HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Finnish hydronymic constructions
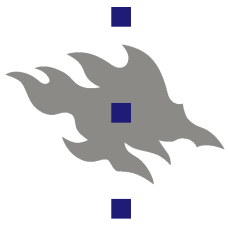
Antti Leino <antti.leino@cs.helsinki.fi>

26th August 2005

Department of Computer Science
Department of Finnish

Research Institute for the Languages of Finland

Background

- Before c. 1970, typologies of Finnish toponyms based mostly on meaning

- Since then, structural analysis of toponyms has been relatively stable

- Typology based on criteria like

  - number of elements

  - inductive vs. 'original'

  - epexegesis, ellipsis

- **Naming patterns** as per Šrámek et al.

  - Ausgangsstellungsmodell: semantic content

  - Wortbildende Modell: syntactic structure
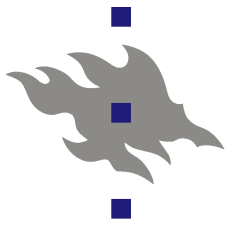
What's new?

- What has changed in the last three decades?

- Computers

  - Electronic corpora allow searches that were too cumbersome with paper files

  - Exploratory data analysis provides methods suitable for such corpora

- Cognitive linguistics

  - Shows some promise for integrating onomastics with mainstream linguistics

 - Explains toponyms at least as well as the traditional approach
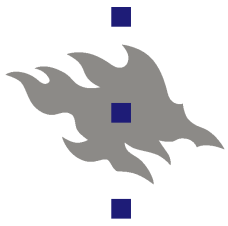
Finnish lake names

- Database of the National Land Survey

- Names that appear on the 1:20 000 Basic Map

	Places	Names
All Finnish names	58 267	25 178
≥ 5 occurrences	29 170	1 492
≥ 20 occurrences	19 230	331
≥ 50 occurrences	12 580	111

- Prior work:

  - Some computer science to get pairs of names that are attracted to each other

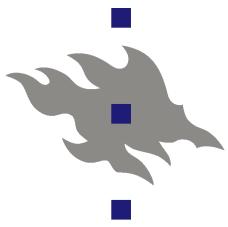 - Interpretation in terms of Construction Grammar
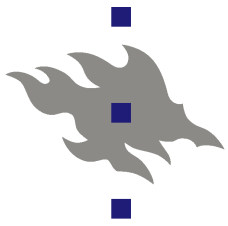
Cognitive linguistics

- No fundamental distinctions syntax~semantics or grammar~lexicon

- A linguistic theory should cope with peripheral phenomena ⇒ toponyms are a good test case

- This work mostly based on Radical Construction Grammar

  - Language is a collection of **constructions**: patterns that join form and meaning

  - Typological / taxonomic approach: a construction is a **generalisation** of more specific linguistic units that are **similar**

  - No syntactic relations: instead semantic and symbolic relations within a construction
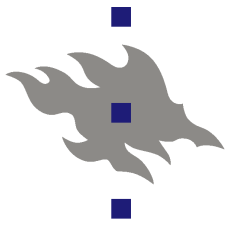
Some refining

- Clustering approach to constructions: they can be viewed as an area around a prototype

- The borders of such an area are blurry

- No sharp division between a schematic construction and a specific construct

  - Any actual utterance can act as a prototype

  - The area around such a prototype is very small and the borders quite sharp, so this is generally quite rare and requires that the new construct is very similar to the old one

  - This is more common with toponyms than in everyday language use

Very crude typology

- Stand-alone names
  - Form does not require the presence of another toponym
  - eg. *Mustalampi* 'Black Lake'
- Inductive names
  - Apparently derived from another name
  - eg. *Pieni Haukilampi* 'Lesser Pike Lake'

	Places		Names	
	Number	%	Number	%
Stand-alone	48 889	84	17 915	71
Inductive	9 378	16	7 263	29
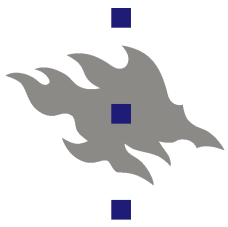
Typical stand-alone name

- Most common construction: identifying element followed by type of place

    - Adjective — mostly a notable feature of the lake

    - Noun — often related to the use, shape or near-by feature

        – Noun in genitive case — often, but by no means always personal names or references to a near-by place

