

FINMT

Machine Translation for Morphologically-Rich Languages

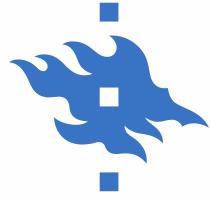
Department of Modern Languages

University of Helsinki

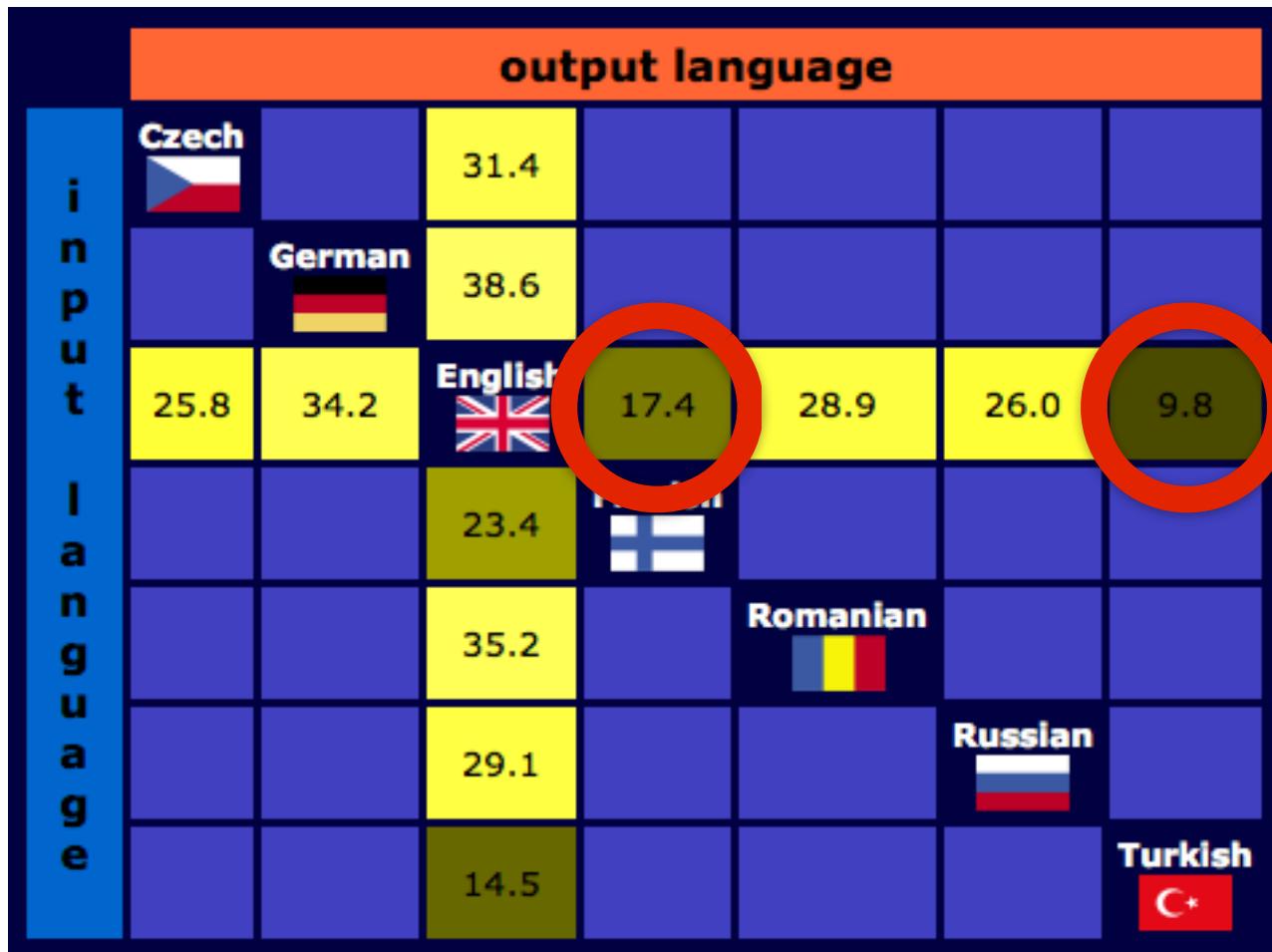
September 12, 2016

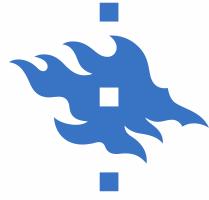
Why are we here?



