



Å TA MÅL AV ET SPRÅK

av André Lynum

Denne artikkelen vil belyse ulike måter å angripe studiet av språk på innenfor teoretisk lingvistikk. Den vil vise at studiet av språk også berører filosofiske problemstillinger, som erkjennelsesteori, og skillet mellom empirisme og rasjonalisme. Det vil presenteres empiriske metoder som et alternativ til rådende metodikk med opphav i generativ lingvistikk. Automatiske maskinlæringsteknikker fra empirisk lingvistikk vil knyttes til teknikker for automatisk forskning innenfor andre felt.

Det er knyttet mange spørsmål til forskning på menneskets språkevne. Hva er den, og hvilke uttrykk av den har man mulighet til å studere? Også erkjennelsesteoretiske problemstillinger melder seg, som hvilke kilder til kunnskap man bør vende seg mot, og i hvilken grad man kan legge abstrakte egenskaper til grunn for sine hypoteser. Hvordan man besvarer disse spørsmålene legger grunnlaget for hvilke hypoteser man kan stille og hvordan man kan gå til verks for å samle støtte for dem. Denne artikkelen vil presentere to ulike svar på disse to spørsmålene og belyse hvordan de former metode og hypotesedannelse. Fokus vil være på nyere metoder motivert av empiriske teknikker, særlig rent datadrevne metoder som har blitt utviklet av lingvist John Goldsmith, og på hvordan de forholder seg til den forståelse av språk som ble utviklet av Noam Chomsky på slutten av 50-tallet og som har vært rådende innenfor generativ lingvistikk siden.

Goldsmiths empiriske teknikker er nært beslektet med statistiske maskinlæringsteknikker, men kan også sees i sammenheng med arbeider innenfor automatisk forskning hvor deler av forskningsprosessen overlates til automatiske prosesser. Goldsmith argumenterer først og fremst for en automatisk og empirisk basert evaluering av hypoteser innenfor visse lingvistiske arbeider, men hans automatiske

utvikling av morfologisystemer fra språkdata kan forståes som et system for automatisk hypotesegenerering og utvikling.

Denne artikkelen vil ta utgangspunkt i generativ lingvistikk siden denne retningen innenfor lingvistikken har hatt til hensikt å beskrive menneskets språkevne på en abstrakt måte. Derfor kan generativ lingvistikk knyttes til viktige filosofiske spørsmål som erkjennelsesteori og skillet mellom rasjonalisme og empirisme. Det er min påstand at de motsetningene innenfor teoretisk lingvistikk som blir beskrevet også representerer en dypere uenighet på disse områdene, og at denne konflikten innenfor lingvistikk representerer et skjæringspunkt mellom to oppfatninger av hva språk som kognitiv egenskap er, og i en større erkjennelsesteoretisk sammenheng hvilke problemer man har med å definere hva et språk er.

GRUNNLAGET FOR MODERNE GENERATIV LINGVISTIKK

Fødselen til moderne lingvistikk sammenfaller med det som har blitt kalt den kognitive revolusjon. Startskuddet var Chomskys kritikk av den psykologiske behaviorismen og dens forsøk på å beskrive den menneskelige språkevne ved bruk av eksternt observerbare stimuli og respons alene (Chomsky 1957). Chomsky formet teoretiske argumenter for at den uendelige variasjonen vi finner i menneskelige språk ikke kan forklares med en seriell modell slik psykologen Skinner hevdet, men snarere krevde en rekursiv struktur¹. Dette ga opphav til Chomskys banebrytende teorier om generativ syntaks som fikk sterk innflytelse på lingvistisk forskning.

