

"Korpusta voi hyödyntää hyvin niin kielten tutkimisessa kuin myös kirjakielten kehittämisessä."

Aleksandra Rodionova

# VepKar vaalii vähemmistökieliä

## Hanke

Karjalan ja vepsän kielen tietokanta sisältää tuhansia tekstejä ja kymmeniä puhenäytteitä, rikasta sanastoa ja kielioppiasioita.

### Marina Tolstyh

Karjalan Sanomat  
Petroskoi  
karjalansanomat@yandex.ru



**Venäjän tiedeakatemia** Karjalan tiedekeskuksen kielen, kirjallisuuden ja historian instituutin lyydin kielen tutkija **Aleksandra Rodionova** korostaa, että karjalan ja vepsän kielen VepKar-korpusta täytyy kehittää eteenpäin.

– Korpusta voi hyödyntää hyvin niin kielten tutkimisessa kuin myös kirjakielten kehittämisessä, Rodionova varmentaa.

Hänestä kielikorpukselta on tullut itämerensuomalaisten kielten tutkimusten tietokanta Luoteis-Venäjällä.

Tällä hetkellä korpus sisältää yli 4 000 tekstiä ja 45 äänitettä karjalan ja vepsän kielen murteilla sekä sanakirjoja, murrekartan ja kielten kielioppiin liittyvää tietoa. Lisäksi vepsäläisiä itkuvirsiä sekä karjalaisia runoja ja tarinoita voi lukea korpuksen sivustolta.

– Sivustolle on jo ladattu seitsemän äänitettä lyydin murteella ja kolme vepsän kielellä. Lisäksi 20 äänitettä voi kuunnella varsinaiskarjalan ja 15 livvinkarjalan murteella, tutkija esittää.

Rodionova kertoi VepKar-korpuksen hankkeesta viime viikolla tutkimuskonferenssissa Karjalan tiedekeskuksessa. Vuosittainen konferenssi kantaa Venäjän suomalais-ugrialaisten kielten tutkijan **Dmitri Bubrihin** nimeä.

– Kielen, kirjallisuuden ja historian instituutti on Venäjällä johtava itämerensuomalaisia kieliä tutkiva keskus, Rodionova painottaa.

**» Tällä hetkellä korpus sisältää yli 4 000 tekstiä ja 45 äänitettä karjalan ja vepsän kielen murteilla sekä sanakirjoja, murrekartan ja kielten kielioppiin liittyvää tietoa.**



Venäjän tiedeakatemia Karjalan tiedekeskuksen kielen, kirjallisuuden ja historian instituutin lyydin kielen tutkija Aleksandra Rodionova (vas.) kertoi VepKar-kielikorpuksen kehityksestä viime viikolla Dmitri Bubrihin nimeä kantavassa tutkimuskonferenssissa. KUVA: KARJALAN TIEDEKESKUKSEN VKONTAKTE-SIVU

## 100 puhenäytettä esillä hankkeessa

Lyhyin korpuksessa esillä oleva äänite kestää 40 sekuntia ja pisin melkein viisi minuuttia. Puhetallenteen ohella on myös samalla murteella kirjoitettu tekstiversio ja sen venäjännös. Käyttäjä saa tietoa myös siitä, missä ja milloin puhe nauhoitettiin sekä ketkä olivat haastateltu henkilö ja haastattelija.

Korpukseseen alettiin lisätä äänitteitä tänä vuonna Venäjän tieteellisen rahaston tukemassa hankkeessa. Tämän ja ensi vuoden aikana on tarkoitus laittaa korpukseseen 100 puhenäytettä.

Esimerkiksi sivustolla voi kuunnella puhetta valdainkarjalan murteella. Noin 90-vuotias **Maria Pavlova** Novgorodin alueen Markovon kylästä kertoi vuonna 1990 haastattelijalle siitä, että hänen nuoruudessaan hänen luonaan kävi kosijoita.

– Tämä on viimeinen nauhoite tuolla murteella. Nyt kukaan ei puhu enää valdainkarjalaisten murretta, Rodionova pahoittelee.