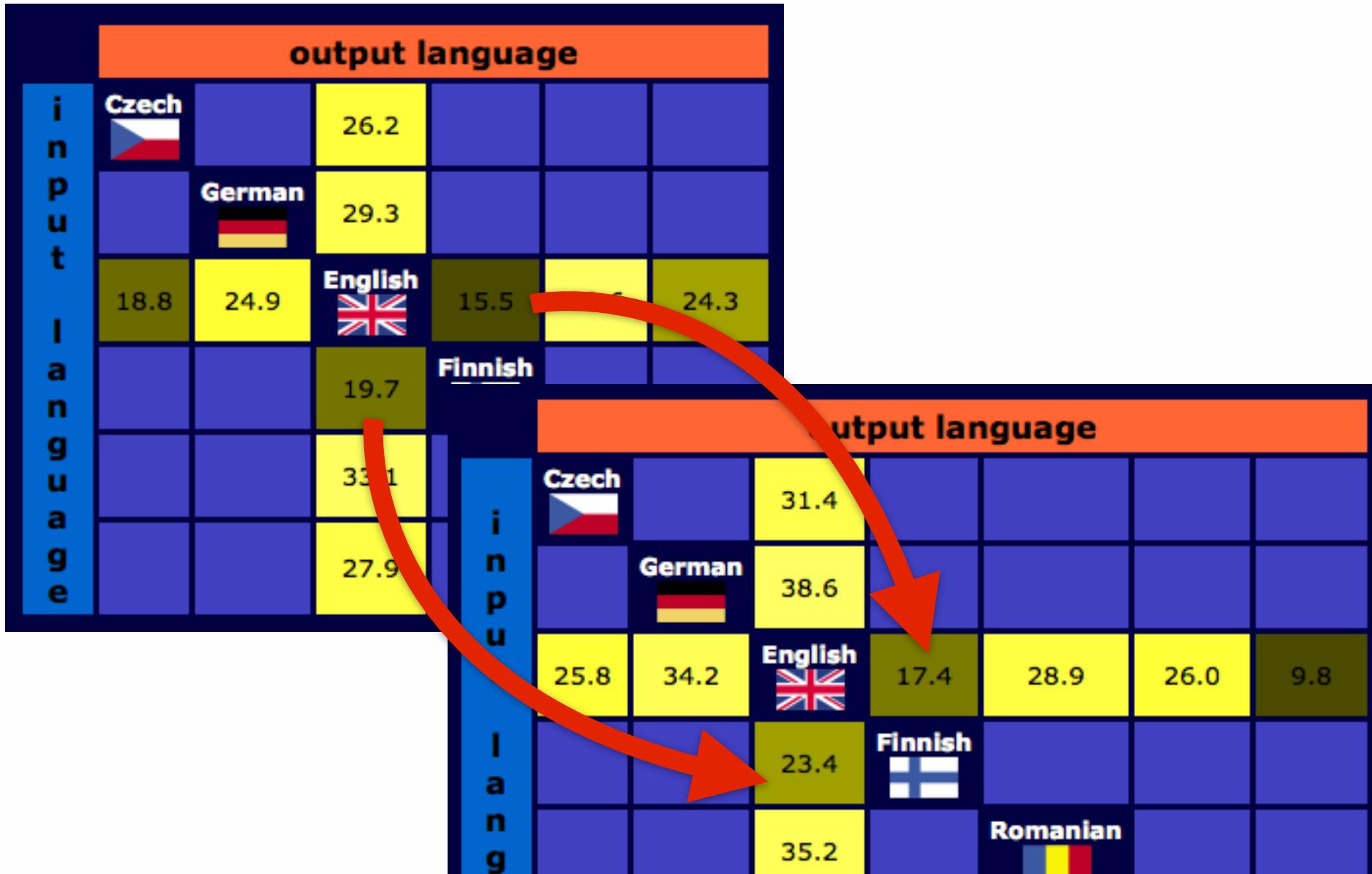


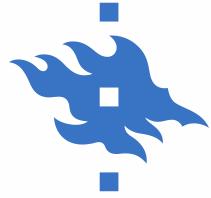
WMT 2016 ...



