Computational Linguistics, Cognitive and Social Sciences

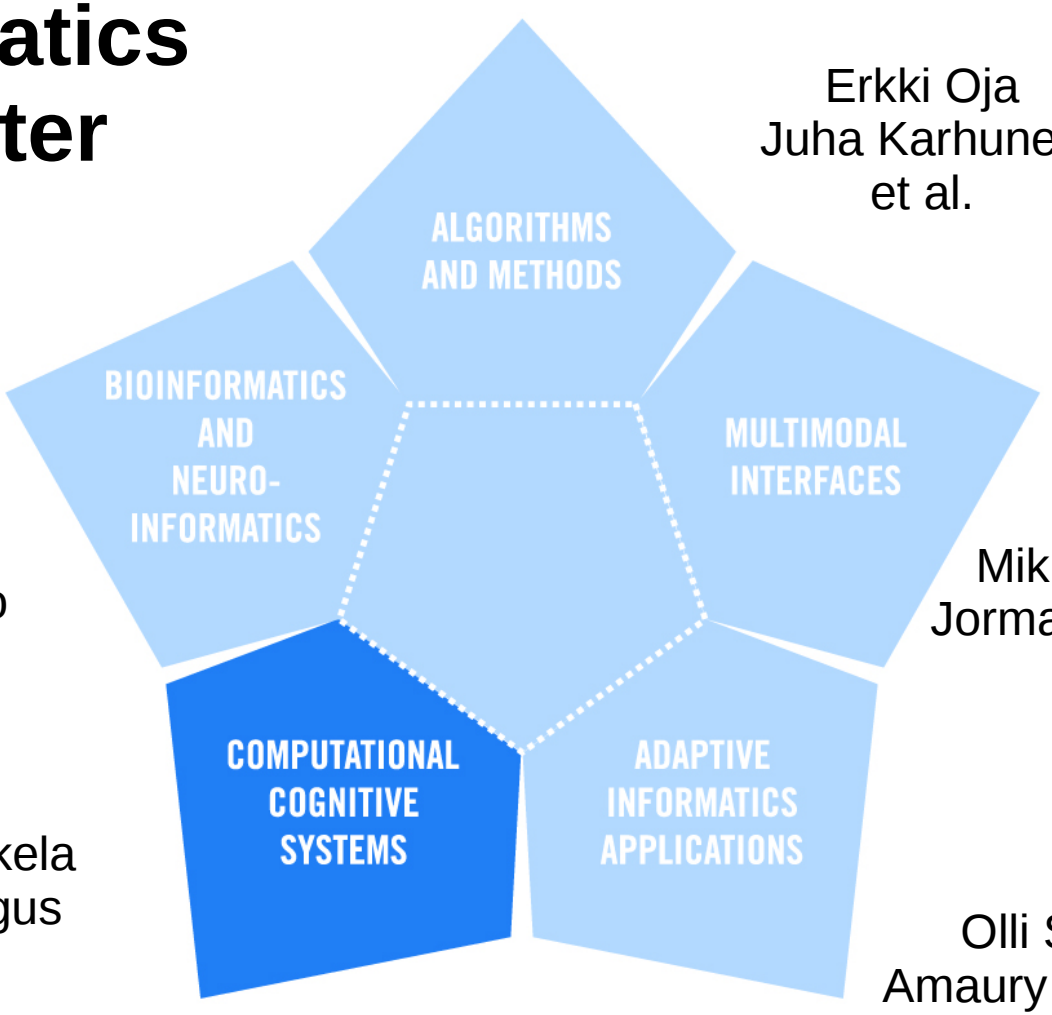
Timo Honkela, Krista Lagus and
the Computational Cognitive Systems group

Aalto University School of Science
Department of Information and Computer Science

Dipoli, Luolamies, Espoo, 31.1.2012

Computational Cognitive Systems

Adaptive Informatics Research Center 2006-2011



Erkki Oja
Juha Karhunen
et al.

Samuel Kaski
Ricardo Vigário
et al.

Mikko Kurimo
Jorma Laaksonen
et al.

Timo Honkela
Krista Lagus
et al.

Olli Simula
Amaury Lendasse
et al.

Definition of Cognitive Systems

- Artificial systems that combine
  - perception,
  - action,
  - reasoning,
  - learning and
  - communication.
- This area of research draws upon biological, cognitive and social system approaches to understanding cognition.

Computational Cognitive Systems

- Computational modeling helps in understanding real-world phenomena and facilitate research and development in diverse areas such as
  - language learning and use,
  - multimodal and multilingual systems,
  - wellbeing research,
  - social and organizational theory, and
  - knowledge-intensive information systems.

Agenda

- Introductions to research themes
- Introductions to scientific articles, blogs, books and some research projects – and the people behind them

Conferences

SCAI 2006

STeP 2008

STeP 2010

Conferences

Timo Honkela, Wlodzislaw Duch, Mark A. Girolami, and Samuel Kaski, editors. Artificial Neural Networks and Machine Learning - Proceedings of ICANN 2011 - 21st International Conference on Artificial Neural Networks, Part I. Springer, 2011.

Timo Honkela, Wlodzislaw Duch, Mark A. Girolami, and Samuel Kaski, editors. Artificial Neural Networks and Machine Learning - Proceedings of ICANN 2011 - 21st International Conference on Artificial Neural Networks, Part II. Springer, 2011.

Jorma Laaksonen and Timo Honkela, editors. Advances in Self-Organizing Maps - Proceedings of WSOM 2011, 8th International Workshop. Springer, 2011.

Tapio Pahikkala, Jaakko Väyrynen, Jukka Kortela, and Antti Airola, editors. Proceedings of the 14th Finnish Artificial Intelligence Conference, STeP 2010, number 25 in Publications of the Finnish Artificial Intelligence Society. Finnish Artificial Intelligence Society, 2010.

Timo Honkela, Mari-Sanna Paukkeri, Matti Pöllä, and Olli Simula, editors. Proceedings of AKRR'08, International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning. Helsinki University of Technology, Espoo, Finland, September 2008.

Tapani Raiko, Pentti Haikonen, and Jaakko Väyrynen, editors. AI and Machine Consciousness — Proceedings of the 13th Finnish Artificial Intelligence Conference STeP 2008, number 24 in Publications of the Finnish Artificial Intelligence Society. Finnish Artificial Intelligence Society, 2008.

Timo Honkela, Ville Könönen, Matti Pöllä, and Olli Simula, editors. Proceedings of AKRR'05, International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning. Helsinki University of Technology, Espoo, Finland, June 2005.

Education

- Statistical Natural Language Processing course
  https://noppa.aalto.fi/noppa/kurssi/t-61.5020/luennot
- Special Seminars on e.g.
  - Statistical Machine Translation
  - Concept Formation
  - Symbol Grounding
  - Computational Pragmatics

  http://cogsys.blogspot.com/2007/01/finnish-swedish-machine-translation.html

  https://noppa.aalto.fi/noppa/kurssi/t-61.6020/luennot
- EIT ICT Labs Wellbeing Innovation Camps
  http://www.cis.hut.fi/wicamp/
- HeCSE and Langnet Graduate Schools
  http://www.joensuu.fi/fld/langnet/english/languagetechnology.html
- Thesis works and individual reports

Networking

- Finnish and international universities and research institutions, e.g.
  - University of Helsinki (computer science [Aapo Hyvärinen, Roman Yangarber], language technology [Lauri Carlson, etc.], cognitive science [Otto Lappi, Anna-Mari Rusanen], philosophy [Tarja Knuuttila]), National Consumer Research Center [Mika Pantzar, Tanja Kotro, Petteri Repo et al.]
  - Stanford, Berkeley, Edinburgh, DFKI, Demokritos, etc.
  - Also: Aalto! Brain research unit [Riitta Salmelin], Semantic computing [Eero Hyvönen, etc.], Media Factory [Juhani Tenhunen, etc.], Department of Media, Department of Design [Turkka Keinonen, etc.]
- Large number of companies

Networking

Paul Kay

George Lakoff

Douglas Hofstadter

Masoud Nikravesh

Peter Norvig

Jim March

Networking

META-NET

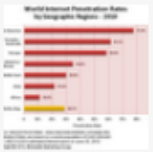
Krister Lindén **Mikko Kurimo**

Multilingual web

Cognitive Systems blog

http://cogsys.blogspot.com

Popular Posts



Content on the Multilingual Web

The Internet world stats site is a fascinating source of information for predicting the trends of web usage in terms of language and locali...

Andrew Chesterman: Translation Strategies

Machine translation is a great challenge. In order to make the scope of the task clearer, our research group invited prof. Andrew Chesterman...

Teemu Leinonen: Designing Learning Tools - Methodological Insights.

Teemu Leinonen from Aalto University School of Art and Design defended today his thesis "Designing Learning Tools - Methodological Insigh...

MultilingualWeb - Where Are We?

Since the beginning of the World Wide Web in the early 90s, English has been the lingua franca of web technologies. Issues related to supp...

META-FORUM 2011

The META-FORUM 2011 is held June 27/28 in Budapest, Hungary. It is organized by META-NET , a Network of Excellence consisting of 47 resea...

Thursday, October 06, 2011

Bernardo Huberman: Social Media and Attention

Bernardo Huberman from Hewlett-Packard Laboratories and Stanford University i
University School of Science. He is giving a joint ICS Forum/Aalto Physics Colloqui
October.

In his talk, Prof. Huberman discussed the explosion of online interactions: the we
metabolism of thought and decision about work, social issues and play. Twitter is
million messages a day. Facebook has more than 800 million users and Google+
40 million users in one month. Rate of sharing content in social media is growing

Huberman further pointed out that attention is limited by brain capacity. Attenti
scarce and valuable resource. Wherever attention flows, issues surface and ideas
discussed - and money often follows. Almost anything else except attention can
manufactured as a commodity. Actually, attention is the coordination mechanism
powers the progress of science. Within academia attention has a symmetry prop
academics seek attention from those who also seek it from them. Topics that get
attention become part of the research agenda of the community. Attention is us
determined professional standing. He also mentioned that productivity can be m
the same way. This raises, however, the question of time delays: the importance
finding or an innovation may be understood only much later. Regardless of this, t
of the work is measured according to more or less recent citations.

Most of the attention that we pay today is reflected and propagated in social me
to a study of 15 million recommendations from amazon.com by Lefkovec, Adami
Transactions on the Web, 2007). The network allocates attention in a highly nonl

Methodological basis (1/2)

- Neural networks and statistical machine learning
  - Unsupervised learning
    (Teuvo Kohonen: Self-Organizing Map;
    Aapo Hyvärinen, Juha Karhunen & Erkki Oja:
    Independent Component Analysis,
    Samuel Kaski: Multi-way, multi-view learning ...)
  - Reinforcement learning
    (Ville Könönen, Paul Wagner)
- Statistical language processing

Methodological basis (2/2)

- Multimodal information processing
  - Jorma Laaksonen, Markus Koskela et al.: image and video processing
  - Mikko Kurimo, Kalle Palomäki et al.: speech processing
- Agent-based simulation
- Probabilistic modeling
- Artificial Immune Systems
  - Matti Pöllä

SOM Theory

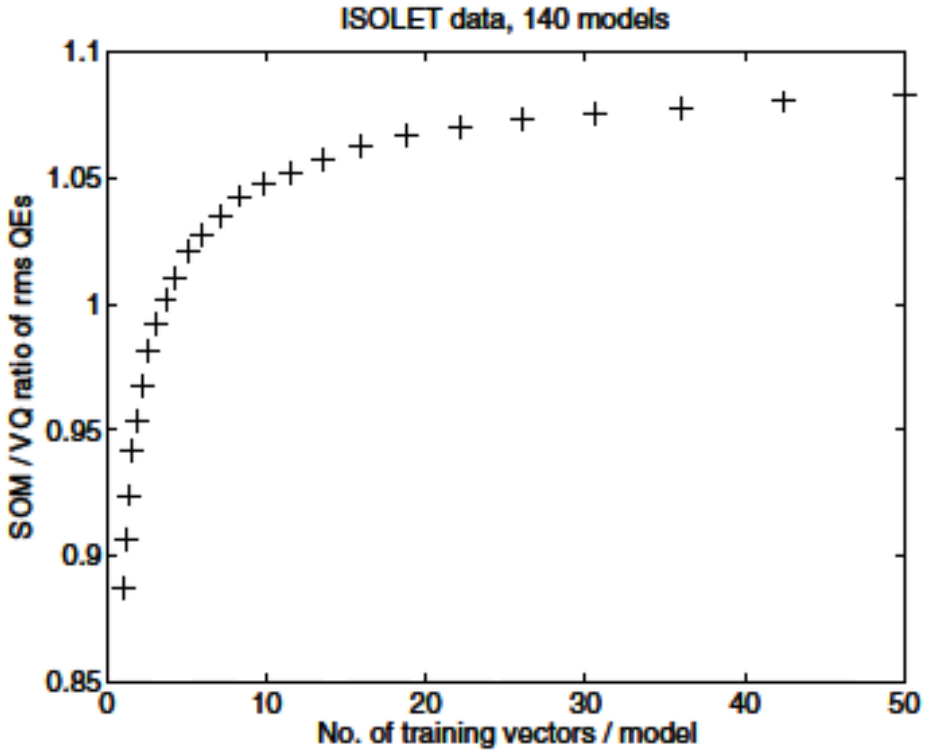


Fig. 5. Ratio of the rms QEs in the SOM and the VQ for the ISOLET data set, as a function of the number of training vectors per model and for 140 models

Teuvo Kohonen, Ilari T. Nieminen, and Timo Honkela. On the quantization error in SOM vs. VQ: A critical and systematic study. In Proceedings of WSOM'09, pages 133–144. Springer, 2009.

