

Rule-based language technology applied to learning Finnish¹

Arvi Hurskainen
Department of Languages
FIN-00014 University of Helsinki, Finland
arvi.hurskainen@helsinki.fi
DOI: 10.13140/RG.2.2.25898.29120

Abstract

Rule-based language technology (RBLT) offers possibilities for developing many kinds of self-tutored language learning systems. The benefits of this technology are especially prominent when learning morphologically complex languages. Most language learning applications in the web are videos, where a person teaches various features of the language – very much as in a normal classroom. There are also various types of games, where certain features are learned by filling in missing slots. All such learning systems are tied up to a certain order in learning.

RBLT offers possibilities for developing such learning systems, where the learner is given almost total freedom to choose what to learn, and in which order to learn. This technology makes the learner free of vocabulary restrictions - the whole vocabulary can be used. The learner is also free of what structures or forms should be learned - the system recognises all word-forms and phrase types.

The learning system can be made to give responses on three types of mistakes, such as typos, word order, and concordance. In addition, various other types of useful information can be given, such as inflection classes, base form of the word, and the inflection stem of the word. Also information on other features of the word can be given.

In this report I will demonstrate how RBLT can be applied on learning Finnish phrase structures and verb forms.

Key Words: *morphological analysis, disambiguation, language learning.*

1 Introduction

In earlier technical reports (Reports No. 3², 9³, and 82⁴) I have described how RBLT can be applied to learning Swahili noun phrases and verb phrases. As other Bantu languages, Swahili has an elaborate noun class system. This makes the construction of the learning system complicated because of the huge number of different structure types.

¹ The report is issued under licence CC BY-NC

² <http://www.njas.helsinki.fi/salama/language-learning-system.pdf>

³ <http://www.njas.helsinki.fi/salama/language-learning3.pdf>

⁴ <http://www.njas.helsinki.fi/salama/rule-based-language-technology-and-self-tutored-language-learning-systems.pdf>

The experience has shown that the coverage and accuracy of the learning system depends entirely on two key components of the system. One is the recall and precision levels of the morphological analyser, and another is the accuracy of the disambiguation system. The wide coverage of the lexicon increases ambiguity, and this again causes problems in disambiguation, especially because the learning phrases are often not full sentences. Reliable disambiguation rules often require access to the whole sentence, and sometimes beyond the sentence boundary. Therefore, there are cases, where it is advisable to remove some rare words from the lexicon that cause such ambiguity that cannot be resolved on the basis of the restricted sequence of words.

Finnish typically inflects to the right. In other words, inflection morphemes are suffixes, and often sequences of suffixes. Finnish has also several inflection classes. For example, verbs in the web applications are often grouped into six main inflection types. This division is, however, too rough, because each of the six classes have many different types of exceptions. Verbs also have inflection patterns according to their stem vowels. Front vowel verbs require inflection with front vowels, and back vowel verbs require inflection with back vowels. The vowels *e* and *i* are neutral vowels and they can occur in both inflection types. Gradation is another feature that causes deviant inflection patterns.

In all, in my morphological analyser I have a total of 218 different types of verbs, each with a unique inflection pattern. In some cases the difference is minimal. There are some inflection classes, which have only one verb. In my system I have omitted ten such cases. These are verbs that unlikely appear in texts. More commonly occurring verbs I have accepted, although they alone form an inflection class. The total number of verbs in the system is 9701. Most of the verbs were extracted from the KOTUS (Kotimaisten kielten keskus) database, and verbs missing in the KOTUS database were added.

Because the learning system based on the RBLT recognizes practically all forms of all words, the learning strategy must be very different from traditional learning systems. Full freedom in what to study is not necessarily only a good thing. Learning should be done systematically, so that all features of the language will be learned. In the Swahili learning system, I implemented also a guided tour through all important phrase structures. It required 43 lessons, each with a number of subtasks.

With Finnish, one possibility is to start training with all verb types. The system allows training with any of the verbs of each type. One can learn with all six persons. Also various other verb forms of each verb can be trained, such as present tense, past tense, conditional and potential forms.

The various forms can be seen in the tags printed in the output. They can be also made more explicit by printing the information in clear text.

The environment based on the RBLT is as a kind of recipe book but without precise recipes. It lists a large amount of food ingredients, but does not give instructions in advance on how to combine them. The user of the recipe book is asked to test which ingredients fit together. If the user selects poisonous or otherwise unfitting ingredients together, the system warns about it and gives instructions on choosing suitable ingredients.