I et forsøk på å studere denne rekursive strukturen konsentrerte Chomsky seg om en abstraksjon av det neu-

rofysiologiske apparatet som faktisk befinner seg i hjernen og som han antok var utenfor rekkevidde å studere. Denne abstraksjonen skulle beskrive en idealisert menneskelig språkevne og hadde til hensikt å minimalisere betydning av sosiopolitiske og normative faktorer slik at man best mulig kunne studere den felles menneskelige språkevnen. Han forstår språkevnen som en relasjon mellom språkbrukernes fysiske hjerne og det idealiserte språket. Som en følge av dette forkaster Chomsky de observerbare språklige ytringer som han kaller E-språk og konsentrerer seg om denne relasjonen som han kaller I-språk, en abstraksjon av den generelle menneskelige språkevnen slik den fremstår som språkbrukerens kompetanse på sitt morsmål (Chomsky 1986).

FOKUS PÅ DEN INTERNE SPRÅKEVNE

Chomsky ser for seg syntaks som et fenomen som både kan isoleres fra annen kontekst og som i stor grad er formet av de underliggende kognitive strukturer som former I-språket. Han velger å konsentrere seg om dette feltet, og han hevder at I-språket nødvendigvis må eksistere siden språket er produktivt på en måte som ikke kan avledes ved analogi eller læring slik mer empirisk orienterte lingvister og språkfilosofene hadde beskrevet den menneskelige språkevnen². E-språket blir betraktet som et slags sekundært og uinteressant fenomen. Studiet av E-språk vil lede til tvilsomme og arbitrære strukturer som ikke er interessante siden det kan utvikles flere grammatikker som beskriver det observerbare materialet i motsetning til studiet av I-språk som vil lede til den riktige underliggende strukturen. Dette fordi Chomsky fremmer en hypotese som hevder at de forskjellige I-språkene ikke er uavhengige, men består av en felles kjernestruktur S_0 og et lært element (Chomsky 1986).

Chomsky reduserer fokuset for lingvistisk forskning til S_0 , den felles kognitive kjernen i I-språket, som han kaller Universal Grammar (UG). Hypotesen om en slik felles kognitiv egenskap er blitt kalt UG-hypotesen og består av de strukturer som igjen legger grunnlag for hvert språks struktur. En velkjent metafor, inspirert av den nå nærmest forlatte prinsipp- og parameterteorien (PPT), er UG— et sett med brytere som styrer strukturen til I-språket. Chomsky beskriver UG som en struktur som setter rammer for hvordan hvert enkelt I-språk kan være formet (Chomsky 1986). UG er dermed en abstrakt teoretisk konstruksjon som legger føringer for strukturen til I-språket, en annen abstrakt konstruksjon. Det er eksistensen til UG som motiverer studiet av I-språket, men hva UG-hypotesen skal inneholde er et omstridt spørsmål. Dette har

konsekvenser for Chomskys syn på språkforskning. En felles styrende struktur for den menneskelige språkevnen avviser all relativisme i forhold til I-språket. En teori om et gitt I-språk må samsvare i sin modell for S_0 i forhold til andre I-språk. Samtidig vil det være mulig for studiet av et I-språk å gi informasjon om strukturen til et annet I-språk. Et relativistisk syn på språk er dermed utelukket under UG-hypotesen, og det følger av den at det er meningsløst å studere enkeltspråk for seg (Chomsky 1986).

Det er spesielt for lingvistikk at lingvisten kan bruke sin egen bedømmelse av språklige ytringer i kraft av å være en innfødt språkbruker av sitt morsmål.

