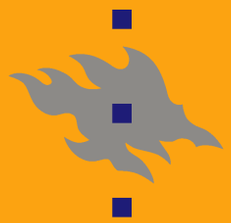


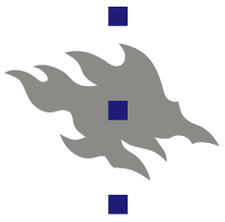
HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

In Search of Naming Patterns: A Survey of Finnish Lake Names

Antti Leino <antti.leino@cs.helsinki.fi>

3rd September 2004

Department of Computer Science



Introduction

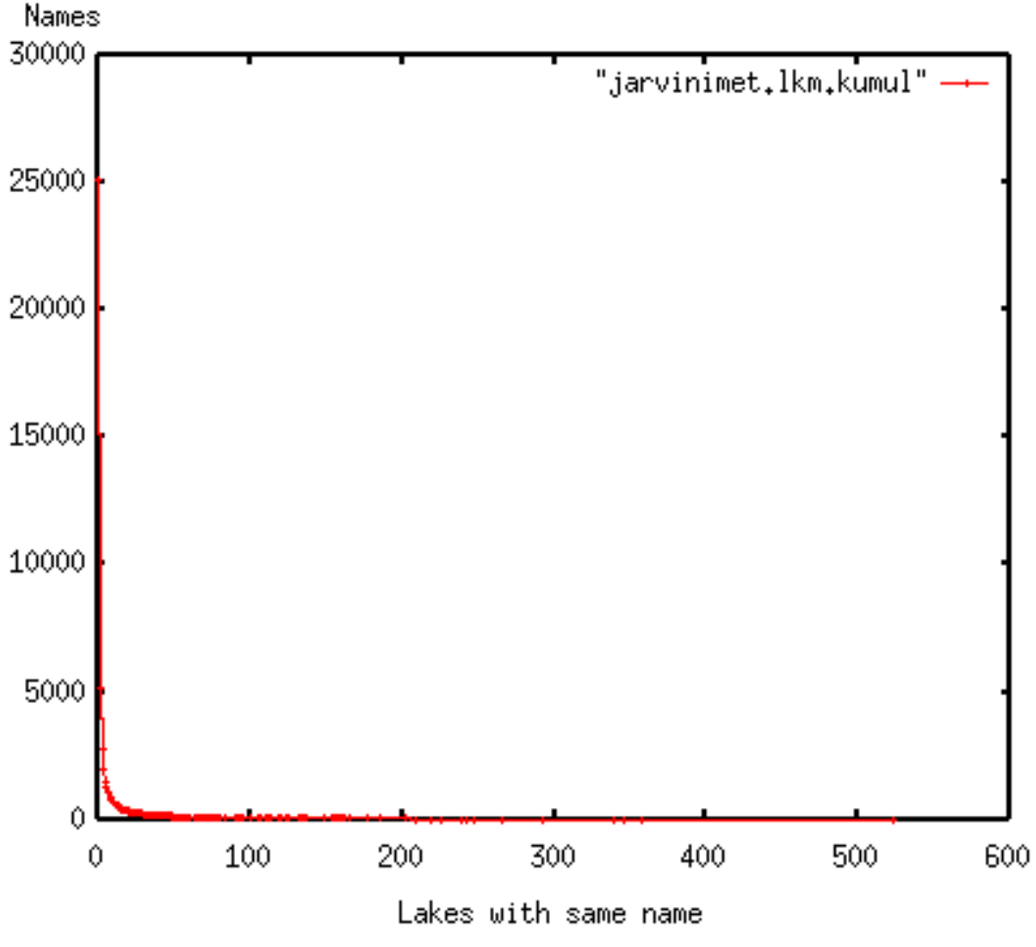
- Patterns and analogy play often an important role in naming
- This happens even when there are other motivations for the name
- How to prove this?
  - Attempt to find groups of names that often appear near each other
  - If such groups are found, see if the constituent names are instances of the same pattern

