

Metodologiske og teoretiske utfordringar for generative dialektologar

Leiv Inge Aa

Tradisjonell dialektologi og generativ lingvistik

Dette nummeret av *Norsk lingvistisk tidsskrift* (NLT) tek føre seg nordisk syntaktisk talemålsforskning. I Noreg strekkjer den vitskaplege talemålsforskninga seg tilbake til Ivar Aasens (1848) første grammatiske storverk. Men sjølv om dialektologien er etter måten gamal, var syntaksen i talemåla lenge berre omtala i mindre studiar og ikkje kartlagt på ein overordna og systematisk måte med ein klar teoretisk ambisjon. Rett nok vart det på byrjinga av 1900-talet gjeve ut nokre avhandlingar om syntaksen i ulike bymål (t.d. Iversen 1918, Svendsen 1931) og utvalde bygdemål (t.d. Heggstad 1916). Leiv Heggstads arbeid om syntaksen i sebyggmålet (setesdalsdialekten) er velkjent, og det vart gjort mange deskriptive arbeid av mindre format oppgjennom 1900-talet.¹ På 1970-talet gjorde Helge Sandøy (1976) ei meir systematisk komparativ vestnordisk undersøking av verbpartikkelkonstruksjonar, og Jan Terje Faarlund gjorde fleire teoretisk-syntaktiske dialektarbeid på den same tida (sjå ScanDiaSyn-bibliografien og jf. Sandøy (1985: 100ff) si dialektsyntaktiske oppsummering). På 1990-talet vart grunnlaget for NORMS² og ScanDiaSyn³ lagt, og desse prosjekta voks for alvor fram utover 2000-talet. Først no vart det på ein meir omfattande og systematisk måte gjort dialektsyntaktisk arbeid i Noreg med ein teoretisk ambisjon lik Faarlund sin, og ulik den ein finn i dei deskriptive tilnærmingane tidlegare. Det verkar rimeleg å dele dialektsyntaksen inn i ein tidsalder før og etter NORMS- og ScanDiaSyn-prosjekta nettopp på grunn av omfanget av og systematikken i det metodologiske og teoretiske arbeidet desse prosjekta førte med seg.

I den chomskyanske tradisjonen vart talemålssyntaks for alvor teoretisk interessant med framvoksteren av Prinsipp og parameter-teorien på 1980-talet,

då ein fann ut at ei samanlikning av systema i nærskyldde språk kunne seie sikrere kva skilnader mellom språk som faktisk var relaterte, enn ei samanlikning mellom heilt ulike språk. Richard Kayne var ein av føregangsfigurane:

In the early to mid-1980s, it became apparent to me that a direct comparison of French and English raised difficult problems to a greater extent than a direct comparison of French and Italian. In essence, in searching for clusters of properties, one must make decisions about what syntactic differences can plausibly be linked to other syntactic differences (Kayne 2000: 4).

Dialektologien har tradisjonelt vore sterkt empirisk retta, og krava til dataarbeidet må seiast å ha vore høgare her enn hjå mange lingvistar innanfor den generative tradisjonen. Vi har såleis vorte introduserte for karakteriserande termar som *knappesammlarar* (dialektologane) og *lenestolslingvistar* (t.d. generativistane) (jf. Christensen 1996: 20). Men straks komparative og seinare mikrokomparative syntaksstudiar vart rekna for fruktbare arbeidsmetodar til å gje større teoretisk innsikt, tvang det generativistane opp av lenestolen i større grad enn tidlegare. Likevel skal det seiast at lenestolen stadig er relevant på den måten at den teoretiske ambisjonen i den generative dialektologien vil vere overordna den reine kartlegginga av empiriske mønster.

Utfordringa med autentiske data

Det auka fokuset på dataarbeid har òg gjeve nye metodologiske utfordringar i det teoretiske syntaksfeltet. Den største utfordringa har kanskje vore den manglande tilgangen på stort nok autentisk materiale til å seie noko *reelt* om talemålssyntaktiske skilnader. Vi kan illustrere det gjennom Øyalf Endresen (1988) si undersøking av den innbyrdes distribusjonen til negasjonar og trykkveike pronomen i Midt-Noreg og Nord-Sverige. Arbeidet hans vart gjort i ei tid før det var utvikla søkbare talemålskorpus her til lands. I staden hadde han eit tusentals hovudsetningar nedskrivne frå eit 200 timars langt lydbandopptak. Ein viktig merknad frå Endresen var at ordstillinga verb–adverb–nomen (V–A–N) som i (1a) (utan subjektsskifte) vart hevda å vere underrapportert både i målprøver og trøndske dialektomtalar i høve til V–N–A-ordstillinga i (1b) (med subjektsskifte).

- (1) a. Da kom itj e.
 b. Da kom e itj.

(1a) har ei markert ordstilling og er mest utbreidd i samband med pronomen som byrjar på vokal (som *e* 'eg' og *a* 'ho'). Endresen opplevde den same V–A–N-underrapporteringa i direkteutspørjing av informantane òg. Dei sa ikkje i spontan tale det dei sjølve meinte dei sa. Såleis konkluderte han med følgjande:

Dette bør sjølv sagt få konsekvensar for det metodiske opplegget ved slike syntaksundersøkingar. Vi kan altså ikkje lite på tradisjonelle målprøver og annan forfatta målførelitteratur. Ein tradisjonell metode med direkte utspørjing av lokale informantar er heller ingen garanti for at vi når fram til den meir umedvetne kompetansen som i praksis styrer språkbruken til folk. Skal vi kunne seie noko heilt sikkert om den faktiske fordelinga av dei to leddstillingssystema i trøndsk, må vi bygge på eit materiale som har grunnlag i spontan tale (Endresen 1988: 54).

Materiale som har grunnlag i spontan tale, er tidkrevjande å skaffe seg, og ikkje minst er det tidkrevjande å finne relevante syntaktiske fenomen i eit autentisk materiale, då fenomenen ein er ute etter å kartleggje, ikkje alltid er frekvente. Utviklinga av talemålskorpus er gjort m.a. for å møte slike utfordringar, dvs. for å få samla store søkbare datamengder på ein og same stad. Det nordiske dialektkorpuset (Johannessen mfl. 2009) vart utvikla utover 2000-talet, og dette korpuset inkluderer i dag 2,8 millionar ord frå samtalar mellom og intervju med 821 informantar frå 228 målepunkt i dei nordiske landa (Johannessen & Hagen 2014: 15). Over halvparten av informantane (440) frå nesten halvparten av målepunkta (111) kjem frå Noreg, det vil seie nokså nær 4 informantar frå kvar stad. Det seier seg sjølv at slike datamengder ikkje er mogleg for éin person å samle til eit tidsavgrensa prosjekt. Dermed har delar av Endresen sine utfordringar vorte møtte med bravur dei siste 30 åra.

Kva kan korpuset (ikkje) hjelpe oss med?

Så lenge taleopptakssituasjonen er truverdig, er det vanleg å rekne korpusdata som autentiske (Johannessen 2003). Likevel er det mange utfordringar knytte til slike data. Det eine problemet er at uansett kor stort korpuset er, så vil det ikkje dekkje alle konstruksjonar i eit språk. I Aa (2015: 81) kan ein sjå at eg ikkje fann nokon trøndske tilfelle i det nordiske dialektkorpuset av verbpartikkelkonstruksjonar med partikkelen distribuert til høgre for nominalet. To døme ville vere som i (2), der % indikerer at korpuset ikkje gjev svar på om konstruksjonane finst i trøndsk.

- (2) a. %Johan kasta hunden ut.
b. %Johan tenkte planen ut.

Jamt over vil direksjonale (2a) vere fullt mogleg i norske dialektar, medan metaforiske (2b) har klare avgrensingar med partikkelen til høgre.⁴ Men kva vil talet *null* treff i korpuset seie? Er det umogleg å distribuere partikkelen til høgre i trøndsk, eller er korpuset ikkje stort nok eller spreidd utover nok i Trøndelag til å finne eit døme? Eit liknande døme har vi når Åfarli (1985) hevdar at (3) er ein ok konstruksjon i halsadialekten.

(3) Måle blå bilen

Dialektkorpuset gjev ingen belegg på at fargeadjektiv distribuert til venstre for nominalet går an elles. Eit korpus gjev såleis ikkje eintydig svar på kva mangel på treff vil seie. Kan vi elles gå ut frå at belegga vi får frå korpuset, er fullgode døme frå ein dialekt, eller er det andre nyansar som ikkje kjem fram? Det er viktig å merke seg at ei setning som opptrer i korpuset, ikkje treng å vere perfekt for andre talarar av det same talemålet eller andre talemål. Vi veit ikkje kor bra setninga vi høyrer, er, eller kva grad av akseptabilitet ho har hjå andre talarar. I akseptabilitetsvurderingar kan vi be informanten gradere døma på ein Likert-skala frå 1 til 5 eller 7, medan korpustreff må vi berre ta for god fisk og håpe på det beste. Men akkurat slik vi ikkje veit sikkert kva fråvær i eit korpus tyder, veit vi altså ikkje sikkert kva eit nærvær heller vil seie, noko som kan synast som eit stort paradoks.

Slik eit talemålskorpus ikkje dekkjer dei graderte nyansane til språkbruket, dekkjer det heller ikkje semantiske nyansar og kontekstuelle skilnader. I Aa (2015: 204ff) diskuterer eg det minimale paret i (4):

- (4) a. Køyre bilen inn
b. Køyre inn bilen

Eg hevdar at dei strukturelle tilhøva talar for ei direksjonal tolking av (4a) (til dømes 'køyre bilen inn i garasjen') og ei *førebu/break in*-tolking av (4b) (til dømes 'køyre inn bilen 100 km før ein rusar motoren'). Men eit viktig poeng er at kontekstuelle tilhøve vil bidra til kva ein vel av den eine eller andre tolkinga, og dei kan til og med overstyre dei strukturelle føringane. Det skal særskilt mykje til at eit korpus vil hjelpe ein med å attestere alle dei moglege tolkingmoglegheitene ein har med dei to strukturane.

I lingvistikken prøver vi å få innsikt i den menneskelege språkkompetansen. Schütze (2011) parafraiserer Chomsky (1965) når han stiller spørsmålet: *Kva inneber kunnskap om språk?* Kva består språkkompetansen vår av, og kva vil det seie at kvar og ein har ein spesifikk kunnskap om morsmålet sitt? Desse spørsmåla fører med seg nokre spørsmål til: Korleis tileignar ein seg språkkompetanse, og kva er tilhøvet mellom språkkompetanse og -bruk (eller perfor-

manse)? Dei vidare spørsmåla er såleis kva metodar ein skal bruke for å svare på dette. Særleg spørsmålet om tilhøvet mellom språkkompetanse og -bruk er interessant og avgjerande når ein jobbar med korpusdata (jf. Johannessen 2003: 147ff). Korpusdata er performanse, og korleis kan dei, eller kan dei i det heile, nyttast til å seie oss noko om kompetansen, eller det vi kallar *I-språk*? Dersom ein er interessert i *I-språk*, så er det individuelle grammatikkar – altså kompetansen til den enkelte språkbrukaren – ein er ute etter å kartleggje. Og dette kan ein seie er inkompatibelt med korpuset som ei oppsamling av fleire språkbrukarars *E-språk* (som ikkje alltid kan sporast tilbake til den enkelte språkbrukaren). Dessutan, om det er dialektar ein undersøker, kan ein seie at *I-språk* er inkompatibel med heile dialektomgrepet, som er eit samleomgrep og ei forenkla generalisering av fleire språkbrukarars talemål (jf. Sandøy 1985: 16). Idiolektar er noko anna, det er ikkje det ein studerer i dialektologi, men det er nettopp idealiserte idiolektar ein kan seiast å studere i generativ grammatikk (Barbiers 2013). Så om ein legg omgrepa ‘individuell’ og ‘internalisert’ i *I-språk* (jf. Chomsky 1986), og det er *I-språk* ein vil undersøkje, inneber både korpusbruk og dialektologi i seg sjølv to generaliseringar av *E-språk* data som potensielt kan verke som alvorlege problem. Fanselow mfl. (2008: 4) beskriv problemet slik:

[A] narrow interpretation of the scope of generative grammar again implies that corpus frequency data are not the kind of object a grammar can explain, since corpus frequencies do not reflect a mental state of a speaker whose internal grammar one wants to describe.

Johannessen (2003) meiner at desse problema med korpusdata må kompenserast for ved at ein kombinerer korpusbruk med andre metodar, som akseptabilitetsvurderingar og introspeksjon.

