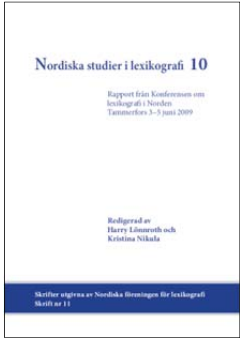


NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Om navn på språk i ordbøger	
Forfatter:	Ari Páll Kristinsson	
Kilde:	Nordiska Studier i Lexikografi 10, 2010, s. 253-262 Rapport från Konferens om lexicografi i Norden, Tammerfors 3.-5. juni 2009	
URL:	http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive	

© Nordisk forening for lexicografi

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 lexicografi (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.

ARI PÁLL KRISTINSSON

Om navn på språk i ordbøker

This article contains some remarks and a preliminary analysis of language names in dictionaries. In addition to a short discussion of lemma selection and orthography, the main focus is on how different definitions, that appear in language name lemmas in monolingual dictionaries, may throw light on the relationship between the culture of which a particular monolingual dictionary is a part, on the one hand, and a particular language on the other. Three definition types are delimited for the present purposes: 1) language X: ‘the language of the people Y’, 2) language X: ‘the [ADJ] language’, 3) language X: ‘the language of state Z / geographical area Z’. Eight language names, Danish, English, Hungarian, Japanese, Quechua, Sami, Swahili, and Tamil were looked up, in four monolingual dictionaries: the Faroese *Føroysk orðabók* (1998), the Icelandic *Íslensk orðabók* (2007), the Norwegian *Bokmålsordboka* (1986), and the Swedish *Svensk ordbok* (1986). One of the conclusions of this investigation is that even if these dictionaries use different definition types, they consistently treat the language names Danish, English, Hungarian and Japanese in the same manner. The three languages that are “culturally more remote” from the Nordic countries, are, along with Sami, never defined with the help of an adjective and never with reference to their official status.

Nyckelord: språknavn, holdninger til språk, definisjonstype

Innledning

Denne artikkelen handler om språknavn som ordbokslemmaer. Kan måten de blir behandlet på, avspeile forskjellig syn blant ordboksredaktørene på forskjellige språk og respektive språkbrukere, etnisitet/nasjonalitet, geografiske områder og land? Finnes det forskjellige prinsipper når man i nordiske enspråklige allmennordbøker beskriver statspråk versus andre språk, eller språk som er vel kjent i Vesten, versus språk fra fjernere kulturer?

Man kan se på spørsmålsstillingen som en språksosiologisk innfallsvinkel til leksikografisk arbeid på et avgrenset område. Ordbøker gjenspeiler alltid meninger og holdninger, i mer eller mindre grad, uansett om redaktørene er oppmerksomme på det, eller ikke. Institusjonaliserte ideologier kommer alltid til uttrykk på en eller annen måte. Derfor er det å forvente at forskjellige holdninger til språk og språkbrukere kan vise seg i ordbøkene, for eksempel

gjennom behandlingen av språknavn. Selve lemmaseleksjonen er naturligvis av interesse i denne sammenheng, men også den måten som språkene eventuelt blir presentert på i lemmaene, ikke minst i definisjonene.

I tillegg til en kort drøftelse av lemmaseleksjon og stavemåte i språknavn handler denne artikkelen om hva slags definisjoner som forekommer i språknavnslemmaer i enspråklige ordbøker. Problemstillingen kan kaste lys over forholdet mellom et bestemt språk/språksamfunn på den ene side, og en enspråklig allmennordbok og dermed den kultur som denne er en del av, på den andre.

Jeg har valgt å avgrense tre definisjonstyper:

- 1) språk X: 'språket til folket Y'
- 2) språk X: '[ADJ] språk'
- 3) språk X: 'språket til staten Z / det geografiske området Z'.