Data

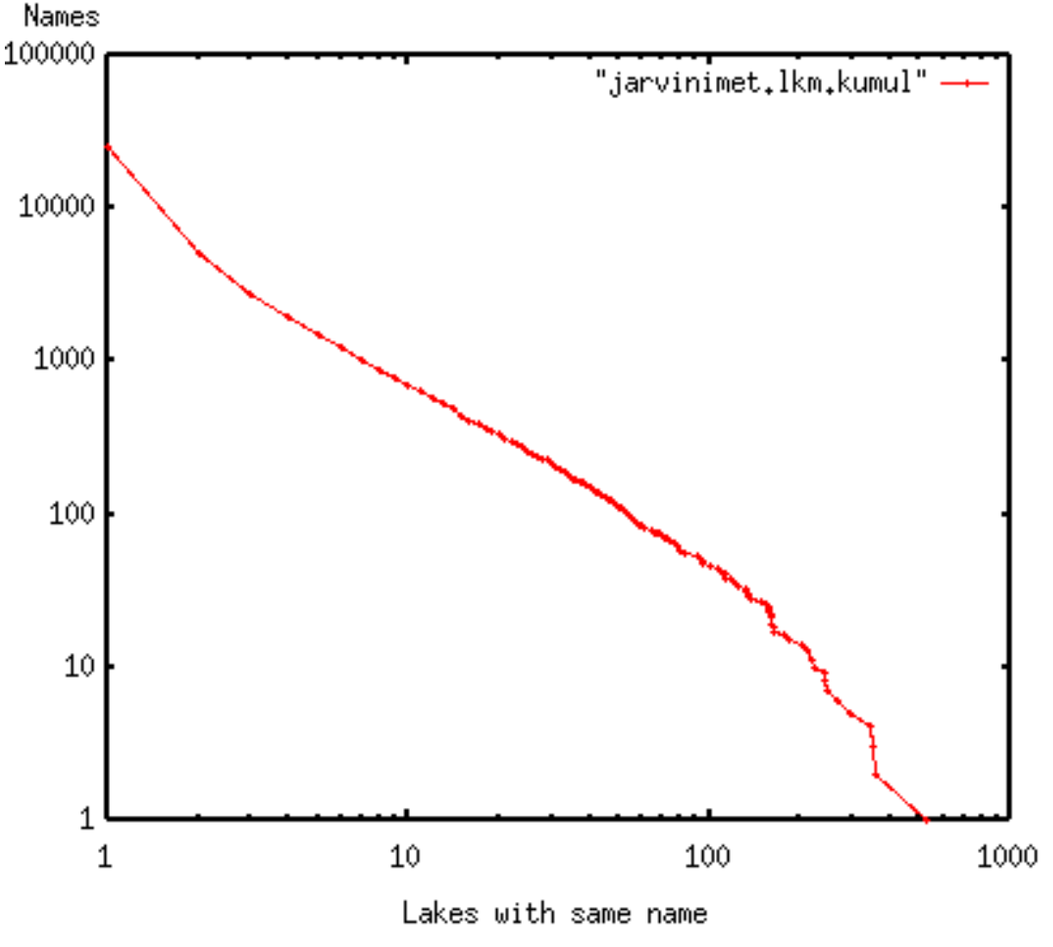
- Place Name Register
  - Used by the National Land Survey of Finland to produce the 1:20 000 Basic Map
  - Lake names

	Occurrences	Names	Named places
All toponyms	≥ 1	303 626	717 747
Lakes	≥ 1	25 178	58 267
This study	≥ 20	331	19 230

Number of similarly-named lakes



Linear scale



Logarithmic scale

Methods

- Spatial Statistics
  - Specifically, small-scale interactions between name occurrences
  - K function: measure of spatial interactions
- Data Mining
  - Branch of Computer Science
  - Attempts to find interesting new information in large corpora of data
  - Tools to find groups of names whose mutual K functions are interesting

K function

- A measure of the attraction / repulsion of point objects
- Rough meaning: if the points were randomly distributed, one would need an area of $K(r)$ to find as many points that really are within radius r of the actual points
- For spatially random data, $K(r) = \pi r^2$
- Spatial attraction: $K(r) > \pi r^2$
- Spatial repulsion: $K(r) < \pi r^2$
- In this study: a variant for two sets of points, each with a spatially non-uniform distribution

Results

- There are groups of names that cluster around each other on a scale of even 1 km or less
- Some of these result from cultural phenomena or natural features
- At least two main types of naming patterns
  - Inductive names
  - Contrastive names
- Both these main types appear to be very productive

Cultural and natural connections

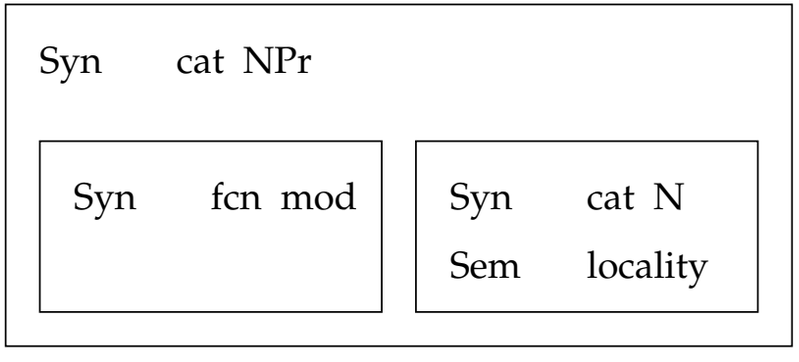
- Cultural connections
  - *Niittylampi* 'Meadow Pond' — *Vasikkalampi* 'Calf Pond'
  - *Myllyjärvi* 'Mill Lake' — *Kirkkojärvi* 'Church Lake'
- Natural connections
  - *Paskolampi* 'Shit Pond' — *Liejulampi* 'Mud Pond'
  - *Kaitajärvi* 'Narrow Lake' — *Hoikkajärvi* 'Thin Lake'
- Not very easy to spot (but then, these are not patterns)

