



UNIVERSITETET I OSLO
DET HUMANISTISKE FAKULTET

Presentasjon av en flerspråklig database med idiomatiske uttrykk

Atle Grønn, prosjektleder RuN, ILOS, UiO

Internasjonal oversetterkonferanse

Klækken 26. mai 2009



Der russisk møter norsk – i språkets grenseland



UNIVERSITETET I OSLO
DET HUMANISTISKE FAKULTET

Where Russian meets Norwegian – languages at the
interfaces

<http://www.hf.uio.no/ilos/forskning/forskningsprosjekter/run/index.html>

Встреча русского с норвежским – языки в
пересечении

<http://www.hf.uio.no/ilos/forskning/forskningsprosjekter/run/russian/index.html>



1: Plan

- 2: om (parallell)korpora
- 3: OMC og RuN
 - 3.1: leksikalske eksempler
 - 3.2: idiomatiske uttrykk
 - 3.3: grammatiske konstruksjoner
- 4: andre anvendelser



2: Korpora

Elektroniske (søkbare) korpora:

- representative data
- metodologisk verktøy for lingvister
- grammatikk & leksikografi



2: Korpora (forts.) enspråklig vs. flerspråklig

- Største enspråklig korpus for engelsk: "google"
- Største enspråklig korpus for norsk: "google"
- Parallellkorpora: forfatter+oversetter som lingvistens informanter.
 - Fordel: kvalitetssikring av data
 - Ulempe: kunstnerisk frihet hos oversetteren (nøytraliseres ved store datamengder)



2: (forts.)

Preview: "miljø" i engelsk/fransk

eksempel 1

Maihaugen, Cappelen, 1989.	kopperstikk og malerier forteller om et fremmedartet miljø
	Movable furniture , chairs , copperplate engravings and paintings all evoke a foreign milieu .
	Des meubles amovibles , des sièges , des estampes et des peintures parlent d' un milieu étranger .



2: (forts.)

Preview: "miljø" i engelsk/fransk

<i>eksempel 2</i>	
Jostein Gaarder, "Maya", Aschehoug, 1999.	sette Ana og José inn i et miljø som de i min erindring alltid vil
	to place Ana and José in an environment with which , as far as my memory is concerned , they will always be irrevocably linked .
	replacer Ana et José dans un contexte auquel ils resteront , dans mon souvenir , à jamais indissociablement liés .



2: (forts.) Hvorfor parallellkorpus?

- Kontrastive språkstudier: ny forståelse av enkeltspråk og forhold mellom språk
- Nye hypoteser om språk – særlig grammatikk – på bakgrunn av kvantitative data
- Praktisk nytte (studenter, oversettere osv.)



2: (forts.) Parallellkorpora ved UiO

- OMC – Oslo Multilingual Corpora

(norsk, engelsk, tysk, fransk ...)

- RuN

(norsk, russisk, engelsk ...)

- Felles grensesnitt: "Glossa" utviklet av Tekstlaboratoriet, UiO (Lars Nygaard).



2: (forts.) OMC – bakgrunn

- 1990-tallet: *The English-Norwegian Parallel Corpus*. Utviklet av Stig Johansson, Knut Hofland og andre.
- Slutten av 90-tallet: fransk, tysk, nederlandsk, portugisisk ...
- Resultat: Oslo Multilingual Corpus
- 2008-2010: RuN-korpuset (russisk)



2: (forts.) RuN-OMC - omfang

- Norsk: 4 millioner ord
- Engelsk: 2,4 millioner ord
- Tysk: 2 millioner ord
- Russisk: 1 million ord
- Fransk: 1 million ord
- Portugisisk: 212 000 ord
- Nederlandsk: 168 000 ord



2: (forts.) Overvekt av skjønnlitteratur

- Bergljot Hobækk Haff
- Bjørg Vik
- Christiane F.
- Doris Lessing
- Erik Fosnes Hansen
- Edvard Hoem
- Ebba Haslund
- Fredrick Forsyth
- Fay Weldon
- Günter Grass
- Gunnar Staalesen
- Lev Tolstoj

OSV. OSV. OSV....



2: (forts.) subkorpus på 5 parallelstilte språk

Eksempel på 5 språk fra Grønn & von Stechow (2008):

Hadde ikke Anders vært i Lofoten, så **hadde** han nok sagt et ord og gitt ham gode råd.

Had Anders not been in the Lofoten Islands, he **would** surely have given his brother good advice.

Wäre Anders nicht auf den Lofoten gewesen, dann **hätte** er wohl ein Wort gesagt und ihm einen guten Rat gegeben.

Si Anders n' **avait** pas été aux Lofoten, il y **aurait** probablement mis son mot et donné quelques bons conseils.

Если бы Андерс не был сейчас на Лофотенах, он **мог бы** дать Нильсу дельный совет.



2: (forts.) Tekniske spesifikasjoner

- Parallellstilling på setningsnivå (semi-automatisk)
- Grammatisk tagging på ordnivå (automatisk)



3a: leksikalske eksempler

eksempel "miljø" oversatt/ikke oversatt med "environment"

"Norge fra
innsiden",
Cappelen,
1986.

Isolasjonen forklarer også en
annen særegenhet ved norsk
miljø

This isolation explains another
peculiarly Norwegian **tradition**



3a: (forts.)

Eksempel på begrenset søk: liksom

<p>Ebba Haslund, "Det hendte ingenting", Aschehoug, 1981.</p>	<p>Så liketil og grei liksom</p>
	<p>Seems so open and natural somehow</p>

3a: (forts.) ”småprate; små.*”

Eksempel på lemmasøk eller trunkert søk

Jostein Gaarder, Maya, Aschehoug, 1999.	for alt jeg visste bare småpratet de om at den ene hadde vondt.
	For all I knew they might just have been chatting about a stomach ache.
	pour autant que je comprenne, ils étaient en train de discuter de choses et d'autres , l'un avait mal au ventre



3b: Idiomer

Eksempel på idiom: "syndebukk"

Bergljot Hobækk Haff, Skammen, Gyldendal, 1996.	Denne harmløse mannen var det vi valgte oss ut som syndebukk på nedre Hov og engang gjorde et pek av aller styggeste slag.
	It was this harmless man that we chose for our scapegoat at Nether Hov, and once played a prank of the nastiest kind imaginable.
	C'est à cet être inoffensif qu'à Hov-le-Bas nous avons un jour décidé de faire une exécration plaisanterie .

3b: (forts.)

*Eksempel på idiom: "tåke"*Knut
Faldbakken,"Adams
dagbok",
Gyldendal,
1978.ble ethvert resonnement oppløst i
en **tåke** av uklare bilderevery bit of reasoning was
dissolved in a **fog** of vague
picturestout raisonnement se dissolvait
dans un **brouillard** d'images floues



3b: (forts.)

Eksempel på idiom: "aldri i livet"

Bergljot Hobækk Haff, Skammen, Gyldendal, 1996.	Aldri, Vemund! Aldri i livet!
	Never as long as you live!
	Jamais, Vemund, jamais de la vie!



3b: (forts.)

Eksempel på idiom: aldri i livet

Bergljot
Hobækk
Haff,
Skammen,
Gyldendal,
1996.

Men utlevere ham til Gestapo,
aldri i livet!

But to deliver him to the Gestapo;
never, never!

Mais, d'ici à le livrer à la Gestapo,
jamais de la vie!



3b: (forts.)

Eksempel på idiom: dit pepperen gror

Bergljot
Hobækk
Haff,
Skammen,
Gyldendal,
1996.

at du ønsket ham **dit pepperen
gror**

that you wished he were **at the
back of beyond**

que tu souhaitais le voir
disparaître au plus vite



3b: (forts.)

Eksempel på idiom: dit pepperen gror

Bergljot Hobækk Haff, Skammen, Gyldendal, 1996.	og ba meg dra dit pepperen gror
	sent me to the back of beyond
	et m'a dit d'aller au diable



3b: (forts.)

