

# Kielipankki ja FIN-CLARIN

Mietta Lennes

Nykykielten laitos, Helsingin yliopisto

[fin-clarin@helsinki.fi](mailto:fin-clarin@helsinki.fi)

**Kielipankki**

[www.kielipankki.fi](http://www.kielipankki.fi)

**FIN-CLARIN**

[www.helsinki.fi/fin-clarin/](http://www.helsinki.fi/fin-clarin/)

# Lyhyt historia

UHLCS (1987)

Yleisen kielitieteen laitoksen korpuspalvelin, HY

Keskitetty virtuaalipalvelu (1996)

[www.kielipankki.fi](http://www.kielipankki.fi)

FIN-CLARIN –organisaatio (2006)

[www.helsinki.fi/fin-clarin/](http://www.helsinki.fi/fin-clarin/)

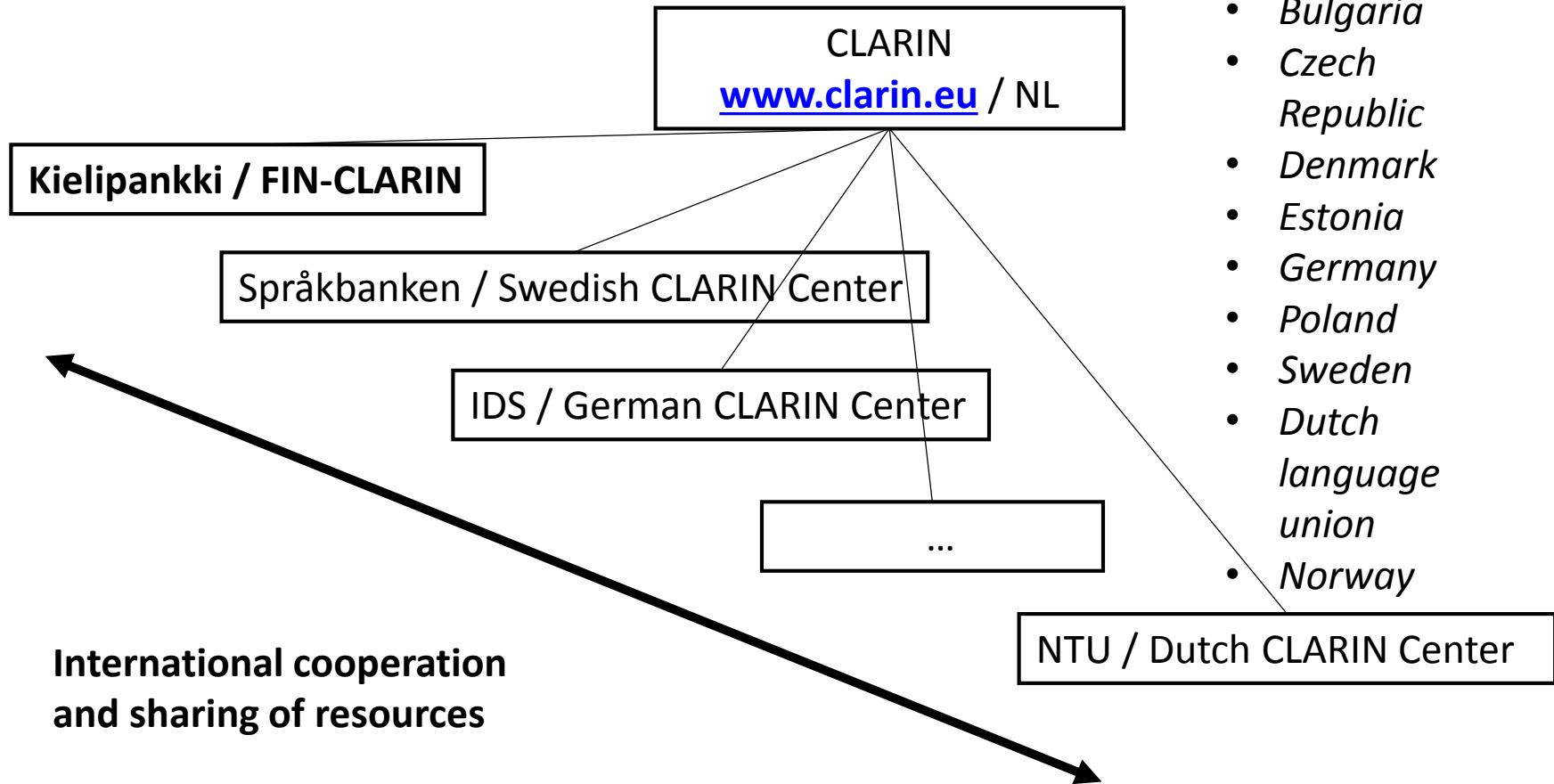
CLARIN ERIC (2012)

[www.clarin.eu](http://www.clarin.eu)

# CLARIN ERIC

European Research Infrastructure Consortium  
founded on February 29, 2012

- ***Netherlands***
- *Austria*
- *Bulgaria*
- *Czech Republic*
- *Denmark*
- *Estonia*
- *Germany*
- *Poland*
- *Sweden*
- *Dutch language union*
- *Norway*



# Missio

Tarjotaan kieliaineistoja tutkiville humanisteille, yhteiskuntatieteilijöille ja tietojenkäsittelytieteilijöille pääsy kielentutkimuksen työkaluihin ja muihin kielivaroihin.

# Visio

Omalla työasemallaan

- **kielentutkijan** pitäisi pystyä hakemaan relevantteja kielellisiä rakenteita ja esimerkkejä
- **yhteiskuntatieteilijän ja historiantutkijan** pitäisi voida löytää dokumentteja sisältöhaakuja käyttämällä
- **tietojenkäsittelytieteilijän** pitäisi pystyä louhimaan tekstiä ja dataa

esimerkiksi käsikirjoitus- ja sanomalehtikokoelmista, Internetin ja eduskunnan keskusteluista tai radio- ja TV-lähetyksistä.