Inductive and Contrastive names

- Inductive
  - *Mäntyjärvi* 'Pine Lake' — *Mäntylampi* 'Pine Pond'
  - *Iso Haukilampi* 'Great Pike Pond' — *Pieni Haukilampi* 'Small Pike Pond'
  - Some overlap with the next category
- Contrastive
  - *Mustalampi* 'Black Pond' — *Valkealampi* 'White Pond'
  - *Valkeajärvi* 'White Lake' — *Mustalampi* 'Black Pond'
  - *Valkeinen* 'The White' — *Mustalampi* 'Black Pond'
  - *Ahvenlampi* 'Perch Pond' — *Haukilampi* 'Pike Pond'

Structure of lake names

- Leaving here the traditional terms of *naming pattern* and *analogy*; instead, changing point of view to that of Construction Grammar[1]
- Name consists of a head denoting the type of place, preceded by a modifier



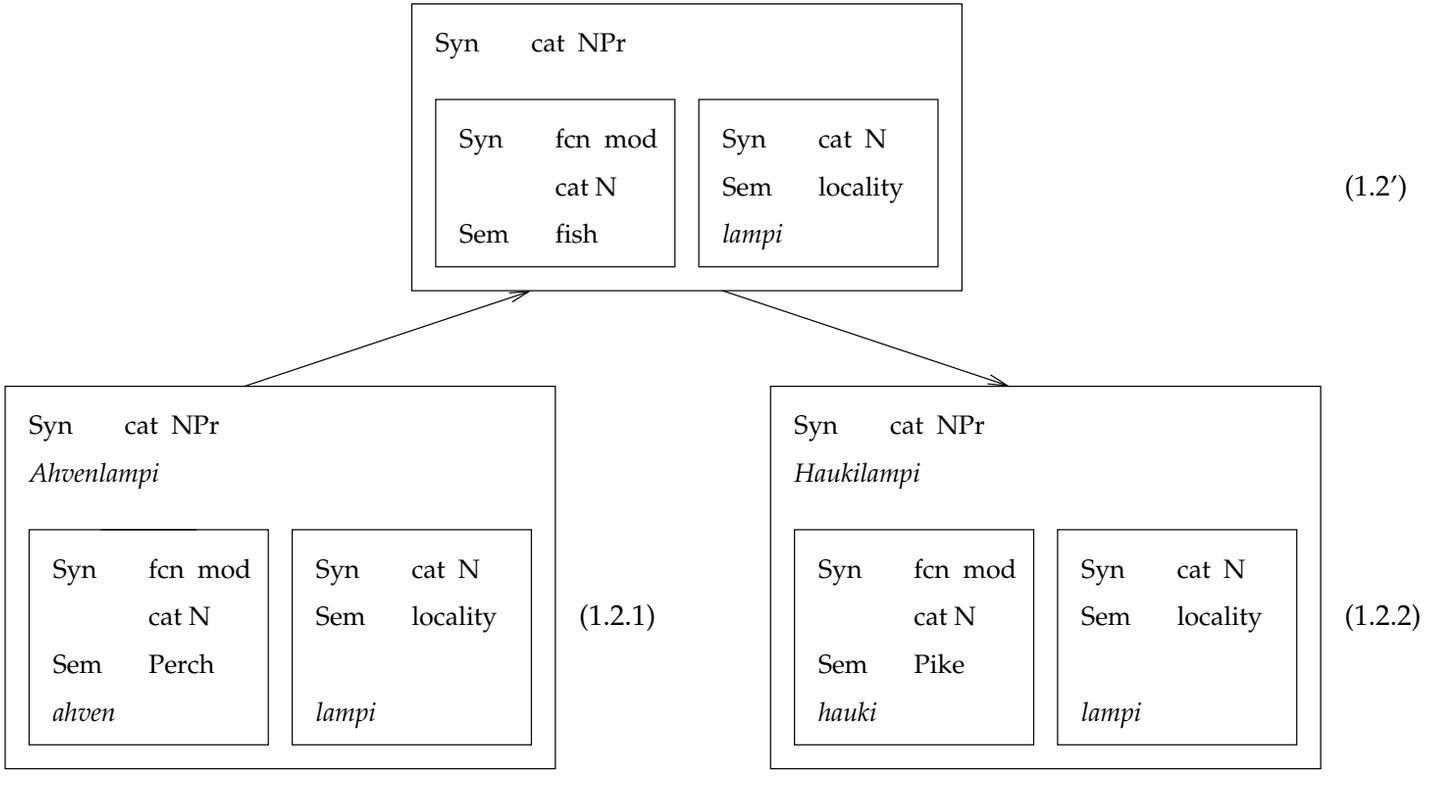
[1] Although a separate school, this is functionally equivalent and ideologically mostly compatible with Cognitive Grammar