2 Phases in constructing the learning interface

It is perhaps not usual to describe all the phases of the process in an application. However, such description is useful for the reader, but also for the developer, because on the basis of

such a description it is easier to memorise the process and make corrections later if needed. The description below contains such phases that are considered critical for understanding the process, and less problematic phases are omitted.

The two basic components are the morphological analyser and disambiguator. I will describe both briefly.

2.1 Morphological analysis

The morphological analyser of Finnish was constructed using the developing environment known as two-level morphology. While normally the analyser is constructed using the lexicon and rule components, in this implementation the rule component was excluded, and all wordforms were described in the lexicon. The result of this was that the lexicon became too large to be run within the system. I had to split the lexicon so that the verbs were separated and included into a different lexicon. Both were separately compiled and then run after each other.

The result when run with the first lexicon without verbs is in (1).

```
(1)
"<*tämä>"
    "tämä" PRON CAP SG DEM1 NOM
    "tämä" PRON CAP SG DEM1 ACC-N
"<pieni>"
    "pieni" A SG NOM
    "pieni" A SG ACC-N
"<kissa>"
    "kissa" N SG NOM
    "kissa" N SG ACC-N
"<naukuu>" Heur
"<>"
    "." **CLB
```

We see that all words except for the verb were analysed. When we rerun this result with the verb lexicon, we get the result as in (2).

```
(2)
"<*tämä>"
    "tämä" PRON CAP SG DEM1 NOM
    "tämä" PRON CAP SG DEM1 ACC-N
"<pieni>"
    "pieniä" V V61-f R:pien IMP SG2
    "pieniä" V V61-f R:pien NEG-PRES
    "pieniä" V V61-f R:pien PAST SG3
    "pieni" A SG NOM
    "pieni" A SG ACC-N
"<kissa>"
    "kissa" N SG NOM
```

```
"kissa" N SG ACC-N  
"<naukuu>"  
"naukua" V V52u-D R:nau PRES SG3  
"<>"  
"." **CLB
```

Now we see that the word *naukuu* was analysed. In addition, the word *pieni* was given also three verb interpretations. Now the string is fully analysed with all ambiguity. The verbs have some strange-looking tags. Their meanings will become clear in later phases of this report.

2.2. Disambiguation

The result will be disambiguated using the Constraint Grammar disambiguator (3).

```
(3)  
"<*tämä>"  
"tämä" +DEM1+SG+NOM+CAP  
"<pieni>"  
"pieni" +A+SG+NOM  
"<kissa>"  
"kissa" +N+SG+NOM  
"<naukuu>"  
"naukua" +V+SG3+R:nau+PRES+V52u-D  
"<>"  
"." **CLB
```

The string was not only disambiguated. The readings were also modified so that they are suitable for further processing. The order of tags was changed, and the tags were joined together using a plus sign.

2.3 Checking for various mistakes

The result will be further processed, so that we can check whether there are mistakes. The first thing to check is whether there are typos (4).

```
(4)  
*iämä pieni ?kssa? naukuu . +DEM1+SG+NOM+CAP +A+SG+NOM Heur  
+V+SG3+R:nau+PRES+[naukua]+V52u-D
```

Question marks were placed around the misspelled word. Then we check whether the order of words is correct (5).

```
(5)  
*iämä pieni ?kssa? naukuu . +DEM1+SG+NOM+CAP +A+SG+NOM Heur  
+V+SG3+R:nau+PRES+[naukua]+V52u-D HEUR3
```

The tag HEUR3 was added telling that the third word of the string is misspelled. We correct the typo (6).

(6)
**tämä pieni kissa naukuu .* +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+SG3+R:nau+PRES+[naukua]+V52u-D DEM+A+N+V_WO

Now a tag describing the order of words was added. The last part _WO indicates that the word order is correct. Let us see what happens when we change the word order (7).

(7)
**tämä kissa pieni naukuu .* +DEM1+SG+NOM+CAP +N+SG+NOM +A+SG+NOM
+V+SG3+R:nau+PRES+[naukua]+V52u-D N+A_!WO

Now the tag N+A_!WO was added telling that the word order on this point is wrong. In the next phase we check whether the concordance is correct (8)

(8)
**tämä pieni kissa naukuu .* +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+SG3+R:nau+PRES+[naukua]+V52u-D DEM+A+N+V_WO CONC4

The tag CONC4 was added telling that the concordance is correct and the string has four words. If the concordance is wrong, the tag will not be printed (9).

