

*Marit Wilhelmsen, Bjørn Are Holth, Øyvind Kleven og
Terje Risberg*

Minoritetsspråk i Norge

En kartlegging av eksisterende datakilder og
drøfting av ulike fremgangsmåter for statistikk om
språk

	Standardtegn i tabeller	Symbol
© Statistisk sentralbyrå	Tall kan ikke forekomme	.
Ved bruk av materiale fra denne publikasjonen skal	Oppgave mangler	...
Statistisk sentralbyrå oppgis som kilde.	Oppgave mangler foreløpig	...
Publisert mars 2013	Tall kan ikke offentliggjøres	:
	Null	-
ISBN 978-82-537-8605-6 (trykt)	Mindre enn 0,5 av den brukte enheten	0
ISBN 978-82-537-8606-3 (elektronisk)	Mindre enn 0,05 av den brukte enheten	0,0
ISSN 1891-5906	Foreløpig tall	*
Emne: Utdanning	Brudd i den loddrette serien	—
	Brudd i den vannrette serien	
Trykk: Statistisk sentralbyrå	Desimaltegn	,

Forord

På oppdrag for Integrerings- og mangfoldsdirektoratet (IMDi), har Statistisk sentralbyrå (SSB) påtatt seg å kartlegge både eksisterende og mulige framtidige datakilder for hvilke språk som benyttes i Norge, og utbredelsen av dem. Bakgrunnen for bestillingen er å fremskaffe kunnskap om nåværende og framtidige tolkebehov i Norge. Dette er nødvendig for å få en hensiktsmessig og behovsstyrt kvalifisering av tolker, og for å vite hvilke språk tolkeutdanningen og statsautorisasjonsordningen bør omfatte.

Arbeidet med notatet er finansiert av IMDi. Ved SSB har Marit Wilhelmsen hatt hovedansvar for arbeidet med notatet. Bjørn Are Holth, Øyvind Kleven og Terje Risberg har vært viktige bidragsytere som forfattere på notatet.

Statistisk sentralbyrå, 25. februar 2013

Hans Henrik Scheel

Sammendrag

I dag finnes det ingen offisiell samordnet språkstatistikk i Norge. Det er heller ingen offentlig registrering av hvilket språk innbyggerne i Norge snakker, eller hvor mange som snakker de ulike språkene.

I dette notatet gjennomgås eksisterende og mulige framtidige datakilder over minoritetsspråk, både internt i SSB og eksternt. Vi konsentrerer oss om de viktigste kildene i henhold til kvalitet og omfang på registreringer. Det er fire administrative systemer/ registre som utpeker seg som relevante når det gjelder eksisterende språkregistrering:

1. Nasjonalt introduksjonsregister (NIR)
2. Grunnskolens informasjonssystem (GSI)
3. Datasystem for utlendingsforvaltning (DUF)
4. Det sentrale folkeregisteret (DSF)

Notatet gir en gjennomgang av disse registrene og beskriver språkvariablene. Deretter vurderes de fire administrative systemene/registrene i forhold til datakvalitet og potensial for å lage språkstatistikk på et befolkningsnivå. Det gjøres en del systematisk registrering av språk i noen etater, og disse kildene gir nyttig informasjon om utbredelsen av minoritetsspråk på et befolkningsnivå. Samtidig er det flere svakheter ved rapporteringen. De kan oppsummeres i fire hovedpunkter: 1) *Relevans*: Registrering av språk er en lite prioritert oppgave i etatene, og dette påvirker kvaliteten. 2. *Datakvalitet*: Det finnes ingen standard for hvordan språk skal rapporteres. 3. *Kodeverk*: Der foreligger ikke et nasjonalt standardisert kodeverk som benyttes på tvers av registrene. 4. *Dekningsgrad*: Samlet gir registrene kun opplysninger om språk for en liten del av befolkningen.

På bakgrunn av denne gjennomgangen er konklusjonen at det er vanskelig å hente ut god og tilstrekkelig informasjon om språksituasjonen i Norge fra de eksisterende registrene. I notatet diskuteres også alternative fremgangsmåter for å samle inn denne informasjonen.

Innhold

Forord	3
Sammendrag	4
Innhold	5
1. Innledning	6
2. Registrering av språk	6
2.1. Hva skal registreres?	7
2.2. Hvordan registrere språk?	7
2.3. Offisielle språk, nasjonalspråk og språkopplæring	8
2.4. Hvordan bør spørsmål om språk administreres?	9
3. Metode	9
4. Relevante datakilder for språk	12
4.1. Nasjonalt Introduksjonsregister (NIR)	12
4.1.1. Vurdering av kvalitet og potensial	13
4.2. Grunnskolens Informasjonssystem (GSI)	13
4.2.1. Vurdering av kvalitet og potensial	14
4.3. Datasystem for Utlendingsforvaltningen (DUF)	14
4.3.1. Vurdering av potensial og kvalitet	16
4.4. Det sentrale folkeregister	16
4.4.1. Vurdering potensial og kvalitet	17
4.5. Sentrale funn	18
5. Andre datakilder	20
5.1. Nasjonale og internasjonale kilder	21
5.2. Utvalgsundersøkelser	22
5.2.1. Vurdering av potensial og kvalitet	23
6. Språkregistrering per i dag – en opptelling av antall språk	23
7. Forslag til videre kartlegginger	24
7.1. Årlige kartlegginger av tolkebehov gjennom spørreundersøkelser hos tolkeformidlere og sektormyndigheter	24
7.2. Forslag til undersøkelse og relevant utvalg for årlige kartlegginger av minoritetsspråk	26
8. Oppsummeringer og konklusjoner	27
Referanser	30
Vedlegg 1. Intervjuguide IMDi	31
Vedlegg 2. Rett og plikt til introduksjonsprogram i henhold til introduksjonsloven	33
Vedlegg 3. Morsmålsopplæring og antall elever fordelt innenfor hver språkgruppe	34
Vedlegg 4. Kodeverket som benyttes av UDI	38
Vedlegg 5. Folkemengde etter fødeland. 2011	50
Vedlegg 6. Opptelling av antall språk fordelt på talere. Kilder: KOST, DUF og internett registreringer (UDI). 2012	54
Tabellregister	61

1. Innledning

I dag finnes det ingen offisiell samordnet språkstatistikk i Norge, og det er ingen offentlig registrering av hvilket språk innbyggerne i Norge snakker, eller hvor mange som snakker de ulike språkene. Fra myndighetenes side er det et behov om å kommunisere med befolkningen på en best mulig måte. En oversikt over befolkningens språkbakgrunn, språkpreferanse og språkkompetanse kan ha en rekke bruksområder.

Statistisk sentralbyrå (SSB) har på oppdrag for Integrerings- og mangfolds_ - direktoratet (IMDi), påtatt seg å kartlegge både eksisterende og mulige framtidige datakilder for hvilke språk som benyttes i Norge, og utbredelsen av dem. Bakgrunnen for bestillingen er å fremskaffe kunnskap om nåværende og framtidige tolkebehov i Norge. Dette er nødvendig for å få en hensiktsmessig og behovsstyrt kvalifisering av tolker, og for å vite hvilke språk tolkeutdanningen og statsautorisasjonsordningen bør omfatte. I denne rapporten vil SSB foreta en sammenstilling og systematisering av relevante datakilder, med en opptelling av språk funnet gjennom analyse av de aktuelle datakildene. Rapporten drøfter hvilke muligheter som finnes for en framtidig kartlegging av språk i Norge og gir anbefalinger om design og opplegg for en slik kartlegging.

Oversikt over hvilke språk som benyttes i befolkningen er en viktig forutsetning for å føre en aktiv språkpolitikk. I Stortingsmelding 35, 2007-2008, fremheves det at en ”helhetlig språkpolitikk i Norge også innbefatter et mangespråklig eller flerespråklig perspektiv”. Selv om staten i særlig grad har ansvar for å styrke nasjonalspråket (nynorsk og bokmål), samisk (sørsamisk, lulesamisk og nordsamisk), de anerkjente nasjonale minoritetsspråkene (kvensk, romani og romanes) og norsk tegnspråk, representerer også alle de nyere minoritetsspråkene et ”språklig mangfold som i dag er med på å utfylle bildet av den totale språksituasjonen i landet. Rett til nasjonalspråk, rett til morsmål og rett til fremmedspråk er trukket fram som tre demokratiske språkrettigheter.” (Språkstatus 2011:67).

Språkrådet har nylig fått utvidet sitt ansvarsområde til også å omfatte andre språk enn norsk. Det skal være Språkrådets oppgave å lage en oversikt over fagmiljø som kan gi nordmenn med minoritetsspråklig bakgrunn informasjon om språket deres, og ”Kulturdepartementet skal vurdere hvilke konkrete oppgaver Språkrådet kan ivareta når det gjelder de nye innvandrerspråkene i Norge.” (Språkstatus 2011:72). Også for dette formålet vil gode tall over hvilke språk som benyttes og deres utbredelse være nyttig.

I rapporten tar vi for oss tre hovedpilarer:

1. **Intern/ekstern oversikt.** Vi gjennomfører en samling og systematisering av eksisterende og mulige framtidige datakilder over minoritetsspråk, både internt i SSB og eksternt. Vi konsentrerer oss om de viktigste kildene i henhold til kvalitet og omfang på registreringen. Det gis også en oversikt over dekningsgraden av språkregistrering i dag.
2. **Kvalitet og potensial.** Vi vurderer ulike datakilder med sikte på relevans og brukeligheten, det vil si at de vurderes i forhold til hvordan de kan benyttes til å kartlegge minoritetsspråk i Norge.
3. **Nye kartlegginger.** Vi foretar en drøfting av andre muligheter for å fremskaffe informasjon om minoritetsspråk i Norge.

2. Registrering av språk

At en offentlig språkregistrering ikke allerede finnes, kan blant annet skyldes en rekke utfordringer knyttet til hvordan informasjon om språkbakgrunn og kompetanse skal registreres.

Det er ingen en-til-en sammenheng mellom fødeland og språk slik at man uten videre kan utlede en persons hovedspråk gjennom å kjenne fødelandet. For enkelte land vil dette være mulig, særlig i en del vestlige land, men ellers kan det være glidende overganger mellom landet og språket/språkene som praktiseres.¹ Hvis en skal utarbeide en oversikt over språk som praktiseres og beherskes i Norge, blir det dermed nødvendig på en eller annen måte å hente ut informasjon om språk fra språkbrukeren selv. En slik registrering reiser en rekke problemstillinger knyttet til hva som skal registreres og hvordan denne registreringen bør foregå.

2.1. Hva skal registreres?

Hva som skal registreres er avhengig av hvilket formål registreringen har. Man kan blant annet ha som mål å avklare språket som personen behersker best, et såkalt "hovedspråk", eller et ønske om å få en bredere kartlegging av personens språkkompetanse. Ut ifra et ønske om bedre kommunikasjon med personen, kan det være ønskelig å vite om personen for eksempel behersker ett av de større (verdens)språkene i tillegg til sitt hovedspråk. Ved kartlegging av språkkompetanse ut over hovedspråket, trengs det også informasjon om språkkompetansen. Vi vil tro at en slik bredere kartlegging av personers språkkompetanse og nivå er aktuelt ut ifra et ønske om å vurdere behov for tolk. Selv om man formulerer spørsmål om språkkompetanse på en god måte, kan det være problematikk knyttet til personens egen forståelse og vurdering av språkkompetanse.

Mer sentralt er en kartlegging av en persons "hovedspråk", forstått som språket en behersker best. Innsamling av slik informasjon kan også være problematisk. "Morsmål" er ofte betegnelsen som benyttes, og forstås som det språket man lærte av sine foreldre eller foresatte som barn. I mange land eksisterer det flere språk som barn lærer seg i oppveksten (LandInfo-4. april 2008). I noen områder av verden vil foreldrene lære barnet flere enn ett språk i den forstand at flere språk snakkes i hjemmet. Samtidig kan barnet mange ganger lære andre språk av andre enn sine foreldre, for eksempel på skolen, av lekekamerater, av naboer eller annen slekt. Resultatet er at enkelte barn kan vokse opp med flere språk som beherskes tilnærmet like godt (Ibid). I en slik situasjon kan det bli misvisende å spørre etter ett bestemt morsmål eller ett første/primær språk. I tilfeller der flere språk læres i barndommen og beherskes godt, kan det være nyttig - ut fra et kommunikasjons- eller tolkebehov - å kartlegge alle språkene som personen mestrer. Det er imidlertid viktig å være klar over at migrasjon kan føre til at en person skifter sitt primære språk gjennom et livsløp, slik at det ikke nødvendigvis vil være språket/ene en lærte som barn som beherskes best senere i livet, og dermed betegnes som "hovedspråk".

2.2. Hvordan registrere språk?

I et datafangstperspektiv er måten man spør om språk på, viktig. Hva språk er, er en kompleks og pågående diskusjon i språkvitenskapen, og det finnes fortsatt ikke én definisjon fagfolk har enes om. En måte å forstå språk på er for eksempel et "ytringssystem hvor de som bruker det forstår hverandre" (LandInfo-4. april 2008:5). Innenfor hvert språk kan det finnes flere dialekter og varianter som ofte er knyttet til bestemte geografiske områder. Det snakkes gjerne om variasjoner og ulike sosiolekter innenfor en dialekt eller et språk. Bildet er komplisert siden det kan være glidende overganger mellom disse tre språkbegrepene. Innenfor ett språk kan ulike dialekter være relativt like, mens innenfor et annet språk kan disse være så ulike at man kan ha problemer med å forstå hverandre (Ibid).

Når vi skal spørre folk om språk, kan det være et problem at det ikke er overensstemmelse mellom den språkvitenskapelige begrepsbruken og begrepene som benyttes i dagligtalen av brukerne. I tillegg er det også problem med ulike stavemåter for det samme språket. Dette kan løses ved å etablere en forhåndsdefinert

¹ Der det er mulig, kan språket utledes fra variabelen *fødeland* i Befolkningsregisteret.

kodeliste over språk i en nedtrekksliste som respondenten kan velge sitt språk fra². Ifølge ”The Concise Encyclopaedia of Languages of the World” (Brown and Ogilvie 2009) eksisterer det i dag drøyt 6 800 språk fordelt på mer enn 250 språkgrupper. Hvis en i tillegg skal inkludere ulike stavemåter for disse språkene, blir listen svært lang. En type utvalg må i tilfelle foretas der en sikrer at de aller fleste aktuelle språk er inkludert, samtidig som listen ikke blir for lang. For de tilfeller der respondenten ikke finner ”sitt” språk i nedtrekkslisten, er det viktig at det også suppleres med en ”annet”-kategori med mulighet til å skrive inn et annet språk.

Hvis en ikke velger et elektronisk spørreskjema, må den samme kodelisten benyttes til koding av språket i etterkant. På den måten vil det være mulig å foreta en tilnærmet riktig plassering av de registrerte språkene ved koding i tråd med kodelisten. Utarbeiding av en kodeliste som både er omfattende nok til å fange opp de fleste aktuelle språk, men samtidig ikke så lang at kodingen blir for tungvint, er sentralt for å kunne registrere språk på en god måte. Det foreligger internasjonale standarder for dette, blant annet en ISO-standard som bør kunne benyttes som grunnlag.

Sammenblandinger av språk – pidgin og kreol

Det finnes ”språk” som ikke har en etablert betegnelse og som vil være vanskelig å kartlegge. I enkelte områder er det sammensmeltinger av språk, såkalte pidgin- og kreolspråk (LandInfo-4. april 2008). Et *pidginspråk* har gjerne et ordforråd basert på et verdensspråk som engelsk, fransk, spansk, arabisk, kinesisk eller en kombinasjon av disse, mens grammatikk og uttale ofte er påvirket av lokale språk. Begrepet pidginspråk brukes om språk som ikke snakkes som førstespråk, men som utvikler seg som nødvendig kommunikasjonsmiddel mellom folk uten annet felles språk.

Kreolspråk defineres som en form for pidginspråk som har utviklet seg til et hovedspråk for en gruppe mennesker. Språket er mer avansert enn typiske pidginspråk (Ibid). Ved å spørre folk om språk, vil denne problematikken komplisere bildet. Det kan være uoverensstemmelse mellom den språkvitenskapelige forståelsen av begrepene pidgin- og kreolspråk, og hvordan brukerne selv betegner språket sitt. De kan bruke begrepene om hverandre, eller benytte helt andre begreper eller betegnelser på ”sitt” språk. I den grad det er ønskelig å kartlegge slike blandingsspråk, er det nødvendig å formulere flere spørsmål rundt dette fenomenet for å få oversikt over kompetanse og praksis hos brukeren.

Lingua franca

På samme måte som pidgin utvikler seg mellom mennesker som ikke har noe felles språk, kan et av verdensspråkene tjene samme kommunikasjonsformål. Et språk som benyttes på denne måten kalles *lingua franca* (LandInfo-4. april 2008). Engelsk er i dag verdens mest utbredte lingua franca. Ut fra et formål om rasjonell og kostnadseffektiv kommunikasjon, vil det være interessant å kartlegge en persons eventuelle foretrukne lingua franca i en framtidig språkstatistikk i Norge.

2.3. Offisielle språk, nasjonalspråk og språkopplæring

Også på et nasjonalt nivå kan bildet over benyttede språk være sammensatt. En stat kan ha ett eller flere offisielle språk som benyttes i nasjonalforsamling, rettsapparatet og i offentlig administrasjon. Samtidig kan staten ønske å løfte fram enkelte andre språk, og gi disse statuser som nasjonalspråk eller regionale språk i landet. Sør-Afrika har for eksempel elleve språk med offisiell status, men ytterligere ni språk har status som nasjonalspråk (LandInfo-4. april 2008:7).

² Dette fordrer en elektronisk administrasjon av spørsmålene (webskjema)

Når det gjelder formell utdanning innen språk, er det forskjeller mellom vestlige land og resten av verden. Vestlige land har ofte formalisert språkoppleringen av sentrale fremmedspråk i skoleverket, mens i utviklingslandene gis det ofte språkundervisning i eventuelle tidligere kolonispråk for landet (Ibid). Flerspråklighet er utbredt, men forståelsen er ofte kun muntlig. Det vil si at språk som beherskes godt muntlig, ikke beherskes like godt skriftlig eller ikke beherskes skriftlig i det hele tatt.

2.4. Hvordan bør spørsmål om språk administreres?

Målsettingen er en entydig forståelse av spørsmålene blant alle respondenter i målgruppa. Samtidig må forståelsen være i tråd med intensjonen med spørsmålene. Dette oppnås ved å tenke grundig gjennom hvilke spørsmål som stilles, formuleringen av dem, rekkefølgen spørsmålene presenteres på, og en vurdering av hvordan spørsmålene best administreres.

I arbeidet med dette prosjektet har det framkommet at mye av den eksisterende kartleggingen ikke har en entydig definisjon. Det er flere ganger uklart om det er morsmål, førstespråk, det språket en behersker best eller foretrukket språk det spørres etter. Samtidig varierer det hvordan informasjonsuthenting er administrert. Noen ganger er det saksbehandler i den offentlige institusjonen som har fylt inn informasjonen. Andre ganger er det respondenten selv som svarer på spørsmål via spørreskjema på papir eller web. I noen tilfeller er det uklart hvordan spørsmålene faktisk er distribuert og hvem som har besvart dem.

Hvordan spørsmålene bør administreres, i en framtidig kartlegging av språk i Norge, avhenger av hva som er hensiktsmessig og praktisk mulig å få til. Hensynet til kostnader vil selvsagt være en viktig faktor. Det optimale er å få registrert og samlet opplysningene om hovedspråk (ett eller flere) i et sentralt register i møtet med ”det offentlige Norge”, med personnummer som identifikasjonsnøkkel. På sikt vil dette være den beste fremgangsmåten for å oppnå god og dekkende oversikt over innvandrerbefolkningens språkpreferanse og kompetanse.

Under kapittel 7 vil vi gjennomgå forslag til alternative fremgangsmåter med sikte på uthenting av språkdata fra tolkeformidlere, sektormyndigheter og fra språkbrukerne selv.