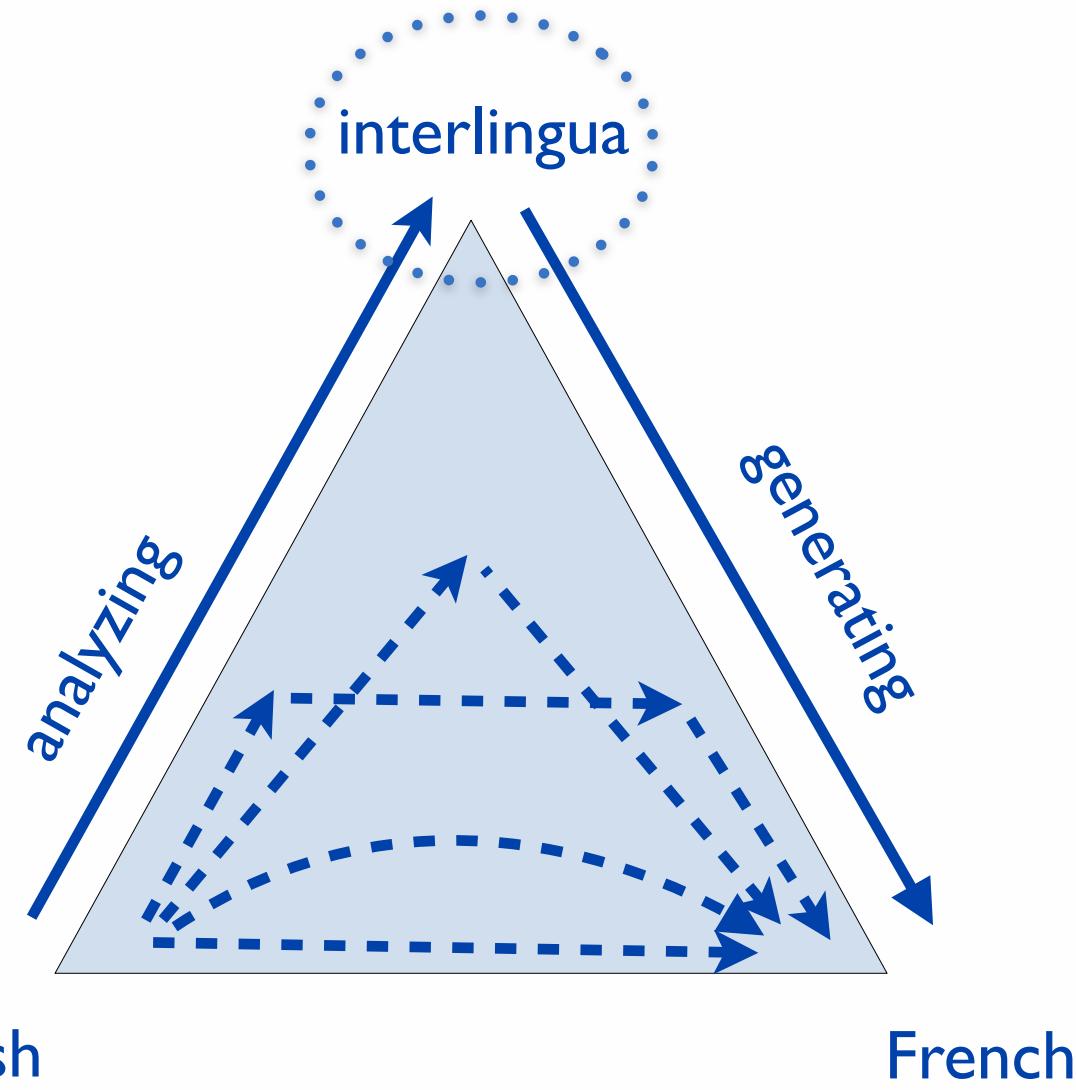


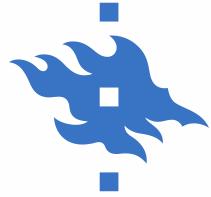
There Is Hope