(9)
**tämä pieni kissa naukuvat .* +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+PL3+R:nau+PRES+[naukua]+V52u-D DEM+A+N+V_WO

The absence of concordance information can be used as a trigger for warning about the mistake in concordance, as we see in the next phase of processing (10).

(10)
Please check concordance!
**tämä pieni kissa naukuvat .* +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+PL3+R:nau+PRES+[naukua]+V52u-D DEM+A+N+V_WO

On this point, also other warnings and instructions are given (11).

(11)
Please check word order!
**tämä kissa pieni naukuu .* +DEM1+SG+NOM+CAP +N+SG+NOM +A+SG+NOM
+V+SG3+R:nau+PRES+[naukua]+V52u-D N+A_!WO

Also instructions on spelling are given (12).

(12)

Check the spelling of the third word!

**tämä pieni kissa? naukuu* . +DEM1+SG+NOM+CAP +A+SG+NOM Heur
+V+SG3+R:nau+PRES+[naukua]+V52u-D HEUR3

In addition to the three types of instructions - typos, word order and concordance - also other types of useful information can be given. Examples on verb forms are below (13).

(13)

Word order and concordance are correct!

- The INFLECTION CLASS of the verb is V52u-D
- The BASE FORM of the verb is [naukua].
- The STEM of the verb is {nau}.
- The verb is in PRESENT tense form.

**tämä pieni kissa naukuu* . +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+SG3

The word order and concordance are considered correct. The inflection class as defined in the morphological analyser is printed. Also the base form and stem of the verb are printed. The verb is here in present tense form. We can test also with other forms (14).

(14)

Word order and concordance are correct!

- The INFLECTION CLASS of the verb is V52u-D
- The BASE FORM of the verb is [naukua].
- The STEM of the verb is {nau}.
- The verb is in PAST tense form.

**tämä pieni kissa naukui* . +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+SG3

Now the verb is in past tense form. Further forms include the potential form (15).

(15)

Word order and concordance are correct!

- The INFLECTION CLASS of the verb is V52u-D
- The BASE FORM of the verb is [naukua].
- The STEM of the verb is {nau}.
- The verb is in POTENTIAL form.

**tämä pieni kissa naukunee* . +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+SG3

Also the conditional form can be reported on (16).

(16)

Word order and concordance are correct!

- The INFLECTION CLASS of the verb is V52u-D
- The BASE FORM of the verb is [naukua].
- The STEM of the verb is {nau}.
- The verb is in CONDITIONAL form.

**tämä pieni kissa naukuisi.* +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+SG3

The result is then finally pruned. In one version the analysis result is retained (17)

(17)

Word order and concordance are correct!

- The INFLECTION CLASS of the verb is V52u-D
- The BASE FORM of the verb is [naukua].
- The STEM of the verb is {nau}.
- The verb is in CONDITIONAL form.

Tämä pieni kissa naukuisi. +DEM1+SG+NOM+CAP +A+SG+NOM +N+SG+NOM
+V+SG3

In another version the analysis result is removed (18).

(18)

Word order and concordance are correct!

- The INFLECTION CLASS of the verb is V52u-D
- The BASE FORM of the verb is [naukua].
- The STEM of the verb is {nau}.
- The verb is in CONDITIONAL form.

Tämä pieni kissa naukuisi.

3 Testing with various kinds of strings

Now when we have gone through all the phases of processing, we can test the learning system with various kinds of strings and mistake types. Instead of using the prompt-based processing, I will use the browser-based interface.

We will test first various kinds of typos. First we have the string without mistakes (19).

(19)

Tämä minun kaunis
kissani nukkuu.

TAG

CLEAR

Word order and concordance are correct!
- The INFLECTION CLASS of the verb is V52-A
- The BASE FORM of the verb is [nukkua].
- The STEM of the verb is {nuk}.
- The verb is in PRESENT tense form.

Tämä minun kaunis kissani nukkuu. +DEM1+SG+NOM+CAP
+POSS+SG1 +A+SG+NOM +N+SG+NOM+POS-SG1 +V+SG3

The string is entered into the left box, and the result is shown in the right box. The entered string is repeated on the low line, together with the interpretation of each word. The report lines are above. On the topmost line, info on the word order and concordance is reported. Then follows a set of info lines on the verb form. They help the learner to construct other verb forms of the same verb, and also of the other verbs of the same inflection class. In the web, there is also a list of Finnish verbs listed according to their inflection class.⁵

Now we make some typos to the string (20).