General research areas

- Language
- Concepts
- Social domain

  - Cognitive point of view defines the basic approach
    - Cognitive linguistics rather than autonomous linguistics
    - Concepts as (socio-)cognitive phenomena
    - Socio-cognitive modeling

Our research on language (1/2)

- We consider natural language also as a **signal** and **dynamic system** at **cognitive** and **social** levels (also in its written form) rather than a symbolic and logical system
- Importance of **embodiment** (cf. e.g. Harnad; Lakoff) and **embeddedness** (cf. e.g. Edelman)
- **Learning** and **pattern recognition** processes are essential; much of the learning is bound to be **unsupervised**

Language

- Thousands of languages (~ 6000); much larger number of dialects; each language has thousands to at least one million lexemes (English); in some languages like Finnish, there are billions of surface word forms
- Many “sublanguages” within one language: doctors, lawyers, scientists and engineers in different areas and disciplines, etc.
- Language is constantly changing at different levels
- Many expressions have a variety of meanings and each individual has a least slightly different interpretation for each expression
- There is a constant attempt to describe with language everything that we learn to know

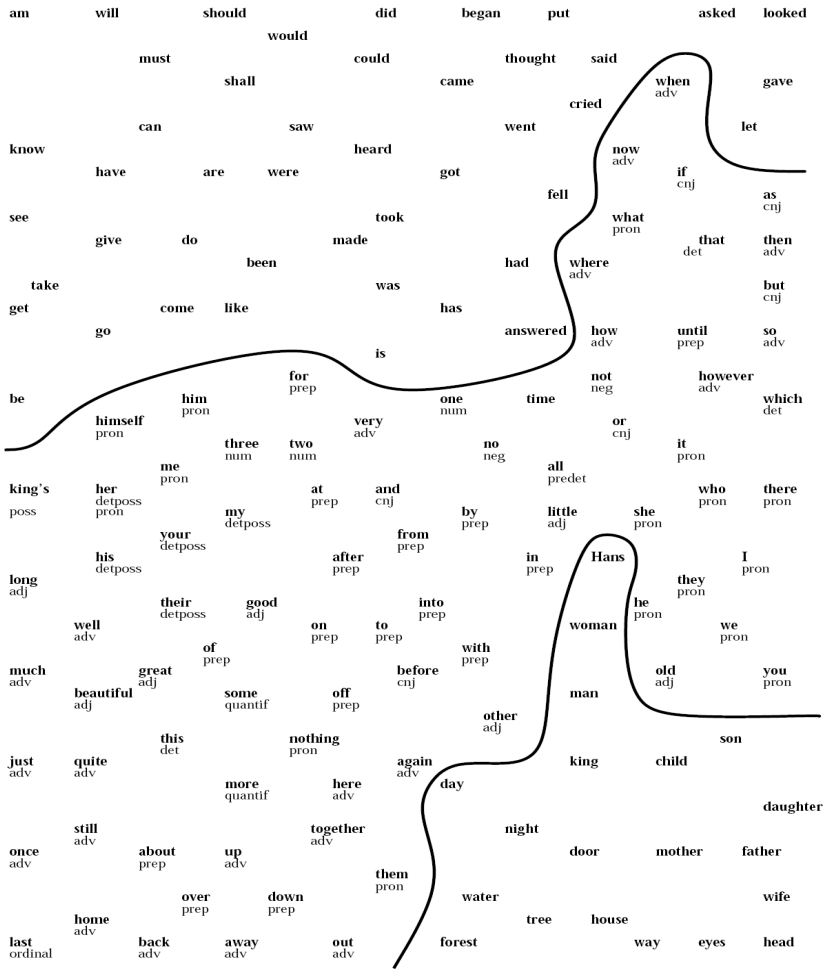
Our research on language (2/2)

- Our focus is in modeling linguistic cognition: we study how language is learned and used by cognitive agents rather than study language as an autonomous system
- Thus, we share many assumptions of cognitive linguistics – with an additional emphasis on modeling learning and lesser reliance on manually encoded representations

→ Unsupervised learning and emergent representations

Pre-history: SOM analysis of Grimm fairy tales

- No prior syntactic or semantic categorization of the words
- Distinct areas of verbs and nouns, clusters of numerals, and possessives emerged
- Distinction between animate and inanimate nouns emerged



(Honkela, Pulkki & Kohonen, 1995)

Morfessor

Mathias Creutz, Krista Lagus, and Sami Virpioja (2006). Unsupervised Morphology Induction Using Morfessor. In Finite-State Methods and Natural Language Processing, Lecture Notes in Computer Science, Volume 4002, pages 300-301, Springer Berlin / Heidelberg.

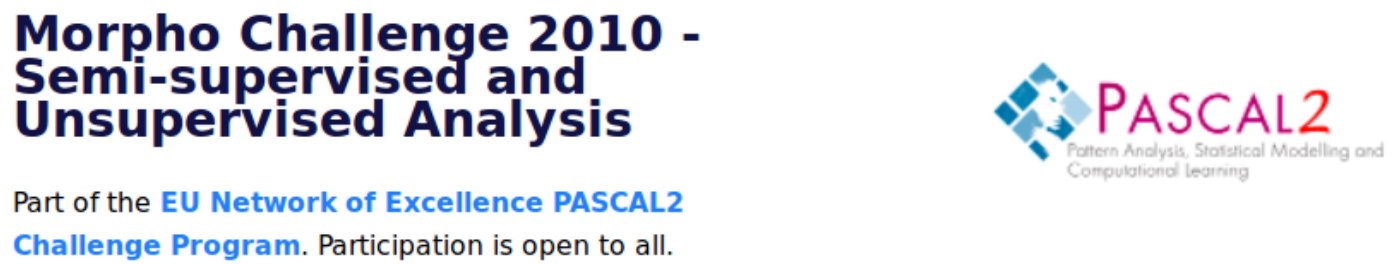
Creutz, M and Lagus, K. (2002). Unsupervised discovery of morphemes. In Proceedings of the Workshop on Morphological and Phonological Learning of ACL-02, pages 21-30, Philadelphia, PA, July 11.

Morfessor - Evaluations

Sami Virpioja, Oskar Kohonen, and Krista Lagus. Evaluating the effect of word frequencies in a probabilistic generative model of morphology. In Proceedings of the 18th Nordic Conference of Computational Linguistics (NODALIDA 2011), volume 11 of NEALT Proceedings Series, pages 230–237. Northern European Association for Language Technology, Riga, Latvia, May 2011.

Sami Virpioja, Ville T. Turunen, Sebastian Spiegler, Oskar Kohonen, and Mikko Kurimo. Empirical comparison of evaluation methods for unsupervised learning of morphology. Traitement Automatique des Langues, 52(2):45–90, 2011.

Morpho Challenge



Mikko Kurimo, Sami Virpioja, Ville Turunen, and Krista Lagus. Morpho Challenge 2005-2010: Evaluations and results. In Proceedings of the 11th Meeting of the ACL Special Interest Group on Computational Morphology and Phonology, pages 87–95. ACL, July 2010.

.. etc.

Allomorfessor

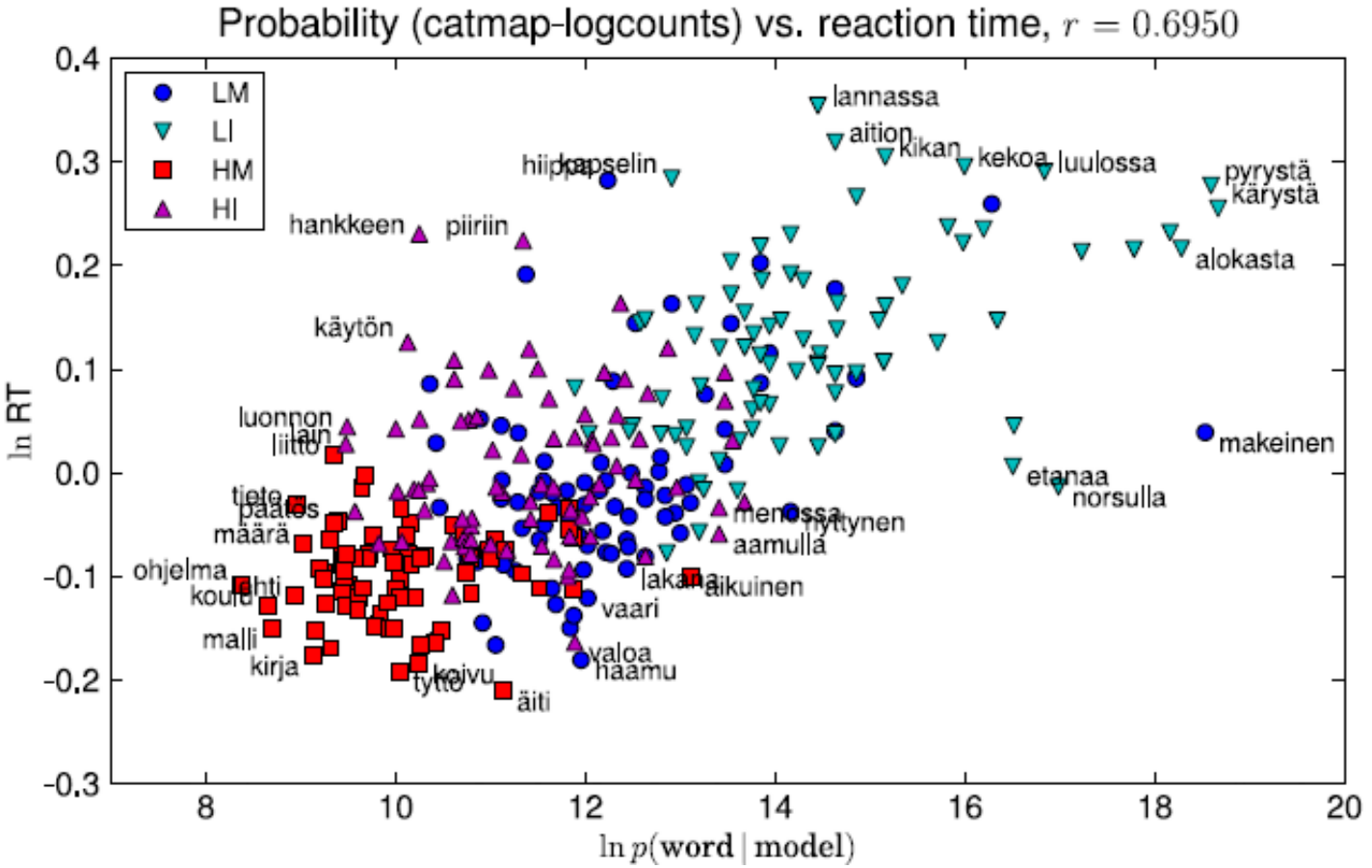
tutkimus, tutkimuksen
kirjoittaa, kirjoitamme
yö, öiden

Oskar Kohonen, Sami Virpioja, and Mikaela Klami. Allomorfessor: Towards unsupervised morpheme analysis. In Working Notes for the CLEF 2008 Workshop, Aarhus, Denmark, September 2008.

