käyttö" C2 N SG ELA
"<yksityismetsien>"
 "yksityis|metsä" C2 N PL GEN
"<yksityisoppitunti>"
 "yksityis|oppi|tunti" C2 C3 N SG NOM
 "yksityis|oppi|tunti" C2 C3 N SG ACC-N
"<yksityistalon>"
 "yksityis|talo" C2 N SG GEN
"<yksityistalouksiin>"
 "yksityis|talous" C2 N PL ILL
"<yksityistilaisuuksiin>"
 "yksityis|tilaisuus" C2 N PL ILL
"<yksityistuntia>"
 "yksityis|tunti" C2 N SG PAR

All the compounds were analysed using the segment-based identification of compounds, except for the compound *yksityiskohdan*, which was analysed as a single string.

There are problematic cases, such as *yksityiskohtaisuus* and *yksityisoppitunti*. The former should be treated as single string, and in the latter case the segments *oppi* and *tunti* should be treated as a single string, because the English gloss is a single word.

How can we do this, because the analyser has split them into segments?

One solution is to list such compounds as single words in the lexicon. Another solution is to translate adjacent segments as a single string.

Here we use the latter solution and treat the strings *yksityis|kohtaisuus* and *oppi|tunti* as single strings, and we translate them accordingly. The result is in (16).

(16)
"<yksityiskodit>"
" { private|koti { house } }" C2 N PL NOM
" { private|koti { house } }" C2 N PL ACC
" { private|koti { house } }" C2 N PL ACC-N
"<yksityiskohdan>"
"yksityiskohta { detail }" N SG GEN
"<yksityiskohtaisuus>"
" { accuracy }" C2 N SG NOM
" { accuracy }" C2 N SG ACC-N
"<yksityiskäyttöön>"
" { private|käyttö { use , using } }" C2 N SG ILL
"<yksityiskäytöstä>"
" { private|käytös { behavior , behaviour , demeanor } }" C2 N SG PAR
" { private|käyttö { use , using } }" C2 N SG ELA
"<yksityismetsien>"
" { private|metsä { forest } }" C2 N PL GEN
"<yksityisoppitunti>"
" { private|lesson }" C2 C3 N SG NOM
" { private|lesson }" C2 C3 N SG ACC-N
"<yksityistalon>"
" { private|talo { house } }" C2 N SG GEN
"<yksityistalouksiin>"
" { private|talous { economics , economy , household } }" C2 N PL ILL
"<yksityistilaisuuksiin>"
" { private|tilaisuus { occasion , opportunity } }" C2 N PL ILL
"<yksityistuntia>"
" { private|tunti { hour } }" C2 N SG PAR

We see that the compounds *yksityis|kohtaisuus* and *oppi|tunti* have got single-word translations.

Some examples have semantically ambiguous translations. For most of them, the first gloss is suitable. For the compound *yksityistalouksiin* we should select the gloss *household*. Because this is a complicated process, we rather choose another solution and use the joint-segment translation, as we did in (15) above.

For the rest we choose the first gloss and prune the reading (17).

(17)
"<yksityiskodit>"
" { private|house }" C2 N PL NOM
" { private|house }" C2 N PL ACC
" { private|house }" C2 N PL ACC-N

"<yksityiskohdan>"
"yksityiskohta { detail }" N SG GEN
"<yksityiskohtaisuus>"
"{ accuracy }" C2 N SG NOM
"{ accuracy }" C2 N SG ACC-N
"<yksityiskäyttöön>"
"{ private|use } C2 N SG ILL
"<yksityiskäytöstä>"
"{ private|behavior }" C2 N SG PAR
"{ private|use } C2 N SG ELA
"<yksityismetsien>"
"{ private|forest } C2 N PL GEN
"<yksityisoppitunti>"
"{ private|lesson }" C2 C3 N SG NOM
"{ private|lesson }" C2 C3 N SG ACC-N
"<yksityistalon>"
"{ private|house }" C2 N SG GEN
"<yksityistalouksiin>"
"{ private household }" C2 N PL ILL
"<yksityistilaisuuksiin>"
"{ private|occasion }" C2 N PL ILL
"<yksityistuntia>"
"{ private|hour }" C2 N SG PAR

Each compound is now translated. We still need to separate the segments as individual words (18).