(20)

Tämä minn kaunis
kissni nukuu.

TAG

CLEAR

Check the spelling of the second, fourth and fifth word!
Tämä ?minn? kaunis ?kissni? ?nukuu? . +DEM1+SG+NOM+CAP
Heur +A+SG+NOM Heur Heur

There are three misspelled words, each surrounded by a question mark. The report tells, that the spelling of the second, fourth and fifth word should be checked. We correct them but make a mistake in one correction (21).

(21)

Tämä minu kaunis
kissani nukkuu.

TAG

CLEAR

Check the spelling of the second word!
- The INFLECTION CLASS of the verb is V52-A
- The BASE FORM of the verb is [nukkua].
- The STEM of the verb is {nuk}.
- The verb is in PRESENT tense form.

Tämä ?minu? kaunis kissani nukkuu. +DEM1+SG+NOM+CAP
Heur +A+SG+NOM +N+SG+NOM+POS-SG1 +V+SG3 HEUR2

Now the warning tells that the spelling of the second word should be checked. The verb is correct, and all relevant info on it is printed.

Next we test the word order (22).

⁵ <http://www.njas.helsinki.fi/salama/finnish-verbs.pdf>

(22)

Minun tämä kissani kaunis nukkuu.	TAG	Possessive pronoun cannot be before the demonstrative pronoun! - The INFLECTION CLASS of the verb is V52-A - The BASE FORM of the verb is [nukkuu]. - The STEM of the verb is {nuk}. - The verb is in PRESENT tense form.
	CLEAR	
Minun tämä kissani kaunis nukkuu. +POSS+SG1+CAP +DEM1+SG+NOM +N+SG+NOM+POS-SG1 +A+SG+NOM +V+SG3		

We are warned that possessive pronoun cannot be before the demonstrative pronoun. No other warnings are reported. We correct the order of pronouns (23).

(23)

Tämä minun kissani kaunis nukkuu.	TAG	Please check word order! Adjective cannot be after noun in this context. - The INFLECTION CLASS of the verb is V52-A - The BASE FORM of the verb is [nukkuu]. - The STEM of the verb is {nuk}. - The verb is in PRESENT tense form.
	CLEAR	
Tämä minun kissani kaunis nukkuu. +DEM1+SG+NOM+CAP +POSS+SG1 +N+SG+NOM+POS-SG1 +A+SG+NOM +V+SG3		

The word order is not yet correct. The warning tells that the order of noun and adjective is not correct in this context. This means that the word order may be correct in another context. We correct the word order (24).

(24)

Tämä minun kaunis kissani nukkuu.	TAG	Word order and concordance are correct! - The INFLECTION CLASS of the verb is V52-A - The BASE FORM of the verb is [nukkuu]. - The STEM of the verb is {nuk}. - The verb is in PRESENT tense form.
	CLEAR	
Tämä minun kaunis kissani nukkuu. +DEM1+SG+NOM+CAP +POSS+SG1 +A+SG+NOM +N+SG+NOM+POS-SG1 +V+SG3		

Now we test with mistakes in concordance (25).

(25)

Nämä minun kolmet kissat nukkuu.	TAG	Please check concordance! - The INFLECTION CLASS of the verb is V52-A - The BASE FORM of the verb is [nukkuu]. - The STEM of the verb is {nuk}. - The verb is in PRESENT tense form.
	CLEAR	
Nämä minun kolmet kissat nukkuu. +DEM1+PL+NOM+CAP +POSS+SG1 +NUM+PL+NOM +N+PL+NOM +V+SG3 **		

The warning tells that the concordance should be checked. However, it cannot give precise instructions, because it is not clear what the typed string should be. We try to make one correction (26).

(26)

Nämä minun kolmet kissat nukkuvat.	TAG	Please check concordance! - The INFLECTION CLASS of the verb is V52-A - The BASE FORM of the verb is [nukkuu]. - The STEM of the verb is {nuk}. - The verb is in PRESENT tense form.
	CLEAR	
Nämä minun kolmet kissat nukkuvat. +DEM1+PL+NOM+CAP +POSS+SG1 +NUM+PL+NOM +N+PL+NOM +V+PL3 **		

It did not solve the problem. We make further corrections (27).