3. Metode

For å fremskaffe relevant informasjon i dette prosjektet, har vi benyttet ulike kvalitative metoder, som dokumentanalyse og kvalitative intervjuer. Siden denne informasjonen ikke forelå i utgangspunktet, har vi vært nødt til å samle den inn. Vi vurderte også om vi kunne samle inn informasjon gjennom et kvantitativt opplegg ved å sende et kort spørreskjema til relevante institusjoner. Det er to grunner til at vi ikke valgte en slik fremgangsmåte.

1. Det forelå ingen oversikt over hvilke institusjoner eller hvilke informanter i institusjonene vi burde sende skjemaet til.
2. Et slikt undersøkelsesopplegg gir ikke mulighet for tilstrekkelig dialog og kommunikasjon. Nettopp det å finne frem til relevante institusjoner, og ikke minst personer i institusjonene, utgjør en viktig del av dette prosjektet. Der hvor kvantitative metoder gjerne er kjennetegnet av standardiserte rutiner, er kvalitative metoder preget av en fortolkende tilnærming. Målsettingen er å oppnå forståelse gjennom nærhet til datakilden ved at fenomenet studeres i dybden. Dybdeintervjuer, observasjon, deltakelse og kvalitativ dokumentanalyse er eksempler på kvalitative metoder (Kvale 2004, Grønmo 1996).

Dokumentanalyse

Dokumentanalyser gjøres for å samle inn relevant informasjon om et felt, og skiller seg fra data vi samler inn selv. Dokumentene er gjerne skrevet for et annet formål enn det vi skal bruke dem til. I denne studien har vi benyttet dokumentanalyse for å lete etter potensielle institusjoner som kan ha registre over språk. Vi har dessuten vært opptatt av å finne frem til de institusjonene vi ønsket å gå inn i for å foreta intervjuer.

Valg av etater

For å finne frem til relevante etater eller institusjoner, snakket vi først med nøkkelpersoner i SSB som kjente til fagfeltet, og som kunne vise oss videre til relevante etater. Vi var også i kontakt med andre institusjoner som har kjennskap til feltet, som Språkrådet, VOX, NAV, UDI, IMDi, og forskere innenfor språkfeltet.

Da vi snakket med de ulike etatene, var vi opptatt av om de kjente til - eller benyttet - administrative registre over språk eller andre saksbehandlingssystemer der språk ble registrert. Det viste seg at flere av etatene hadde registrering av språk i større eller mindre grad, og at de fleste ønsket å få en bedre oversikt over hvilke fremmedspråk som brukes, og hvor mange det er som bruker dem. Vi fant ut at det er to etater, UDI og IMDi, som skiller seg ut med en grundig og systematisk registrering av språk. Registreringen som skjer i disse to etatene dekker all registrering som finner sted i de andre etatene. Dermed er det noe dobbeltregistrering; det vil si at samme opplysningene registreres av ulike etater.

Valg av informanter

Det viste seg å være vanskelig å finne frem til personer som kjente til bruken av saksbehandlingssystemet rundt språk i de ulike etatene. Det er to grunner til det. For det første er institusjonene ofte spesialiserte og hierarkisk oppbygd, slik at det kan være vanskelig å finne frem til rett informant. Det andre er at språk ikke er hovedanliggende, men et biprodukt av institusjonens virkeområde. Det kunne derfor være vanskelig å finne frem til personer som følte seg kompetente til å snakke med oss. Vi fikk bestemt inntrykk av at språk ikke var prioritert i etatene. Vi gjennomførte semistrukturerte intervjuer med Språkrådet, UDI og IMDi. Hos Språkrådet ble intervjuene gjennomført med direktør og fagansvarlig. Dette var hensiktsmessig da intervjuet dreide seg om minoritetsspråk på et overordnet nivå, og Språkrådets rolle i forhold til dette. Språkrådet registrerer bruken av bokmål og nynorsk i det offentlige, men driver for øvrig ikke med noen form for registrering av språk. Men de ser nytten av et eventuelt statlig språkregister.

Det ble gjennomført besøk i UDI og IMDi, og her intervjuet vi personer med ansvar for den registreringen som gjøres.

Intervjuer

Vi valgte å benytte semistrukturerte intervjuer som intervjuform. Dette er en mellomting mellom strukturerte intervjuer - hvor intervjuet i liten grad avviker fra en oppsatt plan - og det som gjerne kalles kulturintervjuer - hvor intervjuet foregår uten mange forhåndsbestemte spørsmål. Vi utarbeidet en intervjuguide med fire hoveddeler:

1. Informasjon om prosjektet, og hensikten med det.
2. Hva blir registrert av språk i etaten?
3. Hvordan foregår registreringen?
4. Hvordan er kvaliteten på dataene?

To personer fra SSB, en intervjuer og en referent, deltok på intervjuene. Intervjuene ble ikke tatt opp på bånd. Underveis var vi opptatt av å la intervjupersonen assosiere fritt, samtidig som vi stilte mange oppfølgingsspørsmål. I tillegg til informasjon om feltet i egen etat, var vi også opptatt av å finne ut om informantene kjente til personer i andre etater som ville være interessante å ta

kontakt med. Denne strategien ledet oss imidlertid ikke til informanter utover de etater og personer vi allerede hadde rekruttert. Vi tok det som en bekreftelse på at vi faktisk hadde fått tak i de personene som kunne bidra med relevant informasjon. Gjennom intervjuene ble det bestemt hvilke registre det var relevant å arbeide videre med. Intervjuguiden fra UDI er lagt ved som eksempel (se neste side). Guiden fra IMDi er lagt ved i vedlegg 1.

Intervjuguide UDI

1) Om prosjektet

I dag er det ingen offentlig registrering av hvilke språk innbyggerne i Norge snakker – ei heller hvor mange talere det er i hvert enkelt språk. IMDi ønsker bedre statistikk på dette området for å kunne estimere framtidige behov for tolker. Statistisk sentralbyrå skal på oppdrag fra IMDi systematisere og analysere tilgjengelige datakilder rundt registrering av språk. Dette er et forprosjekt, hvor vi går igjennom hva som finnes av språkregistrering, andre datakilder som kan benyttes, og hva som gjenstår for å kunne utarbeide god statistikk om språk. Prosjektet skal munne ut i en rapport der vi skal framstille:

- En systematisering av eksisterende og eventuelt kommende datakilder
- En analyse av datakildene
- En opptelling av språk funnet gjennom analysen av de aktuelle datakildene
- Forslag til undersøkelse og utvalg

2) Hva registreres av språk?

- Slik vi har forstått det registreres språk for alle skjema for opphold (asyl, familie, arbeid m.m.) - hva skjer med denne informasjonen?
- Hvorfor registreres ikke språk på enkelte oppholdsgrupper?
- Er det komplisert å få til en registrering av denne informasjonen?
- Hva registreres i DUF?
- Hva registreres i systemet for asylintervjuet?
- Språk som er relevant for integrering.
- Hva slags språk registreres (morsmål, det man kan kommunisere på)?

3) Hvordan foregår registreringen? Kan dere ta oss gjennom gangen i registreringen.

- I hvilke systemer registreres de ulike språkvariablene?
- Hvordan bruker dere denne informasjonen?
- Hvor stor er dekningen av språkregistrering?

4) Hvordan er kvaliteten på dataene?

- Kan dere sjekke ut kvaliteten på DUF mot kvaliteten på det som ligger i systemet dere bruker under asylintervjuet?
- Hvordan er dekningen på registreringen?

4. Relevante datakilder for språk

Det finnes en del offentlige systemer hvor opplysninger om språk blir registrert. Generelt ser det imidlertid ut til at det har vært liten interesse for denne type opplysninger, noe som gjenspeiler seg i kvaliteten på opplysningene som er registrert. I forarbeidet til denne rapporten innhentet vi informasjon om språkregistrering ved flere ulike institusjoner, direktorater og etater.

Fire administrative systemer/register utpeker seg som relevante når det gjelder eksisterende språkregistrering:

1. Nasjonalt introduksjonsregister (NIR)
2. Grunnskolen informasjonssystem (GSI)
3. Datasystem for utlendingsforvaltning (DUF)
4. Det sentrale folkeregisteret (DSF).

Vår gjennomgang tyder på at disse fire registrene dekker den språkrapporteringen som finnes i Norge, og at de gir det beste grunnlaget for å utarbeide statistikk over forekomst og utbredelse av ulike språk i den norske befolkningen. I dette kapitlet gir vi en oversikt over de relevante datakildene, og beskriver språkvariablene som blir registrert. Deretter vurderer vi datakvalitet og potensial for å lage språkstatistikk på befolkningsnivå.

4.1. Nasjonalt Introduksjonsregister (NIR)

Nasjonalt introduksjonsregister (NIR) forvaltes av Integrerings- og mangfoldsdirektoratet. Registeret er opprettet på bakgrunn av forskrift av 20. april 2005 om et nasjonalt personregister for introduksjonsordning og opplæring i norsk og samfunnskunnskap for nyankomne innvandrere, jmfør lov om introduksjonsordning og norskopplæring for nyankommende innvandrere, også kjent som Introduksjonsloven. Registeret skal omfatte personer som er i målgruppen for introduksjonslovens ordninger, det vil si personer som har rett og/eller plikt til, eller kan tilbys, introduksjonsprogram, samt personer som har rett og/eller plikt til norskopplæring (se vedlegg 2).

Registrering av språk i Nasjonalt Introduksjonsregister (NIR)

Nasjonalt Introduksjonsregister (NIR) inneholder flere språkvariabler; morsmål og anbefalt opplæringspråk 1 og 2. Det er kommunene som er ansvarlige for å oppdatere informasjon i NIR, og som legger inn informasjon om språk. Siden det ikke er obligatorisk å legge inn informasjon om språk, er det stor variasjon mellom kommunene når det gjelder utfylling av disse opplysningene. Kommuner som har egeninteresse av å fylle ut informasjonen, fyller oftere ut enn andre. Dette fordi de bruker opplysningene til for eksempel planleggingsarbeid (tilrettelegging av samfunnskunnskap på andre språk), eller for å lage statistikk (møte med IMDI 27. mars 2012).

Variablene er ikke godt nok definert i systemet. Det er uklarheter om hva slags språk som skal fylles ut, for eksempel når det gjelder variabelen ”opplæringspråk1”. Er det et språk personene forstår godt, er det morsmålet, eller er det språket personen ønsker å få tilbudt samfunnsfagsopplæringen på? Det har ikke vært gitt tilstrekkelig sentral veiledning rundt hvordan disse variablene skal fylles ut, og hva som ligger i begrepet. Dermed er det opp til den enkelte hvordan variablene tolkes, og vi antar at det varierer fra person til person hva som registreres (møte med IMDI 27.mars 2012).

I mars 2012 inneholdt NIR opplysninger om omtrent 45 000 personer som har rettigheter og/eller plikter til norskopplæring og eventuelt introduksjonsprogram. Av disse 45 000 var variabelen om språk fylt ut for omtrent 6 500 – altså rundt 14 prosent. Disse er igjen fordelt på 158 forskjellige språk. Da det ikke er benyttet et

standard kodeverk, forekommer enkelte språk flere ganger fordi de er lagt inn med ulike stavemåter.

Alle opplysninger om språk i dette registeret er lagt inn av kommunene. Det hentes ikke inn opplysninger fra andre registre, som for eksempel språkopplysninger fra systemene til UDI, som sannsynligvis ville styrket kvaliteten på registeret.

4.1.1. Vurdering av kvalitet og potensial

NIR gir noe informasjon om minoritetsspråk. Hovedutfordringen er at registeret kun omfatter en liten del av innvandrerbefolkningen, og at den prosentvise andelen av disse som er registrert med et språk, også er liten (ca. 14 prosent). Det er også noen utfordringer i hvordan språket faktisk registreres i registeret. Vi har gjennomgått problemene med begrepsbruken knyttet til språk, og vist til at språksituasjonen ofte kan være sammensatt i enkelte land. Dette påvirker datakvaliteten. Med datakvalitet tenker vi på forhold som gjør at vi er usikre på hva dataene faktisk dekker. I NIR er det kommunene som fyller ut opplysninger om språk. Samtidig foreligger det ingen veiledning for hva slag informasjon som skal fylles ut. Det er derfor noe usikkert hva disse variablene faktisk viser. Etersom IMDi forvalter dette registeret selv, ligger det praktisk til rette for å forbedre kvaliteten.

Forslag til fire tiltak som vil forbedre datakvaliteten:

- Gjøre utfylling av feltene ”morsmål”, ”opplæringspråk1” og ”opplæringspråk2” obligatorisk for kommunene.
- Gi kommunene bedre opplæring og veiledning i hva slags informasjon som skal fylles ut i disse feltene.
- Utvikle et felles kodeverk som er lik kodinger som gjøres av andre etater.
- Hente inn informasjon fra andre kilder, for eksempel UDI.

4.2. Grunnskolens Informasjonssystem (GSI)

Grunnskolens Informasjonssystem (GSI) er en rapporteringsløsning som samler inn data om grunnskoler. Det samles inn drøyt hundre opplysninger, blant annet informasjon om: elevtall, årstimer, ressurser, spesialundervisning, språklige minoriteter, målform, fremmedspråk m.m. GSI ble etablert i 1992, og det gjennomføres årlige datainnsamlinger i oktober hvert år. Til sammen er det omlag 5.000 grunnskoler som leverer data gjennom GSI. GSI eies av Utdanningsdirektoratet (UDIR), men samarbeider med Statistisk sentralbyrå (SSB) om datainnsamlingen. Det er disse to som setter premissene for hvilke data som samles inn i GSI (www.wis.no/gsi). Det innrapporteres fra hver enkelt skole via en felles database (www.wis.no/gsi), og det er rektor som har ansvar for innrapporteringen.

Registrering av språk i GSI

Informasjon om språkopplæring samles inn årlig og omfatter opplæring i fremmedspråk for samtlige elever (engelsk/fransk/tysk m.m.), og hvor mange elever som mottar morsmålsopplæring og/eller tospråklig faglæring eller tilrettelagt opplæring etter språk. Det er kun elever med enkeltvedtak fra skolen som registreres i GSI.

I GSI er ikke informasjonen individbasert, det vil si at den ikke er knyttet til et personnummer. Derfor vil informasjon som hentes ut fra GSI alltid være på et aggregert nivå, det vil si på skolenivå. Dette til tross for at informasjonen foreligger på individnivå ved de enkelte skolene. Innsamlingshjemmelen i opplæringsloven er per i dag ikke presis nok til at UDIR ønsker å endre datafangsten fra en summarisk til en individbasert rapportering. Kunnskapsdepartementet (KD) har i de senere år sett på mulige løsninger for å kunne innføre individbasert rapportering, men foreløpig er det ikke funnet noen løsning på dette.

Morsmålsopplæring defineres som ”den undervisning i elevens morsmål som gis elever med annet morsmål enn norsk og samisk utover det ordinære undervisningstimetallet.” Tospråklig fagopplæring defineres som ”den opplæring som skjer

innenfor det ordinære undervisningstimetallet, og hvor elevens morsmål benyttes i opplæringen (for eksempel "matematikk på urdu") enten alene eller sammen med norsk." (www.wis.no/gsi).

Informasjon om språk i GSI framkommer ved innskrivning til skolen. Her kan foresatte oppgi at barnet har et annet morsmål (det viktigste hjemmespråket) enn norsk og samisk. Foresatte kan dermed kreve morsmålsopplæring, tospråklig faglæring eller tilpasset opplæring på sitt morsmål.

I perioden 2002-2011 ble det framsatt totalt 210 116 krav om morsmålsopplæring, tospråklig faglæring eller tilrettelagt opplæring. Dette er fordelt på 179 forskjellige språk. (Se vedlegg 3, tabell som viser antall elever fordelt på de ulike språkgruppene samlet, og fordelt på årene 2002 til 2011).

4.2.1. Vurdering av kvalitet og potensial

GSI inneholder viktig informasjon om minoritetsspråk i Norge. For det første inneholder registeret informasjon om hvilke morsmål det gis opplæring i, for det andre hvor mange som mottar slik opplæring. Som kilde for utarbeidelse av en språkstatistikk, er det knyttet noen utfordringer til GSI. Det største problemet er at informasjonen ikke er knyttet opp til et fødselsnummer, noe som gjør at GSI ikke kan kobles med andre individbaserte registre for å etablere statistikk som kan si noe om språksituasjonen i Norge. Derfor er det ingen mulighet for å koble opp informasjon om nær familie, og i statistisk sammenheng gjøre antagelser om at andre i samme familie/husholdning snakker samme språk.

Selv om informasjonen hadde vært individbasert, ville det være feilkilder i GSI. Dette handler i første rekke om hvordan informasjonen samles inn, og hva som ligger i begrepet morsmål. Opplysninger blir kun registrert i tilfeller hvor elevens foresatte har bedt om morsmålsopplæring. For familier som har et annet morsmål, men av ulike årsaker ikke ønsker morsmålsopplæring, registreres ikke elevens morsmål i systemet. Samtidig kan det være elever som kommer fra flerspråklige hjem hvor det kan være usikkerhet om hva som er morsmålet. Det gjelder flerspråklige hjem hvor norsk kan være ett av språkene, og flerspråklige hjem hvor de foresatte har forskjellig morsmål, men ingen av dem er norsk.

4.3. Datasystem for Utlendingsforvaltningen (DUF)

Datasystem for Utlendingsforvaltningen (DUF) er et felles verktøy for saksbehandling, statistikk og oppfølging av hele utlendingsforvaltningen: Utlendingsdirektoratet (UDI), Politiet og Utlendingsnemnda (UNE). I dette systemet registreres data om personer som søker opphold i Norge. DUF ble tatt i bruk i 2003 og erstattet det gamle systemet FREMKON. Det er Utlendingsdirektoratet (UDI) som har forvaltningsansvar. DUF inneholder opplysninger om utlendinger som oppholder seg i Norge, personer og søknader/saker som angår andre utlendinger som for eksempel visumsaker.

(http://www.difi.no/statskonsult/publik/rapporter/fulltekst/r2002_11.pdf/
www.udi.no).

Registrering av språk i DUF

I DUF registreres det hvilket språk søkeren forstår, og det er også mulig å registrere flere språk per person. Det er ingen klare retningslinjer for hva slags språk som skal registreres, dermed kan dette være morsmål, språk personen forstår eller snakker godt, eller liknende. For hvert av språkene kan det krysses av for om personen kan lese, skrive og/eller snakke språket. Det er også en mulighet for å markere hvilket av språkene personen foretrekker. I ca. 50 prosent av tilfellene er det registrert opplysninger om hvilket språk personen foretrekker.

Informasjon om språk fylles ut når dette er relevant for integreringsarbeidet, altså når en asylsøker får innvilget opphold. Det er hvert enkelt asylmottak som er ansvarlig for å oppdatere DUF med denne informasjonen. Dermed finnes det relativt god dekning av språkinformasjon for asylsøkere som har fått opphold.

For personer som har søkt om opphold av andre årsaker, for eksempel arbeid, familie eller lignende, er informasjonen mer mangelfull. Dersom personen ikke kvalifiserer til introduksjonsordning (jmfør kapittel 4.1), vil det ikke være registrert noen opplysninger om språk. Det samme gjelder personer som kommer til Norge fra andre EØS/EU-land (møte med UDI 27.mars 2012).

Tabell 1 viser andel med registrert språk etter innvandringsårsak. Tabellen viser klart at i DUF finner vi informasjon om asylsøkere med opphold, men få opplysninger om personer som har søkt om opphold av andre årsaker.

Tabell 1. Andel med registrert språk etter innvandringsårsak i DUF.

Innvandringsårsak	Prosentandel registrert med språk
Asyl	89
Familie	6
Andre	5
Antall	371 581

Kilde: DUF

Ikke alle som er registrert i DUF får innvilget opphold i Norge. Derfor vil vi i de videre analysene kun konsentrere oss om personer som har fått tildelt et fødselsnummer i Folkeregisteret. En person som får innvilget opphold og skal være i Norge mer enn seks måneder, blir registrert med et nummer i Folkeregisteret (<https://www.skatteetaten.no/Alt-om/Folkeregistrering/>). Når vi kobler DUF til Folkeregisteret finner vi igjen 113 425 personer med opplysninger om språk. Det vil si at vi har språkopplysninger på omtrent 30 prosent av dem vi redegjorde for i tabell 1.