Denne oppdelingen legger jeg til grunn for en analyse av hvordan åtte språknavn, *dansk, engelsk, japansk, quechua, samisk, swahili, tamil* og *ungarsk*, blir behandlet – dersom de forekommer – i fire nordiske enspråklige allmennordbøker: *Bokmålsordboka* (1986) = BO, *Føroysk orðabók* (1998) = FO, *Íslensk orðabók* (2007) = ÍO og *Svensk ordbok* (1986) = SO.

Lemmaseleksjon, stavemåte, definisjoner

Redaktørene av de fire ordbøkene har naturligvis måttet treffe valg om hvilke språk som skulle tas med. BO har tatt med alle de åtte språkene. SO har alle, bortsett fra quechua. ÍO har ikke med quechua, swahili og tamil. Sett fra nordisk kultursynspunkt regnes disse språkene som fjerne. De står heller ikke i FO, og FO har faktisk heller ikke med samisk og ungarsk. Det kan være interessant å legge merke til at lemmaseleksjonen ikke nødvendigvis henger direkte sammen med geografisk avstand. FO og ÍO har altså japansk, men ikke for eksempel swahili, slik at det snarere er snakk om kulturell avstand og den betydning som språkene og språkbrukerne har for den kulturen som ordboken er en del av. I lemmaseleksjonen må man vel også ofte ta stilling til noen spørsmål som kan være problematiske eller omstridte, for eksempel spørsmålet om hva som er et språk, og hva som er en dialekt. Det kan også være et problem å finne frem til pålitelige kilder om noen av de språkvarietetene man ellers ville vært interessert i å beskrive.

Når det gjelder ortografi, har BO oppslagsformene *quechua, swahi'li, tami'l*, og SO har *swahi'li* eller *suahili* og *tami'l* (' er aksenttegn). Man bruker altså ortografi og oppslagsformer som ligger nær utbredt internasjonal skrivemåte i BO

og SO, når det gjelder disse tre språknavnene, og det er antakelig i tråd med den mest vanlige måten i for eksempel mange utenlandske stednavn, og eventuelt i andre fremmede navn som forekommer i disse ordbøkene. Til gjengjeld oppgis navnene på de europeiske språkene og japansk i mer tilpassede former.

Jeg kommer tilbake til definisjonsproblematikken litt senere, men innledningsvis kan det nevnes at det her naturligvis reiser seg en del spørsmål. Skal man for eksempel legge mest vekt på at beskrivelsen tar utgangspunkt i språkbrukerne/språksamfunnene, språkernes geografiske område, språkernes status (statsspråk, sidestilte statsspråk, språk uten offisiell status, minoritetsspråk, religionsspråk osv.) eller muligens språktype (typologisk beskrivelse)?

Når man har et lemma som *dansk* i en ordbok, forventer den som bruker ordboken, vel ikke en forklaring på selve språkbegrepet (hva betyr *språk*?). Det begrepet kan beskrives ut fra tre synsvinkler: språk som tegnsystem, språk som samlede muntlige og skriftlige tekster i en språkvarietet, og språk som et felles språklig gode som en gruppe mennesker kan bruke til å identifisere seg med og bruke felles, og som har overordnede samfunnsfunksjoner i tillegg til kommunikasjonsfunksjonen. Når det gjelder disse tre hovedkriteriene for begrepet 'språk', finner ordboksbrukeren forhåpentlig frem til dem under lemmaet *språk*. Under *dansk* forventer man derimot ikke en beskrivelse av det danske språks forskjellige natur som formsystem, som tekster på dansk, og som samfunnsinstitusjon. Oppgaven er snarere å definere *dansk* i motsetning til *engelsk*, *ungarsk* og så videre og det er det egentlige temaet for resten av denne artikkelen.

Tre definisjonstyper

I innledningen ovenfor presenterte jeg en avgrensning av tre typer definisjoner (jf. tabell 1):

Tabell 1. Tre definisjonstyper

Type 1 språk X: 'språket til folket Y'	Type 2 språk X: '[ADJ] språk'	Type 3 språk X: 'språket til staten Z / det geografiske området Z'
eks. "finnenes språk"	eks. "finsk språk"	eks. "Finlands språk"

Den første typen henviser til språkbrukerne. Denne type brukes for eksempel i det følgende, fra *Webster*: ”Japanese [...] the language of the inhabitants of Japan”.