Oskar Kohonen, Sami Virpioja, and Krista Lagus. Semi-supervised learning of concatenative morphology. In Proceedings of the 11th Meeting of the ACL Special Interest Group on Computational Morphology and Phonology, pages 78–86, Uppsala, Sweden, July 2010. Association for Computational Linguistics.

Sami Virpioja, Oskar Kohonen, and Krista Lagus. Unsupervised morpheme analysis with Allomorfessor. In Multilingual Information Access Evaluation I. Text Retrieval Experiments: 10th Workshop of the Cross-Language Evaluation Forum, CLEF 2009, Corfu, Greece, September 30 – October 2, 2009, Revised Selected Papers, volume 6241 of Lecture Notes in Computer Science, pages 609–616. Springer Berlin / Heidelberg, September 2010.

Morphological processing by the brain and machine



Sami Virpioja, Minna Lehtonen, Annika Hultén, Riitta Salmelin, and Krista Lagus. Predicting reaction times in word recognition by unsupervised learning of morphology. In ICANN 2011, volume 6791 of Lecture Notes in Computer Science, pages 275–282. Springer Berlin / Heidelberg, June 2011.

Likey: Language-independent Keyphrase Extraction

Aalto University

1. the	1276847
2. of	1067918
3. and	817852
4. in	625330
5. to	357453
6. for	225307
7. is	205723
8. on	162509
9. **research**	157251
10. be	151475
11. with	136854
12. will	135992
13. as	122707
14. are	116508
15. by	113878
16. **university**	98003
...	

1. the	2023617
2. of	945622
3. to	883206
4. and	717718
5. in	611421
6. that	473739
7. a	445775
8. is	445119
9. **we**	305590
10. for	296092
11. **i**	290412
12. this	286924
13. on	274614
14. it	251343
15. be	246917
16. are	197082
...	

Mari-Sanna Paukkeri, Ilari T. Nieminen, Matti Pöllä, and Timo Honkela. A language-independent approach to keyphrase extraction and evaluation. In Coling 2008: Companion volume: Posters, pages 83–86, Manchester, UK, August 2008. Coling 2008 Organizing Committee.

Mari-Sanna Paukkeri and Timo Honkela. Likey: Unsupervised Language-Independent Keyphrase Extraction. In Proceedings of the 5th International Workshop on Semantic Evaluation (SemEval), pages 162–165, Uppsala, Sweden, July 2010. Association for Computational Linguistics.

Learning taxonomies

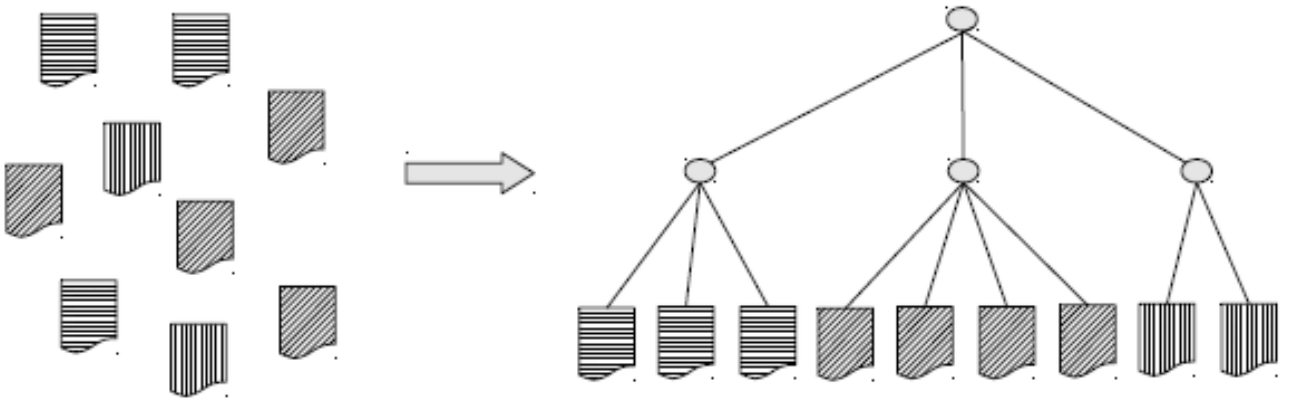
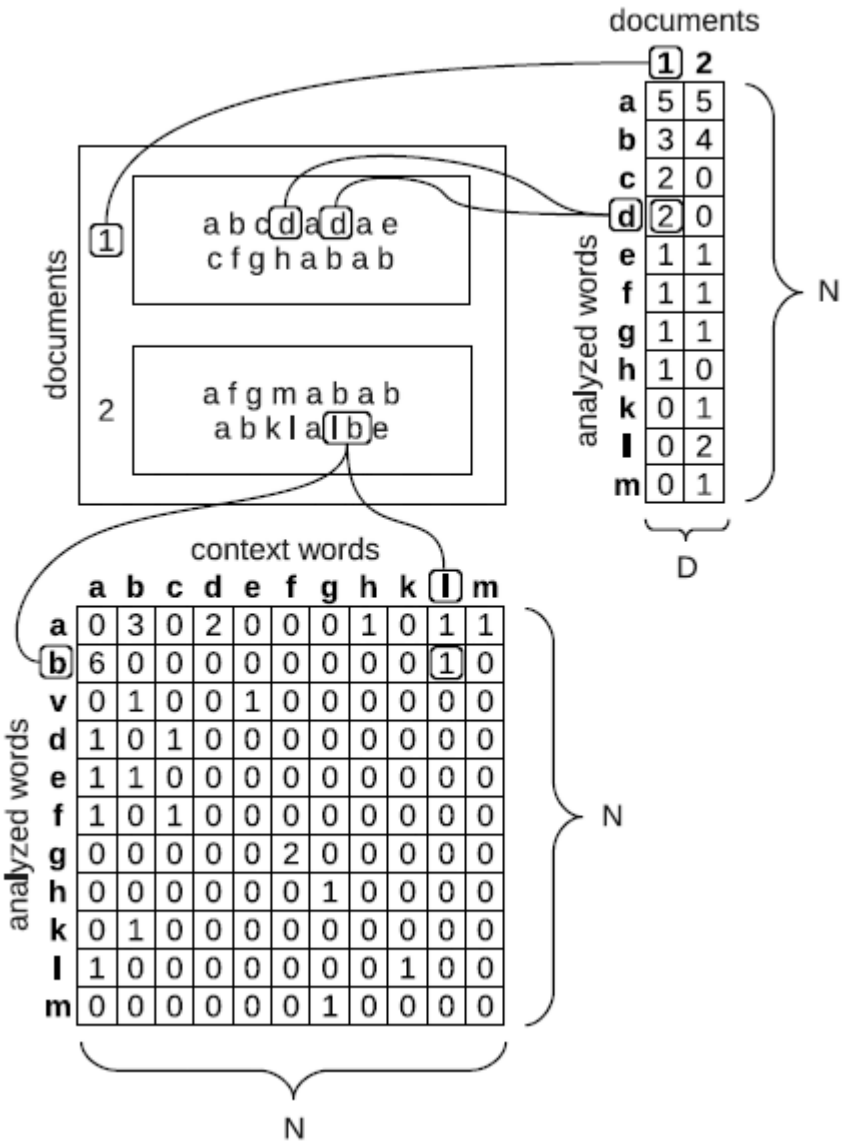


Figure 1: A schematic illustration of taxonomy learning.

Mari-Sanna Paukkeri, Alberto Pérez García-Plaza, Víctor Fresno, Raquel Martínez Unanue and Timo Honkela (2012). Learning a taxonomy from a set of text documents. Applied Soft Computing, 12(3), pp. 1138-1148.

Building vector spaces



Evaluating vector space models

Aalto University

Sami Virpioja, Mari-Sanna Paukkeri, Abhishek Tripathi, Tiina Lindh-Knuutila, and Krista Lagus. Evaluating vector space models with canonical correlation analysis. Natural Language Engineering, to appear. Available on CJO 2011.

WordICA

Table 2. *The most representative words for the first five features, in the order of representativeness, top is highest.*

1	2	3	4	5
or	is	paper	science	networks
and	are	information	university	systems
is	have	it	engineering	learning
are	has	papers	research	models
have	i	system	psychology	processing
has	we	work	neuroscience	algorithms
use	they	networks	technology	recognition
...	...	...	...	...

Timo Honkela, Aapo Hyvärinen, and Jaakko Väyrynen. WordICA - Emergence of linguistic representations for words by independent component analysis. Natural Language Engineering, 16(3):277–308, 2010.

Jaakko J. Väyrynen, Lasse Lindqvist, and Timo Honkela. Sparse distributed representations for words with thresholded independent component analysis. In Proceedings of IJCNN'07, pages 1031–1036, 2007.

Learning Constructions and Grammar Inference

Aalto University

Krista Lagus, Oskar Kohonen, and Sami Virpioja. Towards unsupervised learning of constructions from text. In Proceedings of the Workshop on Extracting and Using Constructions in NLP of the 17th Nordic Conference on Computational Linguistics (NODALIDA), Odense, Denmark, May 2009. SICS Technical Report T2009:10.

Oskar Kohonen, Sami Virpioja, and Krista Lagus. A constructionist approach to grammar inference. In NIPS Workshop on Grammar Induction, Representation of Language and Language Learning, Whistler, Canada, December 2009. Extended abstract.

Analyzing Complexity of Languages

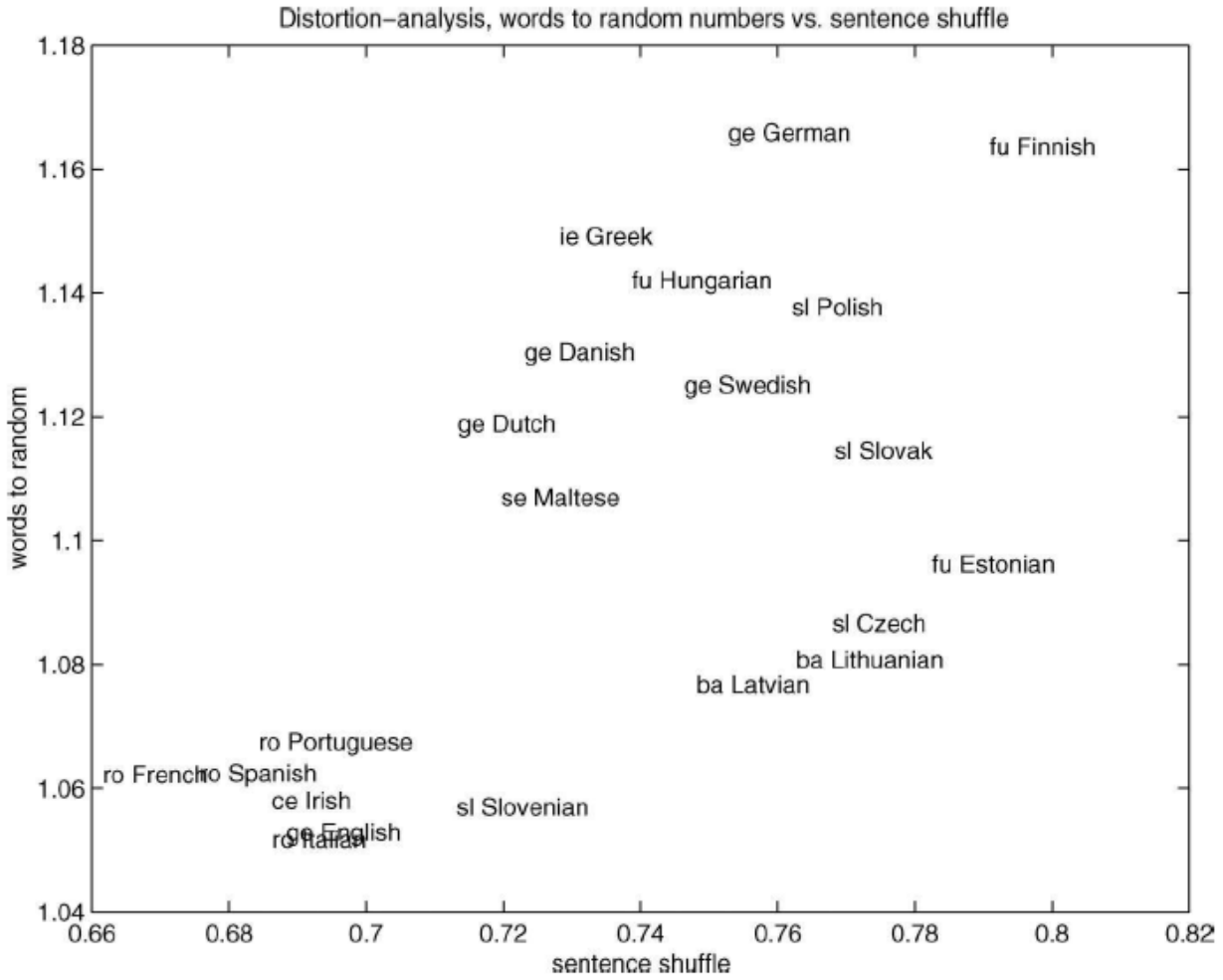
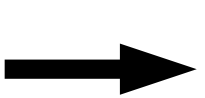


Fig. 4. Language map: morphology vs. word order information.