Bidraga i dette nummeret

I dette nummeret av NLT vil du sjå at fleire av bidragsytarane nettopp kombinerer metodar for å kome dei ovannemnde utfordringane i møte. Bidraget til Elisabet Engdahl og Anu Laanemets er den mest reindyrka korpusstudien i heftet, rett nok av *skriftspråk*. Her blir upersonlege passivkonstruksjonar i skandinaviske undersøkte, og korpustreffa avslører mønster som ikkje var venta på førehand. Mellom anna blir Holmberg (2002) sin parameter om kongruens og ordfølgje dels avvist då nynorsk og svensk, som begge har partisippkongruens, syner ulike ordfølgjer. I staden samsvarer den nynorske syntaksen med bokmål.

Johan Brandtler og David Håkansson undersøker eit interessant og avstikkande fenomen i finlandssvenske dialektar, nemleg negerte deklaratve setningar med V3-ordstilling. Teoretisk konkluderer dei med at negasjonen må reanalyserast frå ein syntaktisk frase til ein syntaktisk kjerne. Med omsyn til datainnsamling og metode støyter forfattarane på nokre problem i arbeidet sitt som vi har skissert ovanfor som typiske når ein skal kartleggje dialektsyntaks. Fordi negative V3-deklarativar er så lågfrekvente, opptrer dei nærmast ikkje i tilgjengelege korpus, og det trengst rett og slett eit ovstort kjeldegrunnlag for å kartleggje fenomenet geografisk og å forstå det betre teoretisk.

I Marit Westergaard sin artikkel spelar òg V3-konstruksjonar ei viktig rolle, men her handlar det om ordstillinga i kv-spørjesetningar i nordnorske dialektar. Utgangspunktet til Westergaard er at nokre spørsmål må ha V2-ordstilling, medan andre tillèt ei veksling mellom V2 og V3. Vekslinga avheng mellom anna av kompleksiteten til spørjeordet og av informasjonsstruktur. I denne studien blir det ikkje nytta korpus; i staden kombinerer forfattaren og medarbeidarane i feltarbeida halv-spontan tale og akseptabilitetsvurderingar. Metodane blir motiverte for å få samla inn dei naudsynte mengdene med relevante data der spontanproduksjonen er avgrensa. Det viser seg at dei to metodane dels gjev ulike svar, og såleis at å kombinere metodar er avgjerande for å få eit best mogleg innsyn i det grammatiske systemet til talarane.

I den siste artikkelen i heftet undersøker Marianne Anderson og Tor A. Åfarli den såkalla lokativalternasjonen på Nordvestlandet, dvs. i Romsdal og på Nordmøre. Forfattarane nyttar mellom anna dativbruken frå dei relevante dialektane til å byggje opp under ein ny-konstruksjonistisk eller eksoskeletal modell (jf. Borer 2005). Dei to leddstillingsvariantane i lokativalternasjonen blir då analyserte som strukturelt heilt ulike, dvs. at dei realiserer ulike tilgjengelege syntaktiske rammer (Åfarli 2007) snarare enn å bli forklarte ut frå den leksikalske argumentstrukturen til dei spesifikke verba som blir nytta. I hovudsak er feltarbeidet gjort i Vestnes kommune, der eldre og yngre informantar (med ulik dativbruk) har vurdert akseptabiliteten til ei rad med setningar.

Sjølv om «generativ dialektologi» kan synast som ein paradoksal term både historisk, metodologisk og med omsyn til kva ein ønskjer å oppnå med arbeidet sitt, er ikkje termen verre å hanskast med enn at slikt arbeid først og fremst gjev interessante utfordringar og fruktbare problemstillingar. Og som bidrağa i dette nummeret syner, kan både metodane og føremålet med arbeidet variere.

Notar

1. Ein omfattande bibliografi over tidlegare målføresyntaktiske arbeid i Noreg finst på heimesidene til ScanDiaSyn: <http://websim.arkivert.uit.no/getfile.php%3fSiteId=150%26PageId=6795%26FileId=48> (lasta ned i juni 2015).
2. <http://norms.uit.no> (lasta ned i juni 2015).
3. <http://websim.arkivert.uit.no/scandiasyn/scandiasyn/index.html> (lasta ned i juni 2015).
4. Sjå til dømes Sandøy (1976), som hevdar at liknande døme av typen (2b) er umoglege i romsdalsdialekten.