Som den beste kilden til studiet av I-språket benytter Chomsky en innfødt³ språkbrukers bedømmelse av velformetheten til en språklig ytring. En slik bedømmelse regnes som direkte evidens, ikke for selve I-språket, men for de begrensninger I-språket legger på språklige ytringer. Chomsky holder eksplisitt døren åpen for annen evidens, og muligheten for at slike bedømmelser blir avleggs hvis et bedre verktøy skulle bli tilgjengelig (Chomsky 1986), men på det nåværende tidspunkt gir disse bedømmelsene en unik innsikt i I-språkets struktur. En teori bør ifølge Chomsky beskrive den mengden av ytringer som er velformede og da per definisjon dens komplementære mengde av ikke-velformede ytringer. Dette skarpe skillet mellom velformede og ikke velformede setninger har vært sentralt innenfor generativ lingvistikk siden, og målet har vært å skape formelle systemer som beskriver denne mengden av velformede ytringer. En innfødt språkbrukers bedømmelse er kilden man bruker for å finne eksemplerytringer, og videre brukes analytiske og teoretiske verktøy til å forme de antatte strukturene (Chomsky 1986). Det er spesielt for lingvistikk at lingvisten kan bruke sin egen bedømmelse av språklige ytringer i kraft av å være en innfødt språkbruker av sitt morsmål. Dette gjør at metodikken innenfor generativ lingvistikk kan ligne på introspeksjon og kritikere har hevdet at fokuset i for høy grad er på forståelse på bekostning av deknning (Goldsmith 2007). Vi kommer tilbake til en mer rigorøs bruk av slike bedømmelser slik de har blitt anvendt av lingvistene Joan Bresnan og Ivan Sag senere i artikkelen.

Det kan også se ut som om UG ikke bare er en hypotese, men at den har en større rolle som motivasjon for generativ lingvistikk som forskningsfelt. For meg som

student var det påfallende hvordan UG presenteres som et fokus for språkforskning i begynnelsen av en lærebok for så å forsvinne nesten helt når man snur seg mot det faglige materialet⁴. Noe lignende kan sees i publiserte artikler hvor UG-hypotesen trekkes fram på et mer abstrakt og retorisk nivå, som i Bresnan (1994), snarere enn som en hypotese med nok innhold til å kunne inngå i direkte støtte for en gitt teori.

KRITIKK AV GENERATIV LINGVISTIKK OG UG-HYPOTESENS INNHOLD

Generativ lingvistikk og Chomskys påstander om grunnlaget for vår språkevrne har vært utsatt for økende kritikk. På den ene siden har man kritikere som ser på språket som et rent lært konstrukt og mener at vår språkevrne er en del av en generell læringsevne. For disse kritikerne kan språket læres fra observert språkbruk alene, og de er skeptiske til et kognitivt apparat dedikert til språk. Denne kritikken er i hovedsak negativ og søker å finne motbevis og kritisere fremlagte bevis for UG snarere enn å legge fram et positivt program for studiet av menneskelig språkevrne (Sampson 2000).

På den andre siden har man Chomsky selv som argumenterer for at det finnes en rik kognitiv struktur med sterk innflytelse på et språks form. En litt spøkefull formulering som er brukt til å beskrive UG er «som kontrollpanelet på et fly», fylt med innstillinger og brytere som styrer språklige fenomener.⁵ Chomskys teoretiske konstruksjon av UG og hans betegnelse av denne egenskapen som en algoritmisk konstruksjon lik en datamaskin har vært utsatt for sterk kritikk innen kognitiv psykologi og neurofysiologi som har mer direkte studier av hjernens oppbygning som mål.⁶

DET EMPIRISKE ALTERNATIV

Lingvistiske empirister ligger mellom disse to ytterpunktene og avviser ikke UG som abstrakt konsept, men ønsker å minimalisere hva som legges i denne hypotesen uten tilstrekkelig evidens fra observerbare data.⁷ En kritiker av generativ lingvistisk metode er Goldsmith. Han fremsetter to argumenter som kritikk av hvordan generative lingvister forsker på den menneskelige språkevrne. Det ene er et forsøk på å redefinere rollen til UG-hypotesen. Det andre er en oppfordring til generative lingvister om å ta observerte data mer på alvor.