Markus Sadeniemi, Kimmo Kettunen, Tiina Lindh-Knuutila, and Timo Honkela. Complexity of European Union languages: A comparative approach. Journal of Quantitative Linguistics, 15(2):185–211, 2008.

Language Identification

Smlouva o ústavě pro evropu
Traktat om en forfatning for europa
Vertrag über eine Verfassung für Europa
Συνθήκη για τη θέσπιση Συνταγμάτος Ευρώπης
Treaty establishing a Constitution for Europe
Tratado por el que se establece una constitución para Europa
Euroopa põhiseaduse leping
Sopimus euroopan perustuslaista
Traité établissant une Constitution pour l'Europe
Conradh ag bunú Bunreachta don eoraip
Szerződés európai alkotmány létrehozásáról
Trattato che adotta una Costituzione per l'Europa
Sutartis dėl Konstitucijos Europai
Līgums par konstitūciju eiropai
Trattat Li Jistabbilixxi kostituzzjoni għall-Ewropa
Verdrag tot vaststelling van een grondwet voor europa
Traktat ustanawiają Konstytucję dla europy
Tratado que estabelece uma Constituição para a Europa
Zmluva o ústave pre Európu
Pogodba o ustavi za evropo
Fördrag om upprättande av en konstitution för europa



?

Tommi Vatanen, Jaakko J. Väyrynen, and Sami Virpioja. Language identification of short text segments with n-gram models. In Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), Valletta, Malta, May 2010. European Language Resources Association (ELRA).

Morphology-aware Statistical Machine Translation (1/2)

Aalto University

Philipp Koehn. Europarl: A Parallel Corpus for Statistical Machine Translation. MT Summit 2005.

Source Language	Target Language										
	da	de	el	en	es	fr	fi	it	nl	pt	sv
da	-	18.4	21.1	28.5	26.4	28.7	14.2	22.2	21.4	24.3	28.3
de	22.3	-	20.7	25.3	25.4	27.7	11.8	21.3	23.4	23.2	20.5
el	22.7	17.4	-	27.2	**31.2**	**32.1**	11.4	26.8	20.0	27.6	21.2
en	25.2	17.6	23.2	-	**30.1**	**31.1**	13.0	25.3	21.0	27.1	24.8
es	24.1	18.2	28.3	**30.5**	-	**40.2**	12.5	**32.3**	21.4	**35.9**	23.9
fr	23.7	18.5	26.1	**30.0**	**38.4**	-	12.6	**32.4**	21.1	**35.3**	22.6
fi	20.0	14.5	18.2	21.8	21.1	22.4	-	18.3	17.0	19.1	18.8
it	21.4	16.9	24.8	27.8	**34.0**	**36.0**	11.0	-	20.0	**31.2**	20.2
nl	20.5	18.3	17.4	23.0	22.9	24.6	10.3	20.0	-	20.7	19.0
pt	23.2	18.2	26.4	**30.1**	**37.9**	**39.0**	11.9	**32.0**	20.2	-	21.9
sv	**30.3**	18.9	22.8	**30.2**	28.6	29.7	15.3	23.9	21.9	25.9	-

Table 2: BLEU scores for the 110 translation systems trained on the Europarl corpus

Morphology-aware Statistical Machine Translation (2/2)

Aalto University

Sami Virpioja, André Mansikkaniemi, Jaakko Väyrynen, and Mikko Kurimo. Applying morphological decompositions to statistical machine translation. In Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR, pages 201–206. Association for Computational Linguistics, July 2010.

Sami Virpioja, Jaakko J. Väyrynen, Mathias Creutz, and Markus Sadeniemi. Morphology-aware statistical machine translation based on morphs induced in an unsupervised manner. In Proceedings of the Machine Translation Summit XI, pages 491–498, Copenhagen, Denmark, September 2007.

Adrià de Gispert, Sami Virpioja, Mikko Kurimo, and William Byrne. Minimum bayes risk combination of translation hypotheses from alternative morphological decompositions. In Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers, pages 73–76, Boulder, USA, June 2009. Association for Computational Linguistics.

MT Domain Adaptation

Marcus Dobrinkat and Jaakko J. Väyrynen. Experiments with domain adaptation methods for statistical MT: From European parliament proceedings to Finnish newspaper text. In Proceedings of the 14th Finnish Artificial Intelligence Conference STeP 2010, number 25 in Publications of the Finnish Artificial Intelligence Society, pages 31–38. Finnish Artificial Intelligence Society, 2010.

MT Evaluation

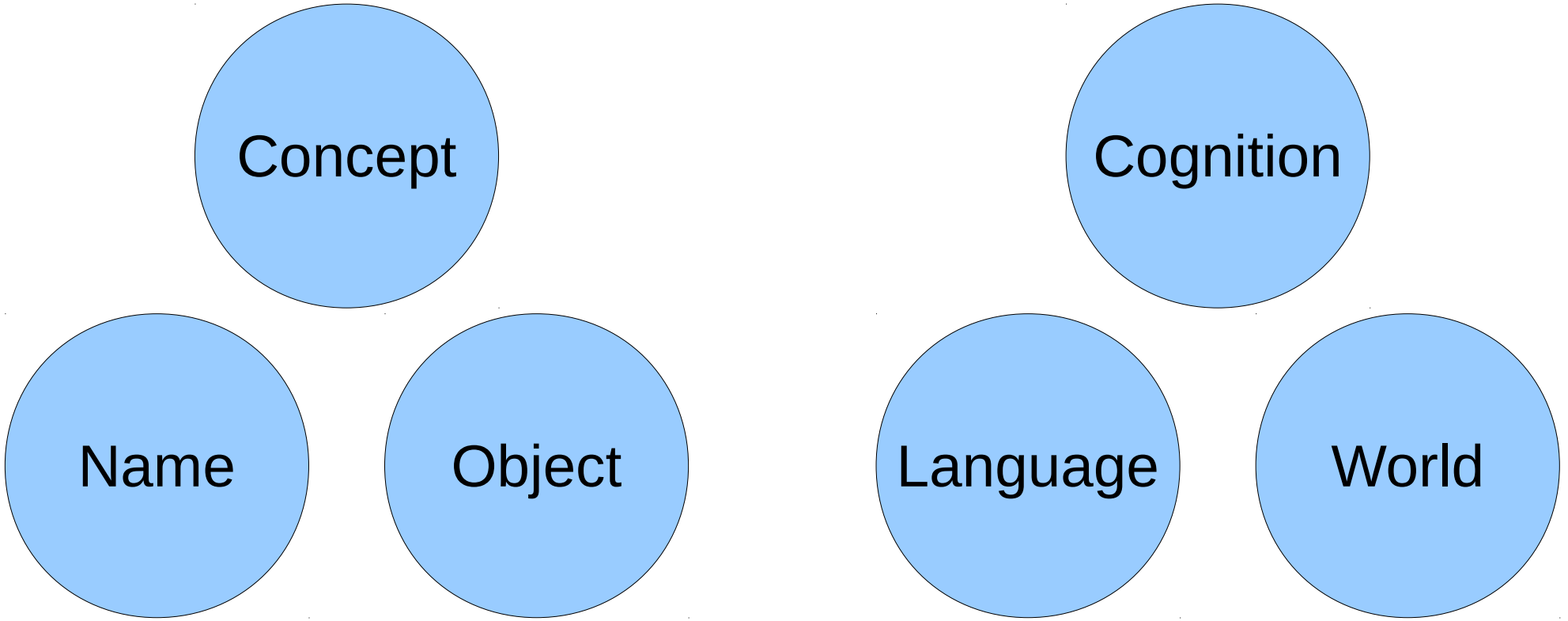
Marcus Dobrinkat, Tero Tapiovaara, Jaakko Väyrynen, and Kimmo Kettunen. Evaluating machine translations using mNCD. In Proceedings of the ACL 2010 Conference Short Papers, pages 80–85. Association for Computational Linguistics, 2010.

Marcus Dobrinkat, Tero Tapiovaara, Jaakko Väyrynen, and Kimmo Kettunen. Normalized compression distance based measures for MetricsMATR 2010. In Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR, pages 343–348. Association for Computational Linguistics, 2010.

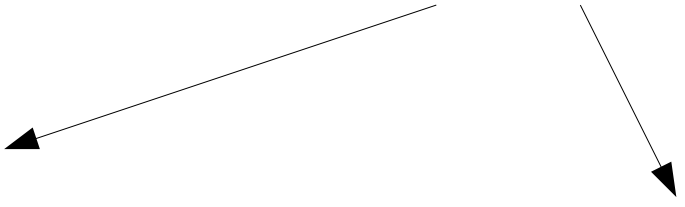
META-NET
Network of Excellence

MultilingualWeb

Concepts

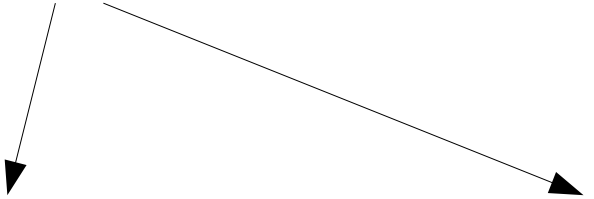


Meaning



Prototypical meaning

Meaning in context



e.g. Semantic web

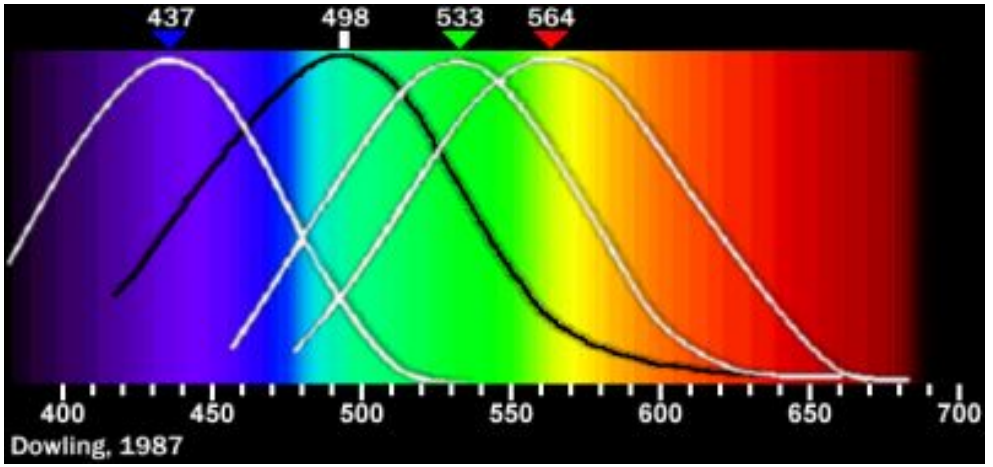
Text context

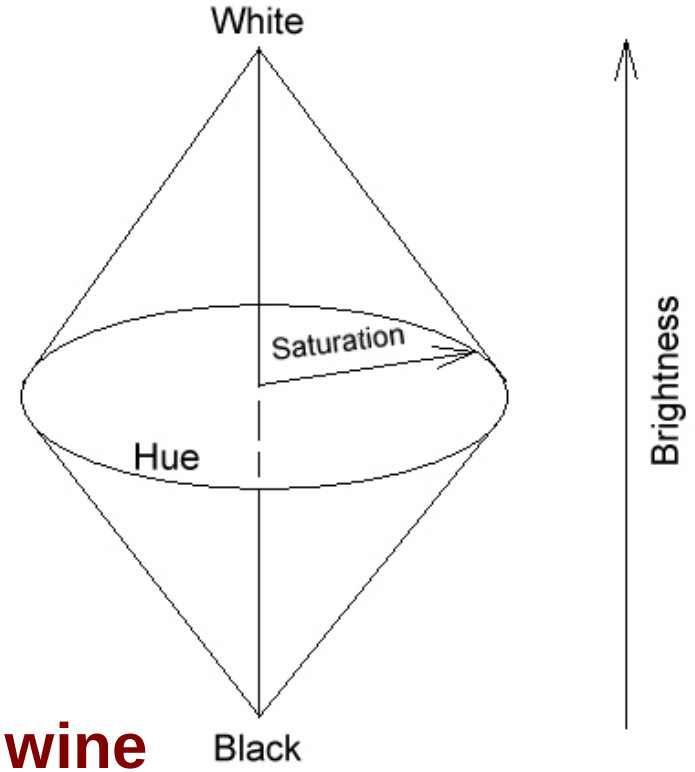
Multimodal context

Eero Hyvönen et al.