(27)

Nämä minun kolme kissaa nukkuvat.	TAG	Please check concordance! - The INFLECTION CLASS of the verb is V52-A - The BASE FORM of the verb is [nukkuu]. - The STEM of the verb is {nuk}. - The verb is in PRESENT tense form.
	CLEAR	
Nämä minun kolme kissaa nukkuvat. +DEM1+PL+NOM+CAP +POSS+SG1 +NUM+SG+NOM +N+SG+PAR +V+PL3 **		

There is still a warning on concordance. We continue with correcting (28).

(28)

Nämä minun kolme kissaa nukkuu.	TAG	Please check concordance! - The INFLECTION CLASS of the verb is V52-A - The BASE FORM of the verb is [nukkuu]. - The STEM of the verb is {nuk}. - The verb is in PRESENT tense form.
	CLEAR	
Nämä minun kolme kissaa nukkuu. +DEM1+PL+NOM+CAP +POSS+SG1 +NUM+SG+NOM +N+SG+PAR +V+SG3 **		

Now the string seems correct, but it still has one concordance mistake. We add a possessive suffix (29).

(29)

Nämä minun kolme kissaani nukkuvat.	TAG	Word order and concordance are correct! - The INFLECTION CLASS of the verb is V52-A - The BASE FORM of the verb is [nukkua]. - The STEM of the verb is {nuk}. - The verb is in PRESENT tense form.
	CLEAR	
Nämä minun kolme kissaani nukkuvat. +DEM1+PL+NOM+CAP +POSS+SG1 +NUM+SG+NOM +N+SG+PAR+POS-SG1 +V+PL3		

Now the string is correct⁶. In speech, the possessive suffix is often omitted, but in correct writing it is not acceptable. Therefore, this learning system warns about omission.

Next, we test with an example with all three kinds of mistakes – spelling, word order, and concordance (30).

(30)

Kauniit nämä kua tuksuu.	TAG	Check the spelling of the third word! - The INFLECTION CLASS of the verb is V52 - The BASE FORM of the verb is [tuksua]. - The STEM of the verb is {tuksu}. - The verb is in PRESENT tense form.
	CLEAR	
Kauniit nämä ?kuat? tuksuu. +A+PL+NOM+CAP +DEM1+PL+NOM Heur +V+SG3 HEUR3		

First, we are reported on spelling mistakes. We make a correction (31).

(31)

Kauniit nämä kukat tuksuu.	TAG	Adjective followed by pronoun cannot initiate a phrase! - The INFLECTION CLASS of the verb is V52 - The BASE FORM of the verb is [tuksua]. - The STEM of the verb is {tuksu}. - The verb is in PRESENT tense form.
	CLEAR	
Kauniit nämä kukat tuksuu. +A+PL+NOM+CAP +DEM1+PL+NOM +N+PL+NOM +V+SG3		

⁶ It is unclear whether the verb should be *nukkuu* or *nukkuvat*. The former one is correct if the phrase does not have a possessive pronoun, but in this example perhaps both alternatives are correct.

Second, the word order is checked. We correct the word order (32).

(32)

Nämä kauniit kukat tuoksuu.

TAG

CLEAR

Please check concordance!
- The INFLECTION CLASS of the verb is V52
- The BASE FORM of the verb is [tuoksua].
- The STEM of the verb is {tuoksu}.
- The verb is in PRESENT tense form.

Nämä kauniit kukat tuoksuu. +DEM1+PL+NOM+CAP
+A+PL+NOM +N+PL+NOM +V+SG3

In the third phase, the concordance is checked. We correct the verb form (33).

(33)

Nämä kauniit kukat tuoksuvat.

TAG

CLEAR

Word order and concordance are correct!
- The INFLECTION CLASS of the verb is V52
- The BASE FORM of the verb is [tuoksua].
- The STEM of the verb is {tuoksu}.
- The verb is in PRESENT tense form.

Nämä kauniit kukat tuoksuvat. +DEM1+PL+NOM+CAP
+A+PL+NOM +N+PL+NOM +V+PL3

Now the string is correct.

4 Learning with longer strings

The basic learning system is so constructed that the number of words in the string is limited. It means that the sequence of words must match with the rules, which guide the learning process. Currently, the maximum number of words in the string is six. Almost all noun phrases in Finnish fall within this limit. At least it is not feasible to train with longer phrases. The learner might be willing to train also with longer strings. The learner can type strings of any length, but the guiding system is applied only to the first six words in the string. For example, if the first six words match with the longest rule in the learning system, this part is treated as a unit of its own, and instructions are given if this part contains any mistakes. Also the rest of the string is printed and analysed, but this part is ignored in further processing.