# Strategia

- Tarjotaan **edistynyt sisältöhaku** data-arkistoihin ja dokumenttikokoelmiin
  - Tutkimusyksiköt, sanomalehdet, radio- ja tv-yhtiöt
  - Kansalliskirjasto, Kansallisarkisto, eduskunta
  - Tutkimushankkeet
- Tarjotaan kotimaisia kieliä tukevia **avoimen lähdekoodin työkaluja**
  - tekstin ja puheen automaattiseen tunnistukseen, tekstin jäsentämiseen ja analyysiin, annotointiin, luokitteluun ja tekstinlouhintaan

# Kielipankki

- **Tekijänoikeuden alaiset aineistot, jotka saattavat myös sisältää henkilödataa**
  - PUB: julkinen
  - ACA: akateeminen
  - RES: rajoitettu
  - Lisäehtoja tarvittaessa: BY, PRIV, NC, INF, NORED, ...
- **Hakuliittymät, työryhmäalustat, virtuaalinen työtila**
  - Tekstiaineistot: [korp.csc.fi](http://korp.csc.fi)
    - >5 miljardia sanaa suomea ja >3 miljardia sanaa ruotsia
    - noin 2 miljoonaa dokumenttia
  - puhetta sisältäviä ääni- ja videoaineistoja [lat.csc.fi](http://lat.csc.fi)
    - >300 tuntia suomea ja >100 tuntia ruotsia
    - noin 600 litteroitua haastattelua
  - Kielipankin virtuaalinen työtila esimerkiksi datan louhintaa varten CSC – Tieteen tietotekniikan keskuksen palvelimilla



# Data Collections

Corpora	2014	2020
<b><i>Text</i></b>		
Newspaper corpus 1790- (National Library)	8 Gw	40 Gw
Discussion fora 2000- (Suomi24)		15 Gw
Literature and manuscripts (Gutenberg, National Archive)	35 Mw	70 Mw
<b><i>Speech</i></b>		
Broadcast news (YLE)		10000 h
Finnish Parliament Debates		10000 h
Dialect corpora (Kotus, UTU, SLS)	500 h	1000 h
Sign language corpora (Aalto)		2000 h
<b><i>Multi-lingual</i></b>		
Multilingual corpora (UTA, EuroParl)	70 Mw	150 Mw
Learners' corpora (UO, JyU)	2 Mw	5 Mw
Open-source lexicons and terminologies (UHEL, UTROM)	200 Kw	400 Kw

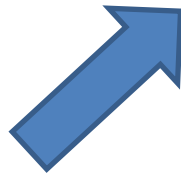
# Language Tools

	Short	Mid	Long
<b><i>User Interfaces</i></b>			
Text corpus interface (UGOT/SE)	x	x	
Speech and video corpus interface (MPI/NL)	x	x	
<b><i>Existing tools</i></b>			
Text processing tools (UHEL/FI, UTU/FI, UGOT/SE)	x	x	
Speech processing tools (AALTO/FI)	x	x	
OCR and Critical edition editor (National Library/FI)	x	x	
Tools for minority languages (UHEL/FI, UTROM/NO)	x	x	x
<b><i>Better tools needed</i></b>			
Hand-writing recognition		x	x
Automated video annotation		x	x
Semantic and pragmatic analysis		x	x

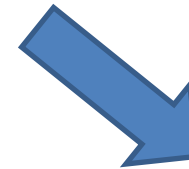
# Kielivarojen “kierrättäminen”

Kielipankki (The Language Bank of Finland)

Kieliaineiston  
automaattinen  
jäsentäminen  
[hfst.sf.org/](http://hfst.sf.org/)



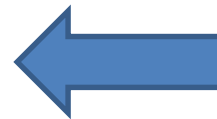
Valinta ja haku  
[korp.csc.fi/](http://korp.csc.fi/)  
[lat.csc.fi/](http://lat.csc.fi/)



Kielivarat

[www.meta-share.org](http://www.meta-share.org)  
[www.clarin.eu/vlo](http://www.clarin.eu/vlo)

Hakutulokset



Tuloksia voidaan jatkoannotoida ja  
tallentaa uudet versiot muiden tutkijoiden saataville



Demo

[korp.csc.fi](http://korp.csc.fi)



129 korpusta valittuina — 4 848 984 674 sanetta

Hakuhistoria

Yksinkertainen Laajene

Etsi

Konkordanssi: osumia sivulla:



Valitse kaikki

Tyhjennä valinnat

- FinnTreeBank: suomen puupankki (3)
- Kansalliskirjaston lehtikokoelman (KLIK) suomenkieliset lehdet (179)
- Kirjallisuutta (4)
- Juridisia tekstejä (3)
- Muita tekstejä (9)
- Puhuttua kieltä (tekstiksi litteroituna) (2)
- Vanhan kirjasuomen korpus (12)
- Varhaisnykysuomen korpus (näytteitä) (3)
- Demo- ja testiaineistoja (1)
- HS.fi [ACA]
- LAS2 [RES]

368 423 332 virkettä valituissa korpuksissa

aset

än perusteella: sana



Yksinkertainen Laaja  
Etsi  
Konkordanssi: osuvia sivulla:

Ei korpuksia valittuna — 0 sanetta

Valitse kaikki  
Tyhjennä valinnat

- FinnTreeBank: suomen puupankki (3)
- Kansalliskirjaston lehtikokoelman (KLIK) suomenkieliset lehdet (179)
- Kirjallisuutta (4)
- Juridisia tekstejä (3)
- Muita tekstejä (9)
- Puhuttua kieltä (tekstiksi litteroituna) (2)
- Vanhan kirjasuomen korpus (12)
- Varhaisnykysuomen korpus (näytteitä) (3)
- Demo- ja testiaineistoja (1)
  - Reitti A-siipeen
  - HS.fi [ACA]
  - LAS2 [RES]

0 virkettä valituissa korpuksissa

**Kansalliskirjaston lehtikokoelman (KLIK) suomenkieliset lehdet**

179 korpusta, joissa:

**5 246 334 710** sanetta  
**392 834 854+** virkettä

Hakuhistoria



Reitti A-siipeen valittu — 218 sanetta

Hakuhistoria

Yksinkertainen Laaj...