Jorma Laaksonen,
Markus Koskela et al.

Subjectivity and contextuality of interpretation





Human vision: rods, cones,...
Physical reasons for color
Contextuality of naming

red wine
red skin
red shirt

→ Hardin

→ Gärdenfors

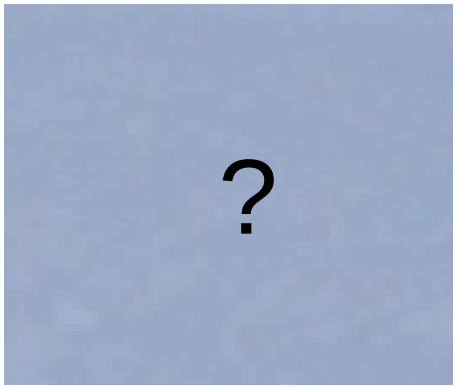
Color naming, example

gray sky

blue water

white snow

blueish cloud



Color naming, example

Aalto University

gray sky

blue water

→ white snow

blueish cloud

Concept Formation and Communication - General Theory

Aalto University

C_i: N-dimensional metric **concept space**

S: symbol space, The **vocabulary** of an agent that consists of discrete symbols

$\lambda : C_i \times C_j \rightarrow R,\ i \neq j$
A **distance** between two points **in** the **concept spaces** of different agents

$\xi: s_i \in S_i \rightarrow C$
An individual **mapping** function **from symbols to concepts**

$\varphi_i: S_i \rightarrow D$
An individual **mapping from** agent i's **vocabulary to** the **signal space** D and an inverse mapping $\varphi^{-1}{}_i$ from the signal space to the symbol space

Observing f_1 and after symbol selection process, agent 1 communicates a symbol s* to agent 2 as signal d. When agent 2 observes d, it maps it to some $s_2 \in S_2$ by using the function $\varphi^{-1}{}_1$. Then it maps the symbol to some point in its concept space by using ξ_2. If this point is close to its observation f_2 in the sense of λ, the communication process has succeeded.

Timo Honkela, Ville Könönen, Tiina Lindh-Knuutila, and Mari-Sanna Paukkeri. Simulating processes of concept formation and communication. Journal of Economic Methodology, 15(3):245–259, 2008.

Bayesian and Self-Organizing Models of Conceptual Systems

Aalto University

Tiina Lindh-Knuutila, Juha Raitio, and Timo Honkela. Combining self-organized and Bayesian models of concept formation. In Julien Mayor, Nicolas Ruh, and Kim Plunkett, editors, Connectionist Models of Behaviour and Cognition II Proceedings of the Eleventh Neural Computation and Psychology Workshop, number 18 in Progress in Neural Processing, pages 193–204. World Scientific, April 2009.

Philosophical Conceptions among Students

Aalto University

Anna-Mari Rusanen, Otto Lappi, Timo Honkela, and Mikael Nederström. Conceptual coherence in philosophy education - visualizing initial conceptions of philosophy students with self-organizing maps. In Proceedings of the 30th Annual Conference of the Cognitive Science Society, pages pp. 64–70, Austin, TX, 2008. Cognitive Science Society.

Pragmatic Web

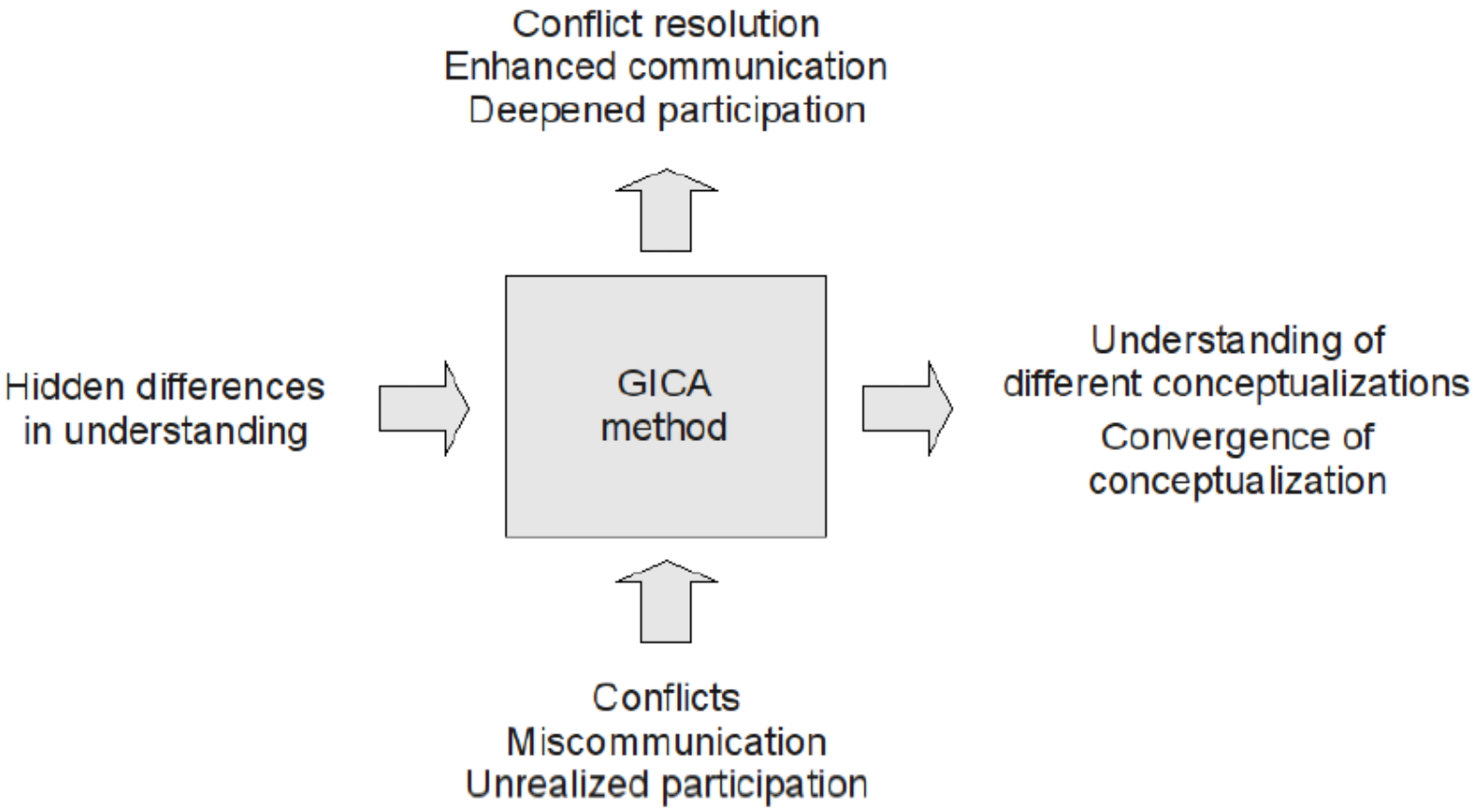
“A potentially highly useful extension of Metaweb is the use of adaptive informatics to build on the communally created, often not mutually compatible resources. If the semantic web involves efforts of top-down knowledge structure and system design, adaptive informatics corresponds to self-organizing processes of technologies that dynamically adapt to evolving requirements of social practices.”

Kai Hakkarainen, Ritva Engeström, Sami Paavola, Pasi Pohjola, and Timo Honkela. Knowledge practices, epistemic technologies, and pragmatic web. In Proceedings of I-KNOW'09 and I-SEMANTICS'09: the 4th AIS SigPrag International Pragmatic Web Conference Track (ICPW 2009), pages 683–694. Verlag der Technischen Universität Graz, 2009.

GICA: Grounded Intersubjective Concept Analysis

Timo Honkela, Nina Janasik, Krista Lagus, Tiina Lindh-Knuutila, Mika Pantzar, and Juha Raitio. GICA: Grounded intersubjective concept analysis - a method for enhancing mutual understanding and participation. Technical Report TKK-ICS-R41, AALTO-ICS, ESPOO, December 2010.

GICA: Grounded Intersubjective Concept Analysis



Why brains?

- What are the central differences between plants and animals?

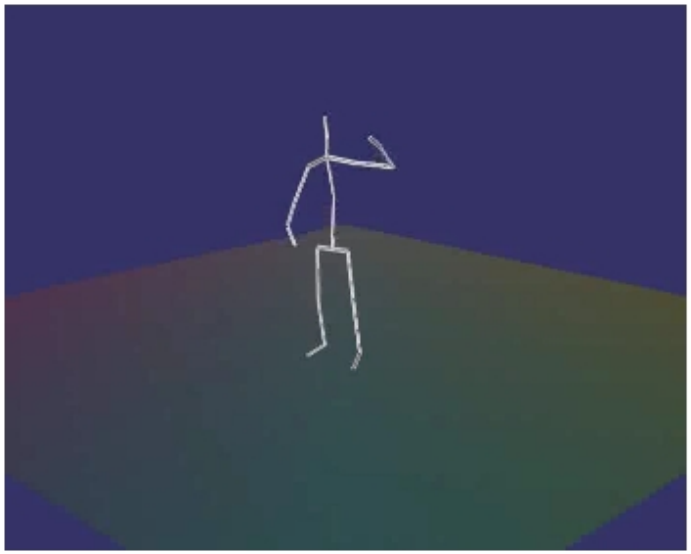
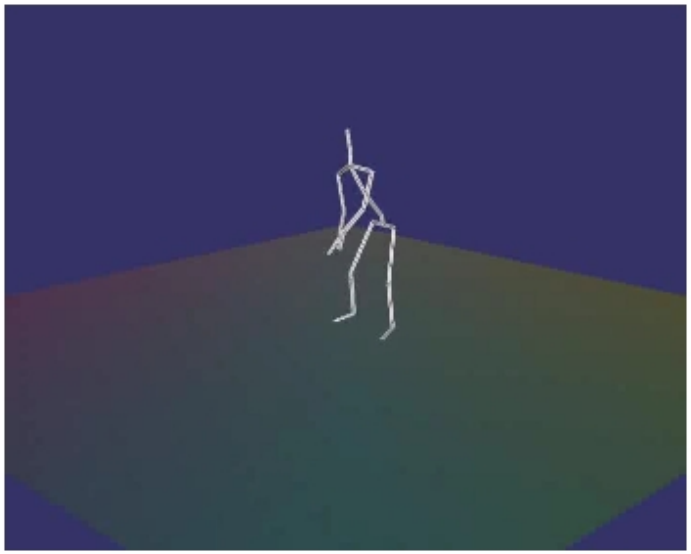
  “The original need for a nervous system was to coordinate **movement**, so an organism could go find food, instead of waiting for the food to come to it.” http://www.fi.edu/learn/brain/