Eksempel på idiom: dit pepperen gror

Knut Faldbakken, "Adams dagbok", Gyldendal, 1978.	at de ønsker bikkja dit pepperen gror
	that they wish the old dog to go to blazes
	on souhaiterait voir le cabot à mille lieues



3b: (forts.)

Eksempel på idiom: "sole seg i ..."

<p>Geir Lundestad, "Hovedlinjer i internasjonal politikk 1945-1990", Universitets- forlaget.</p>	<p>kunne de gamle europeiske stormakter ennå sole seg i glansen fra fortida.</p>
	<p>In one area the old European major powers could still bask in the glory of the past.</p>



3b: (forts.)

Eksempel på idiom: "sole seg i ..."

Mette
Newth,
Bortførelsen,
Tiden forlag,
1987

Hun **solte seg i** all
oppmerksomheten.

She **basked in** all this attention.



3b: (forts.)

Eksempel på idiom: "sole seg i ..."

Knut Faldbakken, "Adams dagbok", Gyldendal, 1978.	Spilte opp til det, og solte seg i det!
	Played up to it, basked in it!



3b: (forts.)

Eksempel på idiom: "get it" (lemma form)

William Boyd, Brazzaville Beach, 1990.	Got it.
	Skjønner.



3b: (forts.)

Eksempel på 5 språk: "bjørnetjeneste" (Herbjørg Wassmo)

han kanskje gjorde sin beste venn en **bjørnetjeneste** ved å gi ham Dina

he might be doing his best friend an **unintentional disservice** by giving Dina to him

il rendait peut-être un **mauvais service** à son meilleur ami en lui donnant Dina

er seinem besten Freund vielleicht einen **Bären dienst** erwies, er ihm Dina gab

выдавая Дину за своего лучшего друга, он оказывает тому **медвежью услугу**.



3c: grammatikk

"holde på" som aspektmarkør i norsk

[BP1N.s659](#)

Losovskij **holdt på å sy** sammen såret

Лозовский **штопал** рану

[OEL2N.s445](#)

Maren **holdt på å falle** da hun så ham

Марен **чуть было не упала**, когда увидела его



3c: (forts.)

Korpusdata:

- "holde på" + aktivitetsverb =

imperfektivt aspekt i russisk

- "holde på" + punktlig verb =

nektelse + perfektivt verb

Hypotese:

- "holde på" er en tvetydig aspektmarkør i norsk.



3c: (forts.) "sequence of tense" – hva sier korpuset?

Eksepsjonell fortid + presens i norsk

Nikolaj Frobenius, "Latours katalog", Gyldendal, 1996.	Var det ikke de gamle romerne som sa at alt er i evig forandring.
	Wasn't it the ancient Romans who said that everything was in a process of eternal flux?
	Les vieux Grecs ne disaient-ils pas que tout est en perpétuel changement?



3c: (forts.) Nakne nominaler på norsk

<i>"kjøre bil"</i>	
Anne Tyler "The Accidental Tourist", 1985.	Kunne man virkelig kjøre bil uten å rygge?
	Could you really drive a car without reversing?
	Konnte man virkelig einen Wagen fahren , ohne den Rückwärtsgang zu benutzen?
	Kon je werkelijk autorijden zonder ooit achteruit te rijden?



3c: (forts.)

Uttrykk med naken nominal – hyppighet i OMC

legge merke	276
få øye	145
finne sted	138
ha lyst	116
være vitne	53
...	...
spise frokost	32



4: Hva mer?

oversettelsesstrategier	
Noen tall fra RuN- korpuset:	Russisk original: 571 138 ord
	Norsk oversettelse: 716 794 ord



4: (forts.)

oversettelsesstrategier	
Noen tall fra RuN- korpuset:	Norsk original: 303 777 ord
	Russisk oversettelse: 297 905 ord



4: (forts.)

- Læremidler: interaktive oversettelsesøvelser i fremmedspråkopplæring