Kielen, kirjallisuuden ja historian instituutin tallennearkistossa on paljon ainutlaatuisia nauhoitteita, jotka täytyy muuttaa sähköiseen muotoon.

– Instituutin arkistossa säi-

## VepKar-korpus

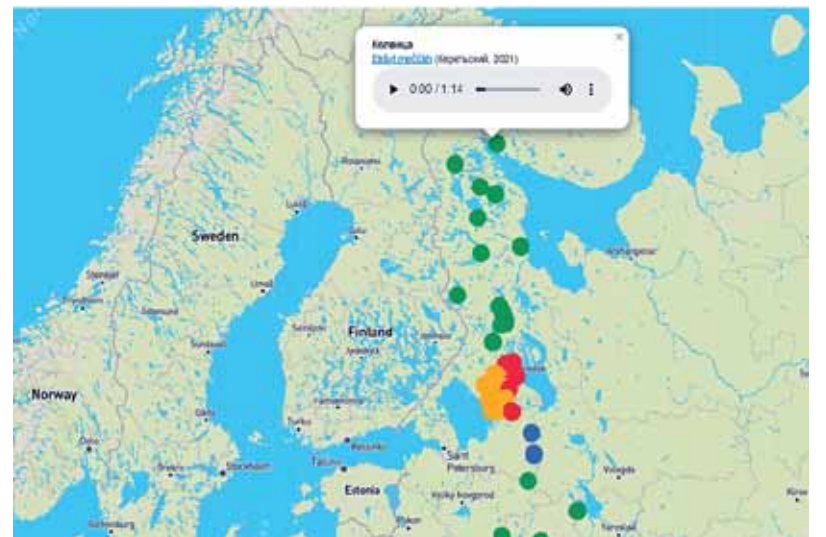
- 4 000 tekstiä karjalan ja vepsän kielen murteilla.
- 45 äänitettä on esillä korpuksessa.
- 100 puhenäytettä ladataan korpukseseen ensi vuoden loppuun mennessä.
- 1 nauhoite on valdainkarjalan murteella, jota ei enää puhuta.

lyy melkein 3 000 karjalankielistä äänitettä. Vepsänkieliset äänitteet kestävät yhteensä noin 400 tuntia. Murresanojen ja ilmaisujen keruu jatkuu yhä, tutkija sanoo. **KS**

## Sähköinen murrekartta on valmis

**Karjalan ja vepsän** tutkijoiden lisäksi korpuksen tekemiseen osallistuu myös Karjalan tiedekeskuksen soveltavan matematiikan instituutin asiantuntijoita. Yhdessä he ovat tehneet ensimmäisen sähköisen murrekartan. Käyttäjän täytyy klikata kartalla pistettä avataksena äänitteen.

– Kaikki korpukseseen ladatut äänitteet ovat saatavilla myös kartalla, lyydin kielen tutkija Aleksandra Rodionova täsmentää.



VepKar-kielikorpukseseen on ladattu murrekartta, josta nyt voi kuunnella 45 lyhyttä puhenäytettä karjalan ja vepsän kielen murteista. KUVA: MURREKARTAN NÄYTTÖKUVA

Kielten tutkijat pyrkivät esittelemään korpuksessa informatiivisimpia ja hyvälaatuisimpia tallenteita. Puhetyytit ovat peräisin viime vuosisadan puolivälistä nykypäiviin.

– Parhaat äänitteet lyydin murteella on korpusta varten jo valikoitu. Edessä on työ varsinais- ja livvinkarjalan murteella tehtyjen nauhoitteiden parissa, tutkija kertoo.

Hänen mukaansa kielikorpuksia tehdään maailmassa kielellisen rik-

kauden säilyttämiseksi ja alkuperäiskansojen kielten oppimiseksi. Esimerkiksi venäjän, brittienglannin ja tšekin kielen korpuksat ovat tunnetuimpia.

– Suomalais-ugrialaisten kansojen keskuudessa tunnetaan Suomen kieli-pankki sekä viron ja unkarin kielen korpuksat. Venäjällä on olemassa udmurtin, marin ja komipermjakin kielen korpuksat, Rodionova sanoo. **KS**