    - Verb stem — usually related to the use of the lake
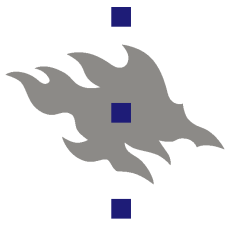
The *identifier* + *type of place* **construction**

lake name	
ROLE identifier SEM descriptive feature	ROLE classifier SEM type of place

Mustalampi

musta ROLE identifier SEM 'black'	*lampi* ROLE classifier SEM 'pond'

Ahvenlampi

ahven ROLE identifier SEM 'perch'	*lampi* ROLE classifier SEM 'pond'

Likolampi

liko- ROLE identifier SEM 'retting'	*lampi* ROLE classifier SEM 'pond'

Ukonlampi

| *ukon* ROLE identifier CASE genitive LXM *ukko* SEM 'old man's'	*lampi* ROLE classifier SEM 'pond'
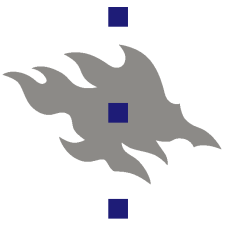
Suffix instead of compounding

- Less common, but still noticeable: identifying element followed by a derivational suffix

  - The identifying can be adjective or noun

  - Diachronically, the *-nen* names are mostly contractions: **Valkeajärvi > Valkeinen*

  - The *-kkV* names are rare and largely opaque; those should perhaps not be classified here

Valkeinen	
valkea ROLE identifier SEM 'white'	*-nen* ROLE classifier

Suolikko	
suoli ROLE identifier SEM 'intestine'	*-kko* ROLE classifier
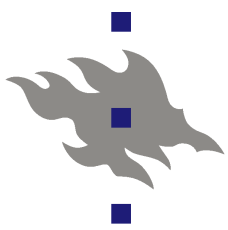
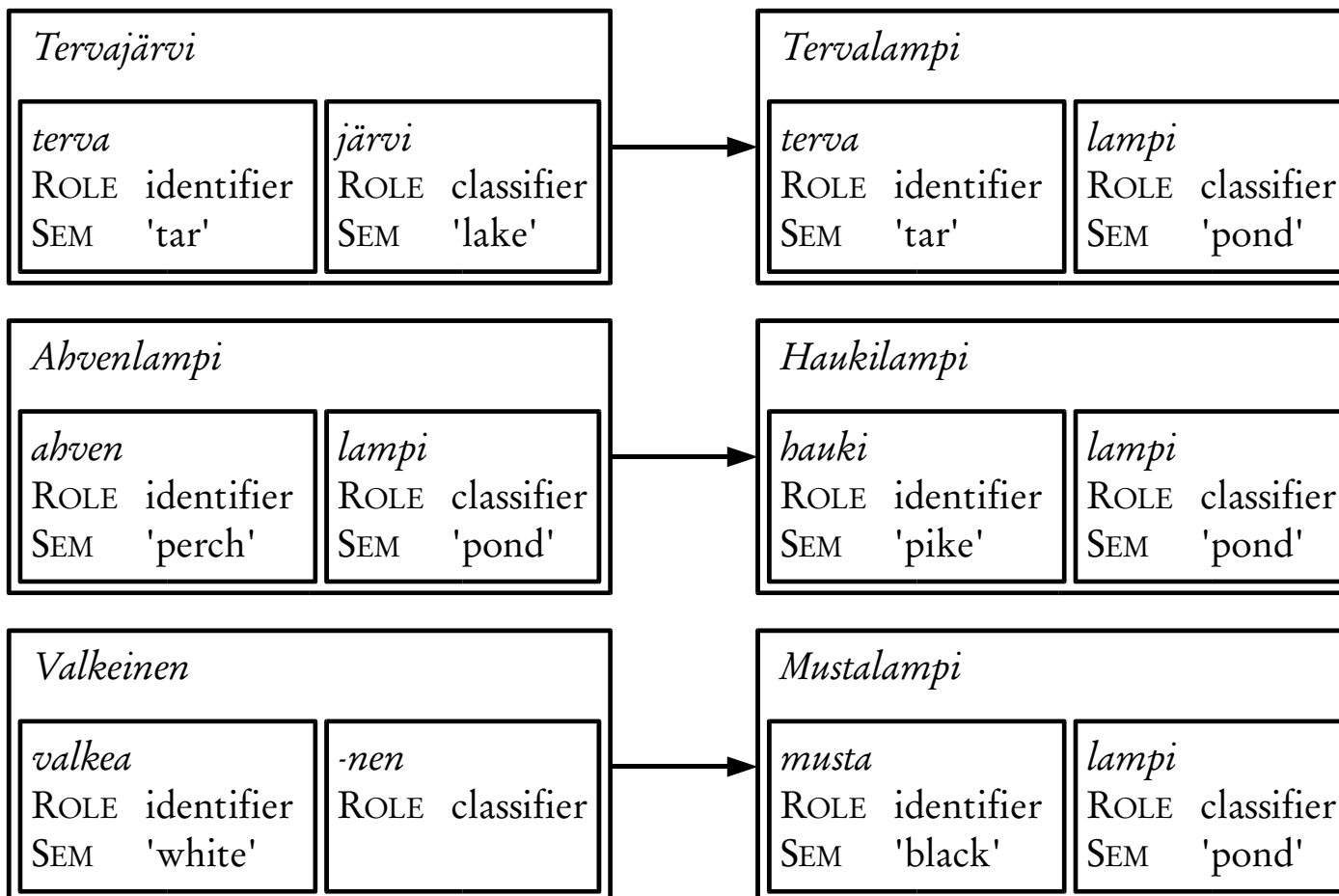## How common are they?