Etsi

Konkordanssi: osumia sivulla:



- ✓ Valitse kaikki
- ✗ Tyhjennä valinnat

- FinnTreeBank: suomen puupankki (3)
- Kansalliskirjaston lehtikokoelman (KLK) suomenkieliset lehdet (179)
- Kirjallisuutta (4)
- Juridisia tekstejä (3)
- Muita tekstejä (9)
- Puhuttua kieltä (tekstiksi litteroituna) (2)
- Vanhan kirjasuomen korpus (12)
- Varhaisnykysuomen korpus (näytteitä) (3)
- Demo- ja testiaineistoja (1)

- Reitti A-siipeen
- HS.fi [ACA]
- LAS2 [RES]

34 virkettä valituissa korpuksissa

set

än perusteella: sana

**Reitti A-siipeen**

Kahdenkeskisen videoidun keskustelun "Reitti A-siipeen" yleiskielistetty litteraatti. Keskustelussa selvitetään reittiä tiettyyn Helsingin yliopiston Metsätalossa sijaitsevaan huoneeseen. Vapaasti käytettäväksi tarkoitettu näyteaineisto.

URN: <urn:nbn:fi:lb-100110012817>  
 Kuvailutiedot  
 Lisenssi: CC-ZERO (CC0)

Saneiden määrä: **218**  
 Virkkeiden määrä: **34**  
 Viimeksi päivitetty: **2014-10-14**

Tukee laajennettua kontekstia.



Reitti A-siipeen valittu — 218 sanetta

Yksinkertainen

Laajennettu

Edistynyt

Etsi  Etsi myös  alkuosa  loppuosa ja  samaista pien- ja suuraakkoset

on sananmuoto

olla (verbi)

*olla"-kysymys (substantiivi)*

*olla&seen (substantiivi)*

*olla&n (adjektiivi)*

*olla# (muu)*

*olla (muu)*

*olla (perusluku)*

*olla (substantiivi)*

*olla (adjektiivi)*

*olla (adverbi)*

Konkordanssi  järjestä korpuksen sisällä: järjestämätön

Tilastoja: laske tilastot tämän perusteella: sana



Yksinkertainen

Laajennettu

Edistynyt

 Etsi 

 Etsi myös  alkuosa  loppuosa ja  samaista pien- ja suuraakkoset

 Konkordanssi: 


 Tilastoja: 

Konkordanssi

Tilastoja

Sanakuva

Tuloksia: 12

Näytä konteksti

REITTI A-SIPEEN

Tänne sinun huoneeseesi	<b>on</b>	hirvittävän vaikea löytää.
No niin varmaan	<b>onkin</b>	.
Minä haluan tulla sisään mahdollisimman nopeasti, minulla	<b>on</b>	kiire, mitä minä...
Fabianinkatu ei	<b>ole</b>	tuolla, kun se on tuolla puolella, jos sinä tulet niin kuin pääove
Fabianinkatu ei ole tuolla, kun se	<b>on</b>	tuolla puolella, jos sinä tulet niin kuin pääovesta.
Eli okei, minä	<b>olen</b>	tuolla puolella.
Sitten, mitä siinä	<b>on</b>	, ne portaat, joissa on ne ikkunat?
Sitten, mitä siinä on, ne portaat, joissa	<b>on</b>	ne ikkunat?
Silloin kun sinä sieltä ovesta tulet, niin sinä	<b>olet</b>	, sinun pitää mennä vasempaan, koska sinä et pääse läpi muut
Okei, eli b-siipi	<b>on</b>	tuolla, ja nyt me olemme a-siivessä?
Okei, eli b-siipi on tuolla, ja nyt me	<b>olemme</b>	a-siivessä?
Nythän se	<b>on</b>	ihän täysin selvää!

 Lataa tiedostona muodossa: [Annot](#) [Ref](#) [Nooj](#)
[JSON](#)

## Korpus

Reitti A-siipeen

URN: urn:nbn:fi:lb-100110012817

Kuvailutiedot

Lisenssi: CC-ZERO (CC0)

## tekstin ominaisuudet

otsikko: Reitti A-siipeen  
 puhunnoksen kesto: 2063  
 puhunnoksen alkuaika: 2879  
 puhunnoksen loppuaika: 4942  
 tekijä: Tero Aalto, Mietta Lennes  
 puhunnoksen numero: 3  
 julkaisuvuosi: 2012  
 puhuja: TA

## sanan ominaisuudet

sanaluokka: verbi  
 perusmuoto: olla  
 perusmuoto (yhdyssanarajat): olla  
 puhemuoto: o  
 morfologinen analyysi:  
 PRS\_Sg3|VOICE\_Act|TENSE\_Prs|MOOD\_Ind  
 dependenssisuhde: kopula

[Näytä video](#)
[Näytä dependenssipuu](#)

## REITTI A-SIIPEN

oneeseesi	<b>on</b>	hirvittävän vaikea löytää.
n varmaan	<b>onkin</b>	.
ti, minulla	<b>on</b>	kiire, mitä minä...
ninkatu ei	<b>ole</b>	tuolla, kun se on tuolla puolella, jos sinä tulet niin kuin pääove
lla, kun se	<b>on</b>	tuolla puolella, jos sinä tulet niin kuin pääovesta.
okei, minä	<b>olen</b>	tuolla puolella.
mitä siinä	<b>on</b>	, ne portaat, joissa on ne ikkunat?
aat, joissa	<b>on</b>	ne ikkunat?
t, niin sinä	<b>olet</b>	, sinun pitää mennä vasempaan, koska sinä et pääse läpi muut
, eli b-siipi	<b>on</b>	tuolla, ja nyt me olemme a-siivessä?
ja nyt me	<b>olemme</b>	a-siivessä?
Nyhdän se	<b>on</b>	ihan täysin selvää!