Kritikere som Goldsmith ønsker ikke å avvise UG-hypotesen, men mener at man bør sette den i et nytt perspektiv. Han hevder at innholdet i UG-hypotesen må revideres og bringes i samsvar med nye metoder som er



Illustrasjon: Ragnhild Aamås

blitt tilgjengelige og med resultater fra andre felt.⁸ Det er dette empirisk orienterte alternativet jeg ønsker å presentere her.

Goldsmith henviser til en allmenn skepsis innenfor forskning til å lene seg på abstrakte konstruksjoner man ikke har direkte evidens for i sine teorier. Det er tilfeller hvor dette har ledet fram, men da har det vært sterkt samsvar mellom observerte data og eksistensen til det abstrakte konstruktet. Goldsmith mener det ikke er grunnlag for å kunne hevde dette om UG og ønsker å fylle UG med et innhold man kan utlede fra et empirisk grunnlag.

Goldsmiths andre argument fremhever at man bør vende tilbake til en metode som tar hensyn til de observasjonene man gjør av språkbruk. Generativ lingvistikk med bakgrunn i Chomskys arbeider har hatt et teoritunget fokus og unnlatt å ta for seg de rammer som kontekst og mer rigorøse eksperimenter kan gi. I likhet med Chomsky ønsker Goldsmith at man innenfor lingvistikken fortsatt betrakter den menneskelige språkevrnen som en funksjonell abstraksjon isolert fra fysiologiske fenomener, uttrykt ved algoritmiske konstruksjoner, og sier seg enig med Chomsky i at dette er et riktig abstraksjonsnivå for studiet av den menneskelige språkevrnen. Men Goldsmith argumenterer for at de observerte data må ha en større innflytelse på evaluering av lingvistiske teorier. En mer empirisk drevet lingvistisk metode kan skape kunnskap om språk som vil være mer relevant i forhold til de framskritt man ser i for eksempel neurofysiologi. Man får stadig bedre verktøy for fysiologiske studier av hjernen, og det er i en lingvists interesse at de modeller man utvikler samsvarer med de oppdagelser som vil komme i kjølvannet av det

te, slik at den lingvistiske forståelse av språk kan forenes med en fremtidig psykologisk eller fysiologisk forståelse.⁹ I motsetning til Chomskys og andres antagelser om den felles menneskelige bakgrunnen for vår språkevne vil empiriske data være et objektivt fundament for lingvistiske modeller og UG-hypotesen.

En mer empirisk drevet lingvistisk metode kan skape kunnskap om språk som vil være mer relevant i forhold til de framskritt man ser i for eksempel neurofysiologi.

Et eksempel på hvordan empiriske eksperimenter kan bidra til arbeide innenfor teoretisk lingvistikk er Joan Bresnans undersøkelser av dativ-alternering hos ditransitive verb. Dette er et svært vanlig språklig fenomen hvor direkte og indirekte objekt bytter plass, og det direkte objektet (kalt dativ) erstattes med en preposisjonsfrase. Et eksempel er dobbel-objekt konstruksjonen «Han ga henne en gave» og den dativ-alternerte konstruksjonen «Han ga en gave til henne». Generative grammatikere har hevdet at noen verb kun tillater den ene av disse konstruksjonene på semantisk grunnlag. For eksempel vil «I dragged John a box» ikke være velformet da «drag» regnes som et *transfer* verb som ikke tillater dativ-alternering. Men man finner moteksempler til dette i observerbar språkbruk, og Bresnan argumenterer for at slike avvik fra generative grammatiske modeller er reflektert i brukernes intuitive språkevner. Konstruksjonen en språkbruker benytter seg av kan forutsies ut ifra språklig kontekst, og hun viser dette ved bruk av psykolingvistiske eksperimenter i Bresnan (2006). Andre generative lingvister har tatt opp lignende empiriske metoder, som Ivan Sags studie av *Wh- superiority effect violations* (SUV), et fenomen hvor engelske *wh*-spørreord plasseres i en setning slik at den vanligvis ikke oppfattes som velformet av generative lingvister. Sag fant at ingen av ytringene med SUV flyttinger ble oppfattet som uakseptable av alle språkbrukerne, og i noen tilfeller var den ugrammatiske setningen mer akseptabel enn den grammatiske. Sag rangerer SUV-fenomen på en skala for språkbrukernes aksept og argumenterer for at denne variasjonen er et resultat av krav til kognitiv prosessering (Sag 2007). Man kan spørre seg om generativ lingvistikks fokus på velformethet som en enten/eller egenskap virkelig er i samsvar med observerbare data, og om begrepet bør utforskes ytterligere hvis man mener dette er tilfelle. Bresnan refererer i Bresnan (2006) til studier som argumenterer for at velformethet ikke er en komplementær