Kjelder

- Barbiers, Sjef 2013: *Microsyntactic Variation*. Dikken, Marcel den (red.): *The Cambridge Handbook of Generative Syntax*. Cambridge: Cambridge University Press, 899–926.
- Borer, Hagit 2005: *Structuring Sense Volume I. In Name Only*. New York: Oxford University Press.
- Chomsky, Noam 1965: *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam 1986: *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Christensen, Kirsti Koch 1996: Mikrokomparativisme og målføresyntaks. Vangsnes, Øystein Alexander (red.): *Nordica Bergensia* 9, 15–23. Universitetet i Bergen.
- Endresen, Øyalf 1988: Innbyrdes posisjon mellom trykklett pronomen og setningsadverbialet *ikkje* i trøndske og nordsvenske dialektar. Bjørkum, Andreas & Arve Borg (red.): *Nordiske studiar. Innlegg på den tredje dialektologkonferansen 1986*. Oslo: Universitetsforlaget, 47–70.
- Fanselow, Gisbert, Caroline Féry, Ralf Vogel & Matthias Schesewsky 2008: Gradience in Grammar. Fanselow, Gisbert, Caroline Féry, Ralf Vogel & Matthias Schesewsky (red.): *Gradience in Grammar. Generative Perspectives*. Oxford: Oxford University Press, 1–21.
- Heggstad, Leiv 1916: Ymist or syntaksen i Sæbyggmålet. *Maal og Minne* 1916, 159–166.
- Holmberg, Anders 2002: Expletives and Agreement in Scandinavian Passives. *Journal of Comparative Germanic Linguistics* 4, 85–128.
- Iversen, Ragnvald 1918: *Syntaksen i Tromsø bymaal: en kort oversigt*. Kristiania: Bymaals-laget.
- Johannessen, Janne Bondi 2003: Innsamling av språklige data. Informanter, introspeksjon og korpus. Johannessen, Janne Bondi, Sonja Erlenkamp,

- Elizabeth Lanza & Arnfinn Muruvik Vonen (red.): *På språkjakt – problemer og utfordringer i språkvitenskapelig datainnsamling*. Oslo: Unipub forlag, 133–171.
- Johannessen, Janne Bondi & Kristin Hagen 2014: Om artiklene i denne boka og *Nordisk dialektkorpus*. Johannessen, Janne Bondi & Kristin Hagen (red.): *Språk i Norge og nabolanda. Ny forskning og talespråk*. Oslo: Novus, 7–25.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Åfarli & Øystein Alexander Vangsnes 2009: The Nordic Dialect Corpus – An Advanced Research Tool. Jokinen, Kristiina & Eckhard Bick (red.): *Proceedings of the 17th Nordic Conference of Computational Linguistics, NODALIDA 2009*, 73–80.
- Kayne, Richard 2000: *Parameters and Universals*. Oxford: Oxford University Press.
- Sandøy, Helge 1976: *Laust samansette verb i vestnordisk: ein samanliknande leddstillingsanalyse for islandsk, færøysk og romsdalsmål*. Masteravhandling, UiO. Tilgjengeleg her: <http://folk.uib.no/hnohs/Publikasjonar.html> (lasta ned i juni 2015).
- Sandøy, Helge 1985: *Norsk dialektkunnskap*. Oslo: Novus.
- Schütze, Carson T. 2011: Linguistic Evidence and Grammatical Theory. *WIREs Cognitive Science* 2, 206–221.
- Svendsen, Martin 1931: *Syntaksen i Stavanger bymål*. Oslo: Aschehoug.
- Aa, Leiv Inge 2015: *The Grammar of Verb-Particle Constructions in Spoken Norwegian*. Doktoravhandling, NTNU.
- Åfarli, Tor Anders 2007: Do Verbs Have Argument Structure? Reuland, E., T. Bhattacharya & G. Spathas (red.): *Argument Structure*. Amsterdam: John Benjamins.
- Aasen, Ivar 1848: *Det norske Folkesprogs Grammatik*. Christiania: Werner & Comp.

Leiv Inge Aa
Høgskulen i Sogn og Fjordane
Pb. 133
6851 Sogndal
leiv.inge.aa@hisf.no