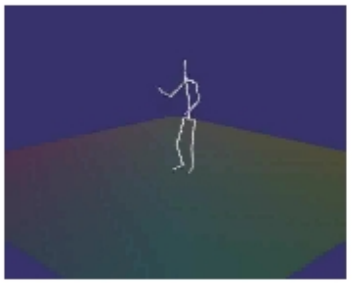
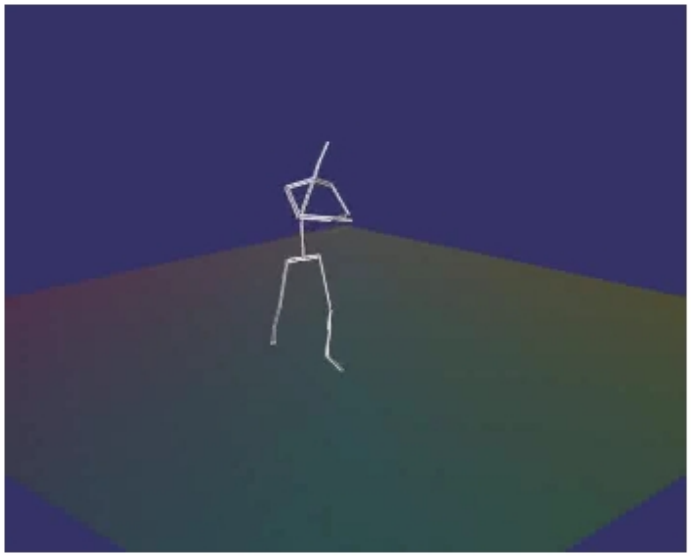
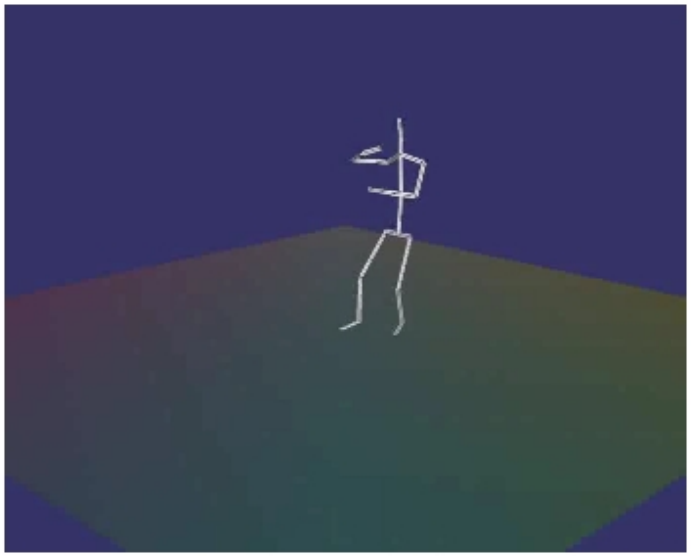
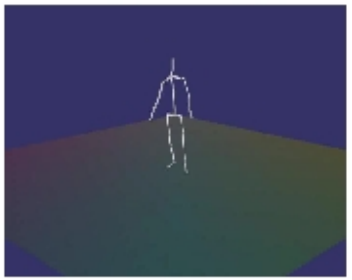
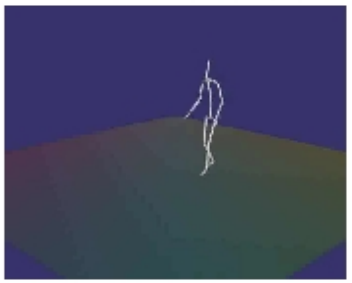
- An extreme example: A sea squirt transforms from an “animal” to a “plant”. It absorbs its own cerebral ganglion that it used to swim about and find its attachment place.

http://goodheartextremescience.wordpress.com/2010/01/27/meet-the-creature-that-eats-its-own-brain/

Project:
Multimodally Grounded Language Technology

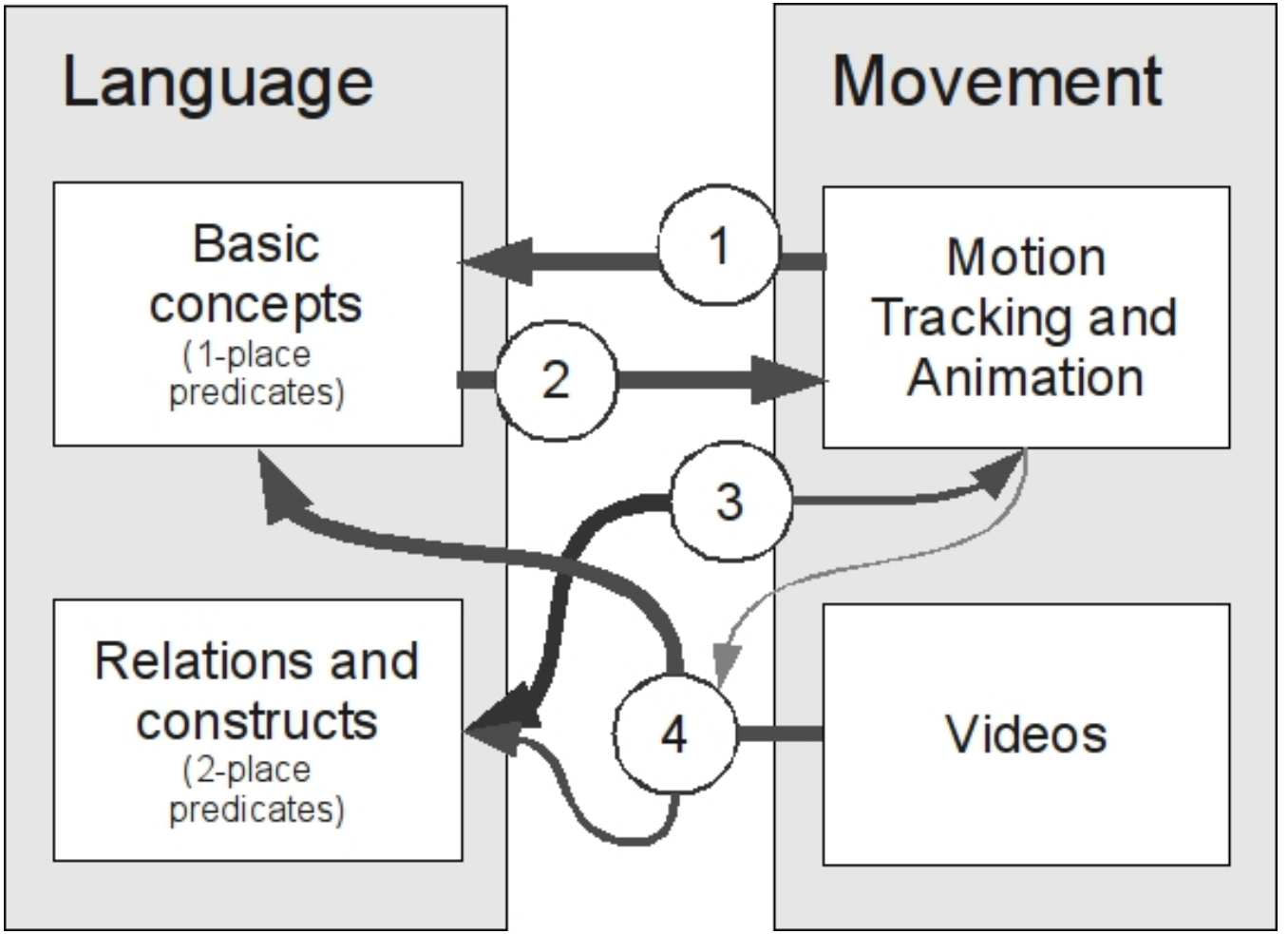
Aalto University

Project:
Multimodally Grounded Language Technology

Aalto University



Poster:

Multimodally Grounded Language Technology

Timo Honkela[1], Oskar Kohonen[1], Klaus Förger[2], Xi Chen[1], Mats Sjöberg[1],Paul Wagner[1,
Jorma Laaksonen[1], Tapio Takala[2], Markus Koskela[1], Krista Lagus[1] and Harri Valpola[1,3]

Aalto University School of Science
[1]Department of Information and Computer Science
[2]Department of Media Technology
[3]ZenRobotics Ltd
Finland

Multimodally Grounded Language Technology is a basic research project funded by Academy of Finland.

Significance of Movement

"The original need for a nervous system was to coordinate **movement**, so an organism could go find food, instead of waiting for the food to come to it."

An extreme example: A sea squirt transforms from an "animal" to a "plant". It absorbs its own cerebral ganglion that it used to swim about and find its attachment place.

Research Objectives

The basic scientific question behind the project is how to computationally model the interrelated processes of understanding natural language and perceiving and producing movement in multimodal real world contexts. Movement is a fundamental part of human activities that ground our understanding of the world.

We are developing methods and technologies to automatically associate human movements detected by motion capture and in video sequences with their linguistic descriptions. When the association between human movement

Welcome to see the posters after the presentations!

Aalto University

Social Dimension

Analysing and modeling human social activities and their results

Map of Finnish Science

Aalto University

Main references:

Honkela, T., and Klami, M. Suomen akatemialle osoitettujen hakemuksien tekstilouhinta [Text mining of applications submitted to the Academy of Finland]. Unpublished report, Helsinki University of Technology Espoo, Finland, 2007.

Mikaela Klami and Timo Honkela. Self-organized ordering of terms and documents in NSF awards data. In Proceedings of WSOM'07. University Library of Bielefeld, 2007.

See also:

Timo Honkela. Kunnioitusta erilaisuutta kohtaan. Tieteessä tapahtuu, 29(4-5):38–39, 2011.

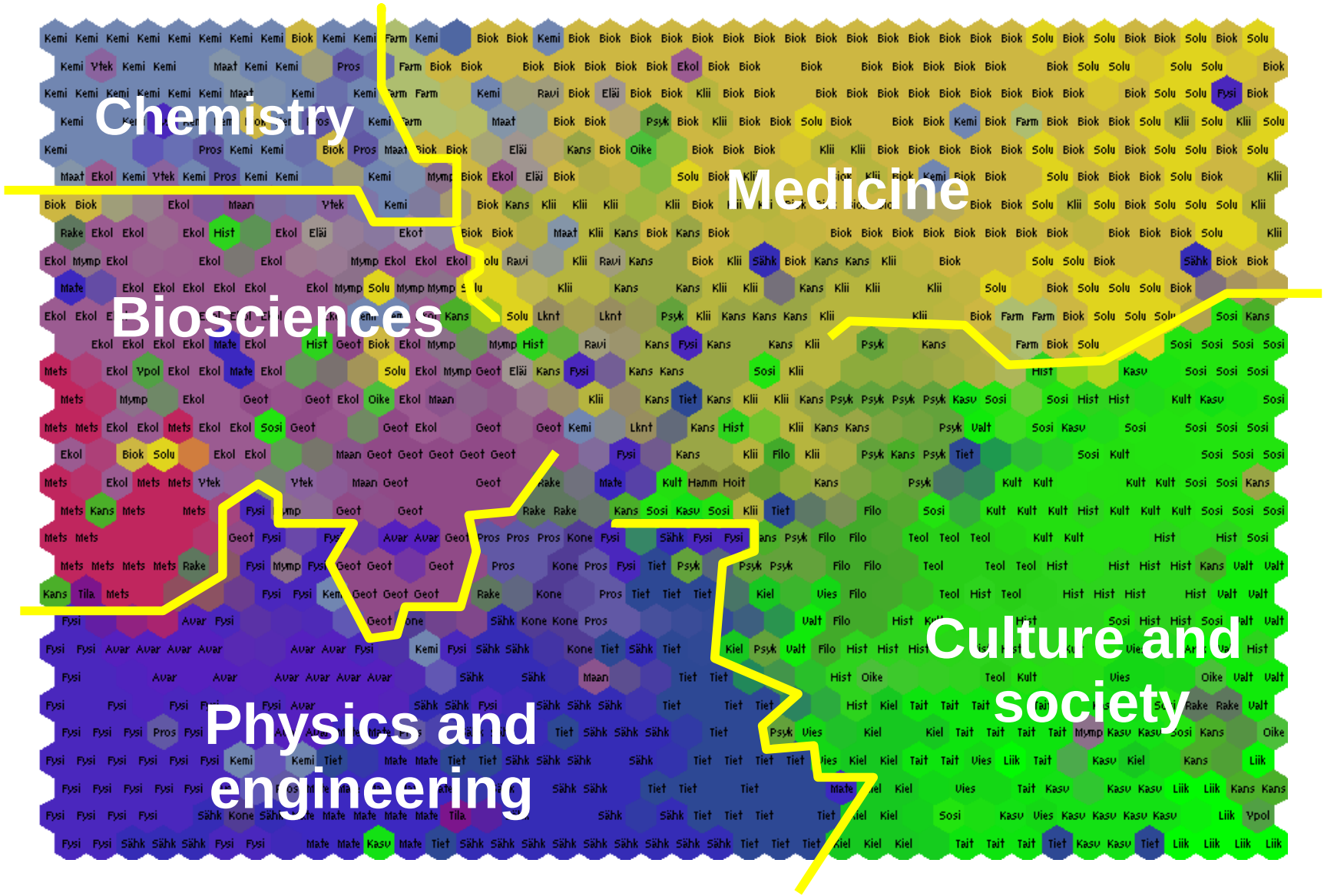
Nina Janasik, Timo Honkela, and Henrik Bruun. Text mining in qualitative research: Application of an unsupervised learning method. Organizational Research Methods, 12(3):436–460, 2009.