egenskap avhengig av syntaks alene, men er kontekstavhengig og probabilistisk (Bresnan 2006).

KVANTIFISERING AV KUNNSKAP

Goldsmith går et steg videre i sin argumentasjon for et empirisk syn på lingvistikk, og presenterer formelle redskaper for bruk i empirisk undersøkelse av språk. For eksempel fremhever han at for syntaks kan probabilistiske grammatikker benyttes som et formelt redskap som kan beskrive frekvensen av fenomener i forhold til data. En probabilistisk grammatikk er en grammatikk med en sannsynlighetsverdi knyttet til dens regler og andre strukturer. Denne sannsynlighetsverdien kan dermed brukes til å beskrive i hvilken grad en grammatikk dekker en samling språklige ytringer.

Goldsmith tar også i bruk rent datadrevne teknikker synlig inspirert av resultater fra maskinlæring. I sin artikkel *An algorithm for the unsupervised learning of morphology* (Goldsmith 1998) beskriver han et system for automatisk læring av morfologisk¹⁰ struktur for en viss klasse med språk. Dette er et forsøk på å kartlegge hvor mye morfologisk kunnskap som kan trekkes ut fra observerte data alene, og hvordan dette forholder seg til den analysen en lingvist har utarbeidet. Dette kan bidra til en kvantifisering av den menneskelige språkegenskap og en objektiv evaluering av lingvistiske teorier på grunnlag av observerte data. Han hevder at så langt man kan se er dekningsgrad og kompleksitet de faktorer man trenger for å kunne rangere to teorier i forhold til hverandre. Jeg vil utdype disse to begrepene i beskrivelsen av systemet under, og også hvordan summen av disse to faktorene danner et objektivt mål for hvor godt en teori beskriver observerte data.

BAYESISK INDUKSJON OG GOLDSMITHS METODE

Goldsmith mener man kan forstå probabilitetsteori som en kvantitativ teori for empirisk evidens og dermed danne et erkjennelsesteoretisk grunnlag for empirisk metode. Han argumenterer for dette på grunnlag av det Bayesiske synet på induksjon, som framholder at induksjon er oppdatering av en a priori tro på kunnskap om en hypotese i lys av nyere observasjoner, og at under en ubegrenset mengde observasjoner vil vår induktive kunnskap om et fenomen nærme seg den reelle tilstanden. Den benytter Bayes formel, en ligning som følger direkte fra definisjonen av en kondisjonell sannsynlighet og de grunnleggende aksiomer for sannsynlighetslære. Ved hjelp av Bayes formel kan man formulere sannsynligheten for en hypotese H gitt evidens D , $\Pr(H|D)$, ved hjelp av sannsynligheten