Lataa tiedostona muodossa: [Annot](#) [Ref](#) [Noo](#)[JSON](#)**Korpus**

Reitti A-siipeen

URN: urn:nbn:fi:lb-100110012817

Kuvailutiedot

Lisenssi: CC-ZERO (CC0)

**tekstin ominaisuudet**

otsikko: Reitti A-siipeen  
 puhunnoksen kesto: 2063  
 puhunnoksen alkuaika: 2879  
 puhunnoksen loppuaika: 4942  
 tekijä: Tero Aalto, Mieta Lennes  
 puhunnoksen numero: 3  
 julkaisuvuosi: 2012  
 puhuja: TA

**sanan ominaisuudet**

sanaluokka: verbi  
 perusmuoto: olla  
 perusmuoto (yhdyssanarajat): olla  
 puhemuoto: o  
 morfologinen analyysi:  
 PRS\_Sg3|VOICE\_Act|TENSE\_Prs|MOOD\_Ind  
 dependenssisuhde: kopula

[Näytä video](#)[Näytä dependenssipuu](#)

- Text
- Grid
- Subtitle
- Waveform
- Timeline**
- Combined

### Video display



0:03 / 1:04

Full Buffer

### Information

General Session Technical

Resource: reitti\_a-siipeen.eaf  
 Media file: reitti\_a-siipeen.mp4  
 Elapsed time: 00:00:02:879

Selected chunk:  
 Begin time: 00:00:02:879  
 End time: 00:00:04:942  
 Text: -

### Mini Data Frame

Tier: none

Font size: 14

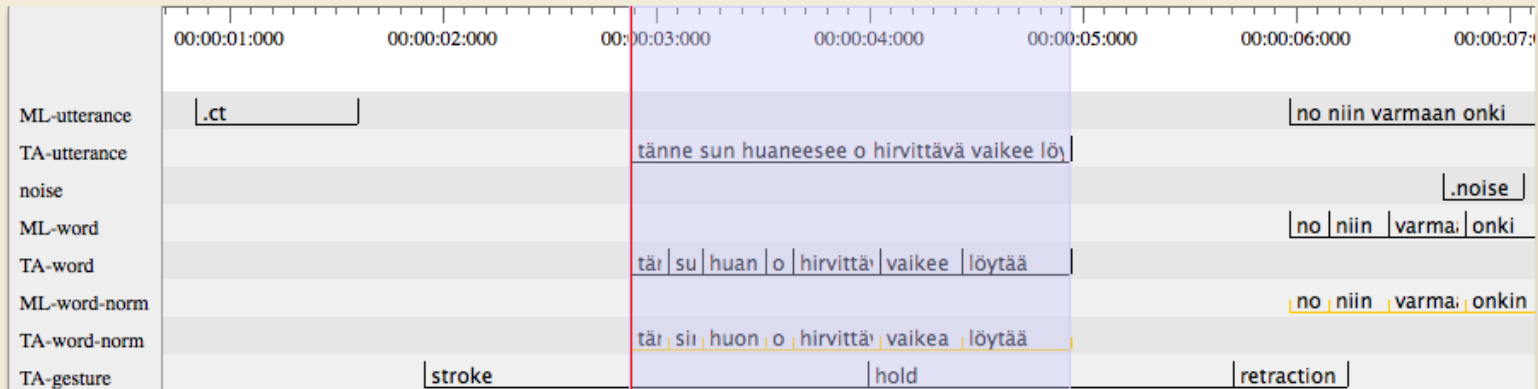
- Play selection
- Clear selection
- Create bookmark

- |< >|
- << >>
- < >
- + -

- Play screen by screen
- Play continually

Tier text font:  
 Lucida Grande

### Timeline



Category	Content
ML-utterance	.ct
TA-utterance	tänne sun huaneesee o hirvittävä vaikee löy
noise	.noise
ML-word	no   niin   varma   onki
TA-word	tär   su   huan   o   hirvittä   vaikee   löytää
ML-word-norm	no   niin   varma   onkin
TA-word-norm	tär   sii   huon   o   hirvittä   vaikea   löytää
TA-gesture	stroke   hold   retraction



Text

Grid

Subtitle

Waveform

Timeline

Combined

Play selection

Clear selection

Create bookmark

&lt;

&gt;

&lt;&lt;

&gt;&gt;

&lt;

&gt;

+

-

 Play screen by screen

Video display

min

Information

min

Mini Data Fram

## Select features to bookmark

Selected region:  00:00:02:879-00:00:04:942Current view:  timelineSelected tiers:  Tier 1: none

Tier 2: none

Tier 3: none

Tier 4: none

Selected media: reitti\_a-silpeen.mp4 (sole choice)

