



(12) PATENT

(19) NO

(11) 324419

(13) B1

NORGE

(51) Int Cl.

G06F 3/048 (2006.01)

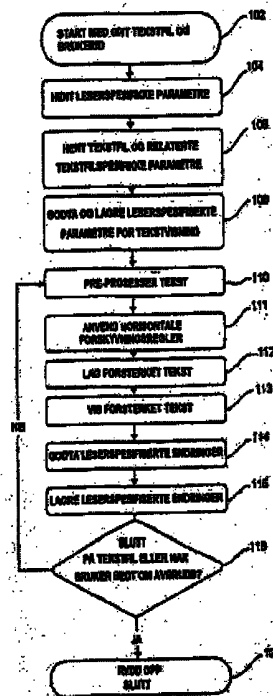
G06F 17/21 (2006.01)

Patentstyret

(21)	Søknadsnr	20003925	(86)	Int.inng.dag og søknadsnr	1999.02.02 PCT/US99/02227
(22)	Inng.dag	2000.08.02	(85)	Videreføringsdag	2000.08.02
(24)	Løpedag	1999.02.02	(30)	Prioritet	1998.02.02, US, 16902
(41)	Alm.tilgj	2000.09.27			
(45)	Meddelt	2007.10.08			
(73)	Innehaver	Randall C. Walker, 923 Tenth Street S.W., Rochester, MN 55902, US			
(72)	Oppfinner	Randall C. Walker, 923 Tenth Street S.W., Rochester, MN 55902, US			
(74)	Fullmektig	Curo AS, Industriveien 53, 7080 HEIMDAL			

(54)	Benevnelse	Fremgangsmåte og apparat for tekstbehandling
(56)	Anførte publikasjoner	JP 5204587 A
(57)	Sammendrag	

Fremgangsmåte for forsterkning av tekst og apparat for presentasjon av tekst for å fremme tekstlesing hos mennesker. Fremgangsmåten omfatter ekstraksjon av tekstspesifikke attributter fra maskinlesbar tekst og endre tekstpresentasjonen i henhold til attributtene. Deler av talen og skille tegn ekstraheres fra ei setning, og det anvendes brytningsregler som bruker deler av talen for å bestemme brytningspunkter, og tekstsegmenter presenteres hver på ei ny linje og med en bestemt horisontal forskyvning basert på de tekstspesifikke attributtene. I en utførelse vises teksten bøyd over kurver med en form basert på tekstinnhold. I en annen utførelse vises relative tekstposisjoner innenfor et hierarki ved bruk av alternerende vertikalt og horisontalt flislage plan. En annen utførelse støtter lesing av tekstsegmenter over motstående sider uten å måtte vente på sideskift. I nok en alternativ utførelse vises tekst som skal kunne leses fra bunn til topp som om det var fra front til bakside. Ord kan vises med farger som gjenspeiler forholdet mellom ordene og de større tekstsegmentene de er en del av.



Fremgangsmåte og apparat for tekstbehandling

Oppfinnelsen angår datamaskin-assistert tekst-presentasjon, nærmere bestemt en forbedret tekstpresentasjons-metode ved bruk av innholdsspesifikke attributter for å fremme leseferdigheten.

Bakgrunn

Formålet med lesing er å gi et inntrykk i leserens hode. Dette er tilfelle enten teksten som leses omfatter en roman eller en informasjonsskjerm om bord i fly. Materiale som presenteres i et ikke-bokstavelig medium leder informasjon som er velegnet for human absorpsjon langt utover den tilsvarende endimensjonale tekst. Også bilder presenterer objekt-attributter slik som farger, relative størrelser, relative posisjoner, mønstre, grupperinger og hierarkier. Et objekt betraktes i en tekstsammenheng. For eksempel vil et bilde av et lite objekt lokalisert i et hjørne sammen med andre gi en forskjellig informasjon enn det samme objektet forstørret, av seg selv, og i sentrum. Begynnende setninger og avsnitt vil alene på grunn av sin posisjon tillegge informasjon som beror på sammenhengen. Bevegende bilder har bevegelse og midlertidig endring som tilleggsattributter. Mye informasjon kommer inn til hjernen visuelt og gjennom mønster-gjenkjenning. Annen informasjon kommer som lydinformasjon og bringer informasjon gjennom tone, varierende modulasjon og volum.

Datamaskin-presentasjon av tekst for lesing av mennesker har mislykkes i å anvende særlig mye av menneskets hjerne. Som et resultat anvendes bare en del av den mulige båndbredde, og datamaskintekst-presentasjon anses ofte som upersonlig, konstant, mekanisk og for å mangle et eller annet. De fleste mennesker vil dersom de får velge foretrekke å lese ei trykket bok i stedet for ei bok på en datamaskinskjerm. Dagens datamaskintekst-presentasjon er dårlig sammenliknet med ei bok, og enda dårligere i forhold til dens potensiale.

Det har blitt utført en del arbeid i forbindelse med datamaskinassistert presentasjon av tekst. Huang beskriver i US patentskrift 4 880 385 en optomekanisk hurtigleseanordning som tillater lesing av tekst trykket på papir gjennom et vindu ei linje om gangen med automatisk frammatning av linjer.

Det er gjort framskritt i framgangsmåter for datamaskinbasert analysering av naturlig språk til delvis tale. Schabes med flere beskriver i US patentskrift 5 475 588 et forbedret analysatorsystem for oppretting av analysator-tre. Black Jr. med flere beskriver i US patentskrift 5 331 556 en framgangsmåte for lagring av deler av taleinformasjon i ei fil sammen med originalteksten for å bedre påfølgende søking i teksten. Okamoto med flere

samt Church beskriver i henholdsvis US patentskrifter 4 661 924 og 5 146 405 framgangsmåter for å gjøre multiple deler av tale klarere. Zamora med flere (US patentskrift 4 887 212), van Vliembergen (US patentskrift 5 068 789), Hempill med flere (US patentskrift 5 083 268) beskriver framgangsmåter for analysering av naturlig

5 tekstspråk.

Gjennom US 4 270 284 er det kjent et system og en metode for presentasjon av tekst i et fremvisningsapparat. I metoden velges deler av teksten og fremvises i forskjellige farger på basis av språkegenskaper.

Gjennom US 5 392 387 er det kjent et system og en metode for forsterkning av

10 dataaksess i en fremvisningsanordning, der fremvisningsanordningen viser et antall indekserte vinduer som benyttes til valg av aktuelle vinduer/sider.

Lesing er en kompleks prosess og det er mange faktorer som bestemmer forskjellen i leseferdighet mellom ulike lesere og selv for samme leser ved ulike anledninger. Disse faktorene omfatter medfødte nevrofysiologiske tilstander slik som dysleksi, men også alder;

15 adferds- og motivasjonsfaktorer; utdanningsnivå og tidligere leseerfaring, og sansbegrensninger. En leser kan også ha forskjellige formål med lesingen som vil påvirke hvordan han/hun vinkler lesestoffet. På tross av det ovennevnte vil datamaskiner på sitt verste vise tekst som endimensjonale "perler på ei snor" og på sitt beste lik todimensjonal tekst presentert i bøker. Det betyr imidlertid lite om teksten presenteres på trykk eller på et

20 elektronisk visningsorgan fordi tekstpresentasjonen i hovedsak er identisk.

Eksempler på den statiske presentasjonen av tekst er mangtallig. Emnesetninger presenteres på samme måte som andre setninger i et avsnitt. Tekniske ord presenteres ikke forskjellig fra ikke-tekniske ord. Adjektiver ser likedan ut som substantiver. Pedagogisk vanskelige uttrykk presenteres på samme måte som enkle uttrykk. Det siste avsnittet i et

25 kapittel presenteres på samme måte som det første avsnittet. Teksten vises i venstrejusterte brutte linjer med konstant bredde som krever en rykkvis øybevegelse fra langt mot høyre til langt mot venstre med jevne mellomrom. Tekst blir brutt midt i en setning, midt i et uttrykk og midt i en tanke, ved bruk av gammeldagse tekstsettingsregler. Slik tekst tvinger øyet til å beveges fram og tilbake over store avstander likt et

30 skrivemaskinhode. Tekst avanserer manuelt inndelt i stumper bestemt av hvor mange linjer som passer inn på skjermen.

Det finnes flere gode grunner til ikke å plassere bare ei enkelt uthevet setning per side i ei bok. Disse årsakene kan uheldigvis ha blitt overført til datamaskinens visning av tekst. Mulighetene for å endre tekstpresentasjonen for å fremme leserens evne til å lese tekst

35 ved utnyttelse av krafta fra digital manipulering ser ut til å ha blitt oversett av de mest kunnskapsrike innen databehandling.

Formål

Det er ønskelig å framskaffe en framgangsmåte for å vise tekstposisjonen i forhold til teksthierarkiene som teksten eksisterer i. En framgangsmåte for visning av tekst på motstående sider som tillater tekstlesing uten å vente på sideskift er også ønskelig. Et tekstvisningssystem som anvender menneskets naturlige evne til og preferanse for å gå framover under observasjon av objekter ville vært nyttig. En framgangsmåte for visning av tekst ved bruk av kvantitativ fargeforandring for å bevirke en transparent presentasjon av de kumulative forhold mellom større setningsledd og ordene innenfor samme ville vært ønskelig.

For å bevare all informasjon i forfattet tekst må tekstpresentasjonen per definisjon systematisk vise sekvensen av ord skrevet av forfatteren i en utvetydig struktur. Det er også ønskelig at tilleggsinformasjonen for skildring i en uthevet tekstmodus for å fremme lesing kan innlemmes i de nødvendige prosessene for tekstpresentasjon i seg selv, (dvs. plassere tegn mot en kontrastbakgrunn i ulike posisjoner i tid og rom) i stedet for å innføre ekstra symboler som krever ekstra opplæring, lesing og tolkning av leseren.

Oppfinnelsen

En framgangsmåte i samsvar med oppfinnelsen er angitt i patentkrav 1. Fordelaktige trekk ved oppfinnelsen er angitt i patentkrav 2-3.

Oppfinnelsen angår en framgangsmåte for å fremme visning av tekst og et apparat for visning av tekst for å fremme lesing av mennesker. Oppfinnelsen omfatter ekstraksjon av tekst-spesifikke attributter fra maskinlesbar tekst, variere eller endre tekstpresentasjonen i henhold til attributtene, og lage et tredimensjonalt (tid- og leseflate) visuelt produkt for å fremme leseytelsen. Den foretrukne utførelsen av oppfinnelsen ekstraherer attributter slik som deler av tale fra en inngående setning og viser den aktuelle setning i strømmende tekstsegmenter nedover og på tvers av skjermen. Segmenteringen og den horisontale forskyvningen bestemmes ved å utøve regler som dels anvender uttale, tegnsetting og leser-preferanser. Fargen på teksten og bakgrunnen kan også varieres avhengig av delen av uttalen og av setningens posisjon innenfor avsnitt og kapitler.

Oppfinnelsen tillegger meningsfylte visuelle attributter for å fremme presentasjon av tekst presentert av datamaskiner langt forbi den statiske og mekaniske presentasjon av kjente systemer. Det dannes et meningsfylt visuelt stikkord som er relatert til innholdet i setningen og dets forhold til den forrige setningen. Dette stikkordet er tilgjengelig før teksten leses, og gir en sammenheng med teksten. Setningsanalysen benytter setningens

tegnsetting og innhold for å danne et system av meningsfylte visuelle stikkord inkludert distinkte setninger eller setningsledd som fremmer raskere gjenkjenning av ordene i en setning og deres betydning. De forsterkede setningene strømmer ned siden i et mønster med meningsfylte setningsledd bestemt av tekstinhold og leser-preferanse, og øyet

5 beveges en kort distanse fra setningsledd til setningsledd.

En setning representerer en fullstendig tanke, og et avsnitt representerer et atskilt emne eller argument. Leseforståelsen forbedres når en setning framtrer om gangen og når overgangen fra ett avsnitt til et annet signaleres visuelt og innbefatter en pause for opptredenen av disse tekstelementene. Oppfinnelsen framskaffer egnete visuelle

10 attributter for å fremme disse visuelle signalene. De visuelle attributtene kan omfatte tekst-segmentering, horisontal forskyvning av en linje i forhold til en annen, tekst- og bakgrunnsfarge, tekstens lysstyrke og animasjon. Animasjonen kan omfatte blinking og slag, tidsbestemt midlertidig progressiv opplysning av tekstelementene samt bevegelse av tekst fra vanlig til strømmende presentasjon (kaskadevisning).

I et eksempel på en lesesessjon vil leseren velge teksten som skal leses, og informerer lesesystemet om teksttypen, f.eks. en roman. Lesesystemet gjenoppretter det egnete miljø som ble lagret tidligere under denne leserens lesing av den valgte teksttypen. Leserens har mulighet til å editere lesereglene og ordsammensetning men gir avkall. Ved igangsetting på begynnelsen av et kapittel der leseren stoppet opp, vises teksten en setning om gangen

20 i en kaskade nedover og på tvers av skjermen. Siden linjene som vises brytes opp i meningsfylte tekstsegmenter i stedet for tilfeldig avslutning ved 80 tegn, er det mulig å lese og forstå et helt segment om gangen der den resulterende øyebevegelsen skjer fra sentrum av ett tekstsegment til sentrum av et annet, nedover og på tvers av skjermen. Setningenes bakgrunnsfarge er en funksjon av setningsposisjonen innenfor avsnittet og

25 avsnittets posisjon innenfor kapitlet. Hver av setningene i et avsnitt kan ha samme nyanse men oppvise tiltakende men gjenkjennbare endringer i fargemetning eller mørkhet mellom setningene, og tiltakende men gjenkjennbare endringer i nyanse mellom avsnittene. Bakgrunnsfargen gir på denne måten en ikke-ordrett posisjonsforandring som signaliserer begynnelsen og slutten av et avsnitt og begynnelsen og slutten av et kapittel.

Setningene blir segmentert i henhold til leserens godkjente regler. Hver preposisjon bryter eller bryter et uttrykk på en forutsigbar måte, som resulterer i to tekstsegmenter som hver ligger på ei ny linje. Innholdet i linja over påvirker den horisontale forskyvning av linja under. Den samme preposisjonen bryter og forskyver teksten horisontalt på samme måte såfremt det er mulig. Som et resultat vil den samme preposisjonen forårsake et

30 liknende mønster som for lengst er gjenkjennelig for leseren. Deler av uttalen påvirker tekstfargen i henhold til leserens godkjente regler. Som et resultat blir både setnings-

35

struktur og deler av språket umiddelbart gjenkjennbare fra mønster- og fargegjenkjenning, selv før et ord blir gjenkjent, og gir på denne måten visuelle ikke-bokstavelige stikkord til leseren.

Setningene rykker fram med en hastighet definert ved bruk av en regel som på forhånd er valgt av leseren. Formelen er en funksjon av teksttype, antall ord, utdanningsnivå og antall stavelser tilstede i linja. Framrykkingshastigheten er litt raskere enn det som av brukeren oppfattes som komfortabel, for å øke forståelsen. En lengre pause mellom setningspresentasjoner signaliserer et nytt avsnitt før det faktisk opptrer. Leseren er på en lettvinnt måte i stand til å samvirke med lesesystemet, holde vanskelige setninger på skjermen i et lengre tidsrom og øke eller redusere hastigheten for presentasjonen.

Forløpet beskrevet foran er et eksempel, og bør ikke anses for å være uttømmende beskrivelse eller begrense omfanget av oppfinnelsen.

Den forbedrete framgangsmåten utnytter det faktum at tekstelementene har ikke-bokstavelige kvaliteter som kan benyttes til å produsere et tekstvisningsprodukt i tid og i rom som er mer meningsfylt og som fremmer leserens evne til å forstå den bokstavelige meningen med teksten i en større grad enn eksisterende formater i datamaskiner eller på papir. Kaskaderingen av uttrykk nedover og på skrå av skjermen tildeler en visuell logisk konsekvens med hensyn til hvordan teksten kunne oppleves dersom den ble hørt fra en opplesning uten at brukeren trenger å lese høyt inni seg. Lesing av en slik forsterket tekst er nesten som å lytte med øyet, men som en totalt visuell pakke som oppleves nesten samtidig.

Et annet aspekt ved oppfinnelsen omfatter registrering av aktiviteter utført av en høytlesende leser, og kan omfatte registrering av uttale, håndbevegelse, øyebevegelse, hodebevegelse og munnbevegelse. Aktivitetene blir fortrinnsvis registrert og kartlagt mot tilsvarende ord i teksten. Tekstvisningen kan varieres i henhold til de registrerte aktivitetene. Tekstvisningsattributtene kan omfatte tekstforskyvning, skrifthøyde, skrifthykkelse, innbyrdes tegnnavstand, innbyrdes ordavstand og linjeavstand.

Et annet aspekt ved oppfinnelsen viser tekstposisjon i forhold til tekstelement-hierarkiene som teksten befinner seg i ved å variere flislegging av plan eller vindu. Vinduene eller planene som representerer de ulike hierarkiene flislegges som en funksjon av rangen av tekst-element innenfor hierarkiet. I en utførelsesform flislegges vindus-settene for tekstelement-hierarkiene vekselvis horisontalt og vertikalt.

Et annet aspekt ved oppfinnelsen viser tekst på motstående sider slik at lesing av tekst over motstående sider kan finne sted uten å vente på ompaginering. En framgangsmåte omplasserer tekstinholdet på hver motstående side mens den motstående siden blir lest, uten å vente på at begge motstående sider skal bli lest.

I et annet aspekt av oppfinnelsen blir teksten vist som den opprinnelige tekstsekvensen, fra bunn til topp, i perspektiv, slik at lesing av tekst fra bunn til topp ser ut som at en leser teksten fra front til bakkant. Ën utførelsesform ruller teksten jevnt framover slik at leseren ser ut til å gå framover gjennom teksten under lesingen.

- 5 I et ytterligere aspekt ved oppfinnelsen blir det foretatt en beregning av forholdet mellom ord og de større tekstsegmentene som de er en del av. Ën utførelsesform begynner med en setning og foretar repeterende inndelinger av setningen i segmenter i henhold til forholdet mellom segmentene og delsegmentene. Forholdet mellom segmentene og delsegmentene brukes til å beregne en fargeverdi. Generasjonen eller rekursjonsnivået i
- 10 segmenterings-algoritmen brukes også til å bestemme fargeverdien. Den beregnete fargeverdien brukes til å vise fargen for hvert ord eller uttrykk.

Ytterligere fordelaktige trekk ved oppfinnelsen vil fremgå av den etterfølgende eksempelbeskrivelsen.

15 **Eksempel**

Oppfinnelsen vil nå bli mer detaljert beskrevet med henvisning til figurene, hvor:

Figur 1 er et høynivå flytskjema for en framgangsmåte som anvendes i Ën utførelsesform av oppfinnelsen,

Figur 2 er en tabell som illustrerer ulike attributter for elementene i en setning,

- 20 Figur 3 inneholder et eksempel på forsterket tekst omfattende kaskader og vinkel som avtar ned mot høyre,

Figur 4 viser setningen i Figur 3 som forsterket tekst inkludert kaskader og vinkel som avtar nedover i retning høyre og venstre,

Figur 5 viser trinn 110 i detalj,

- 25 Figur 6 viser ord-settene som etablerer et hierarki med sekundære brytningspunkter,

Figur 7 illustrerer brytningspunkter i en setning,

Figur 8 viser trinn 136 i detalj,

Figur 9 viser setningen i Figur 7 presentert som forsterket tekst,

- 30 Figur 10 viser PISHM-diagram der tiden som kreves for å lese er en funksjon av setningens vanskelighetsgrad og kompleksitet,

Figur 11 illustrerer setningen i Figur 7 med visuell uttale av framhevede stavelser,

Figur 12 er et diagram som viser gate og spor som benyttes i beregning av horisontal forskyvning,

Figur 13 viser ei kurve egnet for tekstbrytning,

- 35 Figur 14 viser reglene for utplassering av tekstsegmenter på kurva i Figur 13 som en funksjon av segment-hodets og segment-endens tekstattributter,

Figur 15 viser et midlertidig resultat fra en tekstbrytningsalgoritme,

Figur 16 viser flere brutte tekstsegmenter formet over kurver i henhold til Figur 14 og 15,

Figur 17 viser en dataskjerm med ei bok på en dataskjerm før åpning,

Figur 18 viser en dataskjerm med boka i Fig. 17 etter åpning, og illustrerer kapitlene i
5 samme,

Figur 19 viser en dataskjerm etter at et av kapitlene i Figur 18 er valgt, hvorved kapitlets
sider er illustrert,

Figur 19A viser en alternativ utførelsesform av en dataskjerm med kapitlene ifølge Figur
18 etter valg av et kapittel, og viser sidene i kapitlet,

Figur 20 viser en dataskjerm med sidene ifølge Fig. 19 etter valg av side, og illustrerer
10 avsnittene på siden,

Figur 21 viser en dataskjerm med avsnittene fra Figur 20 etter valg av ei setning, og
viser setningene i samme,

Figur 22 viser en dataskjerm med en kaskadevisning av ei setning på ei side valgt i Figur
15 21,

Figur 23 viser en framgangsmåte for visning av tekst over flere visnings-sider,

Figur 24 viser en annen framgangsmåte for visning av tekst over flere visningssider,

Figur 25 viser en framgangsmåte for visning av tekst som muliggjør lesing fra frontside til
bakside, fra bunn til topp, og som kan manipuleres til å opptre bevegende i retning
20 framover,

Figur 26 viser teksten i Figur 25 etter fremadrettet bevegelse av teksten, og

Figur 27 et tredigram av en setning inndelt for tilordning av farge og prosa-maling.

Figur 1 viser et høynivå flytskjema av en framgangsmåte som anvendes i en
25 utførelsesform av oppfinnelsen. Trinnene i Figur 1 og i alle slike figurer kan naturligvis
splittes i flere ekstratrinn eller forenes til færre trinn og fremdeles beskrive samme
utførelsen.

I trinn 202 starter leseanordningen og leseren autentiseres (ID) og tekstfil-identifikasjon
leses. I en foretrukket utførelse framskaffes leserens ID fra operativsystemet. I en annen
30 utførelsesform blir leseren spurt om identifikasjon. I en utførelsesform blir det antatt at den
forrige leseren er i gang med en ny sekvens. Gitt leserens identitet blir leserespesifikke
parametere innhentet i trinn 104. Parametere er diskutert i detalj nedenfor. Gitt tekstens
identitet blir tekstfila og de tekstspesifikke parametrene innhentet i trinn 104. I en
foretrukket utførelsesform har parametrene forhåndsdefinerte verdier og krever ingen inn-
35 verdier fra leseren. I den mest foretrukne utførelsen tillates innregistrering av brukerens
foretrukne parametere, hvorved verdiene aksepteres og lagres som i trinn 108. I en

utførelsesform tillates brukeren å trekke ned menyer som aktiverer dialogbokser, inspisere standardparametere og endre dem ved bruk av standard brukergrensesnitt inkludert numerisk editering, teksteditering, rullestang og knapper. I en foretrukket utførelsesform tillates også endring av globale parametere som gjelder for alle lesere av teksten, forutsatt at det er gitt tilstrekkelige rettigheter. Leseren tillates å sette parametrene bare for den gjeldende teksten, all tekst av samme type som gjeldende tekst, eller all tekst som leses av denne leseren.

Leserspesifikke parametere omfatter: betraktning av feltdimensjoner, fargepallett for bakgrunn og tekst, inntasting av ordsett for regler, minimal uttrykkslengde, maksimal linjelengde, minimal og maksimal tekstsegment-vekt, helningsvinkler, regler for horisontal forskyvning, preferanser for horisontal justering, størrelse på mellomrom mellom uttrykk, forflytningshastighet, intervall-forhold mellom setninger og mellom avsnitt, definisjoner for merking, spesifikasjoner for animasjon, definisjon av parametere for å avlede innholds-spesifikke presentasjonsintervaller, definisjon av parametere og veiing for å avlede uttrykk/tekst-spørsmål, densitet, kompleksitet, innhold, og identiteter for skript for spesielle hendelser.

Betraktning av feltdimensjoner

Betraktning av feltdimensjoner kan spesifisere størrelsen av vinduet for betraktning av tekst i de to X- og Y-dimensjonene og en tredje dimensjon, tid. Dette feltet kan omfatte hele skjermen eller en mindre del av skjermen. Tidsdimensjonen kan omfatte varighet for de tidsbestemte animasjonene, slik som blinking.

Fargepallett

Fargepallett for bakgrunn og tekst tillater spesifisering av foretrukket farge for bakgrunn og tekst som varierer med teksten. I en foretrukket utførelsesform varierer nyansen og lysstyrken for bakgrunnsfargen med både setningsposisjon innenfor et avsnitt og avsnittsposisjon innenfor et kapittel. I den mest foretrukne utførelsen settes lysstyrken på ett nivå for den første setningen i et avsnitt og et andre nivå for den siste setningen i avsnittet, der lysstyrken for hver setning i mellom disse tiltar fra første nivå til andre nivå. I en foretrukket utførelse settes bakgrunnsnyansen til en første verdi for det første avsnittet i et kapittel og en andre verdi for det siste avsnittet i et kapittel, der nyansen i hvert avsnitt tiltar fra den første verdien til den andre verdien. I den mest foretrukne utførelsen er den første nyansen grønn og den andre nyansen fiolett.

I den foretrukne utførelsen kan tekstfarge variere med ordets kategoriske og kontinuerlige attributter. En kategorisk attributt er en funksjon av kategorien som ordet

faller inn i. I den mest foretrukne utførelsen omfatter kategorier deler av uttale, slik som verb og tekniske ord. Kontinuerlige attributter omfatter de som kan måles slik som utdanningsnivå, antall stavelser og ordlengde. I en foretrukket utførelse har ulike deler av uttalen ulike tekstfarger. I den mest foretrukne utførelsen varieres tekstfargen med

5 bakgrunnsfarge for å gi optimal leser-bestemt kontrast.

Det kan brukes ord-sett for å spesifisere noen deler av uttalen. For eksempel "jeg" er et pronomen, "hadde" er et verb, "en" er en artikkel, og "bruk" kan være et substantiv, verb eller adjektiv. Delene av språket trenger bare være sannsynlige deler av språket slik som med ordet "bruk", som kan være substantiv, verb eller adjektiv. Et eksempel på deler av

10 språk i en enkel setning er vist i Figur 2. Flertallet av mulige deler av språkkolonnen 150 er vist i form av den tydeliggjorte språkkolonnen 152.

Ordsett er inngangsverdier i foretrukne utførelser ved bruk av to metoder. Den første anvender et strukturert tekstformat slik som SGML ("Standard Generalized Mark-up Language"), og tillater import og eksport av store ordsett og ordbøker fra diverse kilder.

15 Standard SGML-analysatorer brukes dermed til å lagre ordsettene i en egnet database. Den andre metoden tillater interaktiv editering av ordsett ved bruk av et hierarki av nedtrekksmenyer og dialogbokser. I en foretrukket utførelse lagres leser-spesifikke ordsett i en separat database atskilt fra ordboka eller glossar-ordsettene, for å tillate separat lagring, editering og kortere aksesstid.

20

Minimum tekstsegmentlengde og maksimal linjelengde

En foretrukket utførelse omfatter en leser-spesifisert minimal tekstsegmentlengde. Et tekstsegment er et stykke tekst som presenteres på ei linje på en skjerm. Den minimale tekstsegmentlengden, som kan betegnes med ord, er det minste som leseren ønsker å se

25 på ei separat linje. En maksimal linjelengde er den maksimale lengden som leseren ønsker å se på ei linje. Lengdene kan måles i antall tegn, stavelser, utdanningsnivå eller enhver sum av produktene av samme. I en utførelsesform er den minimale linjelengden ti tegn og den maksimale linjelengden trettifem tegn.

30 Vekt av uttrykk eller setning

Uttrykksvekt er en avledet attributt for et uttrykk (tekstsegment eller potensielt tekstsegment) som gir et mål for mengden materiale i et uttrykk. I en utførelse er uttrykksvekten ganske enkelt antall ord i et uttrykk. I foretrukne utførelser omfatter uttrykksvekten uttrykks-densitet og uttrykks-kompleksitet. Uttrykksdensitet kan omfatte

35 antall tekniske ord eller antall ord som overstiger et visst vanskelighetsnivå.

Uttrykkskompleksitet kan omfatte antall stave-likheter mellom ordene i et uttrykk, antall

tvetydige ord og total vekt av leserens vekt-spesifiserte ord.

Minimal og maksimal tekstsegmentvekt

En foretrukket utførelse omfatter en leser-spesifisert minimal og maksimal
5 testsegmentvekt. Tekstsegmentvekten er mengden tekst som leseren ønsker å se på ei
separat linje. I foretrukne utførelser er tekstsegmentvekten summen av uttrykksvekt
innenfor tekstsegmentet.

Regler for horisontal justering

10 Regler for horisontal justering spesifiserer den horisontale justering av ei linje med tekst i
forhold til linja over. Justering kan omfatte justeringsmåten for linja eller uttrykket som
posisjoneres, dvs. venstrejustert, høyrejustert eller sentrert. Justering kan også omfatte
tekstdelen av linja over fra hvilken justeringa måles, dvs. hele linja med tekst mot et uttrykk,
punktet for den delen som brukes, f.eks. den lengst til venstre, lengst til høye eller i midten.
15 Horisontal justering måles i én utførelsesform ganske enkelt innenfor linja som
posisjoneres i stedet for i forhold til linja over.

I en foretrukket utførelse er det første uttrykket på ei linje sentrert, målt fra sentrum av
det siste uttrykket i linja umiddelbart over. I en annen utførelse er hele linja med tekst
sentrert under sentrum av linja med tekst over. I nok en utførelse blir tekstsegmentets
20 'tyngdepunkt', beregnet ved bruk av vanskelighetsgraden for hvert ord, anvendt som
tekstsegment-sentrum for justerings- eller innrettingsformål.

Nedstigningsvinkel

En nedstigningsvinkel er graden av horisontal forskyvning av hver ny linje framkommet
25 ved endring av den horisontale posisjonen ved bruk av de horisontale justeringsreglene
foran. Per definisjon presenteres hvert tekstsegment på ei ny linje. I en foretrukket
utførelse spesifiseres nedstigningsvinkelen i antall tegn. Nedstigningsvinkelen og
horisontal justering bestemmer i det minste delvis 'tekst-kaskade' ned og over skjermen i
foretrukne utførelser. Et eksempel på kaskadetekst er vist i Figur 3 og 4. En
30 nedstigningsvinkel kan være null, som ikke betyr annet enn at tekstsegmentets horisontale
posisjon bestemmes av de horisontale justeringsreglene alene. En nedstigningsvinkel kan
være venstrerettet, der linja under skal skyves til venstre i forhold til linja over, eller
høyrerettet, der teksten skyves mot høyre. Figur 3 illustrerer bare høyrerettede
nedstigningsvinkler.

35 I én utførelsesform er nedstigningsvinkelen konstant for hver ny linje. I en foretrukket
utførelse er nedstigningsvinkelen en funksjon av tekstsegmentvekten for linja over. I nok en

foretrukket utførelse vil horisontale justeringsregler spørre etter sentrert justering under sentrum av hver linje umiddelbart over, og nedstigningsvinkelen beregnes for å presentere en generelt rett vei, når alle tekstlinjene vises, fra linjesenter til linjesenter, fra øvre venstre til nedre høyre på visningsflata.

- 5 I en foretrukket utførelse inkluderer inngangsverdiene for nedstigningsvinkelen tekstattributtene i linja over. I en foretrukket utførelse inkluderer inngangsverdiene årsaken til brytning av linja over, dvs. det primære brytningspunkt, sekundære brytningspunkt, eller sammenbruddsregel, I en foretrukket utførelse spørres det etter en mer positiv nedstigningsvinkel når linja umiddelbart over brytes på grunn av et primært brytningspunkt i
- 10 stedet for et sekundært brytningspunkt. I nok en foretrukket utførelse omfatter inngangsverdiene tekstsegment-vekten for den aktuelle linja og linja over.

En kan se at den horisontale justeringsreglemen kan spørre etter venstrejustering og måling av horisontal forskyvning fra venstremargen, samt null nedstigningsvinkel, og resultere i venstrejustert tekst på hver linje.

- 15 Det bør være nærliggende at horisontal tekstposisjonering kan utføres på en rekke ekvivalente måter som beskrevet i eksemplet foran. Spesielt kan det utføres beregninger av tekstposisjon ved først å justere/innrette og deretter skifte, eller først skifte og deretter justere med samme resultat.

20 Lengde på mellomrom

- I en utførelsesform er mellomrom forbundet med bruddpunkter hvis posisjon har blitt bestemt men som på grunn av andre regler forblir på samme linje og ikke forårsaker brytning eller linjeskift. Et mellomrom på null eller flere blanke legges til etter et bruddpunkt der bruddpunktet har mislyktes i å forårsake ny linjedannelse. I en foretrukket utførelse er
- 25 lengden på mellomrommet en leserdefinert parameter, der en mellomrom-lengde lik null fører til at det ikke dannes noen mellomrom. Mellomrom tillater et visuelt stikkord med hensyn til eksistensen av uttrykk selv der uttrykkene ikke har forårsaket ny linjedannelse.

Framrykkingshastighet

- 30 Framrykkingshastighet spesifiserer varigheten av og tidsintervallene mellom presentasjon av en tekstgruppering og den neste. I en foretrukket utførelse presenteres en setning på skjermen. I en foretrukket utførelse kan varigheten være en funksjon av tekstens omtrentlige veide antall ord, utdanningsnivå, uttaletid, antall uttrykk, antall stavelser eller uttrykksvekt. Tidsintervaller kan omfatte intervallet mellom setninger og
- 35 avsnitt. I en foretrukket utførelse er tidsintervallet mellom setninger og avsnitt forskjellig. På denne måten kan den konstante framvisningshastigheten for tekst i dagens systemer

erstattes med en hastighet avhengig av hvor teksten ligger og hva innholdet i teksten er.

Stasjoner

Stasjoner er skjermposisjoner som teksten vises i. Stasjoner kan være identifiserte vindu
5 som teksten vises inni eller punkter der tekstpresentasjonen begynner. En foretrukket
utførelse omfatter en aktiv tekststasjon der teksten som leses blir presentert. En annen
foretrukket utførelse omfatter en pre-lesestasjon som viser tekst som er klar til lesning og
en post-lesestasjon som viser allerede lest tekst. I en utførelsesform reguleres all tekst
presentert ved alle stasjoner på visningsflata av den forventede lese- eller uttaletid. I en
10 utførelse tillates bare omlag ett minutt med materiale på skjermen på noe tidspunkt.

Animasjon

Animasjon er den tidsavhengige presentasjon av tekst. Eksempler på animasjon
omfatter blinking, oppløsning og sveiping, det vil si opplysning av suksessive tekstdeler ved
15 en spesifikk hastighet. I en foretrukket utførelse gjennomføres en modifisert rulleeffekt ved
å omstokke tekst som skal leses til flate rader ved en pre-lesestasjon ved toppen av
skjermen, stable tidligere leste rader ved en post-lesestasjon ved bunnen av skjermen, og
kaskadere den aktive setningen ved en aktiv stasjon over midten av skjermen. I en
utførelsesform er det forskjell mellom teksten som leses nå, og tekst som skal leses, lest
20 tekst og/eller bakgrunnsfarge.

Merking

Merking omfatter angivelse av hvorvidt en setning skal merkes med en definisjon eller en
figur relatert til ordet. I en foretrukket utførelse spesifiserer et ordsett hvilke ord som skal
25 merkes med en definisjon. Når et merket ord velges av leseren, kan det vises en definisjon,
en figur eller annen elektronisk illustrasjon av ordet. I en foretrukket utførelse vil et merket
ord, når det blir valgt, vises et sprettopp-vindu som inneholder definisjonen eller figuren.
Foretrukne utvalgsmetoder omfatter bruk av ei mus og høyre Shift-tast.

30 Lese-hendelse

En lesehendelse omfatter inntreden og forsvinning av en setning og enhver ekstra tid før
inntreden av den neste setningen. En lesehendelse inneholder en rekke lese-øyeblikk. Et
leseøyeblikk er en frosset inntreden av tekst overfor leserens øye som prøver å forstå
teksten. Mange lesere vil i leseøyeblikket finne øyet sentrert på et uttrykk, der uttrykket er
35 alene på ei linje og leseren forstår uttrykket alene.

Lesehastighet

Presentasjonsintervallet er varigheten av en setnings-presentasjon på skjermen.

Presentasjonsintervallet kan være regulert av bruker eller bestemt automatisk.

Brukerstyring kan i en utførelse være ved hjelp av museklikk. I en annen utførelsesform

5 kan brukerstyring skje via deteksjon av øyebevegelse eller av endringer i hjernebølger.

Hjernebølge-deteksjon virker ved å skape en ny setningsforespørsel når hjernebølgene tilsvarende aktivitetendring i hjernebølger forbundet med fullført lesing av setningen.

Lesehastighet måles i tekstinnhold per tidsenhet. Slike tekstinnhold-enheter kan måles i lengdeenheter slik som antall ord, stavelser, uttaletid eller ethvert av målene for

10 uttrykkslengde eller -vekt eller enhver sum eller produkt av disse, som diskutert foran.

Tekstinnhold-enheter kan også måles i kompleksitets-enheter slik som de diskutert foran med hensyn til uttrykkskompleksitet eller -vekt eller enhver sum eller produkt av samme.

Tekstinnhold-enheter blir i en foretrukket utførelse beregnet for ei setning som

15 enheter med leserregulerte presentasjonsintervaller, og for å sette

tekstpresentasjonsintervallet for automatisk regulerte tekstpresentasjonsintervaller. På denne måten kan presentasjonsintervall gjøres innholds-avhengig og kan følge den sannsynlige leserens foretrukne presentasjonshastighet.

20 Spesielle hendelser

Spesielle hendelser er tidsavhengige hendelser som opptrer under lesing. Spesielle

hendelser omfatter et visuelt eller hørbart stikkord som gir et vedvarende signal. For

eksempel når 75 % av det innholdsavhengige presentasjonsintervallet har løpt ut, vil det

opptre et visuelt stikkord på skjermen. Spesielle hendelser kan arrangeres med et spesielt

25 hendelses-skript. Komplekse spesielle hendelser omfatter en serie av spesielle hendelser.

Visuell uttale

Visuell uttale omfatter tidsmessig alternering av farge- eller lysendringer for individuelle

uttrykk, ord og stavelser, men ikke nødvendigvis alle disse. En utførelsesform angir vekt på

30 lengre og mer vanskelige ord som krever lengre tid å uttale ved hjelp av utheving. Visuell

uttale er et eksempel på en kompleks spesiell hendelse.

Nedtrekking av kaskaden

I en utførelsesform vises tekst ved en pre-lesestasjon i en prosentandel av det

35 innholdsavhengige presentasjonsintervallet, for deretter å presenteres i en

setningskaskade i gruppering på ett tekstsegment eller ei linje. Nedtrekking av kaskaden er

et eksempel på en kompleks spesiell hendelse.

Ikke-linær tekst

Lineær tekst presenteres og leses fra begynnelse til slutt. Ei bok beregnet for å leses fra
 5 begynnelse til slutt er et vanlig eksempel på lineær tekst. Ikke-lineær tekst omfatter
 presentasjon og lesing av tekst i en rekkefølge forskjellig fra lineær. Et eksempel på ikke-
 lineær tekst er hypertext, der visse ord presenteres på en slik måte at de indikerer en link,
 f.eks. tekst innrammet i klammeparentes eller farget blå. Presentasjon av ikke-lineær tekst
 kan være både leserbestemt og automatisk. Et eksempel på dette er automatisk
 10 presentasjon av hypertext i en av første dybdeorden, etterfulgt av den første hypertext-
 tlinken, og ignorere sykluser. Et annet eksempel er en breddepresentasjon, som
 presenterer det første nivået av tekst, etterfulgt av eventuelle hypertext-linker fra
 hovednivået, etterfulgt av eventuelle hypertextlinker fra det første nivået, som fortsetter
 inntil en har nådd det laveste nivået eller leseren bryter inn. Nok et annet eksempel er
 15 klipping av enten dybde- eller bredde-presentasjonen i første nivå for å inkludere bare
 deler som har visse nøkkelord av interesse.

Magiske lesebriller

I en utførelsesform kan tekstpresentasjonen begynne ved å presentere vanlig uformatert
 20 tekst. Når leseren bestemmer seg for å se tekst presentert som forsterket tekst, vil leseren
 utpeke en del av teksten for forsterkning ved å plassere et symbol slik som et "magisk
 lesebrille-symbol" over en viss del av setningen. Denne setningen og etterfølgende
 setninger presenteres som forsterket tekst. Denne forsterkede tekstpresentasjonen
 fortsetter inntil leseren deaktiverer de magiske lesebrillene. Disse magiske lesebrillene kan
 25 tjene som en alternativ innfallsport til en forbedret tekstbehandler.

Framgangsmåte

I Figur 1, trinn 110, blir teksten pre-prosessert. Trinn 110 er vist i detalj i Figur 5. Teksten
 analyseres (grammatisk) for å identifisere avsnitt, setninger, ord og tegnsetting. Avsnitt
 30 kan identifiseres med blanke linjer, avsnittsmarkører, innrykkstegn, tabulator tegn eller
 enhver annen egnet karakteristikk i teksten. Setninger kan identifiseres ved bruk av
 grammatikalske regler inkludert punktum, avstand, stor forbokstav i første ord og
 forkortinger eller mangel på samme. I en foretrukket utførelse under lesing av en god tekst
 vil følgende signalisere slutten av ei setning: et punktum, et spørsmålstegn eller
 35 utropstegn, enten alene eller etterfulgt av et punktum, etterfulgt av to blanke eller
 avsnittsslutt.

I trinn 124 blir setningen splittet i ord og tegnsetting. Opprinnelig forfatter-spesifisert uthevning, f.eks. kursiv eller understreking, bevares i foretrukne utførelser. En vanlig leksikal scanner slik som Lex (varenavn) anvendes i en utførelsesform der slutten av et ord angis med grammatikalske regler med hvite blanke eller tegnsetting. En annen utførelsesform anvender en håndskrevet leksikalsk analysator. En utførelsesform lagrer formateringstegn slik som tabulatorer og innrykk som skilletegn. Posisjonen for et ord lagres fortrinnsvis som en attributt hos ordet for å tilby linker til og søking i det opprinnelige arbeidet.

En foretrukket utførelse tillater også at grupper av ord "klemmes sammen" og kan gjenkjennes som ei ordgruppe. I en utførelsesform gjenkjennes slike ordgrupper av den leksikale scanneren. I en annen utførelsesform gjenkjennes slike ord av en preprocessor som arbeider før den leksikale scanneren for å sikre gjenkjenning som et uttrykk i stedet for individuelle ord. Sammenklemt ord, for eksempel "prinsen av hvalene", kunne gjenkjennes som et enkelt uttrykk og blir fortrinnsvis ikke oppdelt i to uttrykk av preposisjonen og plassert på to linjer.

I trinn 126 slås ordene opp i ordbøker, glossar og tabeller for å bestemme ordattributter. Teksten behandles ytterligere for å bestemme kategoriske og kontinuerlige attributter. I en foretrukket utførelse omfatter viktige kategoriske attributter deler av uttale, og viktige kontinuerlige attributter omfatter ordposisjon, utdanningsnivå, uttaletid og antall stavelser, posisjon, lyd og vokalt uttrykksnivå. En identifisering av deler av uttalen med 100% nøyaktighet ville kreve omfattende programmering for å bestemme det reelle innholdet i teksten. Slik nøyaktighet er ikke påkrevd for å praktisere oppfinnelsen siden feil har mindre konsekvens fordi leseren er et menneske og ikke en maskin. De mulige delene av uttalen bestemmes først ved å slå opp ordet i ei ordbok eller glossar. Denne ordboka eller glossaret trenger bare ha de sannsynlige delene av uttalen for et ord, ikke en definisjon. For eksempel kan ordet "farm" på engelsk være et substantiv, et verb eller et adjektiv. En foretrukket utførelse lagrer delene av uttaleattributter ved bruk av et bitkart for å bevare de ulike mulige delene av uttale. En utførelse lagrer eksplisitt en tvetydighetsattributt, som indikerer hvorvidt ordet fremdeles har flere mulige uttaler. En annen utførelse bruker eksistensen av flere enn en enkelt mulig del av uttale som en tvetydighetsindikator.

I en foretrukket utførelse eksisterer standard-deler av uttale i ei ordbok og kan slås opp. I den mest foretrukne utførelsen kan en legge til et ordsett for å omdefinere eller supplere standardsettet. I en annen utførelsesform spesifiseres tekniske ord av brukerdefinerte ordsett.

Figur 6 illustrerer ni ordsett som spesifiserer deler av uttale arrangert i et hierarki. Slike ordsett og hierarkier brukes også som inngangsverdier til brytningsregler, beskrevet

nedenfor. En foretrukket utførelse sjekker ordsettene illustrert i Figur 6 ved å begynne med klasse 3, subklasse 1. Søk termineres straks et ord eller et uttrykk er funnet. Gitt det begrensede vokabularet og den statiske natur ved ordsettene i Figur 6, anvender den mest foretrukne utførelsen en rask håndskrevet analysator for å søke etter ordene og uttrykkene i Figur 6. Foretrukne utførelser omfatter analysatorer skrevet i C, C++, Perl, kompilert AWK, AWK og AWK til C, C med regulære uttrykksfunksjoner, eller ethvert annet egnet språk. En analysator-utførelse anvender YACC.

I en utførelsesform et ordboka ei kommersielt tilgjengelig ordbok på et elektronisk medium slik som CD-ROM. Standardordboka analyseres for ordattributter slik som deler av uttale og antall stavelser. Siden orddefinisjoner i mange utførelser ikke er påkrevd kan en lagre en rekke ord med ledsagende antall stavelser og deler av uttale. I den mest foretrukne utførelsen lagres de vanligste og sist brukte ordene i minne med kort aksesstid slik som statisk RAM. I utførelser der ordbøker skal lages for hånd foretrekkes en rask metode ved bruk av hashing, kollisjons-deteksjon og blokker. I utførelser der ordsettene er bestemt før lesing foretrekkes perfekt hashing uten blokker.

I nok en utførelsesform er nivået for uttalevekt avledet som en attributt dels avhengig av delen av uttale. I den mest foretrukne utførelsen er uttalevekt kategorisert som primær, sekundær og ingen. I en utførelsesform hentes også uttaletid og faktisk lyd, f.eks. fra ei lydfil, fra ordboka eller glossaret og lagres som attributter for ordet.

I trinn 128 søkes det i leser- og tekst-spesifikke ord. I en foretrukket utførelse lagres de lesterspesifikke ordsettene i en database atskilt fra ordboka eller glossaret, for å tillate separat lagring og raskere uthenting. I en utførelse sjekkes de lesterspesifikke ordsettene før ordbøkene, og ordbøkene sjekkes bare dersom nødvendige ord og attributter ikke finnes i de lesterspesifiserte ordsettene. Foretrukne utførelser anvender et hierarki med databaser for oppslag av ord.

I trinn 130 blir flere deler av språket gjort utvetydig. I en utførelse anvendes en mikroelementarbok for å bestemme de sannsynlige delene av språket. En mikroelementarbok anvender tilgrensende eller nærliggende ord for å gjøre en mer nøyaktig bestemmelse av de mest sannsynlige delene av språket for et ord. For eksempel vil ordet "farm" i uttrykket "en farm i" sannsynligvis være et substantiv siden det foranlediges av en artikkel og etterfølges av en preposisjon. Et annet eksempel: dersom et ord enten kan være et substantiv eller et verb og ordet innledes med "kunne", "vil", "skal" eller "til", er ordet sannsynligvis et verb. Dersom ordet "bygger" var foranlediget med "vi" er ordet sannsynligvis et verb. I en annen utførelse utføres all utvetydiggjøring ganske enkelt ved å velge den statistisk mest sannsynlige bruken av ordet. I nok en annen utførelse skjer det ingen automatisk utvetydiggjøring, bare manuell utvetydiggjøring ved menneskelig

editering. I en foretrukket utførelse lagres tvetydighetsattributten for hvert ord, som indikerer hvorvidt det fremdeles vil eksistere flere mulige varianter etter utvetydigging. I nok en utførelse lagres ingen tvetydighetsattributt men avledes fra eksistensen av flere mulige varianter lagret for et ord. I en utførelse konkluderes tvetydighet fra den visuelle skjermen med stripe- eller alternerende tekstfarge forbundet med hver del av språket. For eksempel dersom verb er oransje og adjektiv er gule, kan dermed et mulig verb eller adjektiv ha alternerende gule og oransje striper eller tekst-tegn.

I en utførelsesform av oppfinnelsen bestemmes deler av språket dels ved søking etter morfem (rot-ord) og tilordning av deler av språkattributter basert på endinger, f.eks. -lig, -ing.

I trinn 132 lagres attributtene definert for ordene og uttrykkene for å danne ei "anriket setning", som sannsynligvis vil forbli uendret leserne imellom. For eksempel vil utdanningsnivået og uttaledelen for et ord forbli uendret for ulike lesere, selv om de kan ønske andre tekstsegment-lengde og presentasjonshastighet. I en utførelse lagres den anrikede setningen i et varig lager slik som ei fil. I en annen utførelse lagres den anrikede teksten på CD-ROM. I en foretrukket utførelse implementeres den anrikede teksten som ei lenket liste med noder, der hver node har ord- og uttryks-attributtene beskrevet foran inkludert ordets posisjon i den opprinnelige teksten.

I Figur 5, trinn 134, bestemmes primære brytningspunkter ved å bruke primære brytningspunkt-regler. Brytningspunkter er tekstinn-delingspunkter lokalisert mellom bokstaver. I en foretrukket utførelse klassifiseres brytningspunkter som primære og sekundære. Primære brytningspunkter bestemmes ved bruk av primære brytningsregler som bestemmer primære brytningspunkt-posisjoner basert på skille-tegn. Figur 7 illustrerer et primært brytningspunkt etter kommaet etter "Afrika". Primære brytningspunkter deler teksten inn i "Superuttrykk". I en foretrukket utførelse er primære brytningspunkter lokalisert ved hvert komma, semikolon, venstreparentes, klammeparentes og krøllparentes. Brytningspunkt-posisjonen kan lagres som en attributt i en node i ei lenket liste med noder for å danne den anrikete teksten.

Sekundære brytningspunkter bestemmes ved bruk av sekundære brytningspunkt-regler. I foretrukne utførelser rangeres sekundære brytningspunkter og regler i et hierarki og sekundære brytningspunktregler godtar deler av uttrykk som inngangsverdi. I den mest foretrukne utførelsen omfatter som regel sekundære brytningspunktregler inngangsattributter for tekstinnholdet i tekstsegmentene og uttrykk blir prosessert. For eksempel kan det spørres etter et sekundært brytningspunkt av et tekstsegment som overstiger leserens maksimale foretrukne tekstsegmentvekt, selv om en ikke har kommet fram til en maksimal tekstsegmentlengde.

Kontinuerlige attributter slik som uttrykkets vanskelighetsgrad, densitet, kompleksitet, kraft og uttaletid kan anvendes som inngangsverdier til en regel for å modifisere rangeringen etablert av en tabell slik som den i Figur 6, ved bruk av deler av uttrykk alene for å bestemme sekundære brytningspunkt-rangeringer. For eksempel vil et tekstsegment med en vekt større enn 35 prosent over tekstmiddel være tilordnet en rangering av klasse 1 uansett rangeringen som det spørres etter med tabellen i Figur 6. I en foretrukket utførelse anvendes uttrykksvekt eller -kraft alene for å bestemme sekundære brytningspunkt-rangeringer i stedet for utelukkende deler av uttrykket.

I en alternativ utførelse spør brytningsregler etter brytning basert på antall tegn på linja, og delene av uttrykket vises ved bruk av farger som tilsvarer et ords uttaledel. Den siste utførelsen trenger nødvendigvis ikke gi fordelene med kaskadering, men tilbyr likevel visuelle stikkord basert på tekstinhold.

Figur 6 illustrerer en tabell som brukes i en foretrukket utførelse for å bestemme sekundære brytningspunkter. For eksempel bestemmer preposisjoner sekundære brytningspunkter av klasse 3 og subklasse 1. I Figur 7 finnes det sekundære brytningspunkter av klasse 3 før preposisjonene "i" og "av". Sekundære brytningspunkter deler superuttrykk inn i "miniuttrykk", som illustrert i Figur 7. Miniuttrykk er relatert til tekstsegmenter idet miniuttrykk ofte er det samme som tekstsegmentene og vises ofte hver på en ny linje. Straks miniuttrykkene er identifisert kan imidlertid regler bestemme at mer eller mindre enn et miniuttrykk skal opptre som et tekstsegment på ei linje.

Primære brytningsregler anvendes først, etterfulgt av sekundære brytningsregler, utført i henhold til rekkefølgen av rangeringen av brytningspunktregler. Et eksempel på rangering av sekundære brytningsregler er vist i Figur 6, etablert av rangeringen av ordsettene som er opphav til de sekundære brytningspunktene. Noen foretrukne utførelser bruker enten uttrykksvekt eller -kraft for å bestemme rangering av sekundære brytningspunkter i stedet for bare å bruke deler av uttrykket. En mest foretrukket utførelse tillater at leseren oppgir en preferanse for deler av uttrykket eller uttrykksvekt/-kraft-bestemmelse av rangering av sekundære brytningspunkter. Noen lesere foretrekker tekstsegmentering basert på struktur, mens andre foretrekker tekstsegmentering basert på kompleksitet eller estimert til for å lese et tekstsegment.

I en foretrukket utførelse anvendes sekundære brytningspunktregler bare inntil en kommer til ei grense. Denne grensa er ofte den minimale linjelengden. En framgangsmåte for å bestemme posisjonen for sekundære brytningspunkter er vist i Figur 5, trinn 136, og i detalj i Figur 8, trinn 136. I en utførelsesform der anvendelse av en sekundær brytnings-regel på et superuttrykk vil resultere i en miniuttrykkslengde som er mindre enn den spesifiserte minimale linjelengden, anvendes ikke brytningsregelen, og det anvendes ingen

andre brytningsregler på dette superuttrykket. På den andre siden, når det på andre måter ikke vil eksistere noen brytningspunkter i ei linje som overstiger den maksimale linjelengden, anvendes en sammenbruddsregel for å tvinge brytningen av teksten i to linjer. Når alle superuttrykkene ikke skal underlegges noen ytterligere brytningsregler er

5 brytningsprosessen ferdig.

I Figur 1, trinn 111, kan parametrene, attributtene og brytningsreglene anvendes som inngangsverdi til de horisontale forskyvningsreglene. De horisontale forskyvningsreglene bestemmer den horisontale posisjonen for tekstsegmentet. I en foretrukket utførelse omfatter horisontale forskyvningsregler både regler for horisontal justering og regler for nedstigningsvinkel. Regler for horisontal forskyvning i denne utførelsen er summen av resultatene fra den horisontale justeringsregelen og regelen for nedstigningsvinkel. I en utførelse som er lett å implementere er ganske enkelt regelen for horisontal forskyvning nedstigningsvinkelen anvendt på det sentrerte tekstsegmentet. En slik utførelse bruker ikke brytningspunktregelen og terminerer de forutgående tekstsegmentene som inngangs-

10 verdier og gir minimal øyebevegelse under lesing av setningskaskaden. En annen utførelse legger til venstre nedstigning for forutgående brytningspunkter av klasse 1, og høyre nedstigning for brytningspunkter av klasse 3. En foretrukket utførelse tillater leserspesifiserte ekstra høyre- eller venstre-forskyvning for brytningspunkter, inkludert leserspesifiserte verdier for primære brytningspunkter, og hver klasse og subklasse for

15 sekundære brytningspunkter. En utførelse lagrer den adderte forskyvning i en tabell i tegn-enheter. Når den horisontale forskyvningen er bestemt gjenstår bare presentasjonen av teksten.

I eksemplet i Figur 9, superuttrykk 1, "Jeg hadde en farm i Afrika" er inndelt fra superuttrykk 2, "ved foten av Ngong-åsene", av det primære brytningspunktet etter komma.

25 Superuttrykk 2 er inndelt i to miniuttrykk av det sekundære brytningspunktet før preposisjonen "av". I utførelsen illustrert i Figur 9 er to miniuttrykk presentert som det samme tekstsegmentet som miniuttrykket "i Afrika" siden miniuttrykket "i Afrika" er under minimumslengden på tre ord spesifisert av leseren. Det første miniuttrykket "ved foten" i ei ny linje er sentrert under tekstsegmentet over og forskjøvet mot høyre av en høyrerettet nedstigningsvinkel. Det siste miniuttrykket "av Ngong-åsene" er forskjøvet mot venstre på grunn av en horisontal forskyvningsregel som medfører venstreforskyvning når den forutgående linja avsluttes med et sekundært brytningspunkt i stedet for et primært ett. I eksemplet i Figur 9 er følgelig den totale horisontale forskyvning bestemt av en

30 kombinasjon av nedstigningsvinkelen og justeringa av linja.

35 Parvise skilletegn inkludert parenteser, klammeparenteser og anførselstegn tjener som skilletegn som bestemmer primære brytningspunkt-posisjoner i en foretrukket utførelse. I

èn utførelse er slike parvise skilletegn inkludert som en inngangsverdi til de horisontale forskyvningsreglene, inkludert horisontal justering og regler for nedstigningsvinkel. For eksempel kan et langt tekstsegment med flere uttrykk opptatt innenfor parenteser ha redusert horisontale forskyvninger og redusert vertikal forskyvning, dvs. mindre linjeavstand enn den andre teksten i det aktive visningsområdet på visningsflata.

I trinn 112 dannes den anrikete teksten. I dette trinnet lages kodene som behøves for å etablere et skikkelig tekstsegment for visning. For eksempel når leserspesifikasjoner krever rødfarging av tekniske ord og den anrikete teksten indikerer et teknisk ord, kan det opprettes en escape-sekvens som i visningstrinnet vil tolkes til å kreve rød tekst.

10 Tilsvarende koding kan kreves for animasjon. Den anrikete teksten kan på dette punktet lagres for senere visning.

I trinn 113 presenteres den anrikete teksten på visningsorganet - ett tekstsegment per nydannet linje. Den anrikete teksten kan også omfatte animasjonen, bakgrunnsfargen, tekstfargen, merking og presentasjonshastighet som diskutert foran. I en foretrukket utførelse vises bakgrunnsfargen som en funksjon av setningens og avsnittets posisjoner.

I en annen utførelse, illustrert i Figur 11, presenteres initialt noe tekst i én farge eller lysstyrke i et tidsintervall etterfulgt av presentasjon med en andre farge eller lysstyrke. I Figur 11 presenteres setningen på tre nydannede linjer indikert med piler 200, 202 og 203. Innenfor disse linjene presenteres all tekst på samme linje, der ordene i Figur 11 vises på ulike linjer for å illustrere den midlertidige endring i initial tekstfarge/belysning 206 til tekstfarge/belysning 207, og initial tekstfarge/belysning 208 til 209. Teksten "Af" ved 206 er den uthevede stavelsen i "Afrika", og er av den grunn initialt vist med farge/belysning 206- Teksten "gong" er også den uthevede stavelsen i "Ngong", og har initial farge/belysning 208, etterfulgt av farge/belysning 209.

25 Den foretrukne utførelsen skissert i Figur 1 tillater at den anrikete teksten fra trinn 132 og forsterket tekst fra trinn 113 editeres. I en foretrukket utførelse brukes en ordpeke-anordning slik som ei mus for å velge en del av teksten som skal editeres. Den anrikete setningen som tilsvarer denne delen av anriket tekst velges internt, og attributtene tilgjengeliggjøres for editering. Dette kan implementeres ved bruk av nedtrekksmenyer som tillater at hver attributt i den anrikete teksten og forsterket tekst kan editeres. Der editering av den anrikete teksten kan påvirke presentasjonen blir teksten preprosessert på nytt i henhold til trinn 110. I editerings-sesjonen i trinn 114 aksepteres leser-anførsler, enten skjulte eller umiddelbart synlige, og kan lagres i de anrikete og forsterkede setningene. Slike anførsler utøver samme funksjon som håndskrevne anførsler i ei bok. De editerbare trekkene omfatter deler av språket, definisjon, farge, tekst, brytningspunkter og horisontal forskyvning.

De editerte attributtene samt visningen i denne utførelsen kan også, i trinn 116, lagres for å bevare endringene. I en utførelse lagres bare de forsterkede setningene. I en foretrukket utførelse lagres både den anrikete teksten og forsterket tekst. I foretrukne utførelser registreres noen editeringer som enten leser-spesifikke eller globale, for presentasjon til alle lesere av teksten. Manuell editering av den forsterkede teksten er særlig nyttig der samme teksten vil bli lest gjentatte ganger av andre, som med ei elektronisk bok.

I en variasjon av utførelsen i Figur 1, er trinnene 113, 114 og 116 utelatt, uten noen menneskelige inngrep og umiddelbar visning. I denne utførelsen dannes anriket og forsterket tekst og lagres for framtidig visning. Den forsterkede teksten kan lagres i standard tekstbehandlingsformat slik som Microsoft Word og Corel Word Perfect. I denne utførelsen kan presentasjons-programvaren være enkel, liten og rask i forhold til programvaren som kreves for oppslag av ord og tekstanalyse. En slik presentasjon er en foretrukket utførelse for massedistribusjon av forsterket tekst for lesing slik som "elektroniske bøker".

I en relatert variasjon av utførelsen i Figur 1 hentes den forsterket teksten i seg selv i trinn 106, uten noen krav om pre-prosessering og det tillates ingen editering. En slik presentasjon er også en foretrukket utførelse for massedistribusjon av forsterket tekst for lesing som "elektroniske bøker".

Med henvisning til Figur 1, trinn 113, kan presentasjonen av hver skjerm med forsterket tekst utøses manuelt som med et museklikk. I en foretrukket utførelse reguleres presentasjonshastigheten av leser-spesifiserte parametere inkludert visningstid for teksten, og ankomsttider mellom setninger og mellom avsnitt. I den mest foretrukne utførelsen måles tekstinnholdet, og tekstpresentasjonsintervallet er avhengig av dette tekstinnholdet. I en foretrukket utførelse anvendes uttaletiden for teksten som et mål på tekstinnhold for å bestemme presentasjonsintervallet. I en annen utførelse anvendes uttrykksvekt for å måle tekstinnhold og bestemme presentasjonsintervall.

En foretrukket utførelse tillater at lesere kan forlenge presentasjonstiden for ei setning samt akselerere og bremse presentasjonshastigheten. Denne hastigheten kan måles i likhet med setningslengden og vanskelighetsgraden for setningen i samsvar med et spesielt presentasjonsintervall. Figur 10 illustrerer et eksempel på et tredimensjonalt plott - et "Historiekart for Presentasjonsintervall", som skisserer tiden som kreves for å lese ei setning som en funksjon av to attributter, setningens vanskelighetsgrad og setningslengde. Tiden som kreves for å lese kan måles med leser-preferanser eller med sporing av øyebevegelse som angitt foran samt hjernebølge-aktivitet. Tiden kan dermed korreleres mot to attributter som i Figur 10, eller ethvert annet antall attributter. Presentasjonsintervallet justeres i en foretrukket utførelse kontinuerlig for å tilsvare den forutsagte

påkrevde lesetiden.

I nok en utførelsesform, som er mest egnet der lesehastigheten måles nøyaktig, varieres inngangsverdiene til brytningsreglene og den resulterende lesehastigheten spores. I en utførelse egnet for dannelse av et optimalt leseprodukt for det store markedet varieres
5 inngangsverdiene for brytningsreglene, og lesehastigheter registreres for en testgruppe. Inngangsverdiene varieres for optimal lesehastighet og forståelse. For eksempel varieres betydningen av sekundære brytningsregler for deler av språket i forhold til uttrykksvekt med det formål å oppnå optimal lesehastighet.

I en annen utførelse analyseres initial setningsvisning som er editert manuelt av leser
10 med hensyn til relativ fordeling av setningsstruktur, deler av språk, lengde og kompleksitet. Etter en slik initial "avstemnings-" eller "trenings-" periode anvendes tilsvarende vektlegging i brytningsregler for automatisk tekstforbedring. Påfølgende editeringer kan anvendes for å raffinere brytningsreglene ytterligere.

I trinn 118 stopper tekstpresentasjonen når det ikke lenger finnes mer tekst eller leseren
15 har bedt om å stoppe. Trinnene 120 sørger for opprydding inkludert lagring av ferdig anriket tekst og forsterket tekst samt lagring av historikk om lese-sesjonen.

Fremgangsmåteeksempel

20 Sammen drag av eksemplet

Den foreliggende oppfinnelsen omfatter en framgangsmåte for brytning og forskyvning av tekst som oppfinnerne anser som en forbedring framfor framgangsmåten beskrevet under overskriften "framgangsmåte" foran.

25 En utførelse av oppfinnelsen klassifiserer skilletegn enten som "serielle" eller som "parentes-dannende". Serielle skilletegn omfatter kolon, semikolon, komma, en enkelt dobbeltstrek "--", og ellipse. Parentes-dannende eller nøstende skilletegn omfatter anførselstegn, parentes-merker, klammer (som i [og]), og dobbeltstreker. Den foreliggende oppfinnelsen arbeider best når klamme-segmenter ikke begynner inne i et klamme-segment og slutter inne i et annet. Dette er normalt tilfellet.

30 Segmenter med parentes-dannende skilletegn kan inndeles serielt med serielle skilletegn og innkapsles ytterligere med indre parentes-skilletegn, eller begge. Det finnes ett tilfelle der to komma i et segment produserer en type parentes: det midtre segmentet av de tre nye segmentene som er dannet er "innkapslet" av de andre to segmentene. Dette er vanlig og kan skille en relativ setning fra en ikke-restriktiv. Dette er også tilfellet for "for
35 eksempel" og andre uttrykk av denne typen er innleiret i større uttrykk. En del innkapslet tekst (i parenteser), f.eks. tekst innenfor anførselstegn, kan spenne over flere setninger.

I framgangsmåten beskrevet nedenfor blir innkapslete segmenter identifisert. Teksten utenfor det innkapslete segmentet undersøkes med hensyn til serielle skilletegn. Disse anvendes først for brytning. Deretter skilles ikke-innkapslet tekst og innkapslet tekst med brytningstrinn som plasserer hver teksttype i atskilte rader. Innkapslet tekst plasseres i sin egen rad med innkapslings-skilletegnene helt ved begynnelsen og slutten av hvert segment. Denne syklusen gjentas etter behov. Et segment er klart til å forlate algoritmedelen som arbeider med "skilletegn-brytning" og begynner med "brytning av ord alene" (ordbasert brytning), når den ikke inneholder andre skilletegn enn de som finnes ved begynnelsen og slutten. Trinn 1 til 15 i algoritmen under arbeider med skilletegn-brytning.

Ordbasert brytning begynner med trinn 16 og vil bryte hvert segment slik at det passer inn mellom maksimums- og minimums-målene. Brytningspunkter i teksten vil være spesifisert av "ledd-ord". Ledd-ordene kan spesifiseres av ei ordliste. I en foretrukket utførelse er ord eller uttrykk i ordlista forbundet med en klassifiseringskode eller kodeord som omfatter informasjon om ordet. I eksemplet illustrert nedenfor omfatter klassifiseringen et tegn-prefiks og en tallrangering eller nivånummer. Tegnprefiksene omfatter "b" (bryt før dette ordet, ikke etter), "a" (reservert for framtidig bruk, bryt etter dette ordet), "x" (definitiv subjektiv pronomens-liste), "y" (forut-bestemmende liste), "z" (etter-bestemmende liste). Eksempler på ord og assosierte koder er oppført i Appendix 1.

I eksemplet nedenfor er nivå-tallene rangert fra 0 til 99. De laveste nivå-verdiene har høyere rangering og håndteres først av algoritmen. Variabelen "n" anvendes til angivelse av denne verdien i trinnene beskrevet nedenfor. Området 10-39 omfatter leddordene "b". Området 40-46 omfatter verb, inkludert 40 (modalt), 41 (har, hjelpeverb), 42 ("å være" - verb), 43 (utvetydige transitive verb), 44 (utvetydige intransitive verb), 45 (tvetydige transitive verb, ord som ikke alltid er verb), og 46 (tvetydige intransitive verb, ord som ikke alltid er verb).

Tegn-prefiks relatert til verb omfatter "c" (infinitiv), "d" (perfektum partisipp), "e" (andre former som ikke tar hjelpeverb), "f" (kan være både infinitiv og perfektum partisipp), og "g" (verbalsubstantiv). Klassifiseringen av verb brukes i en forbedret verb-håndteringsmetode beskrevet i et annet avsnitt.

"Gate" viser til en enhet med horisontal forskyvning som underinndeles i en under-enhet med horisontal forskyvning benevnt som "spor". I en foretrukket utførelse er hver gate 4 blanke tegn bred, underinndelt i 4 spor som hver er 1 blankt tegn bredt. Gatene er nummerert 1 til 5 og sporene merket A, B, C og D. I en foretrukket utførelse begynner gate 1, spor A 2 blanke fra venstremargen, etterfulgt av spor B, C og deretter D, etterfulgt av gate 2, spor A, osv.

"X" (stor X) betegner generasjonstallet, dvs. rekursjonsnivået i algoritmen. Ei setning

deles for eksempel inn i et segment og et segment 2, der segment 2 deles ytterligere i segment 2.1 og 2.2, og segment 2.1 deles ytterligere inn i segmenter 2.1.1 og 2.1.2. Segmentene 2.1.1 og 2.1.2 er i generasjon 3. X anvendes ofte til bestemmelse av den grove horisontale forskyvning, dvs. gatetallet.

- 5 "x" (liten x) betegner indeksverdien eller avviket inn i setningen eller setningsfragmentet ved et spesielt ord. x økes/redueres i ord, ikke tegn.

Betegnelsen "n" viser til leddord-nivået. I eksemplet nedenfor varierer leddord-nivået fra 1 til 99.

- 10 Med henvisning til Figur 12 omfatter en tabell 320 ei tittelrad 322 som ramser opp blanke fra venstremarg for hver underliggende kolonne. Ei tittelkolonne 324 inneholder ei liste med leddordnivå. Radene inndeles i en serie gater 328 som hver er underinndelt i 4 spor. En serie X'er, for eksempel ved 330, betegner startposisjonen for et tekstsegment, gitt leddordnivået og gata.

- 15 X'ene i tabellen viser gate og spor som et segments første tegn vil plasseres i som en funksjon av leddordet som er brukt til å lage brytningen som produserte segmentet. Hvert leddordnivå, f.eks. b10, b20, kan plasseres i ei vilkårlig gate fra gate 2 og oppover. Gata som brukes bestemmes av hvilken generasjon setningsbrytning gjorde brytning for segmentet opptrer i. Innenfor denne gata (f.eks. gate 3 for en tredjegerasjons brytning) bestemmes deretter et spor av leddord-nivået, som vist i tabellen. På denne måten vil en
20 tredjegerasjons brytning laget med et b20-nivå ord plasseres i gate 3 spor B, og en tredjegerasjons brytning laget med et b70-ord vil plasseres i gate 3 spor D.

- Nivåene vist i tabellen ved nivå b85 vil oppvise brytning BÅDE før OG etter det aktuelle ordet – ved å plassere ordet i seg selv på sin egen linje. Instruksjonene på tabelltilstanden fastslår hvor b85-ordet bør plasseres og hvor segmentet som etterfølger b85-ordet bør
25 plasseres. Kriteriene for et b85-ord er ikke i leddordlista men har i stedet blitt definert som etvært ord som overstiger 12 tegn. Som posisjonen i leddordlista medfører anvendes b85-ord til brytning bare dersom det ikke finnes alternativer lengre opp i lista i segmentet som undersøkes for brytning. Det er ikke beskrevet eksplisitt andre steder i algoritmen.

- 30 Framgangsmåte

Den rekursive algoritmen startes ved å sette $X=1$. På denne måten vil i begynnelsen den første gata benyttes, og algoritme-generasjonen er 1.

Skilletegn-brytning

- 35 1. I tillegg til skilletegn som kan være tilstede helt ved slutten eller begynnelsen av ei setning/segment, finnes det andre skilletegn i setningen/segmentet? Ja: gå til trinn 2, Nei:

gå til trinn 16.

2. Finnes det parentesdannende skilletegn i setningen/segmentet? Ja: gå til trinn 3, Nei: gå til trinn 4.

3. Finnes det serielle skilletegn på utsiden av den (i parenteser) innkapslete tekst i denne setningen/segmentet? (Dersom ei setning begynner uten et anførselstegn men har det andre medlemmet av et par anførselstegn senere i setningen [dvs. at anførsel begynte i ei tidligere setning og slutter i denne], er segmentet fra begynnelsen av setningen til mellomrommet som etterfølger det andre medlemmet i paret også et "innkapslet" segment; en tilsvarende regel gjelder dersom ei setning inneholder det første men ikke det andre av anførselsmerke-paret.) Ja: gå til trinn 4. Nei: gå til trinn 9.

4. Finnes det kolon/semikolon i den ikke-innkapslete tekst fra trinn 3? Ja: gå til trinn 5, Nei gå til trinn 6.

5. Bryt setningen/segmentet etter hvert kolon/semikolon og plasser teksten som etterfølger hvert kolon/semikolon (som kan inneholde innkapslet tekst) på ei ny rad og ordne hver ny rad ved posisjon Gate X, Spor A - gå deretter til trinn 6.

6. Finnes det komma i den ikke-innkapslete tekst? Ja: gå til trinn 7. Nei: gå til trinn 8.

7. Bryt setningen/segmentet etter hvert komma og plasser teksten som etterfølger hvert komma (som kan inneholde innkapslet tekst) på ei ny rad; ordne den nye raden med tekst som kommer etter det første komma ved Gate X, Spor D; ordne den nye raden med tekst som etterfølger det andre komma ved Gate X, Spor B; ordne eventuelle ekstra rader med tekst som etterfølger oddetalls komma ved Gate X, Spor D; ordne eventuelle ekstra rader med tekst som etterfølger liketalls komma ved Gate X, Spor B. Gå deretter til trinn 15.

8. Finnes det noen skilletegn av typen enkle "dobbeltstreker" (--), eller ellipser (. ..) i den ikke-innkapslete teksten? Ja - bryt setningen/segmentet umiddelbart før hver enkel "dobbeltstrek" eller ellipse: plasser hvert nytt segment som inneholder den enkle "dobbeltstrek" eller ellipse og tekst fjerntliggende til samme (som også kan inneholde innkapslet tekst) på ei ny rad; ordne hver ny rad ved Gate X, Spor C. Gå deretter til trinn 15.

9. Finnes det noe tekst mellom begynnelsen av setningen/segmentet (som i seg selv kan være et parentes-dannende skilletegn) og det første "intra-setning/segment"-parentesdannende skilletegn? Ja: gå til trinn 10, Nei - hold det første parentesdannende skilletegnet ved samme posisjon (ikke bryt eller ordne ved dette skilletegnet; for eksempel kan et parentesmerke føre til og konkludere ei setning eller segment men etterfølges umiddelbart av et anførselstegn - disse ville dermed etterlates samlet ifølge dette trinnet); og gå til trinn 11.

10. Bryt setningen/segmentet umiddelbart før det første parentesdannende skilletegnet

(som ifølge trinn 9 ikke har tekst i umiddelbar nærhet); plasser dette første "intra-setnings/segments" parentesdannende skilletegnet og all tekst fjernliggende fra samme på ei ny rad; ordne denne nye raden ved Gate X, Spor D. Gå deretter til trinn 11.

11. Finnes det noe sammenfallende parentesdannende skilletegn i setningen/segmentet, (dvs. det andre medlemmet av det parentesdannende paret)? Ja: gå til trinn 12, Nei: da finnes det andre medlemmet av det parentesdannende skilletegnet i en etterfølgende setning; gå til trinn 15.

12. Finnes det noe tekst fjernliggende fra det andre medlemmet (som i trinn 11) av det parentesdannende skilletegnparet? Ja: gå til trinn 13. Nei: det andre medlemmet av det parentes-dannende skilletegnparet er ved slutten av denne setningen/segmentet; gå til trinn 15.

13. Bryt setningen/segmentet umiddelbart etter dette andre medlemmet av det parentesdannende paret; ordne den nye raden med tekst ved posisjonen Gate X, Spor B; gå til trinn 14.

14. Finnes det andre parentesdannende skilletegn i teksten fjernliggende fra det andre medlemmet av det forutgående parentesdannende paret? Ja: gå til trinn 10 og gjenta trinn 10-14 for dette neste innkapslete segmentet fjernliggende fra (ikke nøstet i) det forrige. Nei: gå til trinn 15.

15. Ta hvert av de nye segmentene dannet i trinn 1 til 14, addér 1 til X og gå tilbake til trinn 1. Gjenta brytnings- og ordnings-instruksjonene ved bruk av Gate X + 1. Gjenta trinn 1 til 15 inntil alle segmentene har skilletegn lokalisert bare ved begynnelsen og/eller slutten av hvert segment. På denne måten blir til slutt alle segmentene sendt til trinn 16 for "ordbasert" brytning alene.

25 Ordbasert brytning

16. Ordbasert brytning begynner med dette trinnet. For enhver verdi X ved trinn 1 der svaret er "nei" og segmentet sendes til trinn 16, bruk den største av: 2 eller X for ei gate for å begynne den ordbaserte brytningen. For eksempel: dersom ei setning som går gjennom syklusen for første gang, (der $X=1$), ikke hadde noen skilletegn mellom dens begynnelse og slutt, og svarer i trinn 1 var "nei"; ikke start ordbasert brytning med X (som er lik 1) men med 2 (som er større), dvs. i gate 2. Dersom det imidlertid var påkrevd med to sykluser gjennom trinn 1-15 for å oppnå et segment klart for ordbasert brytning, deretter, ved begynnelsen av den tredje syklusen, der X nå er lik 3, og svaret til trinn 1 er "nei", så begynner ordbasert brytning i trinn 16 ved gate $X=3$. Dette vil plassere ordbaserte brytninger ei gate til høyre for den siste gata som ble brukt for skilletegnbrytning. Etter bestemmelse av hvilken gate ei setnings/segments førstegenerasjons ordbasert brytning

vil forsynes med merke, gå til trinn 17.

17. Overstiger setningen eller segmentet den maksimale radverdien (f.eks. antall ord eller tegn i segmentet), eller overstiger den tilgjengelig plass til høyre? Nei: stopp, Ja, gå til trinn 18.

5 18. Finnes det ord i segmentet som kan finnes i leddordlista ved nivå $[n]$ (f.eks. nivå b10 når $n=10$)? Nei: gå til 18A, Ja: gå til trinn 19.

18A. Gjenta spørsmålet for leddordlistenivå $[n+1]$ (f.eks. nivå b20, deretter b30 osv.); dersom det ikke finnes ord i segmentet som kan finnes ved ethvert nivå i leddordlista, opp til $n=95$, gå deretter til trinn 29. NB: ord merket 99 med prefiks w, x, y og z brukes i

10 verbløkka, men anvendes ikke direkte som leddord til brytning.

19. Finnes det flere enn et ord i segmentet ved leddordnivå $[n]$? Ja: bruk det første ordet som opptrer i sekvensen i segmentet og gå til trinn 20. Nei: gå til trinn 21.

20. Overstiger segmentet dannet med bruk av ordet i trinn 19 den minimale radverdi? Nei: gjenta spørsmålet ved bruk av det neste ordet i sekvensen av segmentet som er på samme leddordnivå $[n]$; dersom ingen av ordene ved leddordnivå $[n]$ danner et segment som overstiger den minimale radverdi - gå tilbake til trinn 19 og gjenta for $[n+1]$. Ja: gå til trinn 21.

21. Bryt setningen ved plassen angitt av ordet (før dersom prefiks b, etter dersom prefiks a) eller etter), og plasser det nydannete segmentet i et av sporene i gate X, eller gate 2, hvilken av disse som måtte være størst. Sporet som skal anvendes i denne gata vil bestemmes av leddordnivået (se Figur 12 beskrevet nedenfor). Gå deretter til trinn 22.

22. Overstiger segmentet den maksimale radverdien eller forbi den tilgjengelige plass til høyre? Nei: stopp, Ja: gå til trinn 23.

23. Finnes det fremdeles nivå som skal sjekkes i leddordlista? Sjekk opp til $n=95$. Ord i ordlista med kodeord med $n=99$ f.eks. w99, x99, y99, z99 anvendes i verbløkka beskrevet i en etterfølgende seksjon men anvendes ikke direkte som leddord for til brytning. Ja: gå til trinn 24, Nei: gå til trinn 29.

24. Finnes det ord i segmentet som kan finnes i leddordlista ved nivå $[n]$ (f.eks. nivå b10 når $n=10$)? Nei: gjenta spørsmålet for leddordlistenivå $[n+1]$. Dersom det ikke finnes ord i segmentet som kan finnes ved noe nivå i leddordlista - gå til trinn 29. Ja: gå til trinn 25.

25. Finnes det flere enn ett ord i segmentet ved leddordnivå $[n]$? Ja: bruk det første ordet som opptrer i sekvensen for segmentet og gå til trinn 26. Nei: gå til trinn 27.

26. Overstiger segmentet dannet med bruk av ordet i trinn 25 den minimale radverdien? Nei: gjenta spørsmålet ved bruk av det neste ordet i sekvensen for segmentet som er ved samme leddordnivå $[n]$; dersom ingen av ordene ved leddordnivå $[n]$ danner et segment som overstiger den minimale radverdien - gå tilbake til trinn 24 og gjenta for $[n+1]$. Ja: gå til

trinn 27.

27. Bryt setningen ved plassen utpekt av ordet (før dersom prefiks b, etter dersom prefiks a), og plasser det nydannete segmentet i gate X+1, eller gate 2+1, hvilken av disse som måtte være størst (dvs. ei gate til høyre for den som er brukt for merking i den forrige ordbaserte brytningsgenerasjon). Overstiger det dannete segmentet den maksimale radverdi eller rager forbi den tilgjengelige plassen til høyre? Nei: stopp, Ja: gå til trinn 28.

28. Finnes det fremdeles nivå som skal sjekkes i leddordlista? Ja: gå tilbake til trinn 23 og utfør enhver mulig brytning ved bruk av gate X+x, eller gate 2+x, merker fra tabellen i Figur 12. Nei: gå til trinn 29.

29. "Bøy" segmentet i to eller flere nydannete segmenter slik at disse nydannete segmentene er mindre enn den maksimale radverdien og plassen tilgjengelig til høyre, og plasser disse nydannete segmentene i en sentrert posisjon under det første segmentet i gruppa.

Tilleggsbrytning for verb.

Denne seksjonen beskriver en tilleggsprosess for tekstbrytning som involverer verb på norsk. Tilleggsprosessen passer inn i den "forbedrede brytnings"-algoritmen i trinn 18, som beskrevet nedenfor. Tilleggsprosessen er ikke påkrevd for å utøve oppfinnelsen men tilbyr ekstra brytningsprosessering for verb som beskrevet nedenfor.

Det er ønskelig at en tekstbrytningsprosess unngår brytning av en streng med sammenhørende verb som bør opptre i samme segment, og at brytning opptrer ved verb når et diskret verbuttrykk er identifisert; dette krever ikke at et verb nødvendigvis blir plassert ved begynnelsen av ei ny rad når det er identifisert, identifisering av ei egnet uttrykksgrense som involverer ett eller flere verb kan føre til at det dannes ei ny rad etter verbet (på engelsk).

Det er også ønskelig at en tekstbrytningsprosess: unngår oppdeling av ordet "å" fra infinitiv-formen (og bare infinitivformen) fra et etterfølgende verb; tilsvarende sammenkople, dersom de kommer foran et verb, ordene "og" og "eller" med verb identifisert for brytning (alle verb-formene, ikke bare infinitivformene); å prosessere spesifikke verb+preposisjon (eller verb+adverb) - kombinasjoner som ett ord slik at etterfølgende brytning vil unngå atskillelse av disse preposisjonene eller adverbene fra kombinasjonen de hører sammen i (f.eks. "de brøt opp fra selskapet" mot "han spaserte oppover åsen"); og tilsvarende prosessere "ikke" - formene av hjelpeverb som ett enkelt ord slik at de kan gjenkjennes som ett enkelt ord ved posisjoner relativt til andre verb (f.eks.

"kunne ikke gå" bør behandles på samme måte som "kunne gå"). Det er også ønskelig at pronomensammenslåinger av hjelpeverb merkes som verb (f.eks. engelsk "you'll", "we've",

"you're").

Verb er komplekse. De kan være hjelpeverb, modale eller ha en ordinær funksjon. De kan være regulære eller irregulære i konjugasjon. De kan være transitive, intransitive eller heller deres evne til å ta et objekt. De kan være tvetydige med hensyn til hvorvidt de virkelig er verb og ikke andre deler av språket. De kan kombineres med spesifikke preposisjoner eller adverb for å bety noe annet enn verbet alene. De har flere former som kan eller ikke kan kombineres med hjelpeverb for å danne ulike tider og som retter seg etter personen og subjektets kasus.

På det nåværende tidspunkt er det ikke praktisk å anwise kaskadeformatering ved bruk av alle disse variablene i alle sine kombinasjoner. Det er imidlertid mulig å klassifisere verb i fire hovedfunksjonsgrupper (modale, har, er, og ordinære); å underklassifisere de ordinære verbene i utvetydige hhv. tvetydige og transitive hhv. sterkt intransitive; og for hver av disse gruppene identifisere opptil fem ulike former (infinitiv, perfektum partisipp, verbalsubstantiv, andre, og kombinerte infinitiv perfektum partisipp).

15

Definisjoner for verbbrytningsmetode

Tegnet "x" (liten x) betegner indeksverdien i setningen eller setningsfragmentet ved et spesielt ord. Tegnet x økes/redueres i ord, ikke tegn.

Tegnet "n" viser i verbløkke-algoritmen til ordet i tekstposisjonen relativt til ordet som undersøkes. Variabelen "n" viser ikke til leddordnivå som i den forrige algoritmen.

20

Introduksjon

For hver av disse funksjonsform-verbkategoriene finnes det basisoperasjoner som vil bestemme om brytning skal finne sted ved et verb, x, (som ellers ble identifisert for brytning med MIN og MAKS - kriterier). Disse operasjonene ser på ordet umiddelbart før (x-1) det aktuelle verbet og bestemmer om det tilhører et sett av ord som er spesifikt definert for denne funksjonsform-gruppa. Basert på dette blir enten den potensielle brytningen kansellert, (og det neste verbet i segmentet som er en kandidat for brytning ved dette leddnivået blir tatt i betraktning), eller den potensielle brytningen holdes åpen til nærmere undersøkelse avhengig av resultatene fra (x+1)-operasjonene.

30

Innholdet av disse settene for (x-1)-operasjonen vil omfatte: ord som vil utvetydiggjøre et verb-substantiv til et substantiv; og ord som er egnete hjelpeverb for det aktuelle verbet. I en påfølgende operasjon vil en potensiell brytning før x utsettes dersom x+1 er et verb. Verbet x+1 undersøkes for å se om ordene i forkant (nemlig x) er i lista si; dersom x ikke er i lista for x+1 utføres brytning før x+1 i stedet for før x (fordi x+1 er et verb som tilhører et separat verbuttrykk). Dersom x er i lista for x+1 - verbet (dvs. at x er et hjelpeverb for x+1)

35

utføres IKKE brytning ved $x+1$, og ordet etter $x+1$ (dvs. $x+2$) undersøkes. Dersom ordet etter $x+1$ ikke er et verb blir den opprinnelige utsatte brytningen for x nå utført.

Denne prosessen nyttiggjør retningsorienteringen i brytningsprosessen: den eksisterende brytningsprosedyren flytter forespørselen for hvert potensielt verb fra

5 begynnelse til slutt i setningssekvensen. Dersom to verb ligger inntil hverandre utføres $x-1$ - operasjonen for det andre verbet i det tilgrensende paret av verb (en operasjon som betrakter det første verbet i paret for å se om det finnes i ei liste spesifikk for for-gruppa for det andre verbet), som vil bestemme om en brytning skal finne sted mellom det første og andre verbet.

10 En $x+1$ - operasjon er påkrevd for hvert verb selv om det er det første ordet forbi den minimale rad-verdi. Årsaken til dette er at det er mulig at hjelpeverbene for et aktuelt verb x ble ført over i tidligere potensielle brytninger fordi de faller innenfor den minimale radverdi. Ved å se tilbake mot begynnelsen av segmentet, vil imidlertid $x-1$ - operasjonen detektere hvorvidt et hjelpeverb for x virkelig opptrådte i forkant (selv om $x-1$ er innenfor den

15 minimale radverdi). I dette tilfellet utføres ikke brytningen ved x .

Selv om x etterfølges av et annet verb i dette tilfellet (eller ethvert tilfelle der $x-1$ er et hjelpeverb for x), vil dette verbet som etterfølger x (ved $x+1$ - posisjonen) ikke ettersøkes aktivt når ($x-1$) operasjonen bestemmer at x er foranlediget av et hjelpeverb for x . I stedet vil brytningsprosessen gå til det neste tilgjengelige verbet i segmentet (brytningsprosessen

20 drives fremdeles av det faktum at segmentet forblir større enn den maksimale radverdien (fordi den potensielle brytningen ved x ikke ble utført)), og prosessen vil dermed finne verbet som nettopp var $x+1$, og analysere det som en ny x (her kalt x^*). Operasjonen x^*-1 vil nå se tilbake på den forrige x og bestemme hvorvidt x^*-1 er et egnet hjelpeverb for x^* .

Operasjonen $x-1$ for nivå 40-verb (modale verb, se nedenfor), og $x-1$ og $x+1$ -

25 operasjoner for verb på nivå 45/46 (tvetydige verb som enten kan være transitive [45] eller strengt intransitive [46]), er noe forskjellig fra de som benyttes på alle andre verb. I tilfellet med modale verb ble algoritmen forenklet fordi bare en infinitivsform av et modalt verb (og det finnes bare ett modalt verb som kommer i infinitivsform, dvs. engelsk "do" - gjøre) kan foranlediges av et annet modalt verb (f.eks. "kunne gjøre"). I tilfellet med verb på nivå

30 45/46 blir den typiske $x-1$ - operasjonen gjennomført i starten men etterfølges av spesialiserte $x-1$ og $x+1$ - spørsmål. Disse vil forsøke å utvetydiggjøre x og vil omfatte verb-verb - grensesnitt på hver side av x .

En forespørsel om en potensiell brytning ved et verb x på et lavere nivå (f.eks. 40) kan muligens ende opp med at det utføres en brytning ved rommet som ligger forut for et verb

35 på et høyere nivå (f.eks. 46) som kommer etter, dersom ett eller flere verb følger x sekvensielt. Straks det er laget en brytning blir >MAKS - spørsmålet stilt igjen. Dersom det

kreves mer brytning vil prosessen igjen først se etter verb på lavere nivå. Dette utfiltrerer de mer elementære verbuttrykkene først (f.eks. de som begynner med et modalt verb, slik som "Jeg kunne ha kjøpt melk"); etterfulgt av verb som involverer hjelpeverb (f.eks. "Vi har gjort spaserturer der i mange år") etterfulgt av utvetydige verb, deretter tvetydige verb.

5 Forhåpentligvis vil segmentene ved hjelp av de andre verb-nivåene bringes ned til mindre enn Max så ofte at trinnene for tvetydige verb ikke vil benyttes særlig ofte. Ved det tidspunkt verb på nivå 45/46 brukes for å initiere en forespørsel om brytning, vil alle verbene på høyere nivå blitt filtrert ut. Dette tillater noen nyttige antakelser som involverer verb-verb grensesnitt som opptrer mellom to ulike 45/46 - verb.

10 Det finnes fire referanselister som anvendes i tillegg til verb-verb - sammenlikninger: en endelig subjektiv-pronomenliste (som vil gjøre et umiddelbart påfølgende tvetydig verb til et definitivt ikke-verb, betegnet ved y99 i ordlista); en etter-bestemmelsesliste (som vil gjøre et forutgående tvetydig verb til ei egnet uttrykksgrænse, betegnet ved z99 i ordlista); og ei verbalsubstantiv hjelpeverb-liste (som kan forhindre brytning i forkant av verbalsubstantiv-

15 formen), betegnet ved w99 i ordlista.

Algoritmene foran for tilordning av gate og spor i de nylig dannete segmentene kan fremdeles anvendes. Gata X i hvilken et nydannet segment plasseres, er avledet fra den progressive brytningshistorikken for et segments generasjons"forløpere". Sporet (a,b,c,d) innenfor hver gate, i hvilken et nydannet segment plasseres, har til nå vært en funksjon for

20 ordets leddordnivå. Fordi alle disse ordene i verbløkket per definisjon er verb, vil sportilordningen være den samme, f.eks. spor b.

For å prøve å transportere forskjellene blant verbløkke-brytningene kan imidlertid sportilordningene modifiseres. Enhver brytning som gjøres i rommet før det forespurte verbet x kan være i spor c. Enhver brytning som gjøres i rommet før et etterfølgende verb

25 etter x (dvs. x+n), når x etterfølges av ett eller flere verb, kan være i spor b DERSOM verbet ved x+n var utvetydig (43 eller 44). Enhver annen brytning i rommene etter x kan være i spor d.

Prosessoversikt

30 Det bygges et leksikon med verb. Mens dette kan utføres manuelt, muliggjør reglene presentert her en automatisering med ei elektronisk ordbok. Hvert verb merkes i henhold til dets funksjon og form.

I den foreliggende leddordliste-terminologien vil verb være ved nivå på størrelsesorden 40. De fire funksjonsgruppene vil derfor være: modal = 40; ha-verb = 41; være-verb = 42;

35 og ordinære verb = 43/44 og 45/46. Gruppene 43/44 er utvetydige verb, dvs. at de alltid er verb. Gruppene 45/46 er tvetydige verb, dvs. det er ikke sikkert at de er verb. Oddetallet i

hvert av disse parene representerer transitive verb; liketallet representerer strengt intransitive verb. Det vil bli gjort søk etter verbene ved nivå 43 eller 44 samtidig i søket etter et nytt kandidatverb for brytning; straks et verb er funnet ved enten 43 eller 44 er operasjonene ulike for 43 hhv. 44. Tilsvarende vil det i begynnelsen samtidig bli søkt etter verb ved nivå 45 og 46, men de håndteres forskjellig straks de er identifisert for potensiell brytning.

"Åpne" verb, dvs. de som ikke innledes med et verb, vil identifiseres for potensiell brytning som en funksjon av antall funksjonsgrupper, som utføres for de andre leddordene i leddordlista. En streng med sekvensielle verb (eller potensielle verb) kan ha flere funksjonsord-nivå: "He should_not (4= have (41) been (42) bringing (43) work (43) home". Trinnene skissert senere vil tillate at operasjonene ($x-1$) og deretter ($x+1$) fortsetter gjennom strengen uavhengig av om leddordnivåene for ordet er i strengen.

En foretrukket framgangsmåte omfatter fem ulike typer verb-prefiks-klassifiseringer, basert på formen av hvert verb. Vi har derfor kalt hver av disse fem gruppene: c (for infinitivformen, f.eks. "å gå", men også "jeg går", osv.), d (for fortidsformen, f.eks. "gikk"), e (for andre former som ikke tar noen hjelpeverb inkludert de modale verbene som bare har denne formen, tredje person entall av regulære verb, f.eks. "han går", men også "han spaserer"), f (for former som både er infinitiv og fortid, som i noen irregulære verb slik som eng: "have cut"), og g (for verbalsubstantiv, f.eks. "walking").

Eksempler på hvordan spesifikke verb vil merkes med denne klassifiseringen er framsatt nedenfor. Ytterligere eksempler er framsatt i Appendix 1.

"Do" c 40, "Did" e 40, "Does" e 40, "Could" e 40, "Might" e 40, "Have" c 41, "Had" d 41, "Having" g 41, "Has" e 41, "Be" c 42, "Been" d 42, "Being" g 42, "Am" e 42, "Is" e 42, "Were" e 42, "Walk" c 45, "Walked" d 45, "Walking" g 45, "Walks" e 45, "Cut" f 45, "Cuts" e 45, "Done" d 45, "Doing" g 45, "Bring" c 43, "Brought" d 45, "Bringing" g 45, "Go" c 45, "Gone" d 45, "Going" g 46, "Goes" e 44, "Went" e 44.

Integrering med eksempelmetoden foran.

De etterfølgende trinn 18 og 18A erstatter trinnene 18 og 18A i algoritmen over i utførelser som bruker verbløkke-algoritmen:

18. Finnes det ord i segmentet som kan finnes i leddordlista ved nivå [n] (b10 til b39, og b50 til b99)? Nei: gå til 18A, Ja: gå til trinn 19.

18A. Gjenta spørsmålet for leddordliste-nivå [n+1]. Dersom det ikke finnes flere leddord ved nivå b39, gå til verbløkket for algoritmer som involverer nivå 40-49 før inntreden i nivå b50. Etter at verbløkket er ferdig for nivå 40-49, gjenoppta brytning ved å starte med nivå b50 i trinn 18. Dersom det ikke finnes noen ord i segmentet som kan finnes ved noe nivå i

leddordlista, gå til trinn 29.

Verbløkke

5 Q1. Finnes det noen ord i segmentet merket 40-46 (med hvilket som helst av tegnene c, d, e, f eller g før samme)? Nei: gå tilbake til trinn 19 i den ordinære algoritmen, Ja: gå til trinn Q2.

Q2. Finnes det et verb merket 40 i segmentet (eller resterende segment)? Nei: gå til trinn Q20, Ja. gå til trinn Q3.

10 Q3. For det første (eller neste) nivå 40 - verbet (x) i segmentet, er x utenfor den minimale radverdien (>MIN)? Nei: gå tilbake til trinn Q2 og inspiser neste 40 - nivåverb i segmentet fjerntliggende fra x. Ja: gå til Q4.

Q4. Er verbet x en "c"-form (dvs. er det merket c40)? Nei: gå til Q6, Ja: gå til trinn Q5.

15 Q5. Er ordet før x (x-1) et nivå 40 - verb? Nei: gå til trinn Q6, Ja: IKKE utfør brytning ved rommet før x. Gå i stedet tilbake til trinn Q2 og inspiser det neste nivå 40 - verbet i segmentet fjerntliggende fra x.

20 Q6. Det vil muligens skje en brytning ved rommet før x, men det kan i stedet være en brytning før et annet verb ved x+n. Sjekk ordet/ordene som etterfølger x (x+n) ved å starte med n=1. Er x+n et verb (dvs. ethvert ord merket 40, 41, 42, 43, 44, 45 eller 46, med hvilket som helst av de etterfølgende tegnene før tallet: c, d, e, f, g)? Nei: gå til trinn Q6A, Ja: gå til trinn Q7.

25 Q6A. Er x en "c"-form eller en "f"-form? Nei: bryt ved rommet før x. Dersom ordet "og" eller ordet "eller" er ordet som kommer før x skal det ikke skje brytning før x. I stedet, dersom ordet "og" eller "eller" er større enn den minimale radverdi (>MIN) skal det brytes før ordet "og" eller og det "eller". Dersom ordet "og" eller ordet "eller" som er plassert foran x er mindre enn den minimale radverdien (<MIN) skal det ikke brytes, gå i stedet tilbake til trinn 2 og sjekk verb i segmentet fjerntliggende fra x.

Plasser det nye segmentet i gate X (relativt til de andre tidligere generasjonene med brytning fra det opprinnelige segmentet/setningen), spor A. Gå deretter til trinn Q19. Ja: gå til trinn Q6B.

30 Q6B. Er ordet x-1 "til" el. eng. "to"? Nei: bryt ved rommet før x. Dersom ordet "og" eller ordet "eller" er ordet som ligger foran x skal det ikke brytes før x. I stedet skal det brytes før ordet "og" eller ordet "eller" dersom "og" eller "eller" er større enn den minimale radverdien (>MIN). Dersom ordet "og" eller ordet "eller" plassert foran x er mindre enn den minimale radverdien (<MIN) skal det ikke brytes, gå i stedet tilbake til trinn 2 og sjekk verb i segmentet fjerntliggende fra x.

Plasser det nye segmentet i spor X (relativt til de andre forutgående generasjonene med

foldning fra det opprinnelige segmentet/setning), spor C. Gå deretter til trinn Q19.

Ja: bryt ved rommet før "to". Dersom ordet umiddelbart før "to" er "og" eller "eller" eller ordet "eller" og ordet "to" begge er større enn den minimale radverdien ($>MIN$) skal det brytes før ordet "og" eller ordet "eller", avhengig av hvilket som kommer først. Dersom
5 ordet "og" eller ordet "eller" eller ordet "to" lokalisert foran x er mindre enn den minimale radverdien ($<MIN$) skal det ikke brytes, gå i stedet tilbake til trinn 2 og sjekk verb i segmentet fjerntliggende fra x.

Plasser det nye segmentet i spor X (relativt til de andre forutgående generasjonene med brytning fra det opprinnelige segmentet/setning), spor C. Gå deretter til trinn Q19.

10 Q7. Er $x+n$ en "e"-form? Nei: gå til trinn Q7A, Ja: gå til trinn Q17.

Q7A. Er $x+n$ en "c"-form? Nei: gå til trinn Q9, Ja: gå til trinn Q8.

Q8. Er $(x+n) - 1$ et nivå 40 - verb? Nei: gå til trinn Q17, Ja: gå til trinn Q18.

Q9. Er $x+n$ en "f"-form? Nei: gå til trinn Q11, Ja: gå til trinn Q10.

15 Q10. Er $(x+n) - 1$ et verb på nivå 40, 41 eller 42? (Legg merke til at et antas at det vil være svært sjelden av slike irregulære verb hvis infinitivform og fortid er de samme, (dvs. "f"-form)), også vil være strengt intransitive, OG at en forfatter vil innlede fortidsvarianten av et slikt verb med et "være"-verb). Nei: gå til trinn Q17, Ja: gå til trinn Q18.

Q11. Er $x+n$ en eller flere av følgende: d41, d42, d43 eller d45? Nei: gå til trinn Q13, Ja: gå til trinn Q12.

20 Q12. Er $(x+n) - 1$ enten et nivå 41 eller et nivå 42 - verb? Nei: gå til trinn Q17, Ja: gå til trinn Q18.

Q13. Er $x+n$ d44 eller d46? Nei: gå til trinn Q15, Ja: gå til trinn Q14.

Q14. Er $(x+n) - 1$ et nivå 41 - verb? Nei: gå til trinn Q17, Ja: gå til trinn Q18.

Q15. Er $x+n$ en "g"-form? Nei: gå til trinn Q17, Ja: gå til trinn Q16.

25 Q16. Er $(x+n)-1$ et nivå 42 - verb eller et verbalsubstantiv - hjelpeverb? (w99)? Nei: gå til trinn Q17, Ja: gå til trinn Q18.

Q17. ["ja" til Q7 eller "nei" til Q8, 10, 12, 14, 15, 16]. Bryt men IKKE ved rommet før x. Bryt i stedet ved rommet FØR $x+n$. Plasser det nye segmentet i spor X (relativt til de andre forutgående generasjonene med brytning fra det opprinnelige segmentet/setningen), spor
30 B. Gå deretter til trinn Q19.

Q18. ["ja" på Q8, 10, 12, 14, 16]. Adder 1 til n og gå tilbake til trinn Q6 ved bruk av $n+1$ i stedet for n. (Det vil si inspiser det neste ordet som kommer etter $x+1$ og se om det er et verb eller ikke, og fortsett tilsvarende).

35 Q19. Finnes det fremdeles segmenter som overstiger den maksimale radverdien (eller den tilgjengelige plassen til høyre)? Nei: stopp, Ja: gå tilbake til trinn Q2.

Q20. