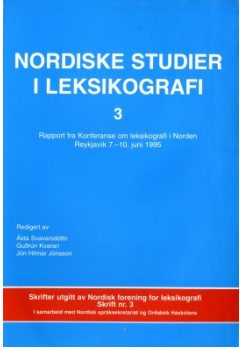


NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Elektronisk sning av Stor norsk-fransk ordbok	
Forfatter:	Tove Jacobsen	
Kilde:	Nordiske Studier i Leksikografi 3, 1995, s. 213-220 Rapport fra Konferanse om leksikografi i Norden, Reykjavík 7.-10. juni 1995	
URL:	http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive	

© Nordisk forening for leksikografi

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Søgbarhed

Artiklerne i de ældre Nordiske studier i leksikografi (1-5) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

Tove Jacobsen

Elektronisk snuing av Stor norsk–fransk ordbok

The Large Norwegian-French Dictionary is computer-readable and coded, and as such it is especially suited as a basis for other projects. By reversing the dictionary by computer, it serves as a base for a French-Norwegian dictionary, and, in this manner the material from an active dictionary is transposed to a passive one.

In this paper, the results of the reversing of the dictionary are described, and the effectiveness and quality obtained by such a method are discussed.

Moderne språkteknologi har gjort at gjenbruk av ordbøker blir mer og mer vanlig. En måte å utnytte en allerede eksisterende tospråklig ordbok på, er å snu den automatisk slik at det opprinnelige utgangsspråket blir målspråket. For at en slik elektronisk snuoperasjon skal være mulig og økonomisk lønnsom, må den eksisterende ordboken være lagret elektronisk. Den bør også være kodet på en enhetlig måte. Disse betingelsene var oppfylt for *Stor norsk-fransk ordbok* (heretter SNFO) da den kom ut i 1991 og senere i 1994. Man mente derfor at det ville være både lønnsomt og faglig interessant å snu SNFO, slik at resultatet kunne bli et råmanuskript for en ny stor fransk-norsk ordbok.

SNFO er lagret på tekstfiler i et postformat. Til den nye fransk-norske ordboken har man bestemt seg for å benytte et SGML-format. Det er flere grunner til det. SGML er et internasjonalt kodesystem som kan brukes både til utgivelser på elektroniske medier og på papir. SGML er også system- og maskinuavhengig. Ut fra gjenbruksperspektivet var det også viktig å ha et ordboksformat som kan brukes igjen ved utarbeidelse av nye ordbøker.

Selve snuoperasjonen ble utført ved hjelp av Pascal-programmering. Deretter foregikk redigeringen ved hjelp av Brief som viste seg å være et meget nyttig tillegg til Pascal-programmene.

Resultatet av en automatisk snuing vil naturligvis være avhengig av hva slags type ordbok man hadde i utgangspunktet. For SNFO's tilfelle ville det si at man hadde en ordbok som primært var laget for brukere med norsk som morsmål og som ville produsere tekster på fransk. Derfor inneholdt ordboksartiklene forholdsvis få oversettelsesekvivalenter, men desto flere eksempler hvor oppslagsordet ble oversatt i kontekst. For den nye ordboken som skal være både for forståelse og produksjon på norsk og fremdeles primært for norske brukere, vil behovet være annerledes. Oversettelsesekvivalentene blir det viktigste og antallet kontekster kan reduseres kraftig.

Eksempel (1) viser første del av en ordboksartikkel fra SNFO. Ved siden av å gi opplysninger om oversettelsesekvivalentene, har redaktørene også tatt med et stort antall eksempler for å vise leseren i hvilke kontekster de forskjellige oversettelsesforslagene kan brukes.