Copy to clipboard

Done

Session

Technical

e: reitti\_a-silpeen.eaf

le: reitti\_a-silpeen.mp4

time: 00:00:02:879

chunk:

time: 00:00:02:879

me: 00:00:04:942

Tier: none

Font size: 14

00:00:04:000

00:00:05:000

un huaneesee o hirvittävä valkee löy

huan | o | hirvittä- | valkee | löytää

tä: sii huon o hirvittä- valkea löytää

|stroke

|hold

<https://lat.csc.fi/ds/annex/runLoader?nodeid=MPI2817%23&time=2879&duration=2063&viewType=timeline>

Yksinkertainen Laajennettu Edistynyt

Etsi olla (verbi) Etsi myös alkuosa loppuosa ja samaista pien- ja suuraakkoset

Konkordanssi: osumia sivulla: 25 järjestä korpuksen sisällä: järjestämätön Tilastoja: laske tilastot tämän perusteella: sana

Konkordanssi Tilastoja Sanakuva

Tuloksia: 12

Näytä konteksti

REITTI A-SIPEEN

Tänne sinun huoneeseesi	<b>on</b>	hirvittävän vaikea löytää.
No niin varmaan	<b>onkin</b>	.
Voitko sinä neuvoa minulle, jos minä seison tuolla Fabianinkadulla ja minä haluan tulla sisään mahdollisimman nopeasti, minulla	<b>on</b>	kiire, mitä minä...
Fabianinkatu ei	<b>ole</b>	tuolla, kun se on tuolla puolella, jos sinä tulet niin kuin pääovesta.
Fabianinkatu ei ole tuolla, kun se	<b>on</b>	tuolla puolella, jos sinä tulet niin kuin pääovesta.
Eli okei, minä	<b>olen</b>	tuolla puolella.
Sitten, mitä siinä	<b>on</b>	, ne portaat, joissa on ne ikkunat?
Sitten, mitä siinä on, ne portaat, joissa	<b>on</b>	ne ikkunat?
Silloin kun sinä sieltä ovesta tulet, niin sinä	<b>olet</b>	, sinun pitää mennä vasempaan, koska sinä et pääse läpi muuten.
Okei, eli b-siipi	<b>on</b>	tuolla, ja nyt me olemme a-siivessä?
Okei, eli b-siipi on tuolla, ja nyt me	<b>olemme</b>	a-siivessä?
Nythän se	<b>on</b>	ihän täysin selvää!

Lataa tiedostona muodossa: [Annot](#) [Ref](#) [Nooj](#)

JSON

Korpus

Reitti A-siipeen

URN: urn:nbn:fi:lb-100110012817

Kuvailutiedot

Lisenssi: CC-ZERO (CC0)

tekstin ominaisuudet

otsikko: Reitti A-siipeen  
puhunnoksen kesto: 3217  
puhunnoksen alkuaika: 30902  
puhunnoksen loppuaika: 34119  
tekijä: Tero Aalto, Mietta Lennes  
puhunnoksen numero: 25  
julkaisuvuosi: 2012  
puhujat: ML

sanan ominaisuudet

sanaluokka: verbi  
perusmuoto: olla  
perusmuoto (yhdyssanarajat): olla  
puhemuoto: oot  
morfologinen analyysi:  
PRS\_Sg2|VOICE\_Act|TENSE\_Prs|MOOD\_Ind  
dependenssisuhde: rinnastus

Näytä video

Näytä dependenssipuu

## META SHARE

entation

Statistics

Route Demo 12  4

View resource name in all available languages

## Reittidemo

<http://urn.fi/urn:nbn:fi:lb-100110012813>ID: <http://urn.fi/urn:nbn:fi:lb-2014101401>

A freely available public demo corpus. Contains a video conversation between two people who are discussing the route to a specific room in Metsätalo at the University of Helsinki, Finland. Link to the corpus: <http://urn.fi/urn:nbn:fi:lb-100110012813>. Link to this metadata file: <http://urn.fi/urn:nbn:fi:lb-2014101401>.

View resource description in all available languages

Vapaasti käytettäväksi tarkoitettu näyteaineisto. Sisältää kahdenkeskisen videoidun keskustelun "Reitti A-siipeen", jossa selvitetään reittiä tiettyyn Helsingin yliopiston Metsätalossa sijaitsevaan huoneeseen. Näyteaineisto löytyy osoitteesta <http://urn.fi/urn:nbn:fi:lb-100110012813>. Tämän tietueen osoite on <http://urn.fi/urn:nbn:fi:lb-2014101401>.

&lt;&lt; Back

Download

Edit Resource

## Distribution

## Availability



Available - Unrestricted Use

## Licence

CC - ZERO

Download location: *hidden*

Distribution Access/Medium: Accessible Through Interface, Downloadable

Execution location: *hidden*text audio video 

## Monolingual text corpus

## Languages

Finnish

## Linguality

Linguality type: Monolingual

## Size

64 Seconds

## Modalities

## Resource Creation

## Resource Creator

Mietta Lennes Tero Aalto Jussi Piitulainen 

## Metadata

Created: 10/14/2014

Last Updated: 10/14/2014

Metadata Language: English (en)

## Distribution

### Availability

Available - Unrestricted Use

### Licence

CC - ZERO

Download location: *hidden*

Distribution Access/Medium: Accessible

Through Interface, Downloadable

Execution location: *hidden*

Licensors:

[Tero Aalto](#)

[Mietta Lennes](#)

Distribution rights holders:

[FIN-CLARIN](#)

### IPR Holder

[Jussi Piitulainen](#)

[Mietta Lennes](#)

[Tero Aalto](#)

## Contact Person

[User support FIN-CLARIN](#)

## Monolingual text corpus

### Languages

Finnish

### Linguality

Linguality type: Monolingual

### Size

64 Seconds

### Modalities

Body Gesture, Spoken Language

## Monolingual audio corpus

### Languages

Finnish

### Linguality

Linguality type: Monolingual

### Size

64 Seconds

### Modalities

Spoken Language, Voice

### Audio Formats

## Resource Creation

### Resource Creator

[Mietta Lennes](#)

