

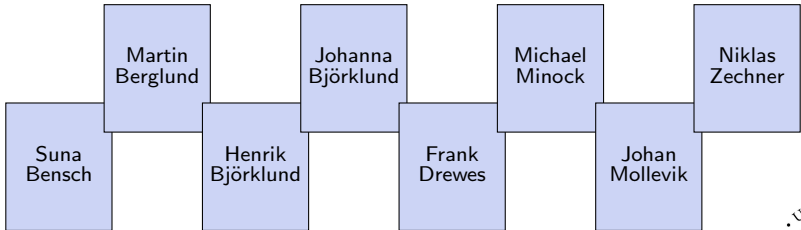
# The Group Natural and Formal Languages

Projects and Research Directions

Frank Drewes  
with input and help by “the group”

CS Research Day





# The Information Age is a Language Age

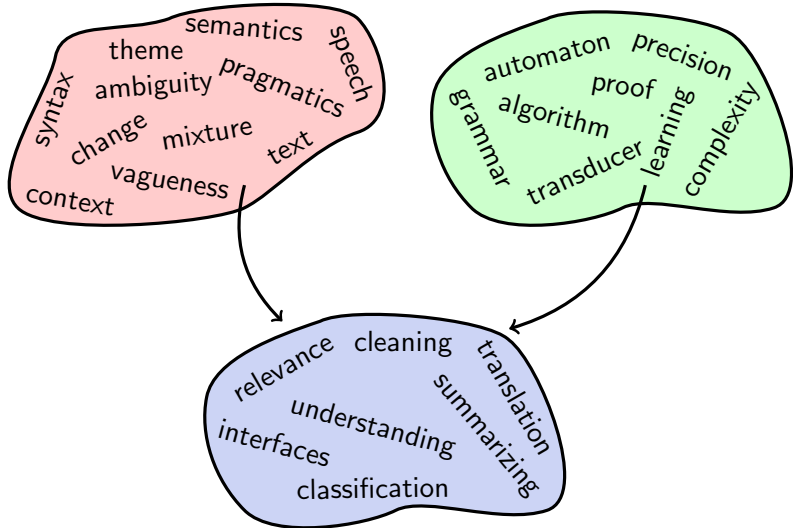
---

Twitter  
scientific journals  
legal documents  
Youtube  
product reviews  
Facebook  
Wikipedia  
Spam  
Commercials

Too much for manual processing!  
Too valuable for being disregarded!

# Natural and Formal Languages

---



We focus on methods that  
exploit structure  
and have a sound theoretical basis.



Some projects. . .



# Syntactic Analysis of Natural Language

---

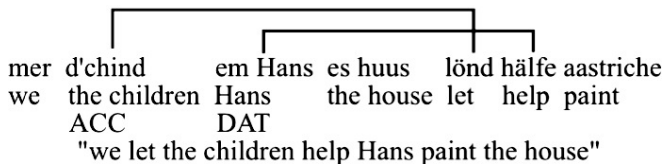
**Financing:** VR project grant (H. Björklund)

**People:** H. Björklund, M. Berglund, F. Drewes, N.N.

- Natural language analysis usually starts with a syntactic analysis.
  - There are many different grammar formalisms for doing this.
  - Efficient parsing is central, but not entirely understood.

What are the parameters that enable efficient syntactic analysis?  
Is there a trade-off between efficiency and descriptive power?







- Financing:** VR project grant  
UmU Young Researcher Award  
FOI & Industrial Doctoral School } all J. Björklund
- People:** J. Björklund, N. Zechner, N.N.

The amount of text makes automatic classification a central issue.

- Do I have an unknown aunt in Nigeria or is it spam?
  - How is our product viewed in social media?
    - What is this text about?
      - Who wrote this text?

Can structure help to answer these questions?

**Niklas Zechner** <[zechner@cs.umu.se](mailto:zechner@cs.umu.se)>

20 november 2012 15:08

Till: Martin Berglund

[Göm detaljer](#)

Kopia: Johanna Björklund , Suna Bensch , Frank Drewes , Henrik Björklund

[0: Lokal inkorg](#)<sup>1</sup>

Re: EU application - progress beyond zero

---

It looks good to me. If I was a component-based text analysis tool, I would notice the following:

- it is written in a very formal style, and looks a lot like an EU application

- parser has some difficulties with phrases like "With all of the above work packages not only will the individually interesting and relevant research result"

- we can identify the author with 90% certainty, from phrases like "to a great part" and "a fair bit"



# Spacebook: Spoken Guidance for Pedestrians

---

**Financing:** FP7 STREP project (M. Minock)

**People:** M. Minock, J. Mollevik  
(L. Edblom, S. Hegner, M. Karlsson, L. Närva, M. Åsander)

**Partners:** 6 universities and one company

Why should I wish to stare at my iPhone when I visit another city?

- I want to be told how to find my way.
  - I want to ask “How old is this bridge?”
    - I want reasonable answers.

Speech-driven hands and eyes free support for pedestrians navigating and exploring the city.



YouTube

SpaceBook - Prototype 1: "Long Thin System" - Initial Street Tests

SpaceBookProject + Prenumerera 5 videoklipp



1:35 / 6:27

Gilla Lägg till i Dela 38 visningar

Publicerades den 28 nov 2012 av [SpaceBookProject](#)

This video shows trials of the SpaceBook Long Thin System. The main aim was demonstrate the integration of modules from different partners. This system was evaluated in Edinburgh, Scotland.

<http://www.youtube.com/watch?v=qQZnwr0yeTE>



# Multi-Dimensional Language Analysis

---

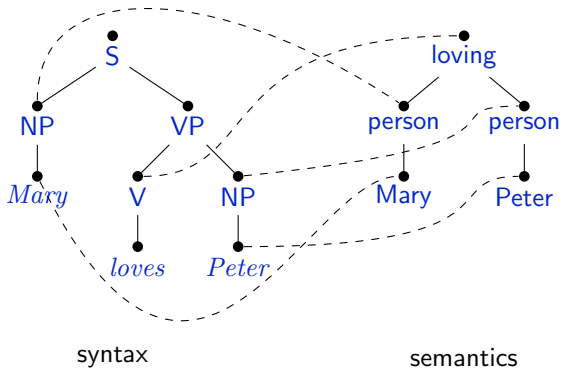
**Financing:** VINNOVA VINNMER grant (S. Bensch)

**People:** S. Bensch, H. Björklund, F. Drewes

- Natural language does not only have the syntactic dimension.
  - Others are phonetics, semantics, pragmatics, . . .
    - Modern linguistic theories consider them as interlinked.
      - We want to develop formalisms that capture this.
        - This could enable better analyses.

We call our formalism **Millstream systems**.





- Shuffled languages
  - Extended formalisms for regular languages
    - Efficient evaluation of XML queries
      - Extensions of context-free graph grammars

...and, of course, Berglund's conjecture!

We will organize 2 international conferences here in Umeå:

- *Non-Classical Models of Automata and Applications* (NCMA 2013), August 13–14, 2013.
- *Conference on Implementation and Application of Automata* (CIAA 2015), 2nd half of August, 2015.





Thank very for listening much you!

[Swiss English]