- A couple of constructions cover most names

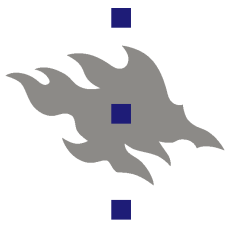|	All				Stand-alone		In inductive	
	Places		Names		Places	Names	Places	Names
		%		%	%	%	%	%
-lampi 'pond'	35 626	61	11 975	48	65	53	37	33
-järvi 'lake'	14 095	24	6 951	28	24	29	25	25
-vesi 'water'	214	0	180	1	0	1	0	0
-nen	2 966	5	1 705	7	4	4	13	13
-kkV	634	1	511	2	1	1	4	4
Other	4 732	8	3 856	15	6	12	20	24

- What about the »Other» names?

  - Common nouns (eg. *Kaakkuri* 'Loon', 8 lakes)

  - Some adjectives (eg. *Hoikka* 'Thin', 26 lakes)

  - Some less common compounds or suffixes (eg. *Peipposenmeri* 'Chaffinch's Sea', 13 lakes)

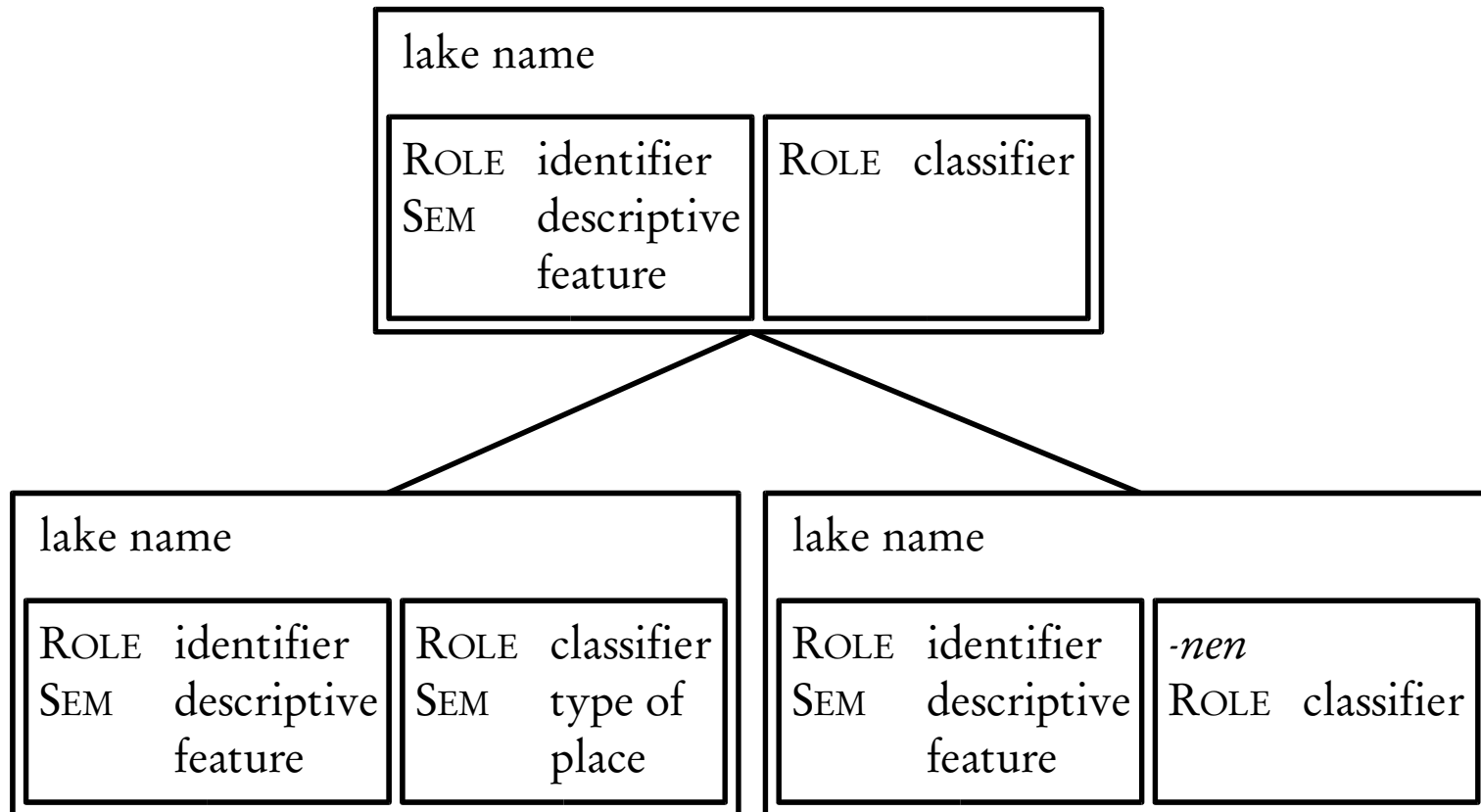 - A few opaque names (eg. *Päijänne*, 21 lakes)
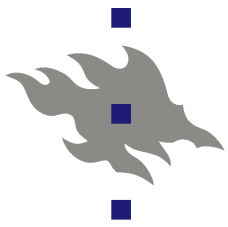
Spatial collocations