[Tero Aalto](#)

[Jussi Piitulainen](#)

## Metadata

Created: 10/14/2014

Last Updated: 10/14/2014

Metadata Language: English (en)

### Metadata Creator

[Mietta Lennes](#)

[Imre Bartis](#)

## Usage

### Foreseen Use

Human Use

Use NLP Specific: Linguistic Research

### Actual Use - Human Use

Use NLP Specific: Linguistic Research



- Latsver
- CLARIN-Terms\_of\_Service.html
- Lat\_Privacy\_Policy.html
- WelcomeToLat.html
- Academic use
- Demo
- PeWi corpus
- Reittidemo
  - Reitti\_A-siipeen
    - reitti\_a-siipeen.aif
    - reitti\_a-siipeen.mp4
    - reitti\_a-siipeen.wav
- Private
- Public
- Restricted use
- Testing



# The Language Bank of Finland

## Language Archive Tools

- Click to open a node.
- Right-click any node to open menu.
- Information page.
- This node is open.
- Log in to access this node.
- Request access to access this node.

[LAT instructions](#)

Korp x IMDI Browser x +

https://lat.csc.fi/ds/imdi\_browser/?openpath=MPI2817%23

IMDI-Browser

show accessibility of resources about manual register user: anonymous login logout

- Latsver
  - CLARIN-Terms\_of\_Service.html
  - Lat\_Privacy\_Policy.html
  - WelcomeToLat.html
- Academic use
- Demo
  - PeWi corpus
  - Reittidemo
    - Reitti A-siipeen
      - reitti\_a-siipeen.eaf
      - reitti\_a-siipeen.mp4
      - reitti\_a-siipeen.wav
- Private
- Public
- Restricted use
- Testing

annotation content search request resource access manage access rights

View Node Create bookmark Download Version info

### Resource Information

URL: [https://lat.csc.fi/corpora/demo/reittidemo/Annotations/reitti\\_a-siipeen.eaf](https://lat.csc.fi/corpora/demo/reittidemo/Annotations/reitti_a-siipeen.eaf)  
PID (URN): <urn:nbn:fi:lb-100110012817>  
Internal Node ID: MPI2817#  
Node type: Written Resource  
Format: text/x-eaf+xml  
File size: 123 KB  
Last modified: Mon May 14 18:10:12 EEST 2012  
MD5 Check sum: unknown

### Access

You can access the resource itself (if you have the privileges), by clicking on the the selected resource icon with the right mouse-button and selecting the "view" option or use your browser (plugin) via this [link](#). You will need to authenticate yourself (again) for downloading the resource when it is protected.

Currently accessible to user anonymous: yes

### Applicable License Agreement(s)

Please find the applicable licenses for this node at the beginning of this node and its preceding parent nodes.

For more information about this resource see the metadata in the containing session.

### Rights

All resources in this archive are Copyright © of their respective holders. Re-distribution of the content in any form, by any means, without the prior permission of the copyright holder is prohibited, unless explicitly stated otherwise in a license. See the metadata descriptions of the resources for contact information or contact [ling@csc.fi](mailto:ling@csc.fi).

### Versioning Information

Node	Date	Accessibility
<a href="#">MPI2817#</a>	Current	true

Yksinkertainen

Laajennettu

Edistynyt

Etsi olla (verbi)

 Etsi myös  alkuosa  loppuosa ja  samaista pien- ja suuraakkoset

Konkordanssi: osumia sivulla: 25

järjestä korpuksen sisällä: järjestämätön

Tilastoja: laske tilastot tämän perusteella: sana

Konkordanssi

Tilastoja

Sanakuva

Tuloksia: 12

Näytä konteksti

REITTI A-SIPEEN

Tänne sinun huoneeseesi	<b>on</b>	hirvittävän vaikea löytää.
No niin varmaan	<b>onkin</b>	.
Minä haluan tulla sisään mahdollisimman nopeasti, minulla	<b>on</b>	kiire, mitä minä...
Fabianinkatu ei	<b>ole</b>	tuolla, kun se on tuolla puolella, jos sinä tulet niin kuin pääove
Fabianinkatu ei ole tuolla, kun se	<b>on</b>	tuolla puolella, jos sinä tulet niin kuin pääovesta.
Eli okei, minä	<b>olen</b>	tuolla puolella.
Sitten, mitä siinä	<b>on</b>	, ne portaat, joissa on ne ikkunat?
Sitten, mitä siinä on, ne portaat, joissa	<b>on</b>	ne ikkunat?
Silloin kun sinä sieltä ovesta tulet, niin sinä	<b>olet</b>	, sinun pitää mennä vasempaan, koska sinä et pääse läpi muut
Okei, eli b-siipi	<b>on</b>	tuolla, ja nyt me olemme a-siivessä?
Okei, eli b-siipi on tuolla, ja nyt me	<b>olemme</b>	a-siivessä?
Nythän se	<b>on</b>	ihän täysin selvää!