Mari-Sanna Paukkeri, Ilari T. Nieminen, Matti Pöllä, and Timo Honkela.
A language-independent approach to keyphrase extraction and evaluation. In Coling 2008: Companion volume: Posters, pages 83–86, Manchester, UK, August 2008. Coling 2008 Organizing Committee.

Matti Pöllä, Timo Honkela, Henrik Bruun, and Ann Russell. Analysis of interdisciplinary text corpora. In Proceedings of Ninth Scandinavian Conference on Artificial Intelligence (SCAI 2006), pages 17–22, Helsinki, Finland, October 2006.

Teuvo Kohonen. Self-Organizing Maps. Springer, 2001.

Map of Finnish Science



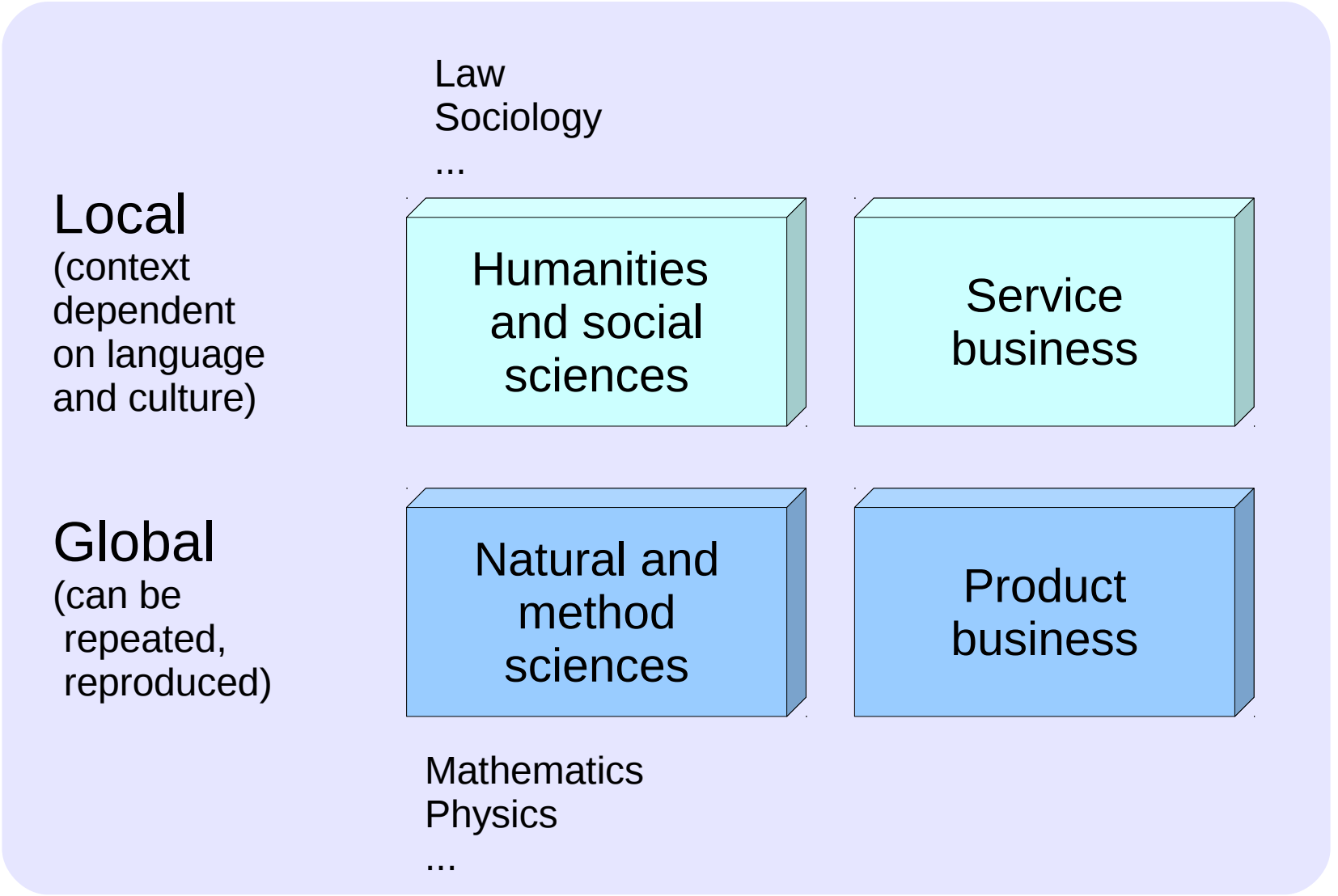
Analysis of Language Use

Matematiikka	95.3
Farmasia	94.1
Kemia	93.7
Fysiikka	93.4
Biokemia, molekyylibiologia, mikrobiologia, perinnöllisyystiede ja biotekniikka	93.4
Solu- ja kehitysbiologia, fysiologia ja ekofysiologia	93.4
Tietojenkäsittelytieteet	93.0
Sähkötekniikka ja elektroniikka	92.8
Ympäristötekniikka	92.7
Geotieteet	92.1
Ekologia, evoluutiotutkimus ja systematiikka	92.1
Kone- ja valmistustekniikka	91.9
Metsätieteet	91.4
Avaruustieteet ja tähtitiede	91.0
Prosessi- ja materiaalitekniikka	90.8
Tilastotiede	90.7
Muu ympäristön ja luonnonvarojen tutkimus	90.1
Kliininen lääketiede	89.6
Ekotoksikologia, ympäristön tila ja ympäristövaikutukset	89.5
Ravitsemustiede	89.3
Psykologia	89.0
Liikuntatiede	88.9
Hoitotiede	88.9
Eläinlääketiede	88.5

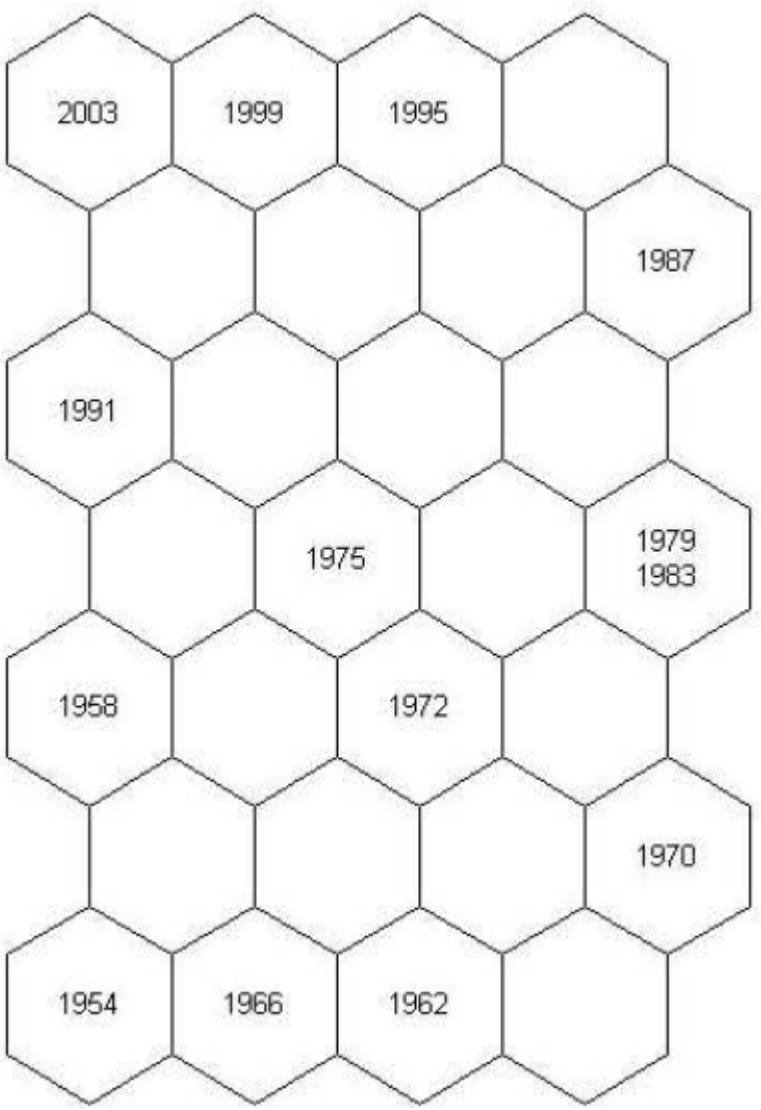
Prosessi- ja materiaalitekniikka	90.8
Tilastotiede	90.7
Muu ympäristön ja luonnonvarojen tutkimus	90.1
Kliininen lääketiede	89.6
Ekotoksikologia, ympäristön tila ja ympäristövaikutukset	89.5
Ravitsemustiede	89.3
Psykologia	89.0
Liikuntatiede	88.9
Hoitotiede	88.9
Eläinlääketiede	88.5
Kansanterveystiede	88.1
Kielitieteet	87.6
Filosofia	87.3
Liiketaloustiede, talousmaantiede ja tuotantotalous	87.2
Hammaslääketiede	86.7
Kansantaloustiede	86.3
Rakennus- ja yhdyskuntatekniikka	85.9
Maatalous- ja elintarviketieteet	85.4
Ympäristöpolitiikka, -talous ja -oikeus	85.3
Maantiede	84.8
Arkkitehtuuri ja teollinen muotoilu	83.7
Viestintä- ja informaatiotieteet	83.1
Kasvatustiede	82.6
Valtio-oppi ja hallintotiede	82.2
Taiteiden tutkimus	81.6
Sosiaalitieteet	80.4
Kulttuurien tutkimus	79.3
Historia ja arkeologia	78.1
Teologia	77.0
Oikeustiede	70.8