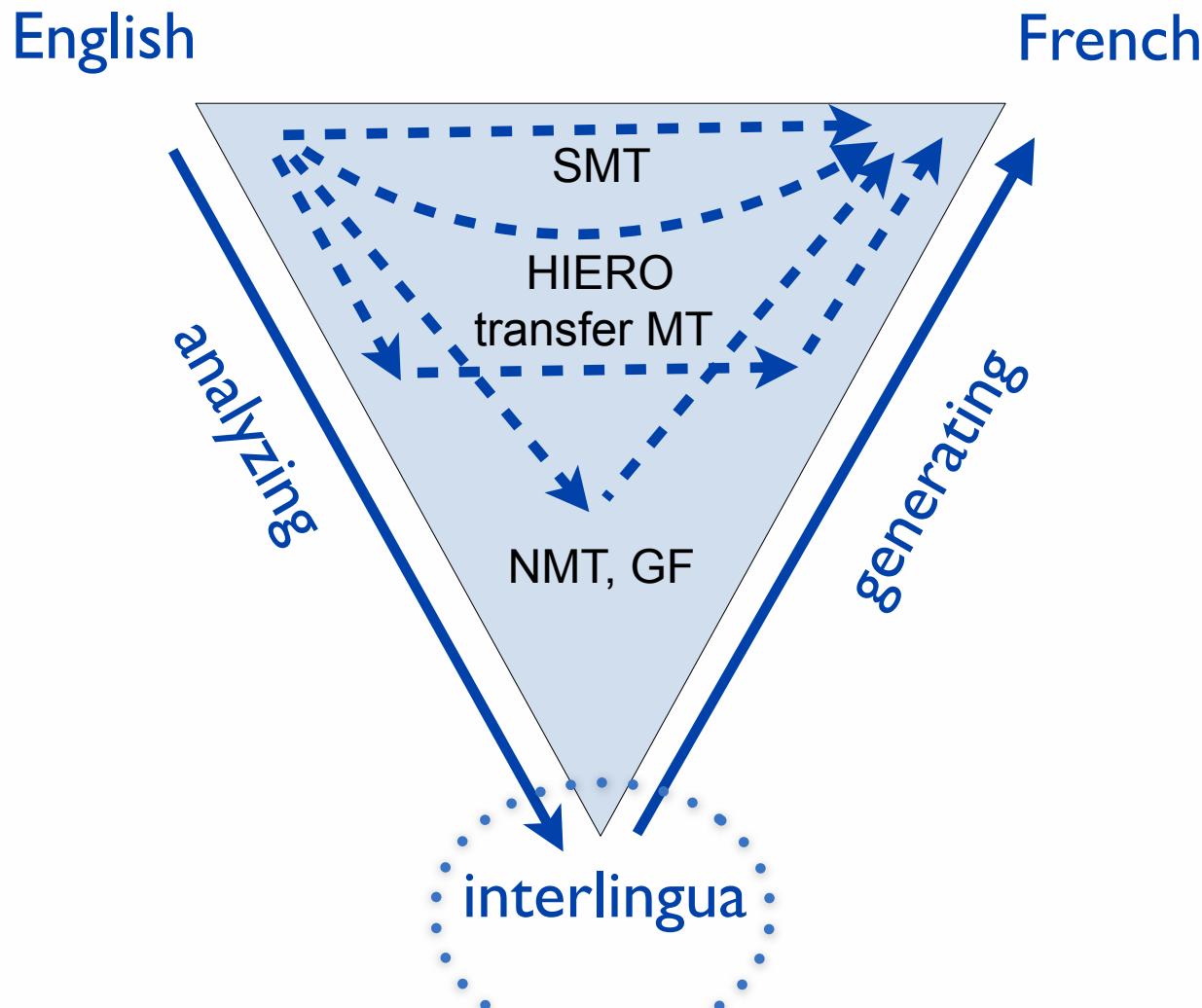


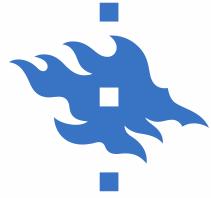
Vauquois Triangle