Som nevnt tidligere registreres det flere språk per søker, og kun i halvparten av tilfellene registreres det hvilket språk søkeren foretrekker. Selv om dette er et uttrykk for det språklige mangfoldet, er det en utfordring når det gjelder å lage statistikk over morsmål på landsbasis. For det første vet vi ikke hvilke av språkene som er morsmålet. Vi kan nok anta at foretrukket språk er morsmålet, men vi kan ikke være sikre på at dette er tilfellet. Dette skaper en viss usikkerhet når det gjelder kvaliteten av registreringen. Vi har lite informasjon om hva slags språk som registreres, og hva som er det faktiske morsmålet.

Andre språkregistreringer i UDI

Det finnes flere systemer utover DUF som registrerer opplysninger om språk i UDI. Det finnes informasjon om språk i UDI sitt interne saksbehandlingssystem for asylsøknader, KOST. Systemet brukes når saksbehandler lager avtale med asylsøkeren om et intervju. Saksbehandleren registrerer hvilket språk intervjuet kan gjennomføres på. Det er viktig at dette er registrert nøyaktig da denne informasjonen benyttes til å innkalle tolk med rett språkkompetanse til intervjuet. Det vil si at vi får informasjon om hvilket språk søkeren kan gjennomføre et langt og krevende intervju på. Dermed kan vi argumentere for at dette er morsmål eller et språk søkeren behersker godt. Registeret ble tatt i bruk i 2002 (møte med UDI 27.mars 2012). DUF og KOST benytter nå det samme kodeverket når de koder språk (se vedlegg 4). Når vi konsentrerer oss om de personene som er registrert med et fødselsnummer i Folkeregisteret, finner vi totalt 31 781 personer med opplysninger om språk fra KOST.

Opplysninger fra KOST, DUF, og eventuelt NIR gir en relativt god dekningsgrad av språkregistrering for asylsøkere. For personer som har kommet til Norge av andre årsaker, og ikke har krav på integreringstiltak, er det lite registrering av språk. Likevel finnes det noen muligheter. UDI startet i 2010 med søknader om opphold på internett, og her er det mulighet for å registrere språk. Det er søkeren selv som registrerer språk, og de skriver inn i et åpent felt hvilket språk de snakker. Dette er ikke et obligatorisk felt, det vil si at søkeren selv velger om hun/han ønsker å fylle inn disse opplysningene. Totalt har rundt 120 000 personer søkt på internett, men foreløpig søker ikke alle, og det er flere som ikke har mulighet til det. Samtidig oppfordrer utenriksstasjonene folk til å søke på internett når det er

mulig. Mange av søknadene på internett er visumsøknader eller liknende som ikke er relevant for dette prosjektet (møte med UDI 27.mars 2012).

Når vi kobler denne filen mot Folkeregisteret står vi igjen med 8 178 personer som er registrert med opplysninger om språk. Her kobles ikke opplysningene om språk mot et standardisert kodeverk. UDI har utviklet et kodeverk som de har ansvaret for, men benytter det ikke her.

Kodeverk

UDI har utviklet et kodeverk som de per dags dato benytter for både KOST og DUF. KOST har tidligere benyttet et annet kodeverk enn DUF, men dette er blitt standardisert nå. Likevel finnes det eksempler på språk i KOST som er kodet etter gammelt kodeverk³. Språkregistreringer utført via internett er ikke kodet etter noe kodeverk. Her er språkene skrevet inn manuelt.

Det er UDI som har ansvaret for det kodeverket som benyttes på deres registre, men de har ikke et nasjonalt ansvar for å utvikle et nasjonalt standardisert kodeverk. I deres kodeverk er det flere mangler. Vår gjennomgang viser at det blant annet inneholder doble språkregistreringer og navn på språkgrupper. UDI anerkjenner at kodeverket er av dårlig kvalitet, og at det trengs et omfattende arbeid for å utbedre og drifte dagens kodeverk.

4.3.1. Vurdering av potensial og kvalitet

Hos UDI finnes det tre sentrale kilder til språk:

1. Registrering av språk gjennom DUF. Dette er et register som er lett tilgjengelig, og som SSB allerede får utlevert for å lage statistikk.
2. Registrering av språk i KOST. Her er det mulig at registreringen er mer nøyaktig, siden det fungerer som et bookingsystem for asylintervjuene.
3. Registrering av språk gjennom søknader på internett. Her er det mulighet for å innhente opplysninger om språk for personer som får opphold av andre årsaker enn asyl.

Det er noen utfordringer med alle disse kildene. For det første har vi et *kvalitetsproblem*. Det er ikke nødvendigvis slik at det er personene eller søkerens morsmål som registreres i de ulike systemene. I KOST har vi argumentert for at kvaliteten på registreringen vil være god, men i de andre systemene har vi mindre kontroll over hvilke språk som faktisk registreres.

Deretter har vi også her et *dekningsproblem* ved at de ulike systemene kun dekker deler av innvandrerbefolkningen. Problemet ligger i at systemene ikke dekker alle typer innvandringsårsaker. Søknader på internett gir noe informasjon om dette, men foreløpig er det mange som ikke har fylt ut denne informasjonen. Dette er et område som lett kan forbedres:

- Vi anbefaler å gjøre språkfeltene obligatorisk, og at søkeren velger et forhåndsbestemt språk fra en liste istedenfor å skrive inn språket selv.
- At man benytter det samme kodeverket for internettsøknader som i KOST og DUF.
- At kodeverket utbedres og sjekkes opp mot internasjonale standarder.

4.4. Det sentrale folkeregister

Det sentrale folkeregister er et administrativt register over alle personer som bor - eller har vært - bosatt i Norge. Folkeregisteret eies av Skatteetaten og danner blant annet grunnlag for befolkningsstatistikken. Både offentlige og private virksomheter

³ I arbeidet med å få en oversikt over antall språk i Norge, måtte vi manuelt omkode språk i KOST til koder som var forenlige med det standardiserte kodeverket

bruker opplysninger fra Folkeregisteret, og er avhengig av at disse er riktige. Den registrerte bostedsadressen avgjør i hvilken kommune man har rett til ytelser og plikter, og hvor skattekort, selvangivelse, skatteoppgjør og andre meldinger fra offentlig myndighet blir sendt. Personer, som har til hensikt å oppholde seg i Norge i seks måneder og har rett til opphold, vil bli registrert i Folkeregisteret med et fødselsnummer (<https://www.skatteetaten.no/Alt-om/Folkeregistrering/>).

Folkeregisteret er grunnlaget for Befolkningsregisteret som er SSBs statistiske populasjonsregister for personer. Befolkningsregisteret er en viktig del av grunnlaget for befolkningsstatistikken. I Befolkningsregisteret er det for eksempel opplysninger om statsborgerskap, fødeland og fraflyttningsland. Opplysninger fra Befolkningsregisteret benyttes til å avlede og beregne variable som brukes i innvandringsstatistikken⁴.

Registrering av språk i Folkeregisteret

Det finnes ingen registrering av språk i Folkeregisteret, men registeret vil være sentralt for å kunne si noe om språk på et befolkningsnivå. Ved å bruke fødselsnummer kan man koble informasjon fra andre registre til Folkeregisteret, for å si noe om dekningsgraden i Norge. Ovenfor forklarer vi hvordan Folkeregisteret er med og danner grunnlaget for befolkningsstatistikken, der det finnes informasjon som kan være nyttig. For det første er det opplysninger om fødeland, det vil si det landet personen er født i. Vedlegg 5 viser den norske befolkningen fordelt på fødeland i 2011. Totalt bodde det 4 985 870 personer i Norge per 1. januar 2012. Av disse hadde 655 170 personer bakgrunn som innvandrere (se tabell 2). Det tilsvarer 13 prosent av den norske befolkningen. Befolkningen var fordelt på 222 ulike land (<http://www.ssb.no/emner/02/01/10/innvbef/index.html>). Det finnes det opplysninger om innvandringskategori og i tabell 2 ser vi på fordelingen av personer etter innvandringskategori A og B.

Tabell 2. Fordeling av befolkning etter innvandringskategori A og B. 1. januar 2012

Innvandringskategori	Antall	Prosent
A: Innvandrere	546 732	83
B: Norskfødte med innvandrerforeldre	108 438	17
Totalt	655 170	100

Tabellen viser at ca. 655 000 personer er registrert i Folkeregisteret som innvandrere eller norskfødt med innvandrerforeldre. Det er denne gruppen vi ønsker å kartlegge morsmålet til. Med opplysninger om fødeland vil det for noen av gruppene være lettere å anta et morsmål, som for eksempel de med bakgrunn fra Polen (62 294 personer). Det er vanskeligere for personer som kommer fra land med flere enn ett språk, eksempelvis personer med bakgrunn fra Irak (28 672).

Til slutt finnes det variable i Folkeregisteret som angir familier. De som er i familie vil dele samme familienummer. Man kan for eksempel anta at personer i samme familie har samme morsmål. Har vi registrert språk på én eller flere personer i familien, kan disse opplysningene overføres til resten av familien.

4.4.1. Vurdering potensial og kvalitet

Folkeregisteret, hvor bosatte er registrert med fødselsnummer, vil være en sentral kilde for å koble sammen informasjon fra andre registre om språk. Det kan hjelpe oss i å få en oversikt over minoritetsspråk på befolkningsnivå. Det finnes variable i Folkeregisteret som kan være til hjelp for å estimere språk:

- 1) Informasjon om landbakgrunn (nyttig for å estimere språk).
- 2) Familienummer som gir oversikt over familier i Norge. Det kan gjøres antagelser om at personer som tilhører samme familie har samme språk.

⁴ I innvandringsstatistikken gjøres en statistikkbearbeiding av opplysninger fra Befolkningsregisteret. Dette gir variable som innvandringskategori og landbakgrunn.

Det vil være flere feilkilder ved å gjøre slike estimeringer. Det er ikke nødvendigvis slik at personer i samme familie eller nær slekt snakker samme språk. Det samme vil være for landbakgrunn. Denne variabelen indikerer det landet man er født i, og det er ikke nødvendigvis slik at personen snakker dette språket. I flere land er det også slik at det snakkes mange forskjellige språk⁵.

4.5. Sentrale funn

Gjennomgangen av de ulike kildene til språkregistrering viser at det gjøres en del systematisk registrering av språk i noen etater. Disse kildene kan til sammen være nyttige for å gi informasjon om utbredelsen av minoritetsspråk på et befolkningsnivå. Samtidig viser vår gjennomgang at det er en del områder som er problematiske. I dette avsnittet sammenfatter vi de funnene vi oppfatter er de mest sentrale fra redegjørelsen ovenfor. Først framstiller vi en tabell over de sentrale registrene. Vi har utelatt Folkeregisteret fra denne tabellen da det ikke inneholder opplysninger om språk, men vil være sentral for å koble sammen informasjon fra andre registre.

Tabell 3. Sentrale registre som registrerer språk

Register	Eier	Innhold av språkopplysninger	Antall	Problemer
Nasjonalt introduksjonsregister (NIR)	Integrerings- og mangfoldsdirektoratet (IMDi)	Registeret består av personer som er i målgruppen for introduksjonslovens ordninger Inneholder: Informasjon om morsmål og anbefalt opplæringspråk	Består av rundt 45 000 personer. Det er utfylt språkopplysninger på 6 500 av disse, fordelt på 158 forskjellige språk	- Svært få av alle i registeret har fått utfylt språkopplysninger - Dekker kun en liten andel av innvandrerbefolkningen - Ingen veiledning for utfylling av språkvariabelen - Ikke obligatorisk - Mangler kodeverk
Grunnskolens informasjonssystem (GSI)	Utdanningsdirektoratet (UDIR)	Registeret består av 5 000 grunnskoler Inneholder: Antall elever som mottar morsmålsopplæring og/eller tospråklig faglæring eller tilrettelagt opplæring etter språk. Det er kun elever med enkeltvedtak	Fra 2002 - 2011 var det totalt 210 116 krav om slik opplæring fordelt på 158 forskjellige språk	- Opplysningene er ikke knyttet til et fødselsnummer og kan ikke kobles med andre individbaserte registre - Får kun informasjon om de som ønsker morsmålsopplæring, og ikke om andre som av ulike årsaker ikke ønsker slik opplæring, men likevel har et annet morsmål - Ingen kodeverk

⁵ Se for øvrig kapittel 2 for ytterligere drøfting av problematikk rundt språkregistrering.

Datasystem for Utlendingsforvaltningen (DUF)	Utlendingsdirektoratet (UDI)	Registeret består av utlendinger som oppholder seg i Norge. Inneholder: Språk personen forstår. Det er mulig å registrere flere språk. Om hun/han kan lese, skrive og/eller snakke språket og hvilket språk er foretrukket, om flere	Ved å koble DUF til Folkeregisteret og fjerne dubletter er det 113 425 personer med opplysninger om språk	- Ingen klar definisjon av språket som skal registreres - En person kan registreres med flere språk. Ikke alltid klart hvilket som er hovedspråket - Dekker hovedsakelig asylsøkere som har fått opphold. Lite informasjon om personer som har kommet til Norge av andre årsaker - Kodeverket har store mangler
Saksbehandlings-system for asylsøknader, KOST	Utlendingsdirektoratet (UDI)	Systemet brukes når saksbehandler lager avtale med asylsøkeren om et intervju Inneholder: Hvilket språk intervjuet kan gjennomføres på	Ved å koble KOST til Folkeregisteret og fjerne dubletter er det 31 781 personer med opplysninger om språk	- Dekker hovedsakelig asylsøkere som er kalt inn til et intervju. Lite informasjon om andre personer som blant annet har kommet til Norge av andre årsaker - Kodeverket har store mangler
Søknader på internett	Utlendingsdirektoratet (UDI)	Personer som søker om opphold på internett. Inneholder: Opplysninger om språk Frivillig å fylle ut	Totalt har rundt 120 000 personer søkt på internett Ved å koble filen mot Folkeregisteret finner vi 8 178 personer med opplysninger om språk.	- Språkfeltet er ikke obligatorisk, og fylt ut i svært få tilfeller - Språket skrives inn av søkeren selv, bør kunne velge fra en liste - Ingen kodeverk - Dekker foreløpig svært få personer

Relevans

Selv om det gjøres en del systematisk registrering av språk i ulike systemer, er dette snarere et biprodukt enn hovedformålet med registreringen. Det virker ikke som språkregistrering er en prioritert del av etatenes virke. Det er sjelden at informasjon om språk blir benyttet av etatene, og dermed er det heller ingen prioritert oppgave å kvalitetssikre registreringen. Unntaket er KOST-registeret.

Det kan dessuten se ut til at de ulike etatene registrerer en del informasjon om språk uten å sjekke hva andre etater har registrert. For eksempel registrerer kommunene informasjon om språk i NIR, samtidig som denne informasjonen allerede finnes i DUF. Det virker ikke som etatene samarbeider om språkregistrering, eller har god innsikt i andre systemer som registrerer språk. Dette kan være en følge av at informasjon om språk ikke er et prioritert område.

Dekningsgrad

Når det gjelder asylsøkere som har fått opphold, er dekningsgraden relativt god, men for personer som har kommet til Norge av andre årsaker, foreligger få eller ingen opplysninger. DUF er det registeret med mest omfattende opplysninger om språk, men da hovedsakelig for personer som har søkt asyl og fått innvilget opphold. Det finnes andre registre som muligens dekker andre grupper i noen grad. For det første har UDI startet med søknader på internett. Her kan søker legge inn opplysninger om språk. Foreløpig gir dette oss få nye opplysninger, men dette kan være en kilde for opplysninger om språk i framtiden. Det andre registeret er GSI, og per i dag inneholder GSI kun opplysninger om morsmålsundervisning på skolenivå, og det er ikke mulig å koble det mot andre individbaserte registre.

Datakvalitet

I redegjørelsen diskuteres kvaliteten på opplysningene som registreres. Det er et problem at ingen av de gjennomgåtte registrene opererer med definisjon av hvilket språk som skal registreres, noe som gjør at datakvaliteten er varierende. Vi vet derfor ikke om foreliggende opplysninger dreier seg om morsmål, et språk personen snakker og forstår godt, eller om det er andrespråk som personen behersker godt - for eksempel engelsk, fransk, spansk. Dessuten blir flere språk registrert per person i enkelte systemer – uten at det er klart hvilken status språket har for personen. Årsaken er at vi vet ikke hvilken status de ulike språkene har for registerpersonene, og datakvaliteten blir mangelfull. For å kunne utarbeide en språkstatistikk, må det være klart hva slags opplysninger vi står overfor, og klart definert om det er snakk om morsmål eller et annet språk personen forstår godt. Dette vil også være avgjørende hvis statistikken skal benyttes til å beregne behovet for tolketjenester innenfor de ulike språkgruppene.

Kodeverk

Det siste funnet vi vil trekke fram fra denne redegjørelsen, er kodeverk. Det er kun i registrene til UDI - KOST og DUF - at det benyttes kodeverk. UDI benytter i all hovedsak samme kodeverk i disse to registrene, og det er UDI som har utviklet kodeverket. Vår analyse viser imidlertid at det hefter store mangler ved kodeverket. UDI anerkjenner også at kodeverket ikke er av god nok kvalitet, og at det er et omfattende arbeid å oppdatere dette. Årsaken er at kodeverket ikke er uttømmende, språk er registrert flere ganger med forskjellige koder, og i enkelte tilfeller er språkgrupper også gitt en kode i kodeverket. De andre registrene som er berørt i denne redegjørelsen benytter ikke kodeverk i det hele tatt. Manglene nevnt over i kodeverket påvirker datakvaliteten fra disse registrene.

Det vil derfor være hensiktsmessig og nødvendig å innføre ett standardisert nasjonalt kodeverk som alle som registrerer språk benytter. I dag har vi ikke klart å finne en slik nasjonal standard, og det er uklart hvem som skulle hatt ansvaret for å utvikle og innføre en slik nasjonal standard. Før man kan få gode data og gjennomføre gode analyser av språk, og språk fordelt på antall brukere i Norge, er imidlertid et slikt kvalitetssikret kodeverk påkrevd. Det finnes internasjonale standarder, og SSB har blant annet benyttet ISO-koder da vi har levert språkdata på internasjonale undersøkelser⁶.

5. Andre datakilder

I dette notatet ser vi også på andre kilder til språkregistrering. I kapittel 5 skal vi gå igjennom andre nasjonale og internasjonale kilder, og se på utvalgsundersøkelser som en kilde til språkstatistikk.

⁶ Dette kodeverket ble benyttet til den internasjonale undersøkelsen om lese- og tallforståelse og den europeiske samfunnsundersøkelsen

5.1. Nasjonale og internasjonale kilder

Per i dag foreligger det lite statistikk over språkstatusen i Norge utover kildene som er nevnt tidligere. Det meste av statistikk om språk handler om nynorsk versus bokmål, samt om utviklingen av samisk språk i Norge. Vi har ingen sammenfattende statistikk over hvor mange språk (og hvilke) som snakkes i Norge. Det samme gjelder for mange andre land. Det finnes lite statistikk om hvor mange språk som snakkes, av hvor store grupper de enkelte språk snakkes, samt hvor høy kompetansen er i de enkelte språkene. Det finnes i dag ingen registerkilder i Norge som gir gode data som kan gi svar på slike spørsmål. Det foreligger noen data fra utvalgsundersøkelser utført av SSB, som kan belyse deler av problemstillingen, men utvalget er for lite og spørsmålene ikke omfattende nok til å kunne gi noe totalt bilde (se avsnittet om undersøkelse om voksne og læring (AES) under).

Samme utfordring finnes i flere land som baserer sine statistikkssystemer innenfor sosialstatistikken på gjenbruk av data fra administrative systemer direkte, eller gjennom aggregerte rapporter som er tatt ut av de samme systemene. Dette gjelder for eksempel alle de nordiske landene, men også i økende grad andre europeiske land.