Den andre typen har som kjerne et ord som betyr ’språk’ og så legger man til et adjektiv. Jævnfør følgende eksempler: ”Japanese [...] The Japanese language” (SOED), ”Japanisch [...] die japanische Sprache” (Duden). Adjektivet henviser egentlig både til språkbrukerne og geografien (eks.: ’japansk’ har både å gjøre med japanere og med Japan).

Den tredje typen henviser til stat og/eller geografisk område, og den finnes i eksempler som disse her: ”Spanish [...] the language of Spain”, ”Basque [...] the language of the Basque area” (CALD). Her legger man straks merke til forskjellen på ”Spain” og ”the Basque area”, men dessuten er definisjonen ”the language of Spain” faktisk tvetydig ettersom *Spain* både kan bety kongeriket Spania og det geografiske området Spania. Definisjonen er heller ikke uttømmende; spansk er jo ”språket i” svært mange andre land. Til gjengjeld er ”the Basque area” nesten utelukkende å tolke som et geografisk område snarere enn som et ’Baskerland’ som en politisk enhet.

Analyse

I dette kapitlet benytter jeg de tre ovennevnte definisjonstypene til å analysere hvordan fire ordbøker – BO, FO, ÍO, SO – behandler åtte språknavn – *dansk*, *engelsk*, *japansk*, *quechua*, *samisk*, *swahili*, *tamil* og *ungarsk*.

BO – Bokmålsordboka (1986)

dansk	dansk språk
engelsk	engelsk språk
japansk	japansk språk
quechua	språket til quechua (I) (det mest utbredte indianerspråket i Sør-Amerika) [quechua (I) ... ”et indianerfolk i Sør-Amerika”]
samisk	finsk-ugrisk språk som særlig tales av innlandssamer i Nord-Skandinavia, Nord-Finland og til dels også av samer i Sovjetunionen
[lappisk]	eldre ord for samisk]
swahi’li	(arab ’som hører til kysten’) bantuspråk (med mange arabiske lånord) som tales på østkysten av Afrika
tami’l	dravidisk språk som særlig blir talt i Sørøst-India (Tamil Nadu) og på Sri Lanka
ungarsk	ungarsk (II) språk, madjarsk (som er en grein av finsk-ugrisk)

Tabell 2. Språknavn i BO

Type 1 'språket til folket Y'	Type 2 '[ADJ] språk'	Type 3 'språket til staten Z / det geografiske området Z'
quechua samisk delvis	dansk, engelsk, japansk, ungarsk	swahili, tamil samisk delvis (quechua)

Tabell 2 viser at dansk, engelsk, japansk og ungarsk bilder en gruppe innenfor type 2 (ungarsk får også tilleggsinformasjon innen parentes om språkfamilien). Tabellen viser også at swahili og tamil tilhører type 3 utelukkende på grunn av den geografiske beskrivelsen, mens det finnes ingen henvisning til det faktum at swahili er offisielt språk i Tanzania og Kenya, og at tamil har vært offisielt i Tamil Nadu siden 1956 (Crystal 1992, Domínguez & López 1995). Geografisk henvisning brukes også delvis i beskrivelsen av samisk (og som tillegg om quechua, i parentes, allikevel uten å nevne quechuas offisielle status i Bolivia og Peru). Type 1 brukes for quechua og delvis om samisk. Beskrivelsen av samisk er altså en blandning av typer 1 og 3, og det gis også tilleggsinformasjon om språkfamilien til samisk. Det er en interessant forskjell på beskrivelsene av quechua og tamil. Quechua blir definert som språket til quechua-folket. Når det gjelder tamil, blir til gjengjeld språkbrukerne (tamilene) definert utelukkende ut fra språket sitt: ”tami ʼl [...] person som har tamil (II) som morsmål”.