Aalto University

Comparison between different kinds of research and business

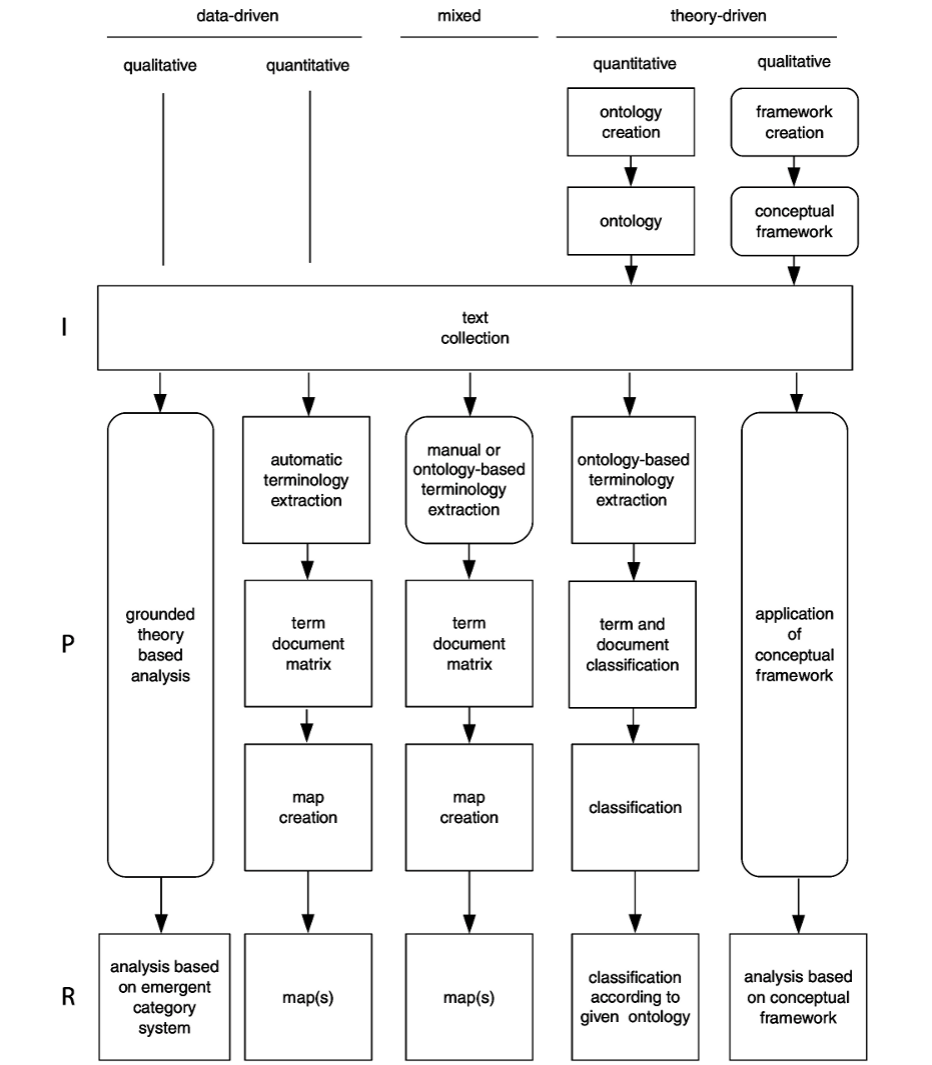


Analysis of parliamentary results



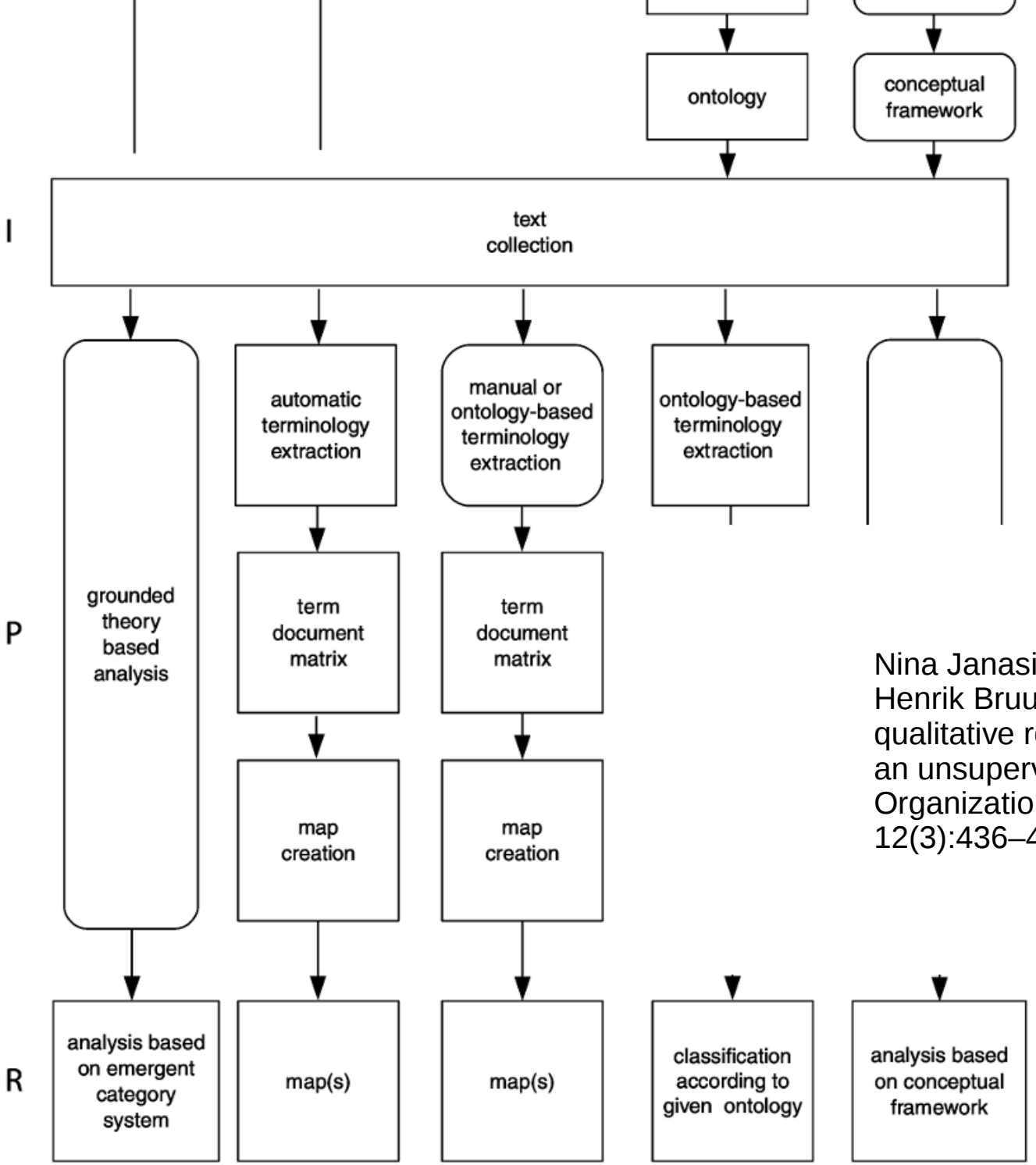
Pyry Niemelä and Timo Honkela. Analysis of parliamentary election results and socio-economic situation using self-organizing map. In Proceedings of WSOM'09, pages 209–218, 2009.

Text Mining for Qualitative Research



Note: I = input; P = process; R = result.

Nina Janasik, Timo Honkela, and Henrik Bruun. Text mining in qualitative research: Application of an unsupervised learning method. Organizational Research Methods, 12(3):436–460, 2009.

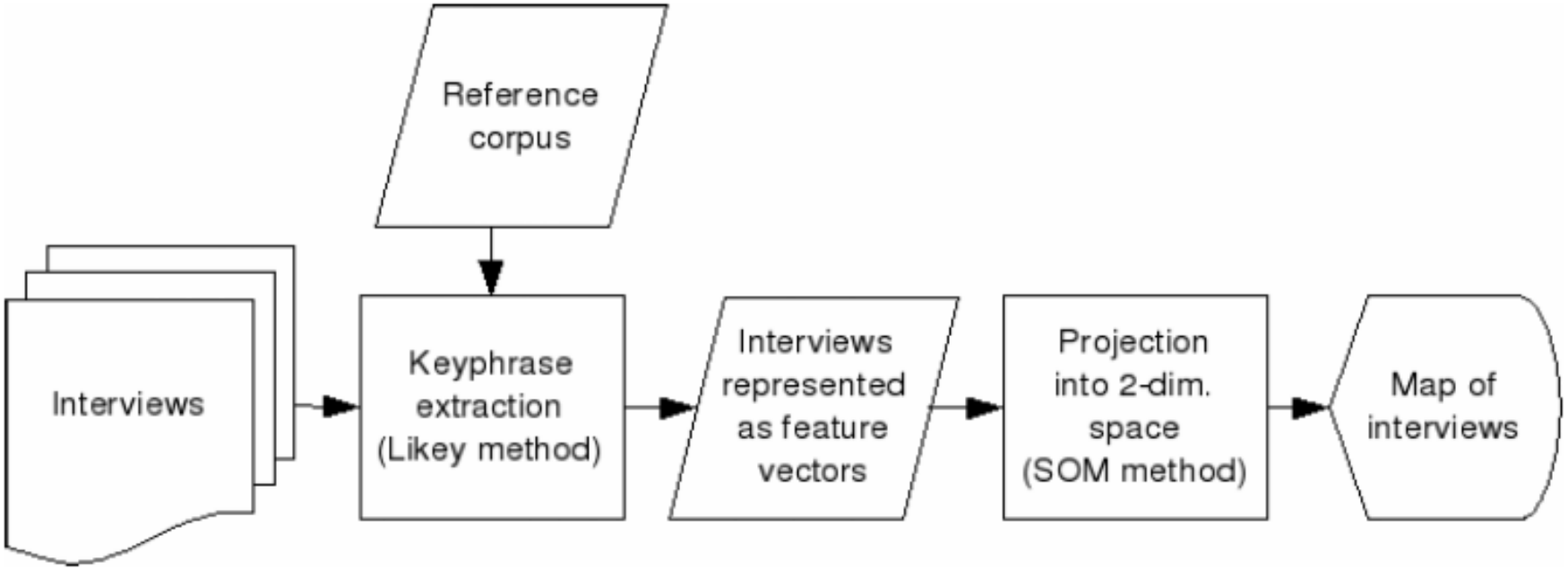


Nina Janasik, Timo Honkela, and Henrik Bruun. Text mining in qualitative research: Application of an unsupervised learning method. Organizational Research Methods, 12(3):436–460, 2009.

Note: I = input; P = process; R = result.

Aalto University

Analysis of Swedish ICT History



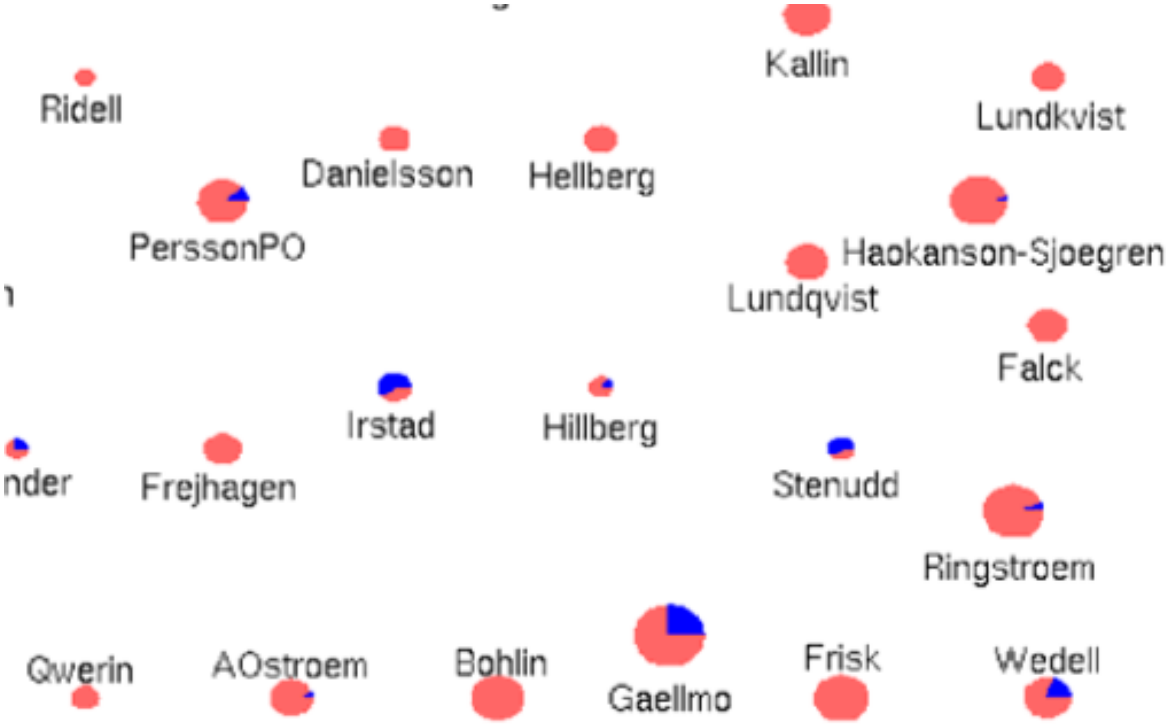
Petri Paju, Eric Malmi, and Timo Honkela. Text Mining and Qualitative Analysis of an IT History Interview Collection. In History of Nordic Computing 3, IFIP Publications, pages 433–443. Springer, 2011.

Analysis of Swedish ICT History

IBM

versus

Ericsson



"We warmly thank Prof. Reijo (Shosta) Sulonen who with his vision brought the authors of this paper together. His discussions with Prof. Yrjö Neuvo originally generated the basic ideas behind this study."

Kulta project:
Modeling Changing Needs of Consumers

Tanja Kotro and Mari-Sanna Paukkeri (2009) Micro democracy - Enhancing Openness in Innovation in an Organizational Context. In Proceedings of OPEN 2009. Helsinki, Finland, October 2009.

Mari-Sanna Paukkeri and Tanja Kotro (2009) Framework for Analyzing and Clustering Short Message Database of Ideas. In Proceedings of I-KNOW'09,the 9th International Conference on Knowledge Management and Knowledge Technologies. Graz, Austria, September 2009.

Mari-Sanna Paukkeri and Aleksi Neuvonen (2007) Privacy in Data Mining of Social Practices (abstract). In Proceedings of The Third Finnish Conference on Cultural and Activity Research (FISCAR'07), pp. 78. University of Helsinki, Finland, September 2007.

Eric Malmi, Juha Raitio, and Timo Honkela. Modeling practice diffusion with an agent-based social simulation framework. In Proceedings of the 6th European Social Simulation Association Conference, ESSA 2009, page 53, Guildford, U.K., September 2009. Extended abstract.

Lasse Lindqvist, Timo Honkela, and Mika Pantzar. Visualizing practice theory through a simulation model. Technical Report E9, Helsinki University of Technology, 2007.

Aalto University

VirtualCoach project: Wellbeing Informatics

Presentation by Krista Lagus

Future

Some institutional form; research area of computational humanities 2015-

Department of Information and Computer Science 2012-2014