I de land som fortsatt gjennomfører store folketellinger ved bruk av tellere (det vil si intervjuere som reiser rundt og intervjuer hele befolkningen eller et representativt utvalg av befolkningen), finnes det riktignok en del statistikk om språk og språkkompetanse. Forrige gang Norge gjennomførte en slik telling var i 1970. Også tellingene i 1980 og 1990 gikk ut til alle personer i Norge, men da som et papirskjema distribuert via posten. Spørsmål om språk og språkkompetanse var ikke med i disse folketellingene. Det skyldes for en stor del at problematikken rundt språk ennå ikke var blitt så aktuell hos oss. Folketellingen i 1990 ble gjennomført som en utvalgsundersøkelse, men heller ikke denne inneholder spørsmål om bruk av fremmedspråk. Det betyr at utover spørsmål om kunnskap og bruk av samisk (innenlands etnisk minoritet), har det aldri blitt stilt spørsmål om bruk og kompetanse i fremmedspråk i norske folketellinger.

I en del andre land hvor det fortsatt gjennomføres folketellinger etter intervjuemetoden (enten fulltelling eller utvalgstillinger), finnes det noe statistikk om språkbruk og språkkompetanse. Et eksempel er fra USA, hvor det spørres om hvilket språk som benyttes i hjemmet, samt at det spørres om kompetanse i engelsk der det benyttes annet språk enn engelsk i husholdningen. Eksempler på tabeller, inklusive definisjoner og omtale av statistikken fra US Census kan man finne her: <http://www.census.gov/hhes/socdemo/language/index.html>

I EU-statistikk har det vært fokus på hvor *mange* språk man behersker, men ikke like mye fokus på *hvilke* språk man behersker. Det har vært mest fokus på hvilke andre EU-språk man behersker.

Internasjonalt er det gjort noen studier for å forsøke å gi et estimat på antall språk som er i bruk. Et eksempel er VALEUR-prosjektet (Valuing All Languages in Europe), som er ett av flere prosjekter gjennomført av The European Centre for Modern Languages (ECML) for å kunne følge opp språkopplæringspolitikken til Council of Europe (CoE). VALEUR-prosjektet leverte sine resultater til CoE i 2007. Dette prosjektet pekte på mange av de samme svakhetene om tilgjengelige data som er påpekt tidligere. I Europa finnes det få fullskala folketellinger som inkluderer spørsmål om språk, noe som gjør det vanskelig å estimere et sikkert tall på antall språk som er i bruk i Europa. Prosjektet beregnet seg allikevel frem til at det var minst 440 språk som blir brukt muntlig i Europa, basert på en egen undersøkelse hvor de fikk rapporter fra utnevnte eksperter i alle de 21 deltakende europeiske landene (Norge er medlem av ECML, men deltok ikke i VALEUR). Usikkerheten rundt tallet er stor, og basert på erfaringer fra andre undersøkelser om mulig underrapportering, skriver de også at det faktiske tallet kan være så høyt som opp mot 600 (kilde: <http://archive.ecml.at/mtp2/publications/Valeur-report-E.pdf>)

5.2. Utvalgsundersøkelser

Statistisk sentralbyrå gjennomfører flere utvalgsundersøkelser der blant annet språk blir kartlagt. Dette kan enten være morsmål og/eller generell språkkompetanse (hvor mange språk personen kan). Noen eksempler på slike undersøkelser er den europeiske samfunnsundersøkelsen (ESS), den internasjonale undersøkelsen om lese- og tallforståelse (PIAAC), og undersøkelse om voksne og læring (AES).

I dette kapittelet diskuterer vi om utvalgsundersøkelser kan gi oss relevant informasjon om språk. SSB har ikke gjennomført utvalgsundersøkelser hvor formålet har vært å samle inn informasjon om hvilke språk innvandrerbefolkningen benytter. Spørsmål om språk har vært stilt i en del undersøkelser rettet mot hele befolkningen. Slike undersøkelser kan ikke gi mye informasjon om språk-utbredelsen i Norge, siden de ulike språkgruppene er for små til at de vil bli godt representert. Vi vil i dette kapittelet diskutere datakvaliteten, og hvorvidt utvalgsundersøkelser kan være et godt alternativ for å få oversikt over utbredelsen av minoritetsspråk på et befolkningsnivå.

Eksempler på spørsmål i utvalgsundersøkelser

I flere undersøkelser SSB har gjennomført har det vært inkludert spørsmål om språk. Her trekker vi fram to eksempler på hvordan spørsmål om språk stilles.

Den internasjonale undersøkelsen om lese- og tallforståelse (PIAAC)

Spørsmålene som stilles om språk er:

- Hvilket språk var det du først lærte i hjemmet i barndommen, og som du fortsatt forstår?
- Hvilket språk var det andre du lærte og fortsatt forstår? (Dette spørres opptil x ganger.)
- Hvilket språk snakker du oftest hjemme?

Undersøkelse om voksne og læring (AES)

Spørsmålene som stilles om språk er:

- Så har jeg noen bakgrunnsspørsmål. Hva er ditt morsmål? Med morsmål menes det første språket du lærte i barndommen, og som du fortsatt forstår.
- Når du ser bort fra morsmålet ditt/morsmålene dine, *hvor mange* andre språk kan du forstå enten skriftlig eller muntlig? Det er ikke nødvendig å forstå språket fullt ut.

Som nevnt over vil ikke disse undersøkelsene gi et godt bilde av minoritetsspråk i Norge, siden for få personer innenfor hver språkgruppe er representert. For eksempel var det i pilotundersøkelsen til PIAAC kun 12,2 prosent som oppga et annet språk en norsk, og disse ble fordelt på 17 forskjellige språk.

En fordel med innsamling av informasjon om språk ved utvalgsundersøkelser er at spørsmålsformuleringen er nøye gjennomtenkt, og man har kontroll på hva slags data som registreres. Vi har vist i gjennomgangen av de ulike registrene, under kapittel 4, at det sjelden foreligger en definisjon av språk i disse. I eksemplene over brukes en definisjonen av morsmål i selve spørsmålet som respondenten svarer på. Morsmål defineres som ”det første språket du lærte i barndommen, og som du fortsatt forstår”. Det samme finner vi ikke i vår gjennomgang av registrene, og dermed kan forståelsen av hva som menes med språk variere. Dette ble grundig diskutert i kapittel 4.

Det at begrepene er godt spesifisert gir en bedre datakvalitet, og det er enkelt for respondenten å forstå hva som menes. For eksempel kan en person fra Tanzania ha lært Chagga i barndommen, samtidig er Swahili det offisielle språket i Tanzania. I et tenkt eksempel lærte respondenten Swahili når hun/han begynte på skolen. Ved

at spørsmålet definerer hvilket språk vi er ute etter, ville vi i dette tilfellet registrert språket hun/han snakket i barndommen, altså Chagga. Som vi har diskutert tidligere kan det være tvetydig hvilket språk som faktisk registreres. Er det morsmålet, det offisielle språket, et annet språk som personen snakker godt (for eksempel engelsk, fransk, eller spansk) eller noe annet?

Dersom en skal kartlegge utbredelse av språk i Norge, kan en utvalgsundersøkelse være et alternativ. I første omgang bør man gå ut til et stort utvalg i populasjonen og stratifisere utvalget etter landbakgrunn, slik at enkelte innvandrergupper blir disproporsjonalt overrepresentert. For å få sikre nok estimater bør en minimum gå ut til ca. 10 000 personer. Frafall er alltid en utfordring ved utvalgsundersøkelser, og vi vet av erfaring at frafallsandelen er høyere i innvandrerbefolkningen,⁷ og at den varierer mellom ulike innvandrergupper. Det er derfor viktig at utvalget er tilstrekkelig stort i de ulike gruppene for å sikre representativitet. Det er hensiktsmessig å benytte informasjon vi har innhentet fra registrene for å sikre at alle grupper representeres. Ved å koble informasjon om språk til fødeland, vet vi noe om hvilke grupper som mangler informasjon om språk.

5.2.1. Vurdering av potensial og kvalitet

Fordelen med å benytte utvalgsundersøkelser til å samle inn opplysninger om språk, er at vi får god kvalitet på den informasjonen vi faktisk klarer å samle inn. Vi har kontroll med definisjonene på variablene, noe som vi ikke har i like stor grad når vi henter opplysninger fra registrene.

På den andre siden er det utfordringer med frafall i utvalgsundersøkelser. Det er derfor viktig at utvalget er stort og disproporsjonalt representert. Vi kan bruke informasjon fra registrene til å designe et utvalg som skal sikre at de fleste grupper representeres. En slik utvalgsundersøkelse vil kunne gi god informasjon om språkfordeling på et befolkningsnivå, men den vil ikke kunne gi like detaljert informasjon om for eksempel antall brukere i de ulike gruppene eller lignende.

6. Språkregistrering per i dag – en opptelling av antall språk

I denne rapporten har vi presentert ulike kilder som oppgir minoritetsspråk. Vi har også gitt en vurdering av kvaliteten på disse, samt i hvilken grad data fra de ulike kildene er egnet til å lage presis statistikk. Formålet med denne gjennomgangen er å si noe om de mest omfattende kildene til språkregistrering som finnes i dag, og vurdere hvilke språkkilder som kan være viktige bidrag til et språkregister på befolkningsnivå. I dette kapittelet skal vi kombinere de kildene vi har tilgjengelig i dag, for å gi en oversikt over antall språk, fordelt på talere. I den sammenheng er det kildene fra UDI som er de mest relevante. Det er flere årsaker til dette:

- 1) Registrering av språk har pågått siden 1990-tallet og registeret gir informasjon om personer som har bodd i Norge en stund.
- 2) Det er disse systemene som helt klart gir best dekning av språkregistrering på et befolkningsnivå i Norge.
- 3) Alle dataene fra DUF og KOST er kodet i samme kodeverk. For de andre registrene brukes ingen kodeverk. Når det gjelder registreringer på internett, er kodingen gjort manuelt i etterkant.

Ved å koble informasjon fra disse tre registrene til Folkeregisteret, får vi en oversikt over antall språk og brukere i Norge. Vi har koblet registrene mot personer som er registrert som innvandrere i Norge. I dette notatet ser vi kun på de som er registrert som innvandrere. Begrunnelsen er at det stort sett er denne gruppen det

⁷ Se blant annet Kultur og mediebruksundersøkelsen blant innvandrere 2008, Innvandrere på internett 2012.

foreligger språkopplysninger om i eksisterende registre. Det gjelder om lag 550 000 personer (jæmfør tabell 2). Der vi har funnet opplysninger i flere registre har vi valgt følgende rangering:

- 1) KOST
- 2) DUF
- 3) Registrering på internett

Rangeringen har bakgrunn i kvalitetsvurderinger av registrene. I kapittel 4 diskuterte vi datakvaliteten på KOST versus DUF. Det kan registreres flere språk i DUF, derfor vil vi velge ett tilfeldig i de tilfellene det er registrert flere språk. I videre analyser kan det være hensiktsmessig å bruke andre kvalitetskriterier for å velge språk. I DUF foreligger opplysninger om foretrukket språk for omtrent 50 prosent av personene, og registeret inneholder variabler om hvilket språk personene snakker, leser og skriver. Ved å koble de tre kildene kunne vi plassere ett språk til totalt 119 147 personer.

Tabell 4 viser fordelingen mellom de tre registrene.

Tabell 4. Fordelingen mellom DUF, KOST og internettregistreringer.

Kilde	Antall	Prosent
Språk fra DUF	80 755	68
Språk fra KOST	31 780	26
Språk fra internettregistreringer	6 612	6
Totalt	119 147	100

Kilde: DUF, KOST og Internettregistreringer

Mesteparten av foreliggende informasjon om språk får vi fra DUF. Hele 68 prosent av alle registreringer kommer fra dette registeret, mens 26 prosent kommer fra KOST. Rundt 6 prosent stammer fra registreringer på internett. Selv om dette antallet ikke er så stort, er det interessant å trekke fram at vi får med 6 612 registrerte personer fra internettregistreringen av totalt 8 178. Det utgjør hele 81 prosent. Det virker som dette registeret gir informasjon om språk på personer som de andre registrene ikke fanger opp.

Totalt klarer vi å plassere et språk på 119 147 personer. Det utgjør omtrent 22 prosent av innvandrerne i Norge. For å gjøre en opptelling av antall språk, og hvor mange talere det er per språk, benyttet vi kodeverket til UDI. Dataene fra internett ble kodet med bakgrunn i dette kodeverket. Totalt finner vi 311 språk i Norge. Vedlegg 6 viser de ulike språkene og antall brukere. Etter å ha fjernet en del åpenbare feilregistreringer, finner vi at det dreier seg om rundt 300 språk. I vedlegget har vi beholdt det slik det ligger i registeret.

Det hefter noen svakheter ved disse tallene. Det første skyldes personer som er registrert med flere språk. Her har vi tilfeldig valgt ett språk som representerer personen. Det er ikke sikkert at dette er morsmålet til personen. Vi får likevel et inntrykk av den språklige sammensetningen i innvandrerbefolkningen, og vi får et inntrykk av dekningsgraden. Den andre svakheten skyldes kodingen. Her er det store utfordringer, og som beskrevet tidligere er kodeverket ennå ikke godt nok. Blant annet blir språkgrupper registrert som språk, og noen språk er oppført flere ganger grunnet ulik stavemåte.

7. Forslag til videre kartlegginger

7.1. Årlige kartlegginger av tolkebehov gjennom spørreundersøkelser hos tolkeformidlere og sektormyndigheter

Både tolkeformidlere og sektormyndigheter vil sitte på informasjon som er egnet til å belyse behovet for tolker innenfor ulike språk. Tolkeformidlere utgjør antagelig

den beste kilden for å få anslått dagens tolkebehov, fordi de har oversikt over hva som faktisk formidles av tolkeoppdrag. I den grad tolkeformidlerne faktisk formidler tilnærmet alle tolkeoppdrag som gjennomføres i Norge, vil en type spørreundersøkelser mot disse gi et godt bilde på omfanget fordelt på språk. SSB har ikke klart å få full visshet i om det gjennomføres tolkeoppdrag utenom registrerte tolkeformidlere. Hvis det er tilfelle, vil man i tillegg være nødt til å inkludere relevante sektormyndigheter i en slik spørreundersøkelse.

Det mest hensiktsmessige vil sannsynligvis være å gjennomføre en slik undersøkelse årlig, med spørsmål om antall oppdrag som er formidlet siste 12 måneder, fordelt på språk.

IMDis ambisjon om å bestemme framtidig tolkebehov, mener vi i utgangspunktet hviler på tre informasjonspilarer:

For det første er det nødvendig å danne seg et godt bilde av behovet, slik det foreligger per dags dato. En spørreundersøkelse mot tolkeformidlere (og eventuelt sektormyndigheter) hvor de bes oppgi antall tolkeoppdrag fordelt på språk siste 12 måneder, vil gi et godt inntak til å beskrive nåværende omfang.

For det andre vil det være behov for å framskrive tolkebehovet basert på dagens situasjon. Herunder må det foretas estimeringer av sannsynlig språkutvikling i populasjonen som har benyttet seg av tolker. På et eller annet tidspunkt har en person tilegnet seg tilstrekkelig språkkompetanse i norsk (eventuelt engelsk), som gjør behovet for tolk overflødig, i alle fall i de fleste aktuelle situasjoner.⁸ Dette kan det gjøres beregninger av på et aggregert nivå. Det må også foretas estimeringer av omfanget av familiegjennforening i den fremmedspråklige populasjonen, fordelt på språk. Det vil gi et bilde av ”nyrekrutteringen” i befolkningsgrupper som vil ha behov for tolk.

For det tredje vil framtidig tolkebehov være avhengig av hvilke personer som i framtiden gis adgang til Norge som for eksempel kvoteflyktninger. Dette er en vanskelig oppgave å forutsi da det vil være avhengig av politiske forhold i andre land, og hvilke folkegrupper Norge blir forespurt om å ta imot flyktninger fra. Når det gjelder omfanget av asylsøkere som vil komme i nærmeste framtid, vil vi tro at det må være UDI som vil være best egnet til å gi denne informasjonen.

For å foreta en årlig spørreundersøkelse mot tolkeformidlere og sektormyndigheter, må en først utarbeide et design som gjør at informasjonen vi får på bakgrunn av undersøkelsen er av tilfredsstillende kvalitet. Da må vi kunne identifisere populasjonen. I dette tilfellet vil det være tolkeformidlere, eventuelt supplert med relevante sektormyndigheter (se redegjøring over). Deretter må det avklares hvem som skal delta i undersøkelsen, det vil si avgrense populasjonen og bestemme utvalget. Videre må innsamlingsinstrumentet spesifiseres: Hva skal vi spørre om, og på hvilken måte. Til slutt må vi sørge for at vi har et utsendings- og innsamlingssystem på plass.

Når det gjelder å avklare populasjonen, finnes det i Nasjonalt tolkeregister en oversikt over tolker, og på Gule Sider en oversikt over tolkeformidlere. Disse kildene kan brukes som et utgangspunkt for å få oversikt over populasjonen med hensyn til tolkeformidlere og i en viss grad tolkere.⁹ Kommunene er også viktige brukere av tolketjenester, slik at alle landets kommuner bør inngå i populasjonen.

⁸ I situasjoner som fordrer svært høy språkkompetanse, for eksempel knyttet til juridiske eller krevende medisinske forhold, vil selv personer med relativ god språkkompetanse kunne ha behov for tolk.

⁹ Registrerte tolker vil imidlertid ikke omfatte alle praktiserende tolker i Norge, Registreringen, slik vi har forstått det, er basert på at tolken selv aktivt går inn og foretar registreringen av seg selv og det/de språk vedkommende tolker på. Vi har ikke klart å få klarlagt hvor stor andel av praktiserende tolker som har registrert seg.

Sentrale statlige institusjoner som IMDi og UDI vil også være en del av populasjonen. Samtidig må det vurderes hva som dekkes opp av tolker gjennom tolkeformidlere, og om disse kan supplere bildet. Hvorvidt hele populasjonen skal delta på undersøkelsen (fulltelling), eller om det skal trekkes et utvalg, avhenger av hvilken nøyaktighet en trenger i estimatene fra undersøkelsen. Benyttes et utvalg, vil nøyaktigheten avhenge av størrelsen på utvalget, samt fordelingen av fenomenet en ønsker å kartlegge. I liknende tilfeller som det vi diskuterer her, hvor populasjonen er relativt liten, kan det i utgangspunktet like gjerne benyttes en fulltelling. Da vil valget mellom fulltelling eller utvalg bli bestemt av om man gjennomfører undersøkelsen som personlig intervju eller som en selvutfyllingsundersøkelse på Web eller papir. Hvis en benytter intervjuere til å foreta data-innsamlingen, er det svært kostnadskrevende, og det vil være gunstig å trekke et utvalg. Hvilket innsamlingsdesign en skal velge avhenger blant annet av hvor kompleks undersøkelsen er, om det er behov for en type dialog i undersøkelsen for å avklare kompliserte spørsmål, samt undersøkelsens lengde. I en kartlegging av tolkeomfang mener vi det vil være hensiktsmessig med en selvadministrert undersøkelse, siden det sannsynligvis vil være få spørsmål som vil trenge personlig rettledning, gitt at spørsmålene er godt gjennomtenkt i forkant.

Vi anbefaler at undersøkelsen gjennomføres som en fulltelling over internett. Undersøkelsen bør være relativt kort. Formålet med undersøkelsen er å kartlegge framtidige tolkebehov, hvor det stilles spørsmål om hvilke språk tolkeformidlerne og sektormyndighetene mener det vil være behov for i fremtiden. Dette kan det imidlertid være utfordrende å spørre om. Spørsmål om forventninger og behov i fremtiden er komplisert for respondenter å svare på, og det er også komplisert å analysere. Før undersøkelsen gjennomføres bør tema, og eventuelt utkast til spørsmål, testes ut ved hjelp av anerkjente testmetoder som fokusgrupper¹⁰ og kognitiv kartlegging¹¹ på potensielle respondenter.