FO – Føroysk orðabók (1998)

danskur	danskt	danskt mál 'dansk språk'
enska		enskt mál 'engelsk språk' [kvæðamál]
enskur	enskt	enskt mál, eingilskt mál (mál í Onglandi, Norðuramerika og víðari) 'engelsk språk, engelsk språk (språk i England, Nord-Amerika m.m.)'
japanskur	japanskt	japanskt mál 'japansk språk'

Tabell 3. *Språknavn i FO*

Type 1 'språket til folket Y'	Type 2 '[ADJ] språk'	Type 3 'språket til staten Z / det geografiske området Z'
	dansk, engelsk, japansk	(engelsk)

Tabell 3 viser at FO behandler språknavnene dansk, engelsk og japansk på samme måte, det vil si med adjektiv som kvalifiserer substantivet *mál* 'språk'. Adjektivene *danskur*, *enskur*, *japanskur* danner oppslagsformene i FO. I tillegg oppgir FO substantivet *enska* 'engelsk' som tilhører språkstilen i gamle færøyske kvad (*kvæðamál*). Under *enskt* 'engelsk' beskrives også engelskens geografiske spredning, som tilleggsinformasjon i parentes, antakelig fordi det dreier seg om større geografisk område enn det adjektivene *enskur* / *eingilskur* henviser til. De øvrige fem språkene finnes ikke i FO. Men det kan nevnes at FO allikevel har adjektivene *lappiskur* / *lappendiskur* / *sámiskur* 'samisk' og *ungarskur* / *ungverskur* 'ungarsk'.

ÍO – Íslensk orðabók (2007)

danska	tunga Dana 'danskenes språk'
enska	tunga Englendinga, Bandaríkjamanna, Kanadamanna, Ástrala og fleiri þjóða 'engelskmennenes, amerikanernes, kanadiernes, australiernes og flere nasjoners språk'
japanska	tunga Japana 'japanernes språk'
samíska, samska [lappneska]	tungumál Sama, lappneska 'samenes språk, lappisk samíska, mál Sama (Lappa), af finnsk-úgrískri málaætt 'samisk, samenes (lappenes) språk, tilhører finsk-ugrisk språkfamilie']
ungverska	tunga Ungverja 'ungarernes språk'

Tabell 4. Språknavn i ÍO

Type 1 'språket til folket Y'	Type 2 '[AD]] språk'	Type 3 'språket til staten Z / det geografiske området Z'
dansk, engelsk, japansk, samisk, ungarsk		

De tre andre språkene finnes ikke i ÍO. ÍO er svært konsekvent i bruken av type 1, det vil sige definisjonen henviser til språkbrukerne, i alle tilfellene. Allikevel legger man merke til en interessant forskjell her: I definisjonene av dansk, engelsk, japansk og ungarsk heter det *tunga* 'språk', mens man under samisk bruker ordet *tungumál* 'språk' eller *mál* 'språk'. Selv om de tre ordene betraktes som synonymer, så har de forskjellige konnotasjoner. Ordet *tunga* knyttes gjerne til diskurs om samfunnssiden av språket, blandt andet om språkets rolle som symbol for felles identitet, mens *tungumál* og *mál* henviser snarere direkte til språket som tegnsystem.

SO – Svensk ordbok (1986)

danska	det offisiella språket i Danmark
engelska	det språk som talas i England samt i mange andra delar av världen
japanska	det offisiella språket i Japan
lapska	samernas språk
swahi 'li äv. suahili	ett språk som är ett viktigt handelsspråk i stora delar av Öst- och Centralafrika
tami 'l	medlem av ett indiskt och ceylonesiskt folkslag äv. om detta folks språk, ett högtstående litteraturspråk
ungerska	det offisiella språket i Ungarn: ~n är avlägset släkt med <i>finskan</i>

Tabell 5. Språknavn i SO

Type 1 'språket til folket Y'	Type 2 '[ADJ] språk'	Type 3 'språket til staten Z / det geografiske området Z'	
samisk, tamil		dansk, japansk, ungarsk	engelsk, swahili

