

Джек М. Рютер (*Хельсинки*)

**Хельсинкиса университетын кыв туялысь
Ижкарын перымса кывъяс симпозиум вылын лыддьомтор**

Финляндияса академия отсӧгӧн Хельсинкиса университе-
тын вӧчӧны ичӧт войтырлӧн кывъяс гижӧд чукӧрлысь проект,
кодӧ лоӧ Войвыв Муясса ӧтувъя проектлӧн юкӧнӧн. Ӑти бок-
сянь, Хельсинкиса университетын эмбсь перымса кывъяс туя-
лысь йӧз. Мӧд боксянь, танӧ эмбсь и йӧз, кодъяс уджалӧны
тӧдан машиналысь кыв туялӧмын. Сӧдзи, лингвистика кафедра-
са профессор Киммо Коскенниемилӧн UNIX вылын лӧсьӧдӧм
TWOL уджтасъяс отсӧгӧн гижӧд чукӧр проектн лои позянлун
артмӧдны и кывчукӧр юкӧн.

Ӑнӧ аслам проектн вӧчӧма нин вель уна – кывчукӧрын дас
сюрс сайӧ кыв да кыв артмӧдан 95 ордым, кодъяс серти петкӧд-
чӧ кывсикасьяслӧн артмӧм да вежласьӧм. Воис кад, кор позьӧ
лыддьӧдлыны компьютерса коми кывкудлысь сьӧкылдлунъяс.

Тӧдан машинаӧ медводз колӧ пыртны газет-небӧг гижӧдъ-
яс. Кӧть ӧнӧ кадӧ газет-небӧгтӧ лӧдзӧны тӧдан машинаӧн,
UNIX-ын уджалӧн ковмис и материалсӧ мӧд ногӧн пуктавны,
медым кужны лыддынысӧ. Та могысь ковмис кириллица шы-
пасъяс пыдди лӧсьӧдны торъя анбур. Эска, мый матысса кадӧ
позьӧ кириллица шыпасъяс лыддыны веськыда UNIX вывсьяс.

Онi верман лыддыны сoмын куим лыдакывъя пасъяс, шуам: а = \340, б = \341, в = \342 да св.

Отарсянь кo, омoль, мoдарсянь кo, бур. Кириллича пыдди артмoдoм системасo позьo лыддыны мушар пасъта быдлаын: а = а, б = б, в = в, г = г да с. в. Эмoсь и сэтшoм пасъяс, кодъясoс ковмис мoд ногoн вoчны: ю = u1, ё = o1, я = a1, i = i2, э = e2, ы = i3, o = e3, ж = z4, ч = c4, ш = s4, щ = s5, ь = w, ь = y.

Водзo сета материалсo кирилличаoн, тoдса гижанногсo кок-ниджык лыддынысo.

Кoть тoдан машинаса материалсo бoсьтан веськыда небoг-газет лэдзаниныс, колo тoдны, кыдзи гижны кывъяссo.

1-кo, эм Молодцовлoн гижанног, код подув вылын кывъяс-сo позьo гижны шуанног серти. Оз ков гижны *область* роч ног, позьo и коми ногoн, *область* артмас, мый Молодцов серти роч кирилличаoн, *морковь* водзтi моз лоас *морков*. Литератураын эм и *гoрд вуж*. Кывчукoрад колo пыртны и выль формаяс.

2-кo, коми газетъясын, рекламаясын эм уна материал, ко-дoс комиoдoма роч да мукoд кыв вылыс, мый вoсна паныдасъ-лo интoм, абу стoч коми кыв: *экономикаса министерство* (колo: *экономика министерство*), *челядь сeянiн* (колo: *челядьлoн сe-янiн*).

3-кo, коми кывйын эмoсь шыяс, кодъяс вежласъoны либo ты-далoны сoмын сэк, кор на бoрын муныс вежлoг либo мукoд си-кас кывйыв заводитчo гора шыoн: *вoв + ыс : вoлыс*; *кыв + ын : кывйын*; *син + ас : синмас* да с. в.

Тадзи кывчукoр гижигoн колo кутны тoд вылын, кутшoм веж-ласъoмъяс паныдасъoны. Сьoкыдлуныяс лoбны *к* да *л* шыяскoд.