Når det gjelder utsending, vil det i forkant av undersøkelsen sendes et brev til respondentene med en invitasjon til å delta, og en url-adresse og et passord. Erfaringsmessig er det i undersøkelser rettet mot virksomheter vanskelig å finne rett/egnet svarperson, derfor bør det også gjøres et forarbeid med å kartlegge hvem i virksomheten brevet med invitasjon bør sendes til. Vi må også her regne med at det vil være en stor andel som ikke svarer ved første gangs utsending. I planleggingen bør en derfor ha foretatt antagelser om størrelsen på denne andelen, slik at en har et egnet opplegg for å sende påminningsbrev til disse. En bør også vurdere å ringe opp respondenter som ikke har svart etter gjentatte purringer.

7.2. Forslag til undersøkelse og relevant utvalg for årlige kartlegginger av minoritetsspråk

Den beste fremgangsmåten, for å få oversikt over den norske befolkningens språkpreferanse og kompetanse, vil være om disse opplysningene kunne hentes fra

¹⁰ Fokusgrupper er en serie med små møter som tar opp ett eller noen få tema i forbindelse med den planlagte undersøkelsen. Dette er målrettede og strukturerte gruppesamtaler, og deltakerne velges ut etter bestemte kriterier. Fokusgruppen ledes av en moderator (møteleder), som benytter en forhåndsskrevet intervjuguide. Tema for diskusjonen i fokusgruppen kan være hvilke begreper deltakerne bruker når de diskuterer temaet undersøkelsen skal handle om, hva deltakerne forstår med de begrepene vi har tenkt å bruke i skjemaet, hvorfor de synes noen spørsmål er lette og andre vanskelige å svare på, osv. Ved å snakke sammen inspirerer deltakerne hverandre til å tenke igjennom forskjellige sider ved det spørsmålet som diskuteres. Ved å lytte til samtalen får vi ideer om hva vi bør spørre om og hvordan vi bør ordlegge oss i skjemaet.

¹¹ Kognitiv kartlegging er en observasjons- og intervjueteknikk som brukes for å kartlegge hvordan svarpersoner oppfatter spørsmålene i et spørreskjema, og hvordan de resonnerer seg frem til et svar. Fokusgrupper brukes gjerne helt i starten av planleggingen av undersøkelsen før vi har begynt å lage spørsmål, mens kognitiv kartlegging er nyttig for å avdekke svakheter etter at det foreligger et utkast til skjema.

administrative registre. Det vil danne et godt grunnlag for en norsk språkstatistikk. Opplysninger innhentet gjennom slike systemer har per i dag, som vi har drøftet over, ikke god nok kvalitet. Dersom man ønsker detaljerte og presise språkdata om utbredelsen av fremmedspråk i Norge, må de eksisterende registrene ha bedre kvalitet. For å få bedre kvalitet må det utarbeides gode spesifiseringer og operasjonaliseringer av hva som skal registreres av språk, hvordan språk skal registreres, og det bør gjøres obligatorisk. Opplysningene vil dessuten være knyttet til personnummer, slik at en også vil ha mulighet til å ta opplysningene inn i Det Sentrale Personregisteret.

Det vil imidlertid være en relativt lang vei å gå før vi har klart å oppnå en fullgod dekning i de administrative registrene. Samtidig vil det meste av denne registreringer foregå i forbindelse med utlendingers etablering i Norge, og det vil derfor være et etterslep med hensyn til språkdata over eksisterende ”fremmedspråklig” populasjon.

Dersom en ønsker oversikt over utbredelsen av språk i Norge, med en kortere tidshorison enn hva en må gjøre regning med for å få registrene fullgode, kan det være et alternativ å gjennomføre en stor utvalgsundersøkelse. Det er samtidig viktig å påpeke at det vil være mangler ved en slik utvalgsbasert språkstatistikk. Det mest iøynefallende problemet er selvsagt at det vi ønsker å måle, altså språkkompetanse, er så sterkt knyttet til muligheten til å kunne delta i undersøkelsen. Sagt på en annen måte, det vil være vanskelig å få i tale nettopp de personene som ikke behersker språket vi henvender oss på (norsk, engelsk eller eventuelle andre språk).

Som et utgangspunkt måtte en ha gått ut til et stort utvalg fra populasjonen. Utvalget måtte stratifiseres etter landbakgrunn, slik at innvandrere blir disproporsjonalt overrepresentert. For å få sikre nok estimater måtte utvalget bestå av et minimum på 15 000 personer. Dette måtte gjøres som en selvutfyllingsundersøkelse, gjerne som en kombinasjon av papir og undersøkelse gjennomført på internett, og undersøkelsen burde oversettes til flere språk for å øke tilgjengeligheten. Det bør være en kort undersøkelse på om lag ett A4-ark. Gangen i undersøkelsen kan skisseres slik:

- Et utvalg på ca. 15 000 trekkes fra SSBs registre over befolkning med landbakgrunn
- Brev med invitasjon sendes via post til de som er trukket ut (språk tilpasset respondentene)
- Deltakerne inviteres til å svare via Web
- Deltakerne får et takkekort/oppfordring om å svare hvis de ikke har svart
- Deltakerne som ikke har svart får undersøkelsen per post

Vi kan anta at ca. halvparten vil svare på en slik undersøkelse, noe som vil gi et utvalg på ca. 7 500. Vi har ikke foretatt nøyaktige variansberegninger. Det kan først gjøres når vi har bestemt hvilke undergrupper vi skal bryte ned statistikken på. En slik utvalgundersøkelse vil gi brukbare data over utbredelsen av de større språkgruppene, men vil sannsynligvis gi vesentlig lavere datakvalitet for mindre språk. Utfordringen er at antall talere i de mindre språkene vil bli forholdsvis få og gi et spinkelt grunnlag for statistikk, samtidig som vi antar at frafallet vil være større innenfor disse små språkgruppene.

8. Oppsummeringer og konklusjoner

Denne rapporten har tatt for seg sentrale kilder for informasjon om forekomst og utbredelse av fremmedspråk i Norge. Formålet har vært å kartlegge eksisterende og mulige framtidige kilder. Vi har vurdert flere ulike administrative systemer hvor opplysninger om språk blir registrert. Fire systemer pekte seg ut som sentrale, og

det er disse som er diskutert i rapporten. Vi mener at disse fire dekker den språkregistrering som finnes i Norge i dag.

I rapporten peker vi på noen svakheter ved disse registrene. De kan oppsummeres i fire hovedpunkter:

1) Relevans

Registreringen bærer preg av å være ad hoc og ser ikke ut til å være noen prioritert oppgave for etatene. Det er dessuten liten etterspørsel etter opplysninger om språk, og opplysningene blir i liten grad benyttet av registrenes eiere. Det virker således som om opplysninger om språk ikke har vært oppfattet som veldig relevante.

2) Datakvalitet

Generelt finner vi at kvaliteten på rapporteringen av språk er dårlig, det finnes ingen "gullstandard" for hvordan språk skal rapporteres. Det er usikkert om opplysninger gjelder morsmål, førstespråk, det språket en behersker best, foretrukket språk, eller annet. Dette gjør at kvaliteten på dataene ikke er særlig god. Vi har i rapporten diskutert noen grep som kan forbedre datakvaliteten i foreliggende registre.

3) Kodeverk

Det foreligger per i dag ikke et nasjonalt standardisert kodeverk som benyttes på tvers av registrene. Registrene fra UDI er faktisk de eneste som i det hele tatt benytter et kodeverk. Dette kodeverket er imidlertid beheftet med store svakheter.

4) Dekningsgrad

Samlet gir registrene opplysninger om språk bare for en liten del av befolkningen – samlet inneholder registrene opplysninger om rundt 22 prosent av innvanderbefolkningen. Det finnes riktignok et potensial i andre registre som favner en videre del enn de 22 prosentene som allerede er dekket. Opplysninger om språk blir for eksempel registrert i GSI når foresatte søker om morsmålsopplæring, men foreløpig er det utfordringer med begge disse registrene.

Med bakgrunn i disse punktene er vår konklusjon at det er vanskelig å hente ut god og tilstrekkelig informasjon om språksituasjonen i Norge i dag fra de eksisterende administrative registrene. Under skisserer vi tre alternativer for hvordan vi kan få bedre opplysninger om språk:

1) Forbedre kvaliteten på eksisterende registre

Det første alternativet er å forbedre kvaliteten på eksisterende registre. Først trengs det å innføre en standard for hvordan man spør om språk, slik at man får gode og sammenlignbare resultater. Deretter må språk prioriteres som en viktig opplysning av registreierne. Det vil si at etatene må vite hvorfor de skal samle inn opplysninger om språk, hva det skal brukes til, hvorfor det er viktig med nøyaktige opplysninger og hvorfor språkstatistikk er viktig. Hvis man klarer å etablere et felles register basert på ett felles kodeverk på tvers av etatene, vil mye være gjort.

Likevel, selv om disse tiltakene gjennomføres, vil det fremdeles være et dekningsproblem. Per i dag foreligger det språkopplysninger på ca. 22 prosent av innvanderbefolkningen i registrene. Det ligger noen muligheter i registrering av språk via internett, men foreløpig er denne andelen svært lav. Det er dessuten mulig å gjennomføre estimeringer på bakgrunn av informasjon vi henter fra befolkningsregisteret (BeReg). For det første kan man estimere språk ut i fra fødeland. For enkelte land vil dette være rimelig greit da de har et offisielt språk, mens for andre områder kan dette være en større utfordring. For det andre kan man gjøre estimeringer ut i fra hvem man er i familie med. Dersom det finnes opplysninger om språk på en person i familien, kan (denne) disse overføres til øvrig familie.

Selv om vi gjør slike estimeringer, klarer vi sannsynligvis ikke å dekke hele populasjonen. Dermed må registrene suppleres med fulltelling og utvalgsundersøkelser. For enkelte grupper kan det tenkes at det holder å gjennomføre utvalgsundersøkelser, mens for andre grupper trengs fulltelling. Avgjørelser på dette må gjøres ut i fra det man vet om språk (er det flere språk i landet?), frafall (er det grupper som er vanskelig å få med?), og språkkompetanse (hvordan skal vi få de som er trukket ut til å forstå og svare på spørsmålene?).

2) Utvalgsundersøkelser

Et annet alternativ er å gjennomføre utvalgsundersøkelser rettet mot innvandrerbefolkningen (jamfør kapittel 5 og 7). Man bør gå ut til et stort utvalg av innvandrerbefolkningen og stratifisere etter landbakgrunn, slik at enkelte innvandrergrupper blir disproporsjonalt overrepresentert.

Det er flere utfordringer ved å gjennomføre utvalgsundersøkelser hvor formålet er å samle inn informasjon om språk i innvandrerbefolkningen. Ut fra tidligere undersøkelser vet vi at frafallet er relativt høyt i innvandrerbefolkningen. Det krever at undersøkelsen må være godt tilrettelagt, og at det legges ned store ressurser for å få tak i personene som er trukket ut. Erfaringsmessig er det et stort problem at kontaktopplysningene om innvandrere er for dårlige, og at vi ikke klarer å finne rette vedkommende (Paulsen og Wilhelmsen 2008). Deretter er det en utfordring å vite hvilket språk vi bør benytte når vi kontakter respondenten, og at respondenten må forstå formålet med undersøkelsen. Det innebærer at det må legges ned et omfattende arbeid med å oversette undersøkelsen til sentrale språk, og utarbeide informasjon om undersøkelsen til sentrale institusjoner og grupper. Ved å gjennomføre en slik utvalgsundersøkelse vil vi få god kvalitet på dataene vi samler inn. Dette ved at vi kan få utviklet gode og entydige spørsmål, blant annet ved å forhåndsteste dem på respondenter. Dette sikrer at spørsmålene blir forstått på samme måte av dem som skal besvare dem.

Hvis man velger å gjennomføre en utvalgundersøkelse, vil dette gi et bilde av språksituasjonen i Norge for de største og viktigste språkgruppene. Man vil altså ikke kunne få en fullstendig oversikt over alle minoritetsspråk som finnes i Norge, men bare over de som har en viss tallmessig størrelse.

3) Registrering av språk i møte med det offentlige Norge

Det siste alternativet ligger utenfor det som er problemstillingen i denne rapporten, og ligger også utenfor ansvarsområdet til Statistisk sentralbyrå. Dermed er det siste alternativet mer et område man kan diskutere relevansen av. Vår gjennomgang har vist at det er svakheter med alle alternativene, og det som foreligger per dags dato. For å få en skikkelig registrering av språk, bør det utvikles en god måte å registrere dette på. Det bør også registreres første gang man møter det offentlige Norge, eller når man får et personnummer. Siden dette ligger utenfor ansvarsområdet til SSB, går vi ikke inn i en diskusjon om dette, men antyder dette som et alternativ for å få til en god språkregistrering.

Alle disse tre alternativene vil gi et bilde av språkfordelingen i Norge i dag, men de gir lite informasjon om hvordan språksituasjonen blir framover. Jamfør det vi har diskutert i kapittel 7 er det vanskelig å si noe om framtiden. Vi har pekt på muligheter til å estimere framskrivninger i forhold til familieegjenforening, men ellers mener vi det er vanskelig å si noe konkret om framtiden. I denne rapporten har vi derfor fokusert på å gi et bilde av språkfordelingen slik den er nå.

Referanser

Brown, K. og Ogilvie, S. (2009). *The Concise Encyclopaedia of Languages of the World*. Elsevier Ltd. Oxford.

Grønmo, S. (1996). "Forholdet mellom kvalitative og kvantitative tilnæringer i samfunnsforskning" i: Holter, H. og Kalleberg, R. (red.). *Kvalitative metoder i samfunnsforskning*. Universitetsforlaget, Oslo.

Guthu, L. og Holm S. (2010). *Innvandrere på nett. En analyse av innvandreres digitale kompetanse*. VOX, Oslo.

Det kongelige barne-, likestillings- og inkluderingsdepartement. (2012). *Lov om introduksjonsordning og norskopplæring for nyankommende innvandrere (introduksjonsloven)*. Rundskriv. Q-20/2012.

Kvale, S. (2004). *Det kvalitative forskningsintervju*. Ad Notam Gyldendal, Oslo.

Paulsen, K.A.S. og Wilhelmsen, M. (2010). *Kultur og mediebruksundersøkelsen blant innvandrerbefolkningen 2008. Dokumentasjonsrapport*. Notater 27/2010. Statistisk sentralbyrå. Oslo.

LandInfo. Utlendingsforvaltningens fagenhet for landinformasjon (2008). *Språkforhold i Afrika sør for Sahara*. Notat. LandInfo - 4. april 2008.

Benyttede nettsider:

<http://www.imdi.no/no/Introduksjonsordningen/Norskopplaring-for-innvandrere>
www.wis.no/gsi

http://www.difi.no/statskonsult/publik/rapporter/fulltekst/r2002_11.pdf/

<https://www.skatteetaten.no/Alt-om/Folkeregistrering/>

<http://www.ssb.no/befolkning/>

Semistrukturerte intervju:

Semistrukturerte intervju gjennomført med UDI 27. mars 2012

Semistrukturerte intervju gjennomført med IMDi 27. mars 2012

Vedlegg 1. Intervjuguide IMDi

1) Om prosjektet

I dag er det ingen offentlig registrering av hvilke språk innbyggerne i Norge snakker – ei heller hvor mange talere det er i hvert enkelt språk. IMDi ønsker bedre statistikk på dette området for å kunne estimere framtidige behov for tolker. Statistisk sentralbyrå skal på oppdrag fra IMDi systematisere og analysere tilgjengelige datakilder rundt registrering av språk. Dette er et forprosjekt, hvor vi går igjennom hva som finnes av språkregistrering, andre datakilder som kan benyttes, og hva som gjenstår for å kunne utarbeide god statistikk om språk. Prosjektet skal munne ut i en rapport der vi skal framstille en:

- 1) En systematisering av eksisterende og evt. kommende datakilder
- 2) En analyse av datakildene
- 3) en opptelling av språk funnet gjennom analysen av de aktuelle datakildene
- 4) Forslag til undersøkelse og utvalg

2) Hva registres av språk i Nasjonalt Introduksjonsregister?

- Introduksjonsloven – samfunnsfag tilrettelagt på et språk man forstår – hvordan får man denne informasjonen?

3) Hvordan foregår registreringen? Kan dere ta oss gjennom gangen i registreringen.

- Hvor får dere data om språk fra? Hvilke kilder? (I forbindelse med at innvandrere som har fått opplæring i norsk og samfunnskunnskap går opp til norskprøver, fyller de ut et såkalt "persondataskjema". Dette administreres av Norsk språktest (Bergen).)
- Kommunene?

4) Hvordan er kvaliteten på dataen?

- Dersom man ikke bor i kommunen man har fått tildelt?
- Hvordan er dekningen av registreringen av de som takker ja til tilbudet?
- Kun asylsøkere som har krav og plikt?

Intervjuguide UDI

1) Om prosjektet

I dag er det ingen offentlig registrering av hvilke språk innbyggerne i Norge snakker – ei heller hvor mange talere det er i hvert enkelt språk. IMDi ønsker bedre statistikk på dette området for å kunne estimere framtidige behov for tolker. Statistisk sentralbyrå skal på oppdrag fra IMDi systematisere og analysere tilgjengelige datakilder rundt registrering av språk. Dette er et forprosjekt, hvor vi går igjennom hva som finnes av språkregistrering, andre datakilder som kan benyttes, og hva som gjenstår for å kunne utarbeide god statistikk om språk. Prosjektet skal kunne ut i en rapport der vi skal framstille en:

- 5) En systematisering av eksisterende og evt. kommende datakilder
- 6) En analyse av datakildene
- 7) en opptelling av språk funnet gjennom analysen av de aktuelle datakildene
- 8) Forslag til undersøkelse og utvalg

2) Hva registreres av språk?

- Slik vi har forstått det registreres språk for alle skjema for opphold (asyl, familie, arbeid m.m.) - - hva skjer med denne informasjonen?
- Hvorfor registreres ikke språk på enkelte oppholdsgrupper?
- Er det komplisert å få til en registrering av denne informasjonen?
- Hva registreres i DUF?
- Hva registreres i systemet for asylintervjuet?
- Språk som er relevant for integrering.
- Hva slags språk registreres (morsmål, det man kan kommunisere på)?

3) Hvordan foregår registreringen? Kan dere ta oss gjennom gangen i registreringen.

- I hvilke systemer registreres de ulike språk variablene?
- Hvordan bruker dere denne informasjonen?
- Hvor stor er dekningen av språk registrering

4) Hvordan er kvaliteten på dataen?

- Kan dere sjekke ut kvaliteten på DUF mot kvaliteten på det som ligger i systemet dere bruker under asylintervjuet.
- Hvordan er dekningen på registreringen?

Vedlegg 2. Rett og plikt til introduksjonsprogram i henhold til introduksjonsloven

Formålet med introduksjonsloven er å styrke nyankomne innvandreres mulighet for deltakelse i yrkes- og samfunnsliv, og deres økonomiske selvstendighet. Rett og plikt til deltakelse i introduksjonsprogrammet (jamfør introduksjonsloven paragraf 2 første ledd) gjelder for nyankommet utlending mellom 18 og 55 år, som har behov for grunnleggende kvalifisering, og som har fått oppholdstillatelse på bakgrunn av en av grunnlagene under:

- Asyl (jamfør utlendingsloven paragraf 28)
- Oppholdstillatelse etter innreisestillatelse som overføringsflyktning (jamfør utlendingsloven paragraf 35 tredje ledd)
- Fornybar oppholdstillatelse etter utlendingsloven paragraf 38 første ledd på grunnlag av søknad om asyl, eller kollektiv beskyttelse i massefluktsituasjon (jamfør utlendingsloven paragraf 34)
- Oppholdstillatelse som familiemedlemmer til personer som nevnt i kulepunktene over (jamfør utlendingsloven kapittel 6). For å være omfattet av denne bestemmelsen er det et vilkår at personene nevnt over ikke har vært bosatt i en kommune i mer enn fem år før det søkes om familiegjeningforening
- Fornybar oppholdstillatelse på selvstendig grunnlag (jamfør utlendingsloven paragraf 53 første ledd bokstav b, paragraf 53 annet ledd eller paragraf 53 tredje ledd). For personer som har fått oppholdstillatelse på selvstendig grunnlag, gjelder retten og plikten til å delta i introduksjonsprogram i to år fra det tidspunkt det er gitt oppholdstillatelse på et av de nevnte grunnlagene i utlendingsloven paragraf 53.