Det framgår av tabell 5 at definisjonene av dansk, engelsk, japansk, swahili og ungarsk tilhører type 3, enten med henvisning til språkernes status som offisielle statsspråk eller til geografien. Samisk forklares med språkbrukerne som utgangspunkt. Samtidig legger man merke til at SO (1986) utelukkende oppgir språknavnet *lapska*. Dette er en påminnelse om at også språknavn kan endres. Swahili får tilleggsinformasjon om språkets funksjon som lingua franca, men man unnlater allikevel å nevne at swahili er offisielt språk i Tanzania og Kenya. Man nevner heller ikke at tamil har offisiell status i India og på Sri Lanka. Tamil defineres her som tamilenes språk. I tillegg kommer det en merknad om tamil som ”høgtstående litteraturspråk”, noe man ikke har valgt å si om for eksempel engelsk og de andre språkene her. Her kan man spørre om ordboksredaksjonen har forutsatt at ordboksbrukerne på forhånd for eksempel vet at ungarsk og engelsk er ”høgtstående litteraturspråk”, men at det må være en (overraskende?) tilleggsinformasjon at tamil også er det. Språknavnet quechua finnes ikke i SO.

Oppsummering

I tabell 6 tar jeg ikke med de eksemplene der en del av språkdefinisjonen stod som tilleggsinformasjon i parentes. (Det dreier seg om quechua i BO og engelsk i FO; jf. tabell 2 og 3.) Samisk vises i tabellen under to typer i BO, ettersom samisk der ble definert med henvisning til begge deler.

Tabell 6 viser at ordbøkene har varierende definisjonsmåter. De to ordbøkene som har færrest språknavn, FO og ÍO, er også de mest konsekvente i beskrivelsen av deres tre, henholdsvis fem, språknavn. SO benytter to definisjonstyper mens BO bruker alle tre typene (til og med brukes det to typer i det samme lemmaet, dvs. samisk).

Tabell 6. Definisjoner av språknavn i fire ordbøker

Type 1 'språket til folket Y'	Type 2 '[ADJ] språk'	Type 3 'språket til staten Z / det geografiske området Z'
BO quechua samisk	BO dansk engelsk japansk ungarsk	BO samisk swahili tamil
ÍO dansk engelsk japansk samisk ungarsk	FO dansk engelsk japansk	SO dansk engelsk japansk ungarsk swahili
SO samisk tamil		

Fra en språksosiologisk synsvinkel kan det kanskje være interessant i tabell 6, at man i disse fire leksikografiske produktene alltid finner språknavnene dansk, engelsk, japansk og ungarsk innenfor samme gruppe: under type 1 i ÍO, under type 2 i BO og FO (ungarsk mangler i FO) og under type 3 i SO. De fire språkene behandles altså på samme måte i alle ordbøkene uansett definisjonstype. Vi har også sett at de tre "fjernere" språkene (i kulturell forstand) defineres (dersom de overhodet forekommer) ut fra språkbrukerne og fra ut geografien (jf. tabell 2 og 5). I definisjonene av dem brukes aldri adjektiv-typen, og det henvises aldri til disse språkenes eventuelle status som offisielle språk.

LITTERATUR

- BO = Bokmålsordboka. Definisjons- og rettskrivningsordbok. Bergen/Oslo/Stavanger/Tromsø 1986.
- CALD = Cambridge Advanced Learner's Dictionary. Cambridge 2008.
- Crystal, David, 1992: An Encyclopedic Dictionary of Language and Languages. Oxford.
- Domínguez, Francesc & López, Núria, 1995: Language International World Directory. Sociolinguistic and Language Planning Organizations. Amsterdam/Philadelphia.

Duden = Duden Deutsches Universalwörterbuch. Mannheim/Leipzig/Wien/Zürich 1989.

FO = Føroysk orðabók. Tórshavn 1998.

ÍO = Íslensk orðabók. Reykjavík 2007.

SO = Svensk ordbok. Göteborg 1986.

SOED = The Shorter Oxford English Dictionary on Historical Principles. Oxford 1978.

Webster = Webster's New Twentieth Century Dictionary. New York 1983.