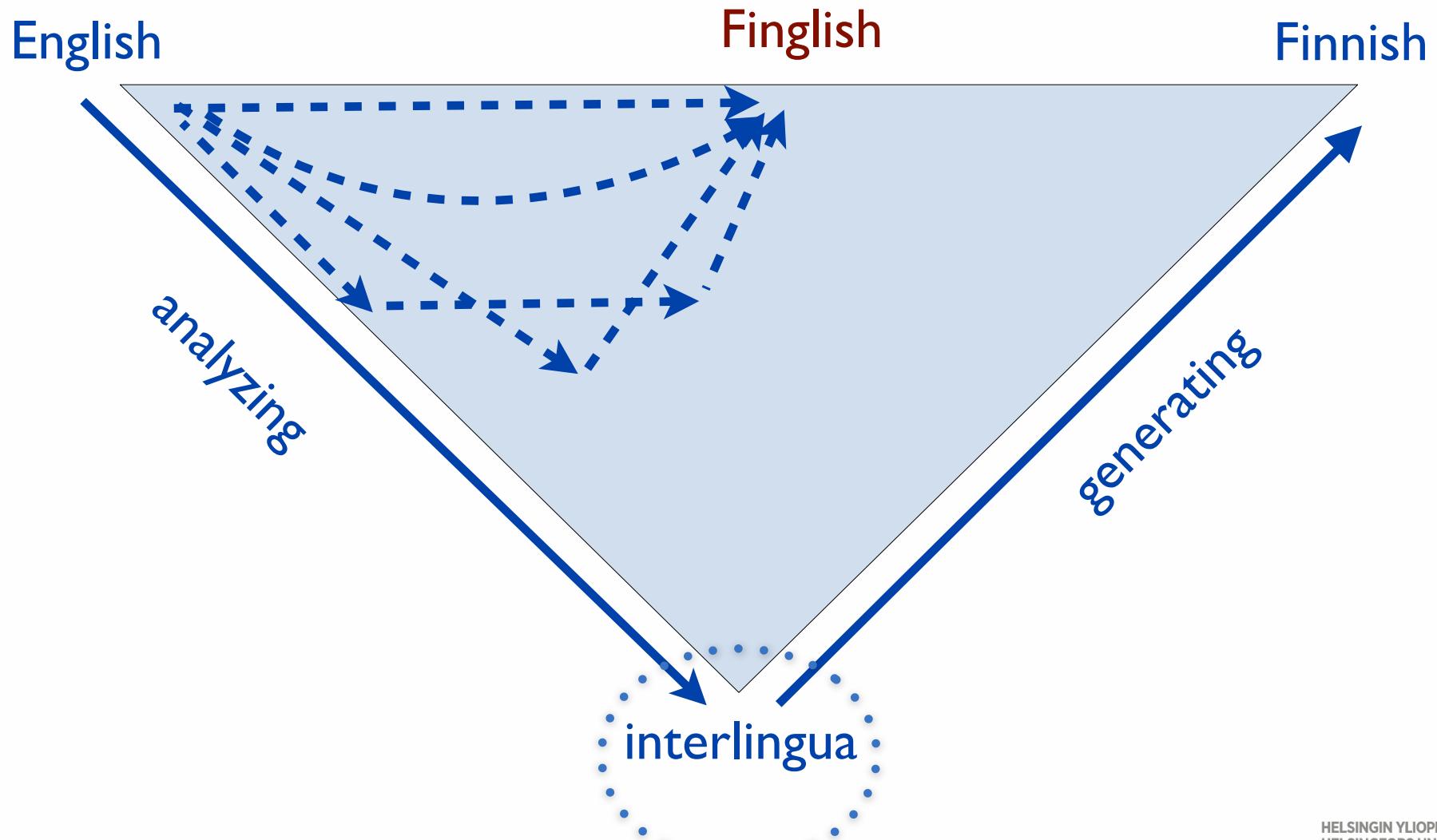


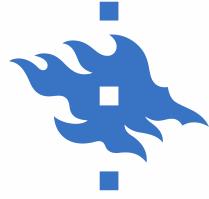
Vauquois Triangle





Vauquois Triangle