An example of a string that exceeds the length of the standard phrase is in (34).

(34)

Nämä minun kolme
kissaani syövät kalaa
ja juovat maitoa.

TAG

CLEAR

Word order and concordance are correct!
- The INFLECTION CLASS of the verb is V64y-f
- The BASE FORM of the verb is [syödä].
- The STEM of the verb is {s}.
- The verb is in PRESENT tense form.

Nämä minun kolme kissaani syövät kalaa ja juovat maitoa.
+DEM1+PL+NOM+CAP +POSS+SG1 +NUM+SG+NOM
+N+SG+PAR+POS-SG1 +V+PL3+TRV
+N+SG+PAR +CC +V+PL3+R:j+PRES+TRV+[juoda]+V64
+N+SG+PAR

The rules for covering the correctness of the string cover the words *Nämä minun kolme hyvää kissaani syövät* in the above sentence. The rest of the string is analysed, but it is not part of the checking routine. The verb *syödä* is part of this routine, and the morphological details of this verb are reported. The verb *juoda* is not part of the checking routine, and the analysis result is printed as such. When we look at the analysis part of these two verbs, we see a difference. For the verb *syödä*, the analysis is +V+SG3+TRV, and all additional information is moved to the report section above. For the verb *juoda*, there is full analysis +V+PL3+R:j+PRES+TRV+[juoda]+V64.

For the sake of clarity, the analysis parts of these two sections are separated in the output.

5 Learning deviant structures

In addition to assisting in learning standard rules of the language, the system also assists in constructing deviant constructions, that is, such constructions that do not follow the general rules. Below are some examples of such constructions (35).

(35)

Kissalla on nälkä.

TAG

CLEAR

Word order and concordance are correct!
- The INFLECTION CLASS of the verb is V67b
- The BASE FORM of the verb is [olla].
- The verb is in PRESENT tense form.
- The subject must be in ADESSIVE form.

Kissalla on nälkä. +N+SG+ADE+CAP +V+SG
+N+SG+NOM

In this phrase, the subject must be in adessive form instead of the normal nominative form. Note that the string to which the rules apply is *Kissalla on*, and the analysis of the rest is printed separately. We test this phenomenon with a longer string (36).

(36)

Minun kissallani on nälkä ja jano.	TAG	Please check concordance! - The INFLECTION CLASS of the verb is V67b - The BASE FORM of the verb is [olla]. - The verb is in PRESENT tense form. - The subject must be in <u>ADESSIVE</u> form.
	CLEAR	
		Minun kissallani on nälkä ja jano. +POSS+SG1+CAP +N+SG+ADE+POS-SG1 +V+SG +N+SG+NOM +CC +N+SG+NOM

There are some modal verbs, which require that the subject is in adessive form. These verbs include *pitää*, *tulee*, and *täytyy*. The two first ones have also other roles, and they must be disambiguated on the basis of the context. Let us take first an example with the verb *pitää* in usual meaning (37).

(37)

Minä pidän paitaa ja housuja.	TAG	Word order and concordance are correct! - The INFLECTION CLASS of the verb is V53-F-f - The BASE FORM of the verb is [pitää]. - The verb is in PRESENT tense form.
	CLEAR	
		Minä pidän paitaa ja housuja. +PRON-PERS+SG1+NOM+CAP +V+SG1+TRV +N+SG+PAR +CC +N+PL+PAR

Here the subject is in nominative form. Next we take an example, where the verb *pitää* is in modal role. First we put the subject into nominative form (38).

(38)

Poika pitää mennä.	TAG	Subject must be in genitive form! - The INFLECTION CLASS of the verb is V67c-f - The BASE FORM of the verb is [mennä]. - The verb is in PRESENT tense form.
	CLEAR	
		Poika pitää mennä. +N+SG+NOM+CAP +V+SG3+S:GEN+O:ELA+VMOD+[pitää]+V53-F-f +V+R:men+INF+VMOD

On the first line there is a warning that the subject should be in genitive form in this phrase. We correct it (39).

(39)

Pojan pitää mennä. TAG CLEAR

Word order and concordance are correct!
- The INFLECTION CLASS of the verb is V67c-f
- The BASE FORM of the verb is [mennä].
- The verb is in PRESENT tense form.