(1)

NB001	øde/legge
NB014	v
NB070	(1)
FR001	(<k>om fysisk skade</k>) casser, briser, détruire, (<k>delvis</k>) abîmer, endommager, (<k>fam</k>) esquinter, (<k>om mat</k>) gâter:
NB410	disse husene ble ødelagt under krigen
FR410	ces maisons ont été détruites pendant la guerre;
NB410	kjøtt ødelegges fort i varmen
FR410	la viande se gâte (<k>el</> s'abîme) vite par temps chaud;
NB410	denne gutten ødelegger alt
FR410	cet enfant casse (<k>el</k>brise) tout, cet enfant est un brise-fer;
NB410	hun fikk ødelagt ansiktet i en kollisjon
FR410	elle a été défigurée dans une collision;
NB410	regnet ødela avlingene
FR410	la pluie a endommagé les récoltes;
NB410	skarpe steiner ødela dekkene
FR410	des cailloux pointus ont endommagé (<k>el</k> abîmé, esquinaté) les pneus;
NB410	ø- sin helse
FR410	ruiner sa santé;
NB410	jeg har ødelagt en tann
FR410	je me suis cassé une dent
NB070	(2)

Etter snuoperasjonen vil hver oversettelsesekvivalent i FR001-feltet generere et nytt oppslagsord med en ny artikkel. For *ødelegge* (1) blir resultatet altså hele syv nye oppslagsord: *casser*, *briser*, *détruire*, *abîmer*, *endommager*, *esquinter* og *gâter*. Hvert av dem får med seg hele settet av opplysninger og eksempler. I (2) ses de første fire nye artiklene:

(2)

(a)

FR001	casser
XX000	<k>om fysisk skade</k>
NB014	v
NB001	øde/legge
FR410	ces maisons ont été détruites pendant la guerre
NB410	disse husene ble ødelagt under krigen
FR410	la viande se gâte (<k>el</> s'abîme) vite par temps chaud
NB410	kjøtt ødelegges fort i varmen
FR410	cet enfant casse (<k>el</k>brise) tout, cet enfant est un brise-fer
NB410	denne gutten ødelegger alt
FR410	elle a été défigurée dans une collision
NB410	hun fikk ødelagt ansiktet i en kollisjon
FR410	la pluie a endommagé les récoltes
NB410	regnet ødela avlingene
FR410	des cailloux pointus ont endommagé (<k>el</k> abîmé, esquinaté) les pneus

NB410 skarpe steiner ødela dekkene
 FR410 ruiner sa santé
 NB410 øde/legge sin helse
 FR410 je me suis cassé une dent
 NB410 jeg har ødelagt en tann

=

(b)

FR001 briser
 NB014 v
 NB001 øde/legge
 FR410 ces maisons ont été détruites pendant la guerre
 NB410 disse husene ble ødelagt under krigen
 FR410 la viande se gâte (<k>el</> s'abîme) vite par temps chaud
 NB410 kjøtt ødelegges fort i varmen
 FR410 cet enfant casse (<k>el</k>brise) tout, cet enfant est un brise-fer
 NB410 denne gutten ødelegger alt
 FR410 elle a été défigurée dans une collision
 NB410 hun fikk ødelagt ansiktet i en kollisjon
 FR410 la pluie a endommagé les récoltes
 NB410 regnet ødela avlingene
 FR410 des cailloux pointus ont endommagé (<k>el</k> abîmé, esquiné) les pneus
 NB410 skarpe steiner ødela dekkene
 FR410 ruiner sa santé
 NB410 øde/legge sin helse
 FR410 je me suis cassé une dent
 NB410 jeg har ødelagt en tann

=

(c)

FR001 détruire
 NB014 v
 NB001 øde/legge
 FR410 ces maisons ont été détruites pendant la guerre
 NB410 disse husene ble ødelagt under krigen
 FR410 la viande se gâte (<k>el</> s'abîme) vite par temps chaud
 NB410 kjøtt ødelegges fort i varmen
 FR410 cet enfant casse (<k>el</k>brise) tout, cet enfant est un brise-fer
 NB410 denne gutten ødelegger alt
 FR410 elle a été défigurée dans une collision
 NB410 hun fikk ødelagt ansiktet i en kollisjon
 FR410 la pluie a endommagé les récoltes
 NB410 regnet ødela avlingene
 FR410 des cailloux pointus ont endommagé (<k>el</k> abîmé, esquiné) les pneus
 NB410 skarpe steiner ødela dekkene
 FR410 ruiner sa santé
 NB410 øde/legge sin helse
 FR410 je me suis cassé une dent
 NB410 jeg har ødelagt en tann