- Pairs of names that systematically appear near each other

- So common that using single names as prototypes cannot be labeled exceptional

Tervajärvi	
terva ROLE identifier SEM 'tar'	*järvi* ROLE classifier SEM 'lake'

→

Tervalampi	
terva ROLE identifier SEM 'tar'	*lampi* ROLE classifier SEM 'pond'

Ahvenlampi	
ahven ROLE identifier SEM 'perch'	*lampi* ROLE classifier SEM 'pond'

→

Haukilampi	
hauki ROLE identifier SEM 'pike'	*lampi* ROLE classifier SEM 'pond'

Valkeinen	
valkea ROLE identifier SEM 'white'	*-nen* ROLE classifier

→

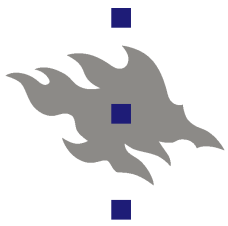| Mustalampi	
musta ROLE identifier SEM 'black'	*lampi* ROLE classifier SEM 'pond'

Top-level generalisation

- Based on cases like the last example, it seems plausible to postulate a family of constructions like
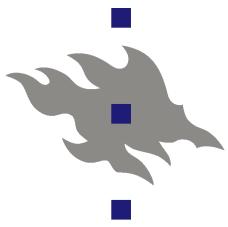
lake name	
ROLE identifier SEM descriptive feature	ROLE classifier

lake name			lake name	
ROLE identifier SEM descriptive feature	ROLE classifier SEM type of place		ROLE identifier SEM descriptive feature	*-nen* ROLE classifier

Inductive names

- Modifier + existing name

Pieni Haukilampi	
pieni ROLE modifier SEM 'small'	*Haukilampi* ROLE head SEM 'Pike Pond'

- Existing name in genitive + type of place

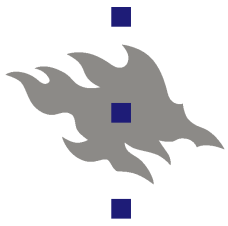| Kalettomanlampi	
Kaleton ROLE identifier CASE genitive LXM *Kaleton* SEM 'Fishless''	*lampi* ROLE classifier SEM 'pond'

These are not always inductive: structurally similar stand-alone names exist as well, especially as a result of epexegesis.

Summary

- Names are modelled after existing ones
  (as we all knew already).

- However, it is often difficult to distinguish between semantic and syntactic patterns.

- There are degrees of productivity.
  Partial productivity is normal.

- **Prototypes** are good.

- However, some concept of **range** is also necessary: one can't go arbitrarily far from the prototype.

- It is possible to find a general linguistic theory that can cover onomastics.
  Names are a proper part of language.

Thank you