• André Lynum

for å observere evidens D gitt hypotesen H, $\Pr(D|H)$:

$$\Pr(H|D) = \frac{\Pr(D|H)\Pr(H)}{\Pr(D)}$$

Denne sannsynligheten, $\Pr(D|H)$, er den delen modellen gjør kjent for oss gjennom observasjoner av data. A priori sannsynlighet for observerte data D er vanskeligere å fastsette, men dette er ikke nødvendig da vi ønsker å finne den mest sannsynlige hypotesen H av mulige hypoteser h, istedenfor å finne $\Pr(H|D)$ nøyaktig. Vi kan da se bort fra nevneren i Bayes formel siden den ikke påvirker forholdet mellom $\Pr(H|D)$ for forskjellige h:

$$H = \underset{h}{\operatorname{argmax}} \Pr(D|h)\Pr(h)$$

Hvor $\Pr(h)$ er a priori sannsynlighet for en hypotese. Dette er en subjektivt estimert størrelse og Goldsmith argumenterer for at den bør favorisere modeller som er målbart kortere, basert på dens algoritmiske kompleksitet. Dette er ikke en gitt forutsetning, men han forsvarer dette på et generelt grunnlag siden en hypotese¹¹ er en abstraksjon, og han tydeliggjør at dette er under den forutsetning at alle andre faktorer holdes fast. Han velger å måle lengden til en hypotese i form av instruksjoner for en universell Turing-maskin (UTM) som er en idealisert datamaskin som kan brukes til å modellere beregnbarhet. Dette gir oss $\Pr(h)$ i formelen over. I denne modellen finner vi ikke bare en kvantifisering basert på observerte data som beskriver hvor god en gitt hypotese er, men også et instrument for å finne en teori for UG basert på observerbare data. Innenfor rammen av en slik maskin kan en lingvist skille ut sin hypotese for UG og igjen formulere spesifikke syntakser G for et gitt språk på grunnlag av denne hypotesen. Dermed er den UG-hypotesen man benytter seg av en del av den totale modellen som evalueres, og Goldsmiths metode omfatter dermed et mål av UG. Goldsmith argumenterer for at man ved denne metoden kan gi UG et innhold forankret i konkrete språkdata, og at dette innholdet vil være mer verdifullt for språkstudier innenfor lingvistikk og på andre felt enn Chomskys strukturerte innhold med opphav i hans teoretiske konstruksjoner.

Goldsmith vier betydelig oppmerksomhet til det å definere en objektiv UTM som ikke favoriserer én teori foran en annen, slik at teorier på en rettferdig måte kan måles mot hverandre. Dette er nødvendig siden Turing-maskinen selv ikke er en del av modellen, og en slik sann-

synlighet vil være avhengig av hvordan denne maskinen er spesifisert.

GOLDSMITHS SYSTEM FOR AUTOMATISK LÆRING OG EVALUERING

Goldsmith har omsatt sine ideer til praksis og utviklet et softwaresystem for automatisk læring av morfologi kalt *Linguistica*¹². Han har tatt for seg morfologi under den antagelse at dette vil være den enkleste delen av menneskelige språk å analysere på denne måten, og at gode resultater på dette feltet er nødvendig for man begynner med mer komplekse problemer, som syntaks. Systemet er beregnet primært på europeisk språk med tradisjonell morfologi.¹³ Formaliseringen av kvaliteten til en teori for morfologi som systemet benytter seg av kalles *description length*. Denne egenskapen skal minimaliseres, og derfor brukes beskrivelsen *minimum description length* (MDL). Den uttrykkes i formelen under:

$$\text{DescriptionLength}(D,H) = \text{length}(H) + \log_2 \frac{1}{\Pr(D|H)}$$

Hvor observerte data D er et språklig korpus og hypotesen H et morfologisk system. Her er *length*-funksjonen spesifikk for systemet. Den baserer seg på antallet ordstammer t, antallet morfologiske suffikser f og signaturer σ som beskriver hvordan stammer og suffikser kan kombineres. Den kan sammenlignes med et komprimeringsprogram som gir den minste mulige beskrivelsen av de observerte data.