We Have To Talk About Better Models

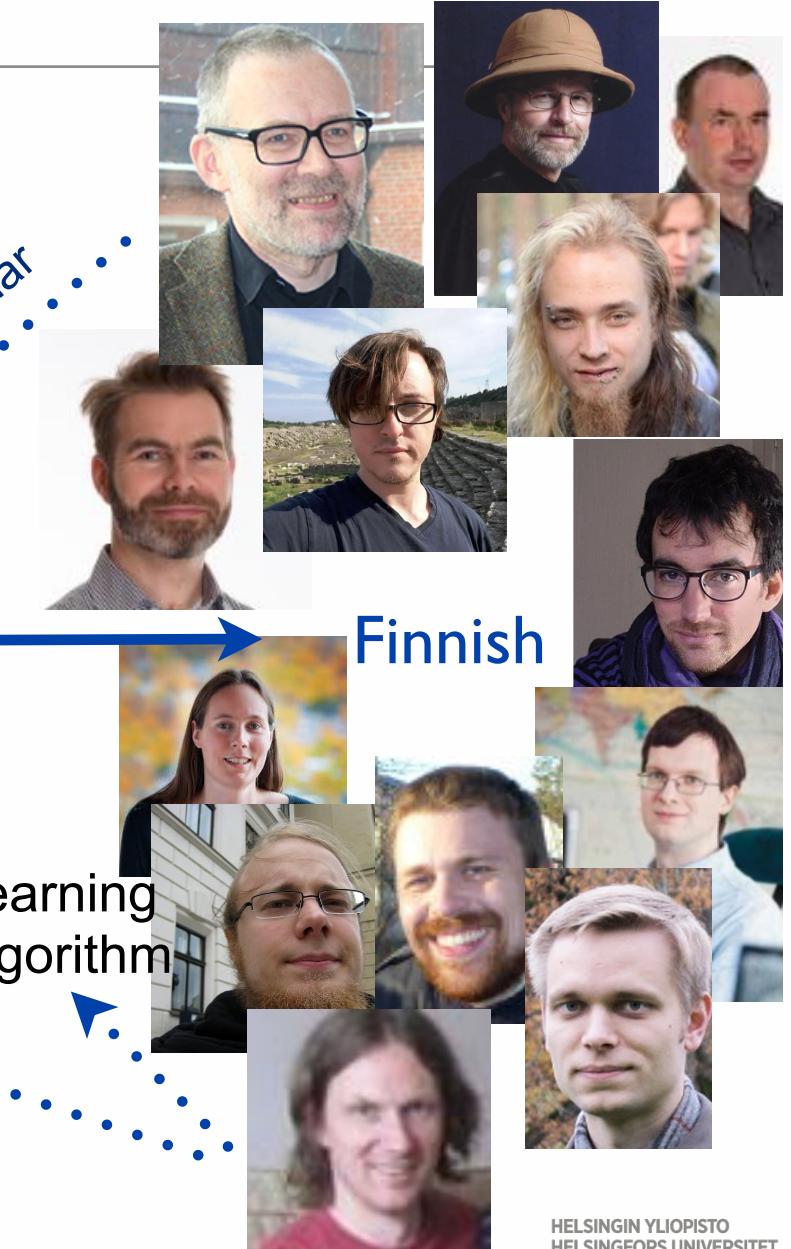
Expert-driven
(instruction-based)