Rett og plikt til deltakelse i introduksjonsprogrammet gjelder bare for personer som er bosatt i kommunen i henhold til særskilt avtale mellom utlendingsmyndighetene og kommunen. Kravet om særskilt avtale gjelder likevel ikke for familiemedlemmer til personer som nevnt i de tre første kulepunktene over, eller for personer som har fått oppholdstillatelse på selvstendig grunnlag. Som nyankommet regnes den som har vært bosatt i en kommune i mindre enn to år når vedtak om deltakelse i introduksjonsprogrammet skal treffes. Nordiske borgere og utlendinger som omfattes av Avtale om Det europeiske økonomiske samarbeidsområdet (EØS-avtalen) faller utenfor introduksjonsordningen (jamfør Det kongelige barne-, likestillings- og inkluderingsdepartement, rundskriv Q-20/2012).

Rett og plikt til norskopplæring

Det å ha rett og plikt til norskopplæring etter introduksjonsloven paragraf 17, vil si at man har krav på gratis norskopplæring fra kommunen maner bosatt i (rett), og at man må ha gjennomført lovpålagt norskopplæring dersom man på et senere tidspunkt skal søke om permanent oppholdstillatelse og norsk statsborgerskap (plikt). (<http://www.imdi.no/no/Introduksjonsordningen/Norskopplaring-for-innvandrere>)

Det er noen personer som har plikt, men ikke rett, til slik opplæring. Dette vil si at de ikke får opplæringen gratis fra kommunen. Personene som har plikt til slik opplæring er:

- Arbeidsinnvandrere utenfor EØS/EFTA- området

Familiiegjenforente med arbeidsinnvandrere utenfor EØS/EFTA- området. Disse personene må ha gjennomført lovpålagt opplæring etter introduksjonsloven dersom de senere ønsker å søke om permanent oppholdstillatelse eller statsborgerskap i Norge, men de må betale for opplæringen selv.

(<http://www.imdi.no/no/Introduksjonsordningen/Norskopplaring-for-innvandrere>).

Som en del av rett og/eller plikt til norskopplæring, inngår 50 timer samfunnskunnskap. Kommunene plikter å tilrettelegge disse kursene på et språk personene behersker. Kommunene får imidlertid begrensede opplysninger om språk fra utlendingsforvaltningen registreringene i NIR ofte er mangelfulle (møte med IMDi 27.mars 2010).

Vedlegg 3. Morsmålsopplæring og antall elever fordelt innenfor hver språkgruppe.

Språk	2002 Elevar	2003 Elevar	2004 Elevar	2005 Elevar	2006 Elevar	2007 Elevar	2008 Elevar	2009 Elevar	2010 Elevar	2011 Elevar	Sum 2002 - 2011 Elever
Norsk	7	0	3	2	21	0	0	0	0	0	33
Afrikaans	1	2	1	10	7	4	7	4	8	3	47
Akan (twi)	6	11	9	7	8	2	1	8	7	19	78
Albansk	1337	1405	1274	1288	1283	1210	1095	1076	932	761	11661
Amharisk	57	59	59	54	66	53	58	70	76	72	624
Arabisk	1680	1777	1737	1830	1940	1939	1913	2271	2423	2334	19844
Armensk	14	15	16	10	21	20	14	12	9	2	133
Azerbaidjansk	4	9	9	6	6	7	4	1	3	2	51
Bahasa indonesia	1	3	6	5	17	10	0	6	7	6	61
Baluchi	0	4	1	3	3	0	1	2	8	1	23
Bemba	1	0	0	0	0	0	0	0	0	1	2
Bengali	23	24	17	22	30	24	21	19	21	29	230
Berbersk	46	52	59	56	36	55	30	32	39	28	433
Bosnisk	949	1002	875	806	754	625	518	505	437	336	6807
Bulgarsk	18	17	17	18	29	26	30	27	34	49	265
Burmesisk	3	1	9	58	114	158	110	128	150	137	868
Chichewa	0	0	0	0	0	0	7	8	0	1	16
Dansk	14	6	8	5	4	7	4	5	3	6	62
Dari	247	343	375	493	526	413	395	513	616	503	4424
Divehi	0	0	0	0	0	1	3	0	1	1	6
Dzhongkha	0	0	0	0	0	0	0	0	2	0	2
Engelsk	520	506	623	654	667	618	480	469	424	380	5341
Estlandsk	13	14	12	17	20	20	13	19	27	33	188
Ewe	0	0	2	3	1	0	0	4	3	0	13
Fang	0	0	0	0	0	0	0	0	1	2	3
Filippinsk	147	174	169	172	207	181	184	190	209	284	1917
Finsk	148	142	122	94	67	66	37	24	22	17	739
Flamsk	1	1	3	2	6	7	5	0	2	0	27
Fransk	89	136	119	135	138	132	136	129	82	76	1172
Fulani	0	0	0	0	0	0	1	0	2	0	3
Færøysk	18	5	0	0	8	4	3	1	0	1	40
Georgisk	5	2	4	4	5	3	1	1	5	6	36
Gresk	24	16	16	21	18	16	11	11	17	22	172
Grønlandsk	0	0	1	2	2	1	0	4	2	1	13
Gujarati	1	2	6	2	3	2	4	3	0	0	23
Hausa	0	0	0	0	0	2	0	0	1	2	5
Hebraisk	18	20	17	17	17	15	10	5	1	1	121
Hindi	66	59	58	42	61	40	32	24	33	19	434
Hviterussisk	6	2	4	6	7	7	4	4	3	3	46
Ibo	0	0	0	0	0	0	4	2	2	5	13
Irsk	3	0	1	0	1	2	1	1	2	1	12
Islandsk	147	116	102	76	68	63	47	123	128	166	1036
Italiensk	12	10	15	17	11	18	22	24	24	16	169
Japansk	7	2	0	3	7	9	7	9	8	8	60
Kantonesisk	43	50	53	51	40	31	30	37	19	26	380
Kasakhisk	3	3	1	1	0	1	1	0	1	0	11
Khmer	10	10	7	9	11	16	20	23	17	17	140
Khoi-san	0	0	0	0	0	0	0	1	0	0	1
Kinesisk	135	152	135	121	158	137	128	114	92	120	1292
Kinyarwanda	12	20	14	16	36	50	37	25	11	11	232

Kirgisisk	0	0	1	1	2	1	0	5	6	2	18
Kirundi	1	22	36	36	60	60	50	62	57	47	431
Koreansk	10	6	11	16	23	11	10	8	12	9	116
Kpelle	0	0	0	0	0	0	0	5	0	0	5
Kreol*	4	1	3	8	1	1	1	2	1	0	22
Krio	1	1	0	1	5	1	7	5	5	2	28
Kroatisk	58	60	46	38	48	48	41	28	20	26	413
Kurdisk	1546	1859	1749	1735	1710	1509	1423	1430	1337	1249	15547
Lao	0	0	0	0	0	0	0	0	0	1	1
Latvisk	10	15	9	13	22	27	35	42	89	134	396
Lingala	9	7	3	2	8	5	6	10	9	9	68
Litauisk	12	19	34	70	137	161	262	300	474	732	2201
Luganda	0	0	1	4	1	4	5	6	2	3	26
Madagassisk	2	2	0	0	1	0	2	1	4	5	17
Makedonsk	4	7	2	3	1	3	1	0	7	3	31
Malayalam	0	0	0	2	2	2	2	2	1	4	15
Malayisk	9	10	18	21	18	22	18	12	6	10	144
Maltesisk	1	1	1	1	0	0	0	0	0	0	4
Mandingo	1	6	4	1	3	5	15	5	9	4	53
Marathi	0	0	0	0	0	1	0	2	1	1	5
Mende	2	6	8	1	5	5	0	2	0	1	30
Moldavisk	2	2	1	1	2	3	2	6	0	0	19
Mongolsk	0	1	2	1	6	3	3	3	0	4	23
Nederlandsk	94	104	118	180	230	243	192	128	140	88	1517
Nepali	13	12	17	19	13	7	13	37	44	45	220
Nyanja	0	0	0	0	3	1	7	9	5	2	27
Oromo	53	65	37	57	58	39	29	62	51	53	504
Pashto	141	183	123	163	197	153	131	158	133	127	1509
Persisk	517	541	503	519	478	407	286	295	248	289	4083
Polsk	185	196	239	353	641	1234	1702	2023	2149	2262	10984
Portugisisk	66	100	95	107	134	151	176	141	129	149	1248
Punjabi	331	238	217	229	241	212	151	161	98	70	1948
Romani	3	6	5	4	8	4	2	7	10	3	52
Rumensk	16	24	24	25	24	27	55	99	146	166	606
Russisk	497	606	726	639	641	619	506	559	496	502	5791
Samisk	6	9	0	0	0	0	0	0	0	0	15
Serbisk	174	140	177	155	153	142	151	143	121	106	1462
Shona	0	1	0	1	1	0	0	0	1	1	5
Sindhi	2	0	0	0	0	0	0	0	0	0	2
Singalesisk	7	3	2	1	3	3	3	2	2	1	27
Slovakisk	9	8	12	10	21	19	17	26	43	40	205
Slovensk	3	3	1	1	0	0	3	2	4	5	22
Somalisk	1644	1891	1983	2175	2388	2460	2570	2633	2584	2553	22881
Spansk	660	603	567	572	528	514	439	404	379	343	5009
Suno	0	0	0	0	0	0	0	0	0	2	2
Susu	0	1	0	0	0	0	0	1	0	4	6
Svensk	11	5	9	5	8	5	4	3	5	4	59
Swahili	58	97	117	149	163	195	185	211	184	148	1507
Syriac - assyrisk	5	4	9	10	11	3	5	13	12	12	84
Tadsjikisk	0	0	1	4	1	0	1	3	10	4	24
Tamil	697	740	717	764	934	853	806	784	621	512	7428
Tatarisk	0	0	0	0	0	0	0	0	1	0	1
Telugu	0	0	0	0	1	1	2	0	0	1	5
Temne	1	0	0	0	0	0	0	0	0	0	1
Thai	233	341	419	531	681	633	639	687	687	640	5491
Tibetansk	0	0	0	0	0	0	0	0	0	1	1

Tigrè	2	4	0	4	2	1	1	1	0	1	16
Tigrinia	36	54	61	82	82	101	135	194	316	375	1436
Tonga	0	0	0	0	1	0	0	0	0	0	1
Thonga	0	0	0	0	0	0	5	0	0	0	5
Tsjekkisk	11	7	11	12	14	18	11	12	8	9	113
Tsjetsjensk	20	69	241	299	355	309	356	389	389	366	2793
Turkmensk	13	12	4	10	4	8	11	21	32	10	125
Tyrkisk	933	915	926	1026	1137	1113	1031	980	824	728	9613
Tysk	251	223	262	262	318	410	472	408	354	302	3262
Ukrainsk	9	12	12	16	29	18	18	19	20	22	175
Ungarsk	19	20	12	12	21	10	19	29	47	36	225
Urdu	2808	2517	2590	2521	2395	2537	2440	2362	2214	1984	24368
Usbekisk	0	1	1	3	6	6	5	5	5	11	43
Vietnamesisk	1575	1585	1483	1551	1496	1434	1272	1141	979	681	13197
Wolof	6	3	3	4	9	4	1	5	9	6	50
Yoruba	2	2	2	0	0	0	0	0	0	2	8
Zulu	1	0	0	0	0	0	0	0	0	0	1
Abkhasisk	0	0	2	0	0	2	1	2	0	1	8
Adja	0	1	2	3	4	0	3	0	5	1	19
Aizo	0	0	0	0	0	0	0	1	1	0	2
Aseri-tyrkisk	2	6	3	3	9	9	8	7	1	2	50
Bamileke	0	0	0	0	0	1	0	0	8	0	9
Bariba	0	0	0	0	0	0	0	0	1	0	1
Basjkirsk	0	0	0	0	0	0	1	0	0	0	1
Berberi- manga	0	0	0	0	0	0	1	0	0	0	1
Bhojpuru	0	0	1	1	1	0	0	0	0	0	3
Bissa	0	0	1	13	1	2	2	3	0	1	23
Cebuano	7	8	9	12	9	11	7	5	3	2	73
Crioulo	0	1	0	0	0	0	0	0	1	1	3
Darginsk	0	0	0	0	0	2	2	0	0	0	4
Diola	0	0	0	0	0	1	0	0	0	0	1
Djerma	0	0	0	3	2	3	0	0	0	0	8
Eshira	0	0	0	0	0	0	0	0	0	1	1
Ga-adangbe	0	1	0	0	0	1	0	0	1	0	3
Galicisk	0	0	0	0	0	0	0	1	1	3	5
Grebo	0	4	0	0	0	0	0	0	0	0	4
Hakka	0	0	0	0	0	0	0	0	0	3	3
Iban	0	0	0	0	1	0	0	0	0	0	1
Ilocano	0	0	0	0	0	0	0	0	0	1	1
Kabardinsk	0	0	0	0	0	0	0	12	0	0	12
Kannada	0	2	0	0	0	0	0	0	0	0	2
Karen	1	0	0	11	41	68	86	131	104	90	532
Kashin	0	0	0	1	0	0	0	1	0	0	2
Khasi	1	0	1	0	0	0	1	0	0	0	3
Kikongo	0	0	0	0	0	0	3	8	8	2	21
Kikuyu	1	1	0	0	0	1	0	0	0	0	3
Kimbundu	0	0	0	0	0	0	0	0	2	0	2
Kongo	0	0	0	7	15	20	23	9	30	14	118
Konkani	0	1	0	0	0	0	0	0	0	0	1
Kono	0	2	2	0	0	0	0	0	0	0	4
Krahn	0	0	0	0	0	0	0	0	1	0	1
Ladino	4	4	0	0	0	0	0	0	0	0	8
Lak	0	0	0	0	0	0	0	1	0	0	1
Limba	0	0	0	0	0	0	0	1	0	0	1
Lorma	0	0	0	0	0	0	0	0	0	1	1
Luo	1	3	8	2	2	0	0	0	0	0	16

Malawi	0	0	0	0	0	0	0	0	1	2	3
Mandé	1	3	5	3	0	2	0	0	0	0	14
Mine	0	0	0	1	1	0	0	0	0	0	2
Moru	2	8	0	0	0	0	0	3	0	0	13
Niloti	5	0	0	0	0	0	0	0	0	0	5
Oriya	0	0	0	0	0	0	0	0	0	1	1
Ovambo	0	0	0	0	0	0	0	0	0	1	1
Pahari	0	0	0	0	2	0	0	2	1	0	5
Romansh	2	2	0	0	0	0	0	3	2	1	10
Sango	0	0	0	0	1	0	0	0	0	0	1
Sanskrit	0	2	1	0	0	0	0	0	0	0	3
Songhai	0	1	0	0	0	0	0	0	0	0	1
Tem	0	0	4	0	0	0	0	0	0	0	4
Tuvinsk	0	0	0	0	0	0	0	0	1	0	1
Ujgurisk	4	6	6	18	21	13	11	13	13	22	127
Umbundu	0	0	0	0	0	0	0	0	0	1	1
Ingusjisk	0	0	0	0	0	0	0	2	6	4	12
Annet	97	97	54	71	147	187	843	103	136	146	1881
Sum	18748	19695	19713	20717	22166	22084	22156	22335	21781	20721	210116

Vedlegg 4. Kodeverket som benyttes av UDI

CODE	NAME	DESCRIPTION
ABK	Abkhazian	Abkhazian
ACE	Acehenese	Achinese
ADA	Adangme	Adangme
AFH	Afrihili	Afrihili
AFR	Afrikaans	Afrikaans
AKK	Akkadian	Akkadian
ALE	Aleutisk	Aleut
ALG	Algonquian languages	Algonquian languages
ANG	English (old)	English (old)
ARA	Arabisk	Arabic
ARC	Aramic	Aramic
ARN	Araua	Araua
ARP	Arapaho	Arapaho
ARW	Arawak	Arawak
ATH	Athapascan	Athapascan
AUS	Australian	Australian
AVE	Avestan	Avestan
AYM	Aymara	Aymara
BAD	Banda	Banda
BAI	Bamileke	Bamileke
BAL	Balutsjisk	Baluchi
BAM	Bambara	Bambara
BAQ	Baskisk	Basque
BAT	Baltic	Baltic
BEL	Hviterussisk	Belarusian
BEN	Bengali	Bengali
BHO	Bhojpuri	Bhojpuri
BIH	Bihari	Bihari
BIN	Bini	Bini
BLA	Siksika	Siksika
BRA	Braj	Braj
BTK	Batak	Batak
BUA	Buryat	Buriat
BUL	Bulgarsk	Bulgarian
CAD	Caddo	Caddo
CAI	Central American Indian	Central American Indian
CAT	Katalansk	Catalan / Valencian
CEB	Cebuano	Cebuano
CHA	Chamorro	Chamorro
CHE	Tsjetsjensk	Chechen
CHG	Chagatai	Chagatai
CHK	Chukchi	Chuukese
CHN	Chinook jargon	Chinook jargon
CHO	Choctaw	Choctaw
CHR	Cherokee	Cherokee
CHV	Chuvash	Chuvash
CHY	Cheyenne	Cheyenne
COM	Coptic	Coptic
COS	Corsican	Corsican
CPF	Creoles and pidgins (french- based)	Creoles and pidgins (french- based)
CPP	Creoles and pidgins (portuguse- based)	Creoles and pidgins (portuguse- based)
CRE	Cree	Cree
CRP	Creoles and pidgins	Creoles and pidgins
CZE	Tsjekkisk	Czech

DAN	Dansk	Danish
DAY	Dayak	Dayak
DEL	Delaware	Delaware
DEN	Dende	Slave
DGR	Dogrib	Dogrib
DIN	Dinka	Dinka
DIV	Divehi	Divehi
DOI	Dogri	Dogri
DRA	Dravidian	Dravidian
DUA	Duala	Duala
DUT	Hollandsk	Dutch
DYU	Dyula	Dyula
DZO	Dzongkha	Dzongkha
EFI	Efik	Efik
EGY	Egyptian	Egyptian
EKA	Ekajuk	Ekajuk
ELX	Elamite	Elamite
ENG	Engelsk	English
EPO	Esperanto	Esperanto
EST	Estisk	Estonian
EWE	Ewe	Ewe
EWO	Ewondo	Ewondo
FAN	Fang	Fang
FAO	Faroese	Faroese
FAT	Fanti	Fanti
FIJ	Fijian	Fijian
FIN	Finsk	Finish
FIU	Finno-ugrian	Finno-ugrian
FON	Fon	Fon
FRE	Fransk	French
FRY	Frisisk	Frisian
FUL	Fullah	Fulah / fra Sierra Leone
FUR	Friulian	Friulian
GAY	Gayo	Gayo / Indonesisk
GBA	Gbaya	Gbaya
GIL	Gilaki	Gilbertese
GLA	Gælisk	Gaelic
GLE	Irish	Irish
GLG	Gallegan	Gallegan
GLV	Manx	Manx
GON	Gondi	Gondi
GOT	Gothic	Gothic
GRC	Gresk (gammelgresk til 1453)	Ancient Greek
GUJ	Gujarati	Gujarati
GAA	Ga	Ga
HAI	Haida	Haida
HAW	Hawaiiisk	Hawaiian
HER	Herero	Herero
HIL	Hiligaynon	Hiligaynon
HIN	Hindi	Hindi
HMN	Hmong	Hmong
HUN	Ungarsk	Hungarian
HUP	Hupa	Hupa
IBO	Ibo	Igbo
IJO	Ijo	Ijo
ILE	Interlingue	Interlingue
ILO	Iloko	Iloko