Pojan pitää mennä. +N+SG+GEN+CAP
+V+SG3+S:GEN+VMOD+[pitää]+V53-F-f
+V+R:men+INF+VMOD

Now the phrase is correct. The info is given for the second verb *mennä*. We can test with the other modal verb *täytyy* (40).

(40)

Poika täytyy lukea. TAG CLEAR

Subject must be in genitive form!
- The INFLECTION CLASS of the verb is V58-D
- The BASE FORM of the verb is [lukea].
- The STEM of the verb is {t}.
- The verb is in PRESENT tense form.

Poika täytyy lukea. +N+SG+NOM+CAP
+V+SG3äy+S:GEN+[täytyä]+V52-F-f +V+R:lu+INF+TRV

There is again a warning that the subject must be in genitive form. In fact, any subjects and any of the last verbs can be used in the structure.

6 Discussion and conclusion

In this report I have shown some examples of how RBLT can be applied to self-tutored learning of Finnish phrases and verb forms. The system gives full freedom to the learner, which in itself might be a problem for many learners. The user of the system must formulate the structures, and also the choice of the words is up to the learner.

One way to start learning is to go through the verb list in the Appendix and train each verb with all person forms, for example, *Minä astun*, *Sinä astut*, *Hän astuu*, *Me astumme*, *Te astute*, *He astuvat*. One can also train with other verb forms, such as, *astuin*, *astuisin*, *astunee* etc.

Verbs of different inflection classes have often very similar inflection in these basic inflection forms. However, they have differences in some other forms, which are not part of this learning system. There is also a classified list of all Finnish verbs⁷, in case one wants to engage into an ultra marathon and learn with all Finnish verbs.

⁷ <http://www.njas.helsinki.fi/salama/finnish-verbs.pdf>

This report has described only a part of learning possibilities of the RBLT approach. There are many more tricky features to learn, such as the object case, which often causes problems even for very advanced learners of Finnish. Perhaps I can return to such problems later.

APPENDIX

Inflection classes of verbs

One verb is selected for representing each inflection class.

V52	astua	step
V52-A	nukkua	sleep
V52-A-f	säikkyä	startle
V52-B	loppua	end
V52-B-f	leppyä	relent
V52-C	asettua	settle
V52-C-f	hedelmöittyä	conceive
V52-D-f	näkyä	be visible
V52-E	saapua	arrive
V52-E-f	yöpyä	overnight
V52-F	ajautua	drift
V52-F-f	erehtyä	mistake
V52-Fo	kutoa	weave
V52-G	hinkua	hanker
V52-H	ampua	shoot
V52-I	paleltua	freeze
V52-I-f	höltyä	loosen
V52-J	hajaantua	disperse
V52-J	jakaantua	split-up
V52-J-f	hiljentyä	calm-up
V52-K	juurtua	root
V52-K-f	nöyryä	humble
V52-f	edistyä	proceed
V52o	hieroa	rub
V52o-A	rikkoa	break
V52o-D	aikoa	intend
V52o-E	leipoa	bake
V52o-G	penkoa	delve
V52o-H	tempoa	wrench
V52o-K	kertoa	tell
V52u-D	liukua	slide
V53	aavistaa	anticipate
V53-C	aiheuttaa	cause
V53-C-f	elättää	sustain
V53-D	purkaa	dissemble

V53-F	armahtaa	pardon
V53-F-f	säikähtää	be frightened
V53-K	sortaa	oppress
V53-f	elää	live
V54-F	huutaa	shout
V54-F-f	löytää	find
V54-I	oivaltaa	realise
V54-I-f	kieltää	deny
V54-J	suurentaa	enlarge
V54-J-f	heikentää	weaken
V54-K	kumartaa	bow
V54-K-f	kiertää	circle
V54-f	lypsää	milk
V55-F	soutaa	row
V55-F-f	liittää	glide
V55-I-f	yltää	reach
V55-J-f	entää	fly
V56	kasvaa	grow
V56-A	virkkää	say
V56-B	tappaa	kill
V56-C	auttaa	help
V56-D	alkaa	begin
V56-F	raataa	labour
V56-J	antaa	give
V57-F	kaataa	fell
V57-K	saartaa	encircle
V58	laskea	count
V58-D	hakea	fetch
V58-E-f	kylpeä	bathe
V58-F	potea	suffer
V58-F-f	päteä	be valid
V58-G	tunkea	cram
V58-L	kulkea	go
V58-L-f	särkeä	break
V58-f	itkeä	cry
V59-J	tuntea	feel
V60-F-f	lähteä	leave
V61	halveksia	despise
V61-A	hankkia	acquire
V61-A-f	leikkiä	play
V61-B	oppia	learn
V61-B-f	nyppiä	pluck
V61-C	moittia	blame
V61-C-f	miettiä	ponder
V61-D	poikia	calve
V61-D-f	pyrkiä	pursue