=

(d)

FR001	abîmer
XX000	<k>delvis</k>
NB014	v
NB001	øde/legge
FR410	ces maisons ont été détruites pendant la guerre
NB410	disse husene ble ødelagt under krigen
FR410	la viande se gâte (<k>el</> s'abîme) vite par temps chaud
NB410	kjøtt ødelegges fort i varmen
FR410	cet enfant casse (<k>el</k>brise) tout, cet enfant est un brise-fer
NB410	denne gutten ødelegger alt
FR410	elle a été défigurée dans une collision
NB410	hun fikk ødelagt ansiktet i en kollisjon
FR410	la pluie a endommagé les récoltes
NB410	regnet ødela avlingene
FR410	des cailloux pointus ont endommagé (<k>el</k> abîmé, esquinaté) les pneus
NB410	skarpe steiner ødela dekkene
FR410	ruiner sa santé
NB410	øde/legge sin helse
FR410	je me suis cassé une dent
NB410	jeg har ødelagt en tann

=

Den snudde ordboken har på denne måten fått 35% flere oppslagsord enn den opprinnelige, og det er klart at man må foreta betydelige flytte- og strykeoperasjoner. I første omgang kan alle eksemplene hvor det nye oppslagsordet ikke er med strykes. Av de gjenblivne eksemplene er det sikkert også endel som bør fjernes fordi kontekster som tidligere nevnt er mindre interessante for den nye ordboken.

For (2c og 2d) vil det etter en første opprydning stå igjen (3 og 4)

(3)

FR001	détruire
NB014	v
NB001	øde/legge
FR410	ces maisons ont été détruites pendant la guerre
NB410	disse husene ble ødelagt under krigen

=

(4)

FR001	abîmer
XX000	<k>delvis</k>
NB014	v
NB001	øde/legge
FR410	la viande se gâte (<k>el</> s'abîme) vite par temps chaud
NB410	kjøtt ødelegges fort i varmen
FR410	des cailloux pointus ont endommagé (<k>el</k> abîmé, esquinaté) les pneus
NB410	skarpe steiner ødela dekkene

=

Spørsmålet blir deretter om disse eksemplene er interessante. Det er de antageligvis ikke, og fra en lang artikkel kan man sitte tilbake med bare oversettelsesekvivalenten *ødelegge*.

Til gjengjeld vil man ved mindre komplekse oppslagsord utrolig lett få en rekke nye oversettelsesekvivalenter fra ett og samme franske ord etter snuingen. Dette kommer tydelig fram i eksempel (5):

(5)

FR001 avalanche
YY001 f
NB014 s
NB001 fonn
NB015 -a, pl -er/fenner
=

FR001 avalanche
YY001 f
NB014 s
NB001 lavine
NB015 -en
=

FR001 avalanche
YY001 f
NB014 s
NB001 ras
NB015 -et
=

FR001 avalanche
YY001 f
NB014 s
NB001 skred
NB015 -et
=

FR001 avalanche
YY001 f
NB014 s
NB001 snøskred
=

Vesentlig er også at ordboksredaktørene automatisk vil få ut en del viktig leksikalsk informasjon om franske uttrykk som man ellers lett ville overse. Som et eksempel kan nevnes det faktum at mange norske adverb kan eller må oversettes til fransk med en bestemt type preposisjonssyntagme, nemlig *avec* + substantiv. Etter snuingen kommer denne kombinasjonen ut som eget oppslagsord, noe som gjør at man ikke kan unngå å bli oppmerksom på denne forskjellen mellom de to språkene. Erfaring har vist at når man skal oversette fra fransk til norsk blir man veldig bundet til den franske konstruksjonsmåten og man kommer ikke på at hele konstruksjonen faktisk tilsvarer et norsk adverb. I eksempel (6) sees en

del slike kombinasjoner (det finnes en hærsikare av dem) som har kommet fram som egne oppslagsord etter sruingen:

(6)

FR001 avec acharnement (<k>el</k> ténacité)
 NB014 adj
 FR070 adv
 NB001 iherdig
 =

FR001 avec admiration
 NB014 adj
 FR070 adv
 NB001 beundrende
 =

FR001 avec agilité
 NB014 adj
 FR070 adv
 NB001 behendig
 =

FR001 avec âme
 NB014 adj
 FR070 adv
 NB001 sjel/full
 =

FR001 avec amertume
 NB014 adj
 FR070 adv
 NB001 bitter
 NB015 pl bitre
 FR410 regretter qch amèrement
 NB410 angre bittert på noe
 FR410 dire qch avec amertume (<k>el</k> amérement)
 NB410 si noe bittert
 FR410 ce remède a un goût amer
 NB410 medisinen smaker bittert
 =

FR001 avec application
 NB014 adj
 FR070 adv
 NB001 flittig
 FR410 travailler avec application (<k>el</k> entrain)
 =

FR001	avec approbation
NB014	adj
FR070	adv
NB001	bifallende
=	
FR001	avec ardeur
NB014	adj
FR070	adv
NB001	idig
FR410	travailler avec ardeur
Nb410	arbeide idig
=	
FR001	avec attention
NB014	adj
FR070	adv
NB001	oppmerksom
NB015	-me
FR410	suivre qch avec attention
NB410	følge oppmerksomt med i noe
FR410	écouter attentivement, prêter une oreille attentive
NB410	høre oppmerksomt etter
=	

Disse kombinasjonene kan selvfølgelig ikke fortsette å stå som oppslagsord, men må flyttes for å stå under oppslagsordene *acharnement*, *admiration*, *agilité* osv. som da vil inneholde eksempler som *avec acharnement*, *avec admiration*.

Selv i større fransk-franske ordbøker som *Petit Robert* (1993) mangler ofte denne typen opplysning. *Avec approbation* og *avec amertume* finnes ikke, mens *avec admiration* finnes kun som definisjon på adverbet *admirativement*.

Slår man opp i en allerede eksisterende fransk-norsk ordbok (Kunnskapsforlagets) mangler alle av de ovenstående uttrykkene, noe som viser hvor nyttig den automatiske snuingen også kan være for kvaliteten på den nye ordboken. Mange opplysninger av denne typen burde gi denne ordboken et forsprang i forhold til ordbøker som er redigert på tradisjonell måte.

Bibliografi

Ordbøker

Grundt, L. O. 1991 og 1994: *Stor norsk-fransk ordbok*. Oslo: Universitetsforlaget.

Kunnskapsforlagets fransk-norske ordbok 1986. Oslo: Kunnskapsforlaget.

Petit Robert 1993 = Josette Rey-Debove/Alain Rey: *Le Nouveau Petit Robert*. Paris: Dictionnaires Paul Robert.

Annen litteratur

- Jacobsen, Tove/Randi Sæbøe 1996: The electronic conversion of a dictionary: from Norwegian-French to French-Norwegian. I: *Proceedings of the Sixth International Symposium on Lexicography* May 7-9, 1994. Tübingen: Niemayer. (under trykking).
- Massereeuw, P. C./I. Serail 1992: DictEdit: a computer program for dictionary data entry and editing. I: *Euralex '92 Proceedings I-II*. Tampere: University of Tampere, 237-247.
- Nygaard, Anders Drejer 1994: Datamatstøttet vending af ordbøger. I: A. Garde/P. Jarvad (red.): *Nordiske Studier i Leksikografi II*. Rapport fra Konference om Leksikografi i Norden 11.-14. maj 1993. København: Leda, 229-237.
- Honselaar, Wim/Marijke Elstrodt 1992: The electronic conversion of a dictionary: from Dutch-Russian to Russian-Dutch. I: *Euralex '92 Proceedings I-II*. Tampere: University of Tampere.