English



MT

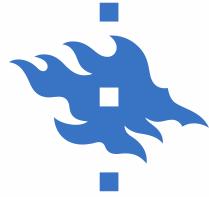
grammar



(example-based)
Data-driven

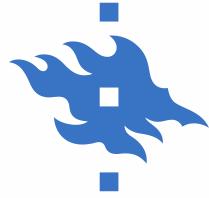
data

learning
algorithm



Workshop Program (Morning)

- 09:30 – 10:00 Coffee
- 10:00 – 10:15 Welcome and workshop goals
- 10:15 – 10:30 Short introductions of participants
- 10:30 – 11:20 Current trends and tricks of the trade in NMT
- 11:30 – 12:30 MT models for Finnish
 - MT in Helsinki/Turku/Uppsala
 - MT at Aalto
 - Experiences from AbuMaTran
 - Helsinki Neural MT
- *Lunch (at [Restaurant Piano](#), Rauhankatu 15)*

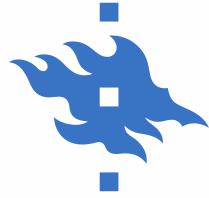


Workshop Program (Afternoon)

- 13:30 – 15:00 Tools, data and resources
 - MT at LingSoft, MT at Yandex
 - GF and Machine Translation
 - Using oMorfi with Apertium and Moses
 - Tools for highly-inflecting languages
 - Mining parallel data from incomparable corpora
- *Coffee*
- 15:30 – 16:30 Tuning, optimising and evaluation
 - Tuning sequence models
 - MT quality and evaluation for Finnish
 - MT from a customer point of view
 - Discussion
- 16:30 – 17:30 Crazy ideas, plans, collaborations
- *Drinks and food (at [Kolme Kruunua](#), Liisankatu 5)*

Welcome to Helsinki!





Introduction of Participants

University of Helsinki

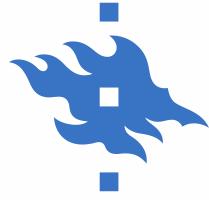
Arvi Hurskainen
Emily Öhman
Jörg Tiedemann
Jyrki Niemi
Krister Lindén
Maarit Koponen
Mathias Creutz
Robert Östling
Senka Drobac
Shanshan Wang
Tuija Kinnunen

University of Turku

Filip Ginter
Jenna Kanerva
Juhani Luotolahti
Hans Moen

Aalto University

Sami Virpioja
Stig-Arne Grönroos



Introduction of Participants

Uppsala University

Sara Stymne

Kvensk Institutt

Sindre Reino Trosterud

University of Gothenburg

Aarne Ranta
Pashant Kolachina

Dublin City University

Antonio Toral

University of Hamburg

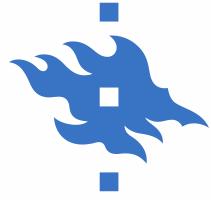
Tommi A Pirinen

University of Edinburgh

Barry Haddow

University of Tromsø

Trond Trosterud
Francis Tyers
Sjur Moshagen



Introduction of Participants

AAC Global Oy

Jussi Rautio

Lingsoft

Sami Virpioja

Tommi Nieminen

Prime Minister's Office

Simio Kankkunen

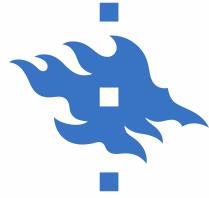
Utopia Analytics

Oskar Kohonen

Yandex

Alexey Baytin

Irina Galinskaya



Discussions I

Frameworks for comparing approaches and implementations

- lots of NMT frameworks
- lots of new models

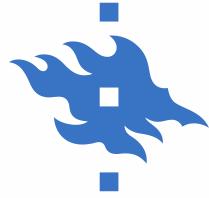
Tricks of the trade, tutorials, trends, resources

- NMT Wiki? Shared resources
- (N)MT Marathon in Helsinki?

Evaluation for morphologically-rich languages

- metrics beyond BLEU
- inflection variants and free word order

Collection of more (clean and noisy) data



Discussions II

Collaboration ideas

- pushing performance for Finnish
- NMT + SMT + expert MT
- industry and academia

Funding and activities

- regular meetings / workshops?
- project applications?
- EU support?