V61-E	sopia	agree
V61-E-f	hiipiä	tiptoe
V61-F	kadehtia	envy
V61-F-f	kiirehtiä	hurry
V61-G	onkia	angle
V61-G-f	tinkiä	bargain
V61-H-f	empiä	hesitate
V61-J	sontia	defecate
V61-L-f	hylkiä	repel
V61-f	etsiä	seek
V62	adoptoida	adopt
V62-f	esitelmöidä	lecture
V63	saada	get
V63-f	jäädä	stay
V63y-f	myydä	sell
V64	tuoda	bring
V64i-f	viedä	move
V64y-f	syödä	eat
V65-f	käydä	go
V66	ilmaista	express
V66-E-f	häväistä	disgrace
V66-G	rangaista	punish
V66-f	ehkäistä	prevent
V67	ahdistella	harass
V67-A-f	leikellä	slice
V67-B	tapella	fight
V67-B-f	hypellä	jump about
V67-C	ajatella	think
V67-C-f	esitellä	introduce
V67-F	huudella	heckle
V67-F-f	riidellä	quarrel
V67-H	ommella	sew
V67-I	sukellella	keep diving
V67-I-f	vihellellä	keep whistling
V67-J	kuunnella	listen
V67-J-f	säännellä	regulate
V67-K	askarrella	be busy
V67-K-f	kierrellä	wander
V67-f	epäillä	doubt
V67b	olla	be
V67c-f	mennä	go
V67n	panna	put
V67r	surra	grieve
V68	ahkeroida	work hard
V68-f	ikävoidä	long-for
V69	ansaita	earn

V69-f	häiritä	disturb
V71-f	tehdä	make
V72	laajeta	expand
V72-A	hoiketa	slim
V72-A-f	heiketä	weaken
V72-B	helpota	become easier
V72-C	loitota	move away
V72-D	aueta	open
V72-D-f	kyetä	be-able
V72-Da	erata	diverge
V72-Do	ulota	extend
V72-E	halveta	cheapen
V72-F-f	edetä	proceed
V72-L	valjeta	dawn
V72-L-f	seljetä	clear
V72-f	hiljetä	quiet
V72a-F-f	mädätä	rot
V73	arvata	guess
V73-A	hakata	beat
V73-A-f	hyökätä	attack
V73-B	harpata	stride
V73-B-f	hypätä	jump
V73-C	mitata	measure
V73-C-f	kytätä	stalk
V73-D	uhata	threaten
V73-D-f	hylätä	abandon
V73-E	luvata	promise
V73-E-f	levätä	rest
V73-F	haudata	bury
V73-F-f	tähdätä	aim
V73-G	hangata	rub
V73-I	vallata	capture
V73-J-f	rynnätä	charge
V73-K	verrata	compare
V73-L-f	peljätä	fear
V73-f	herätä	awaken
V74	erota	separate
V74-B	upota	sink
V74-C	lotota	play lotto
V74-D-f	keretä	reach in time
V74-E	turvota	swell
V74-E-f	hävetä	be ashamed
V74-K	irrota	detach
V74-L	haljeta	split
V74-L-f	teljetä	bar
V74-f	älytä	realise

V74e-A	poiketa	deviate
V74e-D	laueta	go-off
V74e-F	todeta	notice
V74e-G	langeta	lapse
V74e-H	kammeta	crank
V74o-D	liota	soak
V74o-F	pudota	fall
V74o-G	lingota	sling
V74u-E	kavuta	clamber
V74u-H	kummuta	spring
V75	haluta	want
V75-B-f	ryöpytä	rush
V75-D-f	selitä	clear
V75-H-f	lämmitä	warm
V75-I	aallota	wave
V75-I-f	hellitä	free
V75-f	hävitä	lose
V76-F	taitaa	master
V76-F-f	tietää	know
V77	kumajaa	boom
V78	kaikaa	echo
V78-f	ryskää	crash