 Lataa tiedostona muodossa: [Annot](#) [Ref](#) [Nooj](#)
[JSON](#)

## Korpus

Reitti A-siipeen

URN: urn:nbn:fi:lb-100110012817

Kuvailutiedot

Lisenssi: CC-ZERO (CC0)

## tekstin ominaisuudet

otsikko: Reitti A-siipeen  
 puhunnoksen kesto: 2063  
 puhunnoksen alkuaika: 2879  
 puhunnoksen loppuaika: 4942  
 tekijä: Tero Aalto, Mietta Lennes  
 puhunnoksen numero: 3  
 julkaisuvuosi: 2012  
 puhuja: TA

## sanan ominaisuudet

sanaluokka: verbi  
 perusmuoto: olla  
 perusmuoto (yhdyssanarajat): olla  
 puhemuoto: o  
 morfologinen analyysi:  
 PRS\_Sg3|VOICE\_Act|TENSE\_Prs|MOOD\_Ind  
 dependenssisuhde: kopula

[Näytä video](#)
[Näytä dependenssipuu](#)

Yksinkertainen

Laajennettu

Edistynyt

 Etsi 

 Etsi myös  alkuosa  loppuosa ja  samaista pien- ja suuraakkoset

 Konkordanssi: 


 Tilastoja: 

Konkordanssi

Tilastoja

Sanakuva

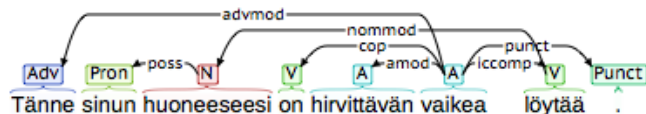
Tuloksia: 12

Näytä konteksti

REITTI A-SIIPEN

Tänne sinun huoneeseesi	<b>on</b>	hivittävän vaikea löytää.
No niin varmaan	<b>onkin</b>	.
sinä haluan tulla sisään mahdollisimman nopeasti, minulla	<b>on</b>	kiire,
Fabianinkatu ei	<b>ole</b>	tuolla
Fabianinkatu ei ole tuolla, kun se	<b>on</b>	tuolla
Eli okei, minä	<b>olen</b>	tuolla
Sitten, mitä siinä	<b>on</b>	, ne po
Sitten, mitä siinä on, ne portaat, joissa	<b>on</b>	ne ikk
Silloin kun sinä sieltä ovesta tulet, niin sinä	<b>olet</b>	, sinun
Okei, eli b-siipi	<b>on</b>	tuolla
Okei, eli b-siipi on tuolla, ja nyt me	<b>olemme</b>	a-siiv
Nythän se	<b>on</b>	ihan t

## Dependenssipuu



## Korpus

Reitti A-siipeen

URN: urn:nbn:fi:lb-100110012817

Kuvailutiedot

Lisenssi: CC-ZERO (CC0)

tekstin ominaisuudet

 puhemuoto: o  
 morfologinen analyysi:  
 PRS\_Sg3|VOICE\_Act|TENSE\_Prs|MOOD\_Ind  
 dependenssisuhde: kopula

Näytä video

Näytä dependenssipuu

Yksinkertainen Laajennettu Edistynyt

Etsi olla (verbi) Etsi myös alkuosa loppuosa samaista pien- ja suuraakkoset

Konkordanssi: osumia sivulla: 25 järjestä korpuksen sisällä: järjestämätön Tilastoja: laske tilastot tämän perusteella: sana

Konkordanssi Tilastoja Sanakuva

Tuloksia: 12

Näytä konteksti

REITTI A-SIIPÉEN

Tänne sinun huoneeseesi	<b>on</b>	hirvittävän vaikea löytää.
No niin varmaan	<b>onkin</b>	.
Voitko sinä neuvoa minulle, jos minä seison tuolla Fabianinkadulla ja minä haluan tulla sisään mahdollisimman nopeasti, minulla	<b>on</b>	kiire, mitä minä...
Fabianinkatu ei	<b>ole</b>	tuolla, kun se on tuolla puolella, jos sinä tulet niin kuin pääovesta.
Fabianinkatu ei ole tuolla, kun se	<b>on</b>	tuolla puolella, jos sinä tulet niin kuin pääovesta.
Eli okei, minä	<b>olen</b>	tuolla puolella.
Sitten, mitä siinä	<b>on</b>	, ne portaat, joissa on ne ikkunat?
Sitten, mitä siinä on, ne portaat, joissa	<b>on</b>	ne ikkunat?
Silloin kun sinä sieltä ovesta tulet, niin sinä	<b>olet</b>	, sinun pitää mennä vasempaan, koska sinä et pääse läpi muuten.
Okei, eli b-siipi	<b>on</b>	tuolla, ja nyt me olemme a-siivessä?
Okei, eli b-siipi on tuolla, ja nyt me	<b>olemme</b>	a-siivessä?
Nythän se	<b>on</b>	ihän täysin selvää!

Lataa tiedostona muodossa: **Annot Ref Nooj** JSON

**Korpus**  
Reitti A-siipeen

URN: urn:nbn:fi:lb-100110012817  
Kuvailutiedot  
Lisenssi: CC-ZERO (CC0)

**tekstin ominaisuudet**  
otsikko: Reitti A-siipeen  
puhunnoksen kesto: 3217  
puhunnoksen alkuaika: 30902  
puhunnoksen loppuaika: 34119  
tekijä: Tero Aalto, Mietta Lennes  
puhunnoksen numero: 25  
julkaisuvuosi: 2012  
puhujat: ML

**sanan ominaisuudet**  
sanaluokka: verbi  
perusmuoto: olla  
perusmuoto (yhdyssanarajat): olla  
puhumuoto: oot  
morfologinen analyysi:  
PRS\_Sg2|VOICE\_Act|TENSE\_Prs|MOOD\_Ind  
dependenssisuhde: rinnastus