Sannsynligheten for observerte data D kalkuleres som summen av sannsynligheten for hvert ord w i D:

$$\log_2 \frac{1}{\Pr(D|H)} = - \sum_{w \in D} \log \Pr(w|H)$$

Og sannsynligheten for hvert ord gitt komponentene i morfologien H beskrevet over:¹⁴

$$\Pr(w) = \Pr(\sigma)\Pr(t|\sigma)\Pr(f|\sigma)$$

Denne beskrivelsen av forholdet mellom observerte data og morfologisk teori danner grunnlaget for den delen av systemet som iterativt genererer og evaluerer morfologiske systemer. Dette ligner på systemer for automatisk forskning hvor det benyttes ulike former for deduksjon til

å automatisk generere potensielle hypoteser som evalueres og som kan danne grunnlag for videre hypotesedannelse. Automatiske prosesser blir stadig vanligere innenfor felt som medisin og genetik, og i enkelte tilfeller har man vært i stand til å automatisere betydelige oppgaver som *the Robot Scientist*, et nesten helautomatisk system for hypotesegenerering og evaluering innenfor genetik, som inkluderer en robot som utfører eksperimentene (King 2004).

KONKLUSJON

I nyere tid har det blitt mulig å ta i bruk sofistikerte metoder og redskaper innenfor probabilitetslære og statistikk, og sammen med fremskritt innenfor datateknologi gjør dette det mulig å benytte empiriske data i en helt ny målestokk. Det er blitt mulig å lage, teste og evaluere hypoteser på svært kort tid. Dette kan sees i sammenheng med nyere utvikling av maskinlæring og maskinledet forskning, og jeg tror det kan være på sin plass å sammenligne Goldsmiths system for automatisk generering av morfologiske modeller med andre systemer for automatisk hypotesegenerering og evaluering. Goldsmiths maskinlæringsteknikker knyttet

til empirisk forankrede evalueringskriterier kan vise veien mot en lingvistisk forskning som bygger sin kunnskap på et fundament av observerte data og at deler av dette fundamentet kan skapes av automatiske teknikker.

Utvikling innenfor moderne lingvistikk er også interessant i et vitenskapsfilosofisk og erkjennelsesteoretisk perspektiv. Forståelsen av hva den menneskelige språkevnen er, former de metoder man tar i bruk for å samle kunnskap om den. Chomskys, i høy grad introspektive metoder, står på mange måter i opposisjon til de empiriske metodene som stadig flere lingvister og andre språkforskere tar i bruk. Man kan også spørre seg om den definering og avgrensning av fokus som har lagt grunnlaget for generativ lingvistikk også har bidratt til å gjøre feltet mer innadvendt og gjort resultatene mindre tilgjengelig for andre forskningsdisipliner som kognitiv psykologi. I så henseende ser jeg på de nyere empirisk orienterte metodene som et viktig perspektiv og et fundament som kan gjøre arbeidet innenfor lingvistikk svært relevante for forskere også utenfor feltet.

Abonner på **FILOSOFISK SUPPLEMENT**

Filosofisk supplement kommer ut fire ganger årlig, og behandler alle slags emner som kan tenkes knyttet til filosofi.

Et årsabonnementet koster kr. 140,-

For å bestille send en mail med navn og adresse til:
filosofisk-supplement@ifikk.uio.no



Mer info finner du også på vår hjemmeside <http://foreninger.uio.no/filosofisk-supplement/>