(18)
"<yksityiskodit>"
"{ private house }" C2 N PL NOM
"{ private house }" C2 N PL ACC
"{ private house }" C2 N PL ACC-N
"<yksityiskohdan>"
"yksityiskohta { detail }" N SG GEN
"<yksityiskohtaisuus>"
"{ accuracy }" C2 N SG NOM
"{ accuracy }" C2 N SG ACC-N
"<yksityiskäyttöön>"
"{ private use } C2 N SG ILL
"<yksityiskäytöstä>"
"{ private behavior }" C2 N SG PAR
"{ private use } C2 N SG ELA
"<yksityismetsien>"
"{ private forest } C2 N PL GEN
"<yksityisoppitunti>"
"{ private lesson }" C2 C3 N SG NOM

"{ private lesson }" C2 C3 N SG ACC-N
"<yksityistalon>"
"{ private house }" C2 N SG GEN
"<yksityistalouksiin>"
"{ private household }" C2 N PL ILL
"<yksityistilaisuuksiin>"
"{ private occasion }" C2 N PL ILL
"<yksityistuntia>"
"{ private hour }" C2 N SG PAR

5 Discussion and conclusion

In this report I have shown how compound words in Finnish can be translated into English in a covering and efficient way. Although the number of compounds is almost endless, it is possible to perform translation using the split-word method. The requirement for translation is that segment boundaries in compounds are marked, and each segment is translated separately. The method is efficient, but its back side is that all segments cannot be translated without context information. The solution suggested here is that problematic cases are translated so that the demarcated segment and its adjacent segment, whether it is before or after it, is translated as a single string. In the translation process, longer strings override shorter strings, so that strings with combined segments will be translated first. We can take such an example as *yksityistalous*. It is analysed in (19).

(19)
"<yksityistalous>"
"yksityis|talous" C1 N SG NOM
"yksityis|talous" C1 N SG ACC-N
"yksityis|talous" C2 N SG NOM
"yksityis|talous" C2 N SG ACC-N

It was analysed via two routes. One found the string *yksityis* in the first slot (C1), and another found it in second slot (C2). The string *talous* was found in the noun lexicon. When we translate the compound using the standard system, we get the result as in (20).

(20)
"<yksityistalous>"
"{ private|talous { economics , economy , household }" C2 N SG NOM
"{ private|talous { economics , economy , household }" C2 N SG ACC-N

If we accept the default gloss economics, we get a wrong translation. Therefore, we must translate the string *yksityis|talous* instead the string *talous* (21).

(21)
"<yksityistalous>"
"{ private household }" C2 N SG NOM
"{ private household }" C2 N SG ACC-N

The string *yksityis|talous* overrides the string *talous*, which is why we get the correct translation.

Another clashing case occurs, when we get the analysis result through the single-word system and the split-word system. An example is in (22).

(22)
"<yksityiskohta>"
 "yksityiskohta" N SG NOM
 "yksityiskohta" N SG ACC-N
 "yksityis|kohta" C1 N SG NOM
 "yksityis|kohta" C1 N SG ACC-N
 "yksityis|kohta" C2 N SG NOM
 "yksityis|kohta" C2 N SG ACC-N

This is an easy case, because the disambiguator removes split-word interpretations, if a single-word interpretation is present (23).

(23)
"<yksityiskohta>"
 "yksityiskohta" N SG NOM
 "yksityiskohta" N SG ACC-N

This will then be translated as a single string (24).

(24)
"<yksityiskohta>"
 "yksityiskohta { detail }" N SG NOM
 "yksityiskohta { detail }" N SG ACC-N

This report is based on experiments made on various kinds of translation problems of compounds. Only nouns were included into the test. The results suggest, however, that also other compound types, such as adjectives, verbs, and adverbs, can be handled in the similar way. However, the full implementation for covering free text requires a substantial amount of work.