Main toponymic construction

- Three main subtypes, depending on the modifier
  - Adjective (eg. *Mustalampi* 'Black Pond')
  - Noun in the genetive case (eg. *Kaakkurinlampi* 'Loon's pond'; rare in names)
  - Noun in the nominative case (eg. *Housulampi* 'Trouser Pond'; often no appellative homonym)
- Even names that have a corresponding common noun construction occur often in pairs $\Rightarrow$ likely did not originate entirely as descriptive designations

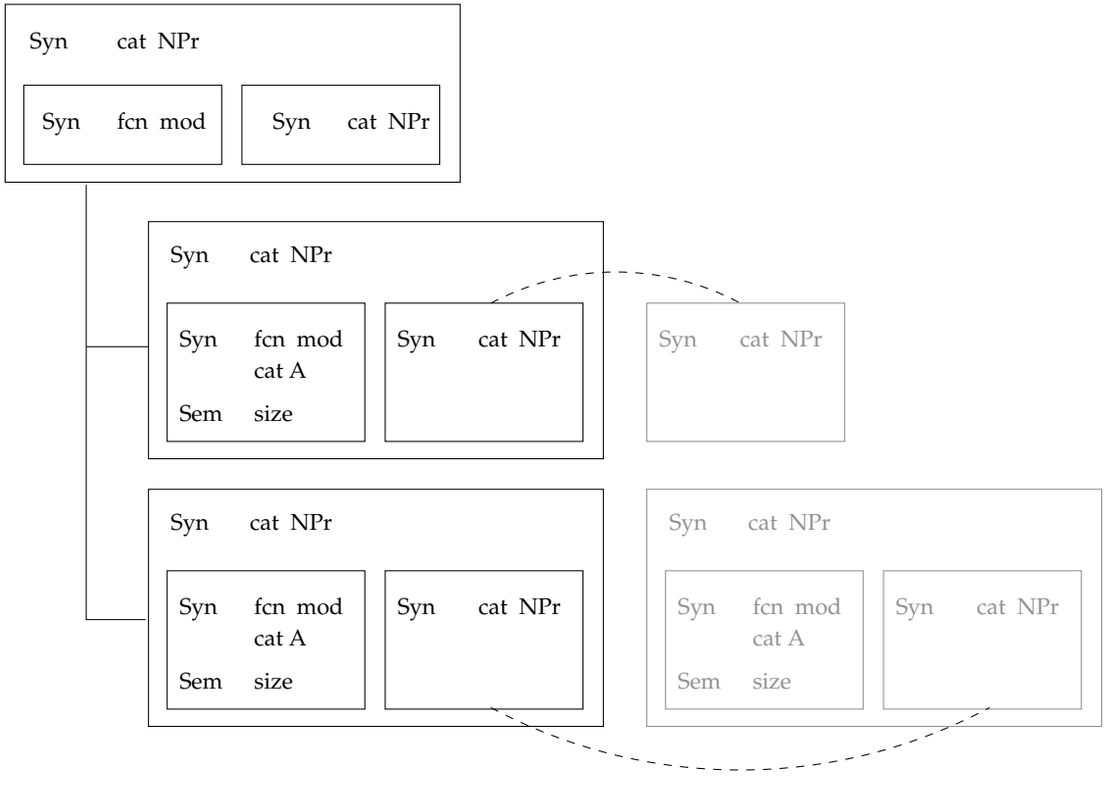
Variation in naming

- Existing name used to give a construction later used for creating a new name
- Beware: somewhat heretical in CG terms



Inductive naming

- The existing toponym is relevant (cf. *Ahvenkorpi* 'Perch Waste' near *Ahvenjärvi* 'Perch Lake')
- Easiest to include it in the construction



Conclusions

- Clearly, the old hypothesis was right: analogy plays an important role in naming lakes
- *Naming pattern* as a term does not feel right
  - Sounds like something that can be clearly defined
  - Most of the »patterns» are rather small and have an ad-hoc nature
  - Not necessarily just the term that needs fixing, but possibly the concept as well
- Construction Grammar / Cognitive Grammar seems a promising approach
  - However, further work is needed on the theoretical side