Näytä video  
Näytä dependenssioppu

A1 fx ## Korp search results

1 ## Korp search results  
 2  
 3 ## Date: 2014-11-10 00:48:25  
 4 ## Korp URL: https://korp.csc.fi/#corpus=reitidemo&search=lemgram%7Colla..vb.1&page=0  
 5 ## Query parameters:  
 6 ## corpora: REITIDEMO  
 7 ## CQP query: [(lex contains "olla..vb.1")]  
 8 ## context: 1 sentence  
 9 ## within: sentence  
 10 ## sorting:  
 11 ## start: 0  
 12 ## end: 24  
 13 ## Total hits: 12

match	word	lemma	lemmaco	pos	msd	deprel	deprel	ref	spoken
# REITIDEMO (URN urn:nbn:fi:lb-100110012817; licence CC-ZERO (CC0): http://creativecommons.org/publicdomain/zero/1.0/legalcode.txt; metadata http://um.fi/um:nbn:fi:lb-2014101401): sentence 1, position 3; text attributes: text_author: Tero Aalto, Mieta Lennes; t									
	Tanne	tänne	tänne	Adv	CASECHA6	advmod	1	1	tänne
	sinun	sinä	sinä	Pron	SUBCAT_13	poss	2	2	sun
	huoneeseen	huone	huone	N	NUM_Sg C7	nommod	3	3	huoneeseen
***	on	olla	olla	V	PRS_Sg3 6	cop	4	4	o
	hirvittävän	hirvittävä	hirvittävä	A	NUM_Sg C6	amod	5	5	hirvittävä
	vaikea	vaikea	vaikea	A	NUM_Sg C0	ROOT	6	6	vaikee
	löytää	löytää	löytää	V	NUM_Sg C6	iccomp	7	7	löytää
	.	.	.	Punct	_	punct	8	8	.
# REITIDEMO (URN urn:nbn:fi:lb-100110012817; licence CC-ZERO (CC0): http://creativecommons.org/publicdomain/zero/1.0/legalcode.txt; metadata http://um.fi/um:nbn:fi:lb-2014101401): sentence 2, position 11; text attributes: text_author: Tero Aalto, Mieta Lennes; t									
	No	no	no	Interj	CASECHA4	intj	1	1	no
	niin	niin	niin	Adv	_	advmod	2	2	niin
	varmaan	varmaan	varmaan	Adv	_	advmod	3	3	varmaan
***	onkin	olla	olla	V	PRS_Sg3 0	ROOT	4	4	onki
	.	.	.	Punct	_	punct	5	5	.
# REITIDEMO (URN urn:nbn:fi:lb-100110012817; licence CC-ZERO (CC0): http://creativecommons.org/publicdomain/zero/1.0/legalcode.txt; metadata http://um.fi/um:nbn:fi:lb-2014101401): sentence 3, position 32; text attributes: text_author: Tero Aalto, Mieta Lennes; t									
	Voitko	voida	voida	V	PRS_Sg2 0	ROOT	1	1	voit
	sinä	sinä	sinä	Pron	SUBCAT_11	nsubj	2	2	sä
	neuvoa	neuvo	neuvo	N	NUM_Sg C1	doj	3	3	neuvoo
	minulle	minä	minä	Pron	SUBCAT_11	nommod	4	4	mulle
	,	,	,	Punct	_	punct	5	5	,
	jos	jos	jos	C	SUBCAT_18	mark	6	6	jos
	minä	minä	minä	Pron	SUBCAT_18	nsubj	7	7	mä
	seison	seisoa	seisoa	V	PRS_Sg 1	ccomp	8	8	seison
	tuolla	tuolla	tuolla	Adv	_	advmod	9	9	tual
	Fabianinkä	Fabianinka	Fabianinka	N	NUM_Sg C8	nommod	10	10	fabianinkadulla
	ja	ja	ja	C	SUBCAT_18	cc	11	11	ja
	minä	minä	minä	Pron	SUBCAT_113	nsubj	12	12	mä
	haluan	haluta	haluta	V	PRS_Sg 18	conj	13	13	haluun

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
	## Korp search results																						
1	## Korp search results																						
2	## Date	2014-11-10 00:54:17																					
3	## Korp URL	https://korp.csc.fi/#corpus=reittidemo&search=lemgram%7Colla..vb.1&page=0																					
4	## Query parameters																						
5	## corpora	REITTIDEMO																					
6	## CQP query	[(lex contains "olla..vb.1")]																					
7	## context	1 sentence																					
8	## within	sentence																					
9	## sorting																						
10	## start	0																					
11	## end	24																					
12	## Total hits	12																					
13																							
14	sentence	Tänne sinun huoneeseesi <<< on >>> hirvittävän vaikea löytää .																					
15	corpus	REITTIDEMO																					
16	URN	urn:nbn:fi:lb-100110012817																					
17	licence	CC-ZERO (CC0)																					
18	licence link	http://creativecommons.org/publicdomain/zero/1.0/legalcode.txt																					
19	metadata link	http://um.fi/urn:nbn:fi:lb-2014101401																					
20	text_author	Tero Aalto, Mietta Lennes																					
21	text_title	Reitti A-siipeen																					
22	text_year	2012																					
23	paragraph_id	1																					
24	sentence_id	1																					
25	utterance_id	3																					
26	utterance_participant	TA																					
27	utterance_begin_time	2879																					
28	utterance_end_time	4942																					
29	utterance_duration	2063																					
30	utterance_annex_link	nodeid=MPI2817%23&time=2879&duration=2063																					
31																							
32	sentence	No niin varmaan <<< onkin >>> .																					
33	corpus	REITTIDEMO																					
34	URN	urn:nbn:fi:lb-100110012817																					
35	licence	CC-ZERO (CC0)																					
36	licence link	http://creativecommons.org/publicdomain/zero/1.0/legalcode.txt																					
37	metadata link	http://um.fi/urn:nbn:fi:lb-2014101401																					
38	text_author	Tero Aalto, Mietta Lennes																					
39	text_title	Reitti A-siipeen																					
40	text_year	2012																					
41	paragraph_id	2																					
42	sentence_id	2																					
43	utterance_id	4																					
44	utterance_participant	ML																					
45	utterance_begin_time	5970																					