IND	Indonesisk	Indonesian
IRA	Iranian	Iranian
ITA	Italiensk	Italian
JPN	Japansk	Japanese
JPR	Judeo-persian	Judeo-persian
KAB	Kabyle	Kabyle
KAC	Kachin	Kachin
KAM	Kamba	Kamba
KAN	Kanarese	Kannada
KAS	Kashmirsk	Kashmiri
KAW	Kawi	Kawi
KHA	Khasi	Khasi
KHI	Khosian	Khosian
KHO	Khotanese	Khotanese
KIN	Kinyarwanda / mulenge	Rwanda
KMB	Kimbundu	Kimbundu
KOK	Konkani	Konkani
KON	Kono	Kongo
KOS	Kosraean	Kosraean
KPE	Kpelle	Kpelle
KUA	Kuanyama	Kuanyama
KUR	Kurdisk (Kurmanji)	Kurdish (Kurmanji)
KA	Kara-kalpak	Kara-kalpak
LAH	Lahnda	Lahnda
LAO	Laotisk	Lao
LAV	Latvisk	Latvian
LEZ	Lezghian	Lezghian
LIT	Litauisk	Lithuanian
LOZ	Lozi	Lozi
LUG	Luganda	Ganda
LUI	Luiseno	Luiseno
LUO	Luo	Luo
MAC	Makedonsk	Macedonian
MAG	Magahi	Magahi
MAI	Maithili	Maithili
MAK	Makasar	Makasar
MAN	Mandingo	Mandingo
MAR	Marathi	Marathi
MAS	Masai	Masai
MDR	Mandar	Mandar
MIC	Micmac	Micmac
MIN	Minangkabau	Minangkabau
MLG	Malagasy	Malagasy
MNI	Manipuri	Manipuri
MOL	Moldavisk	Moldavian
MON	Mongolsk	Mongolian
MWR	Marwari	Marwari
NAH	Nahuati	Nahuati
NAU	Nauru	Nauru
NAV	Navajo	Navajo
NBL	Ndebele south	Ndebele south
NDO	Ndonga	Ndonga
NEP	Nepalesisk	Nepali
NIA	Nias	Nias
NIU	Niuean	Niuean
NOR	Norsk	Norwegian
NUB	Nubisk	Nubian

NYA	Nyanja / Chicheva (chewa, sheva)	Malawi, Zambia, Zimbabwe, Mozambique
NYN	Nyankole	Nyankole
NZI	Nzima	Nzima
OJI	Ojibwa	Ojibwa
ORI	Oriya	Oriya
OSA	Osage	Osage
PAG	Pangasinan	Pangasinan
PAL	Pahlavi	Pahlavi
PAN	Punjabi	Panjabi
PAP	Papiamentu	Papiamentu
PER	Persisk	Persian
PHN	Phoenician	Phoenician
PLI	Pali	Pali
POL	Polsk	Polish
PON	Pohnpeian	Pohnpeian
POR	Portugisisk	Portugese
PRA	Prakit	Prakit
PRO	Provensialsk	Provensal
QPA	Pashto	
PUS	Pashto	Pushto
PAA	Papuan	Papuan
QAD	Adja	
QAS	Aseri-tyrkisk	
QAZ	Azzari	
QBD	Baadjunne	
QBM	Buskmann	
QBR	Barawa	
QBS	Burushaski	
QCG	Kongo	
QCH	Chuang	
QCK	Creek	
QCN	Chaldeon	
QCS	Chickasaw	
QCR	Crow	
QDM	Udmurt	
QDR	Dari	
QES	Eskimo	
QEV	Evenki	
QFL	Flamsk	
NNB	Kikandi (kikande, nande)	Congo (COD)
CIT	Chittagonian	Burma
ANU	Anuak / Jambo, Nuro	Sudan, Ethiopia
NUS	Nuer / Naath	Sudan, Ethiopia
MRH	Mara	Burma og India
ZOM	Zo, Zomi, Zorni	Burma og India
ZYP	Zyphe, Zophei	Burma
CZT	Zotung, Zoby, Banjogi	Burma
CTD	Chin, Tedim	
DAO	Chin, Daai	Sa Mi Chin/ Paletwah Daai Matu/Burma
MWQ	Chin, Mun	Mindat Daai/ Kanpalet Daai Matupi / Burma
FLM	Falam, Hallam, Chin	Burma
MNW	Mon	Burma
		Assyrian Neo-Aramaic, Iraq, Armenia, Iran, Georgia
AII	Assyrisk	
MLY	Malay/ Indonesia, Thailand	
SCC	Serbokroatisk	Serbian
SEL	Selkup	Selkup

SHN	Shan	Shan
SIN	Singalesisk	Sinhalese
SIT	Sino-tibetan	Sino-tibetan
SLV	Slovensk	Slovenian
SNA	Shona	Shona
SNK	Soninke	Soninke
SON	Songhai	Songhai
SPA	Spansk	Spanish
SRR	Serer	Serer
SSW	Swazi	Swati
SUS	Susu	Susu
SUX	Sumerian	Sumerian
SYR	Syrisk	Syriac
TAM	Tamil	Tamil
TEL	Telugu	Telugu
TET	Tetela	Tetum
TGL	Tagalog	Tagalog
TIG	Tigre	Tigre
TIV	Tiv	Tiv
TLI	Tlingit	Tlingit
TOG	Tonga (nyasa)	Tonga (nyasa)
TSI	Tsimshian	Tsimshian
TSO	Tsonga	Tsonga
TUR	Tyrkisk	Turkish
TVL	Tvalu	Tvalu
TYV	Tuvinian	Tuvinian
UKR	Ukrainsk	Ukrainian
URD	Urdu	Urdu
VAI	Vai	Vai
VIE	Vietnamesisk	Vietnameses
WAK	Wakashan	Wakashan
WAR	Waray	Waray
WEL	Walisisk	Welsh
YAO	Yao	Yao
YID	Jiddisk	Yiddish
YPK	Yupik	Yupik
ZEN	Seneca	Zenaga
ZND	Zande	Zande
ZUN	Zuni	Zuni
QBA	Bari	Sudan
QGO	Gola	fra Sierra Leone og Liberia
AJG	Adja (Aja-Gbe)	
ADY	Adygei / Adyghe	
BSS	Akose (Bakossi)	bnt-bss
ANY	Anyin (Agni)	
HAZ	Azargi (Hazargi)	
BRI	Bakweri (Mokpwe)	
DOT	Barawa (Dass)	
BFA	Bari	
BYN	Bilin, Blin (Bilen)	
BOS	Bosnisk	Bosnian
CIC	Chickasaw	
CNH	Chin, Haka	
YUH	Kinesisk (Kantonesisk)	Chinese (Cantonese)
COP	Koptisk	Coptic
CRO	Crow	
DAL	Dahalo	

PRS	Dari	(Farsi, Eastern)
DAI	Day	
MYV	Erzya (Mordvin)	
EVN	Evenki	
GLK	Gilaki	
DAF	Gio (Dan)	
GRE	Gresk	Greek
HAV	Havu (Kihavu)	
HKE	Hunde (Kihunde)	
INH	Ingusjisk	Ingush
ISO	Isoko	nic-iso
KSS	Kisi, Southern	
GJK	Koli, Kachi	
KNO	Kono	
KRW	Krahn	
KRI	Krio	
KRO	Kru	
BDF	Kurdisk (Behdini)	Kurdish (Behdini)
KDB	Kurdisk (Sorani)	Kurdish (Sorani)
LAJ	Lango	
LIA	Limba	nic-lia
LOK	Loko	
LOM	Loma (Lorma)	
LRI	Luri / Lori	
LWO	Luwo	
MHI	Ma'di	
MLQ	Malinke (Maninkakan)	
MBX	Mari	
MHO	Mashi (masi)	Zambia, Congo
MIS	Miscellaneous languages (Døv)	
MOH	Mohawk	
MGD	Moru	
MKW	Munukutuba	
NAQ	Nama (Hottentot)	
SBA	Ngambay	
NGJ	Ngie	
QUT	Quiche', West Centra	
SBZ	Sara Kaba	
SRC	Serbokroatisk	Serbo-Croatian
WEN	Sorbisk	Sorbian / Tyskland
KDH	Tem (Kotokoli)	nic-kdh
LUA	Luba-Kasai (luba-luba, tshiluba, western luba, luva)	Congo (COD)
TUG	Tunia (Tounia)	
UDM	Udmurtisk	Udmurt
UND	Udefinert	Undetermined
VLA	Flamsk	Vlaams (Flamand)
DJE	Zarma	
III	Sichuan Yi	
LTZ	Letzeburgesch; Luxembourgish	
BSK	Burushaski	
DRO	Matu	Daro-Matu
MLS	Masalit	Sudan
YYY	Lautu / Chin Miso	Burma
KSW	Karen S'gaw	
KJP	Karen Pwo, østlig	Burma, Thailand
PWO	Karen Pwo, vestlig	Karen Pwo (Western) / Burma

PDU	Kayan / Padaung	Burma
TAP	Taabwa / Rungu	Kongo, Zambia
KNG	Kongo/Kikongo	Kongo/Angola
ALN	Albansk	Albanian
ARB	Arabisk	Arabic
AZJ	Aserbajdsjansk	Azerbajjani
BCC	Baluchi	
APA	Apache languages	Apache languages
ARM	Armensk	Armenian
ART	Artificial	Artificial
ASM	Assamesisk	Assamese
AVA	Avarisk	Avar / Avaric
AWA	Awadhi	Awadhi
AZE	Aserbajdsjansk	Azerbajjani
BAK	Bashkir	Bashkir
BAN	Balinesisk	Balinese
BAS	Basa	Basa
BEJ	Beja	Beja
BEM	Bemba	Bemba
BER	Berbersk	Berber
BIK	Bikol	Bikol
BIS	Bislama	Bislama
BNT	Bantu	Bantu
BRE	Bretonsk	Breton
BUG	Buginese	Buginese
BUR	Burmesisk	Burmese
CAR	Carib	Carib
CAU	Caucasian	Caucasian
CEL	Celtic	Celtic
CHB	Chibcha	Chibcha
CHI	Kinesisk	Chinese
CHM	Mari	Mari
CHP	Chipewyan	Chipewyan
CHU	Church slavic	Church slavic
CMC	Chamic	Chamic
COR	Cornish	Cornish
CPE	Creoles and pidgins (english-based)	Creoles and pidgins (english-based)
CUS	Cushitic	Cushitic
DAK	Dakota	Dakota
GOR	Gorontalo	Gorontalo
GRB	Grebo	Grebo
GRN	Guarani	Gurani
GWI	Gwich'in	Gwich'in
HAU	Hausa	Hausa
HEB	Hebraisk	Hebrew
HIM	Himachali	Himachali
HIT	Hittite	Hittite
HMO	Hiri motu	Hiri motu
IBA	Iban	Iban
ICE	Islandsk	Icelandic
IKU	Inuktitut	Inuktitut
INC	Indic	Indic
IPK	Inupiaq	Inupiaq
IRO	Iroquoian	Iroquoian
JAV	Javanesisk	Javanese
JRB	Judeo-arabic	Judeo-arabic
KAL	Kalmyk	Kalaallisut

KAR	Karen	
KAU	Kanuri	Kanuri
KAZ	Kasakhisk	Kazakh
KHM	Khmer	Khmer
KIK	Kikuyu	Kikuyu
KIR	Kirgisisk	Kirgiz / Kirghiz
KOM	Komi	Komi
KOR	Koreansk	Korean
KRU	Kurukh	Kurukh
KUM	Kumyk	Kumyk
KUT	Kutenai	Kutenai
LAD	Ladino	Ladino
LAM	Lamba	Lamba
LAT	Latin	Latin
LIN	Lingala	Lingala
LOL	Mongo	Mongo
LUB	Luba	Luba-katanga
LUN	Lunda	Lunda
LUS	Lus, Dullien, Lushai	Burma og India
MAD	Maduresisk	Madurese
MAH	Marshall	Marshall
MAL	Malayalam	Malayalam
MAO	Maori	Maori
MAY	Maya (quiche)	Malayalam
MEN	Mende	Mende
MKH	Mon-khmer	Mon-khmer
MLT	Maltisk	Maltese
MNO	Manobo	Manobo
MOS	Mossi	Mossi
MYN	Mayan	Mayan
NAI	North American Indian	North American Indian
NDE	Ndebele north	Ndebele north
NEW	Newari	Newari
NIC	Niger-kordofanian	Niger-kordofanian
NSO	Sotho, northern	Sotho, northern
NYM	Nyamwezi	Nyamwezi
NYO	Nyoro	Nyoro
OCI	Occitan	Occitan
ORM	Oromo	Oromo
OSS	Ossetisk	Ossetic / Ossetian
PAM	Pampanga	Pampanga
QPP	Papago	
PAU	Palauan	Palauan
PHI	Philipine	Philipine
QKR	Kirundi	
QKV	Kihavu	fra Zaire
QMA	Mandailing	Indonesisk
QMG	Malgche	
QMK	Marshallesisk	
QMO	Moru	
QMW	Mohawk	
QNK	Nakhi	
QOS	Ostiak	
QPD	Pidgin engelsk	
QSR	Sorbisk	
QSX	Sioux	
QTJ	Tsjekkisk	

QTT	Tsjetsjensk	
RAJ	Rajisthani	Rajisthani
ROA	Romance	Romance
ROM	Romani	Romany
RUS	Russisk	Russian
SAG	Sango	Sango
SAI	South American Indian	South American Indian
SAS	Sasak	Sasak
MHV	Arkanese / Magh, Maghi, Marma (arkanesisk) / Burma	
TVN	Tavoy	SIL-kode
BOD	Tibetansk	Tibetan
BXM	Buriat	Buriat
BMB	Bembe (Kibembe)	Den dem. rep. kongo, Tanzania
CES	Tsjekkisk	Czech
CKB	Kurdisk (Sorani)	Kurdish (Sorani)
CMN	Mandarin kinesisk	Chinese mandarin
CRK	Crow	
DEU	Tysk	German
DKS	Dinka	
ELL	Gresk	Greek
EUS	Baskisk	Basque
GAX	Oromo	
GBO	Grebo	
GYA	Gbaya	
HRV	Kroatisk	Croatian
HSB	Sorbisk	Sorbian/ Tyskland
HYE	Armensk	Armenian
ILK	Ilongot / Bukalut, Lingotes	Filippinene
ISL	Islandsk	Icelandic
KAT	Georgisk	Georgian
KHK	Mongolsk	Mongolian
KMR	Kurdisk (Badini)	Kurdish (Behdini)
KPV	Komi	
KSV	Kikusu / Kusu	Kongo DR
LRC	Luri / Lori	Luri
MHR	Mari	
MKD	Makedonsk	Macedonian
MNK	Mandingo	
MYA	Burmesisk	Burmese
NGU	Nahuatl	
NLD	Hollandsk	Dutch
PBT	Pashto	Pushto
QUH	Quecha	Quechua
RMY	Romani, sigøynersk	Romany
RON	Rumensk	Romanian
SEE	Irokesisk	Iroquian Seneca
SHK	Shiluk / Colo / Dhocolo / Chulla	Sudan
SLK	Slovakisk	Slovak
SME	Samisk	
SRP	Serbisk	Serbian
SWH	Swahili	
SYC	Syrisk	Syriac
UNM	Delaware / Unami	
UZN	Uzbekisk	Uzbek
VLS	Flamsk	Vlaams / Flamlend
XPE	Kpelle	
XXX	Flere språk	

YDD	Jiddisk	Yiddish
ZNE	Zande	
LUE	Luvale / Luena	DR Congo, Angola
YOM	Yombe / Kiyombe	DR Congo, Congo, Angola
BNX	Bangubangu	DR Congo
KCW	Bwari / Kabwari	Dem. rep. Congo
bft	Balti/Baltistani	Pakistan
SEZ	Chin, Senthang	Haka, Chin Hills
CNK	Chin, Khumi	
MID	Mandeisk / Sabeisk	Mandean
LUP	Lumbu	Ilumbu, fra Gabon, men snakkes også i Kongo
RUD	Kirundi (rundi)	Språk fra Burundi
KYU	Karenni	Språk fra Myanmar
BLK	Pao	Myanmar
PCM	Pidgin	Nigeria
ISH	Esan	Nigeria
PES	Farsi	Iran/Persisk
PGA	Juba arabisk / Pidgen Arabic	(Sudanese Creole) Sudan
AVN	Avatime / Sideme	G(h)ana
FLR	Fuliiru / Fuliro, Kifuliiru	DR Congo
SSY	Saho	Saho
ACH	Acoli	Acholi
AFA	Afro-asiatic	Afro-asiatic
AKA	Akan	Akan
NYJ	Nyanja / Chicheva (chewa, sheva)	Malawi, Zambia, Zimbabwe, Mozambique
MDW	Mbosi (mboshi, mboshe, mbochi)	Congo
RAW	Rawang	Hkanung, Krangku, Taron, Ch'opa, Chiutse, Kiutze /Burma, India
RHG	Rohinga	Burma
ALB	Albansk	Albanian
AMH	Amharisk	Amharic
KHY	Kele / lokele, ekele, kili, likelo, yakusu	DR Kongo
FUC	Pulaar	Senegal (Vest-Afrika)
CSY	Chin, Siyin	
SDH	Kurdisk (dial Feyli)	Kurdish (dial Feyli) Baada/ Baaden/ språk i Etiopia/ Eritrea ISO 639-3
KUN	Kunama	
QFN	Fulani	
QFR	Færøysk	
QFX	Fox	
QGN	Ganda	
QHL	Hollandsk	
QHT	Hottentot	
QIS	Visayan	
QIY	Yi	
QJD	Jiddish	
QKB	Kibemba	fra Zaire
QKE	Kreolsk	
QKH	Kihunde	Fra Congo
QKI	Krio	
QKK	Kikongo	
QKL	Kaldeisk	
QKN	Kantonesisk	
QKO	Kobol	
QKP	Koptisk	
QKT	Katalansk	
QLR	Lori	

QLX	Luxemburgsk	
QMB	Mbunda	
QMD	Madi	
QMI	Miao	
QML	Malinke	
QMR	Mordvin	
QMS	Mashi	fra Congo
QMY	Maylaylam	
QNN	Nenets	
QOV	Ovimbundu	fra Angola
QRN	Ruanda	
QSB	Sibo	
QSV	Svartfot	
QTC	Teochew	
QTS	Tshiluba	fra Zaire
QUE	Quechua	Quechua
RAP	Rapanui	Rapanui
RAR	Rarotongan	Rarotongan
ROH	Rhateo-romansk	Raeto-romance
RUM	Rumensk	Romainian
RUN	Rundi	Rundi
SAD	Sandawe	Sandawe
SAH	Yakut	Yakut
SAL	Salishan	Salishan
SAM	Samaritan aramic	Samaritan aramic
SAN	Sanskrit	Sanskrit
SAT	Santali	Santali
SCO	Scots	Scots
SCR	Kroatisk	Croatian
SEM	Semitic	Semitic
SID	Sidamo	Sidamo
SIO	Siouan	Siouan
SLA	Slovakisk	Slavic
SLO	Slovakisk	Slovak
SMI	Samisk	Sami
SMO	Samoisk	Samoan
SND	Sindhi	Sindhi
SOG	Sogdian	Sogdian
SOM	Somalisk	Somali
SOT	Sotho	Sotho
SRD	Sardinian	Sardinian
SSA	Nilo-saharan	Nilo-saharan
SUK	Sukuma	Sukuma
SUN	Sundanesisk	Sudanese
SWA	Swahili	Swahili
SWE	Svensk	Swedish
TAH	Tahitisk	Tahitian
TAI	Tai	Tai
TAT	Tartar	Tatar
TEM	Timne	Timne
TER	Tereno	Tereno
TGK	Tadsjikisk	Tajik / Tadzjik
THA	Thai	Tthai
TIB	Tibetansk	Tibetan
TIR	Tigrinia	Tigrinya
TKL	Tokelau	Tokelau
TMH	Tamashek	Tamashek

TON	Tongangisk	Tonga
TPI	Tok pisin	Tok pisin
TSN	Tswana	Tswana
TUK	Turkmensk	Turkmen
TUM	Tumbuka	Tumbuka
TUT	Altaic	Altaic
TWI	Twi	Twi
UGA	Ugrritiumbe	Ugrritiumbe
UIG	Uighur	Uighur
UMB	Umbundu	Umbundu
UZB	Uzbekisk	Uzbek
VEN	Venda	Venda
VOL	Volapuk	Volapuk
VOT	Votic	Votic
WAL	Walamo	Walamo
WAS	Washo	Washo
WOL	Wollof	Wolof
XHO	Xhosa	Xhosa
YAP	Yapese	Yapese
YOR	Yoruba	Yoruba
ZAP	Zapotec	Zapotec
ZHA	Zhuang	Zhuang
ZUL	Zulu	Zulu
AAR	Afar	Afar
QKS	Kissi	fra Liberia
QGI	Gio	fra Liberia
QLO	Loko	fra Sierra Leone
QTO	Tounia	fra Tchad
QSA	Sara	fra Tchad
QDA	Day	fra Tchad
QNG	Ngambaye	fra Tchad
QKA	Karen	fra Myanmar
GEO	Georgisk	Georgian
GER	Tysk	German
GEZ	Geez	Geez
999	Ugyldig språkkode	

Vedlegg 5. Folkemengde etter fødeland. 2011.