NOTER

- 1 Chomsky utdyper dette i hans berømte kritikk av Skinners «Verbal Behaviour» Chomsky (1957)
- 2 Chomsky refererer til Quine, Bloomfield Og C.F. Hockett sine argumenter i Chomsky (1986).
- 3 Jeg mener med innfødt språkbruker det begrepet som på engelsk kalles *native speaker* og refererer til en språkbruker som har det aktuelle språket som morsmål.
- 4 Min lærebok i introduksjon til generativ syntaks var Carney (2002)
- 5 Goldsmith sammenligner UG med kontrollpanelet til et fly i Goldsmith (2007). Analogien md UG som et panel med brytere er svært populær blant beskrivelser av prinsipp og parameter teorien. Læreboken Carney (2002:130) inneholder til og med en tegning av et lysbryterpanel som illustrasjon på hvordan UG former språk.
- 6 ibid 5.
- 7 Beskrivelsen av og argumentasjonen for de empiriske metodene i det som følger er hentet fra (Goldsmith 2007) hvor ikke annet er referert eller kommer frem av teksten.
- 8 Chomsky har utformet et svar til dette hvor han skisserer kontaktpunkter mellom generativ lingvistik og biologisk evolusjon (Chomsky 2002) og dette svaret har igjen blitt kritisert (Pinker 2004).
- 9 For eksempel reflekterer Marcus, Vouloumanos og Sag på hvordan UG og generativ grammatikk kan knyttes til områder i hjernen som Broca's område og refererer til ytterligere materiale (Marcus 2003).
- 10 Morfologi er læren om ords indre struktur som f. eks. bøyingsparadigmer for verb og substantiv i norsk.
- 11 Jeg bruker ordene hypotese og modell om spesifikke modeller for syntaks og morfologi i mer eller mindre spesifikk betydning.
- 12 Beskrivelsen av systemet er hentet fra Goldsmith (1998) og teksten i dette avsnittet refererer til denne artikkelen.
- 13 Her menes morfologi slik vi finner den i norsk, engelsk, tysk og andre beslektede europeiske språk stort sett bestående av stammer og suffikser.
- 14 I Goldsmith (1998) uttrykt som summen av logaritmen av sannsynlighetene.

LITTERATUR

- Bresnan, J. 1994, "Locative inversion and the architecture of universal grammar", *Language* 1, bd. 70, s.72-131
- Bresnan, J. 2006, *Is syntactic knowledge probabilistic? Experiments with the English dative alternation*. Tilgjengelig fra: <http://www.stanford.edu/~bresnan/bresnan-lingevid-corrected.pdf>
- Carney, A. 2002, *Syntax a generative introduction*, Blackwell publishing, England.
- Chomsky, N.1957, "A Review of B. F. Skinner's Verbal Behavior", *Readings in the Psychology of Language*, Prentice-Hall, 1967, s.142-143. Tilgjengelig fra: <http://www.chomsky.info/articles/1967----.htm>
- Chomsky, N. 1986, *Knowledge of language. It's nature, origin, and use*, Praeger publishers, kap. 1 og 2, hentet fra HUMIT4700 kompendium, høst 2006, Unipub, Oslo.
- Goldsmith, J. 1998, *An Algorithm for the Unsupervised Learning of Morphology*. Tilgjengelig fra: <http://hum.uchicago.edu/%7Ejagoldsm/Papers/algorithm.pdf>
- Goldsmith, J. 2007, *Towards a new empiricism*, ver. 1.6. Tilgjengelig fra: <http://hum.uchicago.edu/~jagoldsm/Papers/empiricism1.pdf>
- Hauser, M. D. ea. 2002, "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?", *Science*, 22, bd. 298, s. 1569 - 1579.
- King, R. D. ea. 2004, "Functional genomic hypothesis generation and experimentation by a robot scientist", *Nature*, 427, s. 247-251.
- Marcus, G. F. ea. 2003, "Does Broca's play by the rules?", *Nature Neuroscience*, 7, bd. 6, s. 651-652. Tilgjengelig fra: <http://lingo.stanford.edu/sag/papers/Marcus+2003.pdf>
- Pinker, S. & Jackendoff, R. 2004, "The faculty of language: what's special about it?", *Cognition*, 2, bd. 95, s. 201-236.
- Sag, I. A. ea. 2007, *Processing Accounts for Superiority Effects*. Tilgjengelig fra: <http://lingo.stanford.edu/sag/squib.pdf>
- Sampson, G. 2000, *There is No Language Instinct*. Tilgjengelig fra: <http://www.grsampson.net/ATin.html>