Выль кывъясас найo озджык вежласъны, шуам *с* бoрын: *воск : воскыс = вос : воскыс = вос : восыс*. Мукoддырйи кывпомса *л* либo *в* шыяс кольoны вежласътoг, шуам: *зал : залыс* либo *ов : овыс*; видз. *ов : оvas = ов : олас*. Та вoсна и колo содтыны кык пас: *б* лыдпасoн кo, *кб* да *лб*.

4-кo, шы вежласъoмъяссыс кындзи эмoсь и мукoд гижан сьo-кыдлуныяс, шуам кык да куим юкoна кывъяс, кодъясoс oтутван-визьoн гижoны: *мам-бать : мам-батьлы = мамы-батьлы*.

Корсюрo кык юкoна кывйын кывйыв содo бoръя юкoн по-мас. Тайo тшoкыджыка паныдасълo коми кывйын. Корсюрo кык юкoна кывйын содтoма кывйывъяс воддза и бoръя юкoн помас, костаныс пуктoма oтутван визь.

Видзёдлам коми гижысь Иван Тороповлысь кыв, кодi уна
нога да аслыспёлёс:

1 *ыши-пошиасьны* (Торопов 1995, 10 лб.)

2 *козья-пи́пуа-пожёма вёрын* (9 лб.)

3 *венё-косьё* (10 лб.)

4 *вердысь-удысьясён* (17 лб.)

1 колё пыртны кывчукөрö. *Ыши-* оз овлы торъя кывйён, став
татшём кык юкёна кывсё пуктам аслас местаё.

2 оз ков пыртны кывчукөрö. Пыртам кё, позьё эськё пукты-
ны нельёд юкён, витёдöс тшöтш – дженъдджыка +а- = 'и': вёр,
көнi эм и коз, и пипу, и пожом и с. в.

3-ыс съёкылдуна. Кыкнан кывйыс овлёны торйён. Кыкнан
юкёныс вежласьёны, мый тöдан машинаыдлы вывті съёкыд гё-
гөрвоны. Сöмын контекст серти тöдан, эм-ö налён торъя вежөр-
тас.

4 пыр кокниджык, кор кык юкёна кадакывсьыд артмё эма-
кыв, сы вöсна мый танi сöмын веськыд юкёныс артмөдöма. Колё-ö
3 да 4 пыртны кывчукөрö, тöдö сöмын велөдчөм коми кывъя
морт. Танi колё кутны тöд вылын, мый тöдан машинаыд сöмын
компьютер – 'артасян'.

Водзё коми кыв правилöяс йывсьыс, кодъясöс лöсьөдöма тö-
дан машиналы:

шыпасчукөр: а б в г д е ё ж з и i й к кб:к л лб:л м н
о ö п р с т у ф х ц ч ш щ ь ы ь э ю я #:0 %- %_:0;

диакритическöй пас: %+;

торъя чукөръяс

! став согласнöй

С = б в г д ж з й к л м н п р с т ф х ц ч ш щ ь ь;

! ь-(небыд пас) кындзи став согласнöй

С1 = б в г д ж з й к л м н п р с т ф х ц ч ш щ ь;

! став согласнöй, кодъяс бöрын вермас лоны сöмын и

С2 = б в г ж й к м п р ф х ч ш щ;

! гора шыпасъяс

V1 = а е ё и i о ö у ы э ю я;

! шыпасъяс

Segm = б в г д ж з й к л м н п р с т ф х ц ч ш щ ь

ь

а е ё и = о ö у ы э ю я;

! кыв су (öтувтöм чукөръясын, торйён гижöм кывъясын)

```

WB = # %- % _;
Гижан правилӧяс
"л:в" л:в <=> V1 _ [WB: | %+: C1:];
"м:0" м:0 <=> н _ [WB: | %+: C1:];
"к:0" к:0 <=> [с | ш | с ь | ч] _ [WB: | %+: C1:];
"т:0" т:0 <=> п _ [WB: | %+: C1:];
"ы:0" ы:0 <=> V1: + C:* _ %+: V1;
"л:0" л:0 <=> л _ ь [WB: | %+: C1:];
"д:0" д:0 <=> д _ [ь | з] [WB: | %+: C1:];
"й:0" й:0 <=> V1 _ [б | в | г | д | ж | з | к | м | н | п | р | т] _ [WB:
| %+: C1:];
"ьяс:ьяс" ь:ь <=> [д: з: | :ч :0*] %+: _ я с;
"ьяс:яс" ь:0 <=> [V1 (й:) | ь] %+: _ я с;
"ь+а:00я, й+а:ъ0я да с. в (1)"
X:Y <=> [\[:з |:ч] | \:д :з] [ь: | й:] V1:0* %+: _;
X (а о у)
Y (я ё ю);
"ь+а:00я да с. в. (2)"
ь:0 <=> _ V1:0* %+: _ [а: | о: | у: | i:];
"й+а:ъ0я да с в (2)"
й:ъ <=> C1 _ V1:0* %+: [а: | о: | у:];
"й+а:00я да с. в. (2)"
й:0 <=> V1 _ V1:0* %+: [а: | о: | у:];
"i:и" i:и <=> [V1 | C2 | ь: | д з] :0* %+: _;

```

Prof. Kimmo Koskenniemi-лӧн уджтасьяс йылысь:

AUTOMATA komitwol.aut

43 Characters in the Alphabet (102)

15 Automata Read In: 979 Transitions (9000)

113 Rows (10000) 124 Columns (5000)

59 Aligned Pairs (155)

DICTIONARY komi.lex

30098 Branches (240000)

96 Sublexicons (1000)

9389 Entries (65000)

85177 string characters (120000)

Кык ашӧр кыв, налӧн вежӧртас, артмӧдӧм, английскӧй вы-
лӧ вуджӧдӧм:

вӧв

"<вөв>"

"вөл#" V ENG [be] V CONNEG SG3

"вөл#" V ENG [be] V CONNEG SG2

"вөл#" V ENG [be] V CONNEG SG1

"вөл#" N ENG [horse] NOM SG

вөвтөмтор

"<вөвтөмтор>"

"вөл+төм+торй#" V ENG [be] V CAR1 DA-торй N
[thing] NOM SG

"вөл+төм+торй#" V ENG [be] V CAR1 DN-торй N
[a little] NOM SG

"вөл+төм+торй#" N ENG [horse] V CAR DA-торй N
[thing] NOM SG

"вөл+төм+торй#" N ENG [horse] V CAR DN-торй N
[a little] NOM SG

Мый могысь поэь вөдитчыны татшөм кывчукөрөн?

От-кө, сійөс поэь используйтны кызди коми-английскöй-фин кывкуд.

Мөд-кө, төдан машинаса материал подув вылын поэь лөсьöдны коми кывлысь стöч гижанног петкөдлысь уджтас, мый сетас отсöг газет-журналын уджалысьяслы, гижысьяслы да уна мукöд йöзлы, кодъяслөн эм та вылö коланлун. Оз поэь вунöдны, мый татшөм уджтас лөб лөсьöдöма абу озырмөдчөм вылö, а сөмын коми кыв сөвмөдөм могысь. Мед көть татшөм уджтас дасьтөм вылö поэяс судзöдны отсöг öтувья удж нуöдан көсйöмьяс серти Рытыввыв Европаысь.

Коймөд-кө, кор лексикаö чукөрмас 60.000 кыв да сійö материалыс шымыртас озыр коми кывья классикьясöс, лөб поэянлун лөсьöдны сідз шусяна торйөдлана кыввелөдөм (Constraint Grammar). Кута лача, мый татшөмсяма öтувья удж вöчны көсйöм ловзяс Комиын.

Рöскöдьясö вермас чинтыны небöг-газет лэдзанін да кыв туялан учрежденнö костын кутшөмкө сёрнитчөм кырымалөм. Шуам, төдан машинаö пыртөм газет гижöдьяс да художествоа уджъяс, туялан текстыяс поэь видзны дискетаяс вылын. Дискетаясөн вөдитчөм тшöтш лөб төдчымөн донтөмджык гижöда материал вылысь пырталөм дорысь. Поэь мөвпавны, мый став коми кывья художествоа гижöдыс төрас некымын CD ROM вылö.