Land	Antall personer	Antall land
Norge	4037249	1
Danmark	53630	2
Grønland	422	3
Finland	13523	4
Færøyene	2027	5
Island	8982	6
Sverige	78830	7
Albania	681	8
Belgia	2374	9
Bulgaria	3329	10
Andorra	9	11
Estland	3315	12
Frankrike	8315	13
Gibraltar	19	14
Hellas	1844	15
Hviterussland	951	16
Irland	1564	17
Kroatia	4027	18
Italia	4344	19
Latvia	5415	20
Malta	143	21
Nederland	13032	22
Liechtenstein	33	23
Luxembourg	155	24
Monaco	20	25
Polen	65294	26
Portugal	1948	27
Romania	6726	28
San Marino	7	29
Litauen	16882	30
Spania	5422	31
Moldova	436	32
Storbritannia	36312	33
Russland	18771	34
Sveits	3562	35
Tyrkia	19206	36
Tyskland	40847	37
Slovenia	339	38
Ukraina	3506	39
Ungarn	4305	40
Østerrike	2774	41
Bosnia-Hercegovina	17095	42
Makedonia	3749	43
Slovakia	2707	44
Tsjekkia	2441	45
Serbia	3639	46
Montenegro	415	47
Kosovo	14197	48
Algerie	2146	49
Angola	585	50
Botswana	159	51
St. Helena	20	52
Det britiske området i Indiahavet	3	53
Burundi	1234	54

Komorene	7	55
Benin	21	56
Ekvatorial-Guinea	2	57
Elfenbenskysten	428	58
Eritrea	8087	59
Etiopia	7126	60
Egypt	1366	61
Djibouti	76	62
Gabon	12	63
Gambia	2152	64
Ghana	2577	65
Guinea	202	66
Guinea-Bissau	48	67
Kamerun	821	68
Kapp Verde	783	69
Kenya	2242	70
Kongo-Brazzaville	207	71
Kongo	2417	72
Lesotho	15	73
Liberia	1246	74
Libya	398	75
Madagaskar	918	76
Malawi	77	77
Mali	71	78
Marokko	10332	79
Vest-Sahara	26	80
Mauritania	28	81
Mauritius	378	82
Namibia	131	83
Niger	68	84
Nigeria	1943	85
Mosambik	225	86
Réunion	4	87
Zimbabwe	365	88
Rwanda	738	89
São Tomé og Príncipe	2	90
Senegal	303	91
Den sentralafrikanske republikk	9	92
Seychellene	32	93
Sierra Leone	740	94
Somalia	28140	95
Sudan	1710	96
Swaziland	30	97
Sør-Afrika	2154	98
Tanzania	1246	99
Tsjad	30	100
Togo	158	101
Tunisia	1797	102
Uganda	1284	103
Zambia	570	104
Burkina Faso	38	105
Afghanistan	12122	106
Armenia	272	107
Aserbajdsjan	518	108
Bahrain	44	109
Bangladesh	1111	110
Bhutan	265	111

Brunei	38	112
Myanmar	3382	113
Sri Lanka	15070	114
De forente arabiske emirater	320	115
Filippinene	22423	116
Georgia	280	117
Taiwan	400	118
Hongkong	1286	119
India	12949	120
Indonesia	1945	121
Irak	28672	122
Iran	19032	123
Israel	1399	124
Japan	1894	125
Jordan	523	126
Kambodsja	567	127
Kasakhstan	863	128
Kina	12074	129
Nord-Korea	53	130
Sør-Korea	8101	131
Kuwait	423	132
Kypros	289	133
Kirgisistan	190	134
Laos	128	135
Libanon	3074	136
Macao	101	137
Malaysia	881	138
Maldivene	10	139
Mongolia	132	140
Oman	40	141
Det palestinske området	3525	142
Nepal	1065	143
Pakistan	35722	144
Øst-Timor	18	145
Qatar	40	146
Saudi-Arabia	340	147
Singapore	874	148
Tadsjikistan	129	149
Turkmenistan	81	150
Usbekistan	392	151
Sør-Jemen	0	152
Syria	2351	153
Thailand	19595	154
Vietnam	22226	155
Yemen	266	156
De amerikanske Jomfruøyene	16	157
Barbados	40	158
Antigua og Barbuda	8	159
Belize	23	160
Bahamas	17	161
Bermuda	18	162
De britiske Jomfruøyene	5	163
Canada	5501	164
Caymanøyene	5	165
Costa Rica	342	166
Cuba	1190	167
Dominica	42	168

Den dominikanske republikk	930	169
Grenada	44	170
Guadeloupe	8	171
Guatemala	558	172
Haiti	54	173
Honduras	200	174
Jamaica	282	175
Martinique	13	176
Mexico	1577	177
Montserrat	2	178
De nederlandske antillene	103	179
Aruba	42	180
Nicaragua	259	181
Panama	84	182
Panamakanalsonen	0	183
El Salvador	260	184
Saint Kitts og Nevis	3	185
Saint Lucia	22	186
Saint Vincent og Grenadinene	7	187
Trinidad og Tobago	735	188
Turks- og Caicosøyene	1	189
USA	33922	190
Puerto Rico	45	191
Argentina	1572	192
Bolivia	651	193
Brasil	5546	194
Guyana	201	195
Chile	11208	196
Colombia	5775	197
Ecuador	917	198
Falklandsøyene	3	199
Fransk Guyana	8	200
Paraguay	208	201
Peru	1984	202
Surinam	78	203
Uruguay	377	204
Venezuela	928	205
Amerikansk Samoa	6	206
Australia	3066	207
Salomonøyene	8	208
Christmasøyene	1	209
Cookøyene	9	210
Fiji	75	211
Vanuatu	4	212
Tonga	14	213
Fransk Polynesia	10	214
Kiribati	2	215
Tuvalu	2	216
New Zealand	1017	217
Mikronesiaføderasjonen	8	218
Papua Ny-Guinea	10	219
Samoa	6	220
Ny-Caledonia	14	221
Palau	4	222
Totalt	4920305	222

Vedlegg 6. Opptelling av antall språk fordelt på talere. Kilder: KOST, DUF og internett registreringer (UDI). 2012

Språk	Antall talere	Prosent	Antall språk
Acehenese	90	0,076	1
Acoli	24	0,02	2
Adja (Aja-Gbe)	9	0,008	3
Afar	5	0,004	4
Afrikaans	5	0,004	5
Afro-asiatic	1	0,001	6
Akan	3	0,003	7
Albansk	4535	3,809	8
Aleutisk	1	0,001	9
Algonquian languages	1	0,001	10
Amharisk	1433	1,203	11
Anuak / Jambo, Nuro	4	0,003	12
Arabisk	11489	9,649	13
Arkanese / Magh, Maghi, Marma (arkanesisk) / Burma	8	0,007	14
Armensk	118	0,099	15
Artificial	1	0,001	16
Aserbajdsjansk	193	0,162	17
Aseri-tyrkisk	3	0,003	18
Assamesisk	3	0,003	19
Assyrisk	16	0,013	20
Australian	1	0,001	21
Avarisk	1	0,001	22
Azargi (Hazargi)	19	0,016	23
Bakweri (Mokpwe)	1	0,001	24
Balinesisk	21	0,018	25
Balti/Baltistani	1	0,001	26
Baluchi	2	0,002	27
Balutsjisk	15	0,013	28
Bambara	3	0,003	29
Bamileke	1	0,001	30
Banda	22	0,018	31
Bangubangu	2	0,002	32
Bantu	9	0,008	33
Barawa (Dass)	1	0,001	34
Basa	16	0,013	35
Bashkir	1	0,001	36
Batak	2	0,002	37
Bemba	40	0,034	38
Bengali	80	0,067	39
Berbersk	20	0,017	40
Bihari	15	0,013	41
Bilin, Blin (Bilen)	139	0,117	42

Bini	5	0,004	43
Bislama	4	0,003	44
Bosnisk	227	0,191	45
Braj	6	0,005	46
Bulgarsk	134	0,113	47
Burmesisk	848	0,712	48
Cebuano	123	0,103	49
Chagatai	6	0,005	50
Chamorro	3	0,003	51
Chin, Daai	26	0,022	52
Chin, Haka	420	0,353	53
Chin, Mun	8	0,007	54
Chin, Senthang	11	0,009	55
Chin, Siyin	2	0,002	56
Chin, Tedim	27	0,023	57
Chinook jargon	27	0,023	58
Chittagonian	3	0,003	59
Chukchi	1	0,001	60
Cree	1	0,001	61
Creoles and pidgins	5	0,004	62
Creoles and pidgins (english-based)	7	0,006	63
Creoles and pidgins (portuguse- based)	2	0,002	64
Dansk	83	0,07	65
Dari	5212	4,377	66
Dinka	6	0,005	67
Dyula	1	0,001	68
Dzongkha	3	0,003	69
Egyptian	3	0,003	70
Engelsk	11130	9,347	71
Esan	1	0,001	72
Estisk	53	0,045	73
Ewe	7	0,006	74
Ewondo	2	0,002	75
Falam, Hallam, Chin	37	0,031	76
Fang	1	0,001	77
Fanti	1	0,001	78
Farsi	4929	4,139	79
Fijian	1	0,001	80
Finsk	11	0,009	81
Flamsk	2	0,002	82
Fon	1	0,001	83
Fransk	1616	1,357	84
Friulian	14	0,012	85
Fulani	5	0,004	86
Fuliiru / Fuliro, Kifuliiru	4	0,003	87
Fullah	9	0,008	88

Ga	3	0,003	89
Ganda	1	0,001	90
Gayo	1	0,001	91
Gbaya	7	0,006	92
Georgisk	51	0,043	93
Gilaki	7	0,006	94
Gio (Dan)	4	0,003	95
Grebo	50	0,042	96
Gresk	6	0,005	97
Gresk (gammeligresk til 1453)	45	0,038	98
Gujarati	6	0,005	99
Haida	2	0,002	100
Hausa	15	0,013	101
Havu (Kihavu)	5	0,004	102
Hebraisk	25	0,021	103
Hiligaynon	22	0,018	104
Hindi	196	0,165	105
Hmong	10	0,008	106
Hollandsk	236	0,198	107
Hunde (Kihunde)	8	0,007	108
Hviterussisk	24	0,02	109
Ibo	31	0,026	110
Iloko	4	0,003	111
Ilongot / Bukalut, Lingotes	86	0,072	112
Indic	2	0,002	113
Indonesisk	159	0,134	114
Ingusjisk	26	0,022	115
Iranian	16	0,013	116
Irish	3	0,003	117
Italiensk	243	0,204	118
Japansk	81	0,068	119
Javanesisk	3	0,003	120
Kabyle	14	0,012	121
Kachin	17	0,014	122
Kamba	3	0,003	123
Kanarese	20	0,017	124
Kantonesisk	5	0,004	125
Karen	270	0,227	126
Karen Pwo, vestlig	5	0,004	127
Karen Pwo, østlig	25	0,021	128
Karen S`gaw	366	0,307	129
Karenni	9	0,008	130
Kasakhisk	25	0,021	131
Kashmirsk	1	0,001	132
Katalansk	5	0,004	133
Kayan / Padaung	7	0,006	134

Khasi	1	0,001	135
Khmer	97	0,081	136
Kikuyu	28	0,024	137
Kimbundu	3	0,003	138
Kinesisk	221	0,186	139
Kinesisk (Kantonesisk)	313	0,263	140
Kinyarwanda / mulenge	266	0,223	141
Kirgisisk	6	0,005	142
Kirundi	15	0,013	143
Kirundi (rundi)	65	0,055	144
Kisi, Southern	11	0,009	145
Konkani	5	0,004	146
Kono	3	0,003	147
Koreansk	59	0,05	148
Kosraean	2	0,002	149
Kpelle	27	0,023	150
Krahn	40	0,034	151
Kreolsk	2	0,002	152
Krio	104	0,087	153
Kroatisk	268	0,225	154
Kru	63	0,053	155
Kumyk	1	0,001	156
Kunama	33	0,028	157
Kurdisk (Badini)	216	0,181	158
Kurdisk (Behdini)	25	0,021	159
Kurdisk (Kurmanji)	5945	4,993	160
Kurdisk (Sorani)	914	0,768	161
Kurukh	1	0,001	162
Lango	4	0,003	163
Laotisk	9	0,008	164
Latin	6	0,005	165
Latvisk	52	0,044	166
Lautu / Chin Miso	40	0,034	167
Lezghian	1	0,001	168
Lingala	135	0,113	169
Litauisk	260	0,218	170
Loma (Lorma)	21	0,018	171
Lori	3	0,003	172
Lozi	1	0,001	173
Luba	11	0,009	174
Luba-Kasai (luba-luba, tshiluba, western luba, luva)	41	0,034	175
Luganda	32	0,027	176
Lunda	1	0,001	177
Luo	11	0,009	178
Luri / Lori	7	0,006	179
Lus, Dullien, Lushai	2	0,002	180

Luvale / Luena	3	0,003	181
Ma'di	1	0,001	182
Madi	4	0,003	183
Maithili	2	0,002	184
Makasar	7	0,006	185
Makedonsk	49	0,041	186
Malagasy	8	0,007	187
Malay/ Indonesia, Thailand	15	0,013	188
Malayalam	67	0,056	189
Malinke (Maninkakan)	9	0,008	190
Maltisk	1	0,001	191
Mandar	2	0,002	192
Mandarin kinesisk	50	0,042	193
Mandingo	123	0,103	194
Manobo	1	0,001	195
Maori	1	0,001	196
Mara	5	0,004	197
Marathi	53	0,045	198
Mari	1	0,001	199
Masai	13	0,011	200
Masalit	5	0,004	201
Mashi (masi)	5	0,004	202
Matu	7	0,006	203
Maya (quiche)	49	0,041	204
Mende	53	0,045	205
Miscellaneous languages (Døv)	1	0,001	206
Mohawk	1	0,001	207
Moldavisk	12	0,01	208
Mon	3	0,003	209
Mon-khmer	1	0,001	210
Mongolsk	32	0,027	211
Moru	4	0,003	212
Mossi	3	0,003	213
Munukutuba	3	0,003	214
Nahuati	1	0,001	215
Ndebele north	1	0,001	216
Ndebele south	3	0,003	217
Nepalesisk	360	0,302	218
Newari	1	0,001	219
Ngambay	4	0,003	220
Ngie	1	0,001	221
Norsk	4306	3,616	222
Nubisk	1	0,001	223
Nyanja / Chicheva (chewa, sheva)	131	0,11	224
Oriya	2	0,002	225
Oromo	719	0,604	226

Ossetisk	4	0,003	227
Pampanga	7	0,006	228
Pao	4	0,003	229
Pashto	1145	0,962	230
Persisk	677	0,569	231
Philippine	719	0,604	232
Pidgin	1	0,001	233
Polsk	1536	1,29	234
Portugisisk	547	0,459	235
Pulaar	4	0,003	236
Punjabi	383	0,322	237
Quechua	3	0,003	238
Quiche´, West Centra	1	0,001	239
Rawang	2	0,002	240
Rohinga	12	0,01	241
Romance	1	0,001	242
Romani	129	0,108	243
Ruanda	2	0,002	244
Rumensk	249	0,209	245
Rundi	338	0,284	246
Russisk	5062	4,251	247
Saho	16	0,013	248
Samoisk	1	0,001	249
Sara Kaba	1	0,001	250
Serbisk	274	0,23	251
Serbokroatisk	12250	10,288	252
Shan	6	0,005	253
Shona	10	0,008	254
Sindhi	7	0,006	255
Singalesisk	87	0,073	256
Slovakisk	51	0,043	257
Slovensk	3	0,003	258
Somalisk	10904	9,157	259
Sorbisk	2	0,002	260
Spansk	2764	2,321	261
Sukuma	1	0,001	262
Sundanesisisk	17	0,014	263
Susu	10	0,008	264
Svensk	465	0,391	265
Swahili	1009	0,847	266
Swazi	3	0,003	267
Syrisk	202	0,17	268
Taabwa / Rungu	11	0,009	269
Tadsjikisk	24	0,02	270
Tagalog	838	0,704	271
Tahitisk	2	0,002	272

Tai	18	0,015	273
Tamil	2339	1,964	274
Tartar	6	0,005	275
Tavoy	2	0,002	276
Telugu	79	0,066	277
Thai	1400	1,176	278
Tibetansk	9	0,008	279
Tigre	241	0,202	280
Tigrinia	5552	4,663	281
Timne	4	0,003	282
Tonga (nyasa)	3	0,003	283
Tshiluba	2	0,002	284
Tsjekkisk	66	0,055	285
Tsjetsjensk	1431	1,202	286
Tswana	1	0,001	287
Tunia (Tounia)	1	0,001	288
Turkmensk	166	0,139	289
Twi	24	0,02	290
Tyrkisk	1564	1,313	291
Tysk	1757	1,476	292
Uighur	531	0,446	293
Ukrainsk	295	0,248	294
Ungarsk	70	0,059	295
Urdu	745	0,626	296
Uzbekisk	160	0,134	297
Vai	2	0,002	298
Vietnamesisk	1462	1,228	299
Visayan	209	0,176	300
Waray	7	0,006	301
Wollof	12	0,01	302
Xhosa	2	0,002	303
Yakut	1	0,001	304
Yombe / Kiyombe	1	0,001	305
Yoruba	27	0,023	306
Zande	1	0,001	307
Zo, Zomi, Zorni	6	0,005	308
Zotung, Zoby, Banjogi	1	0,001	309
Zulu	4	0,003	310
Zyphe, Zophei	63	0,053	311
Totalt	119075	100,000	311

Tabellregister

Tabell 1.	Andel med registrert språk etter innvandringsårsak i DUF.	15
Tabell 2.	Fordeling av befolkning etter innvandringskategori A og B. 1. januar 2012..	17
Tabell 3.	Sentrale registre som registrerer språk	18
Tabell 4.	Fordelingen mellom DUF, KOST og internettregistreringer.	24

B Returadresse:
Statistisk sentralbyrå
NO-2225 Kongsvinger

Statistisk sentralbyrå

8/2013

Minoritetsspråk i Norge

Avsender:
Statistisk sentralbyrå

Postadresse:
Postboks 8131 Dep
NO-0033 Oslo

Besøksadresse:
Kongens gate 6, Oslo
Oterveien 23, Kongsvinger

E-post: ssb@ssb.no
Internett: www.ssb.no
Telefon: 62 88 50 00

ISBN 978-82-537-8605-6 (trykt)
ISBN 978-82-537-8606-3 (elektronisk)
ISSN 1891-5906

ISBN 978-82-537-8605-6



9 788253 786056



Statistisk sentralbyrå
Statistics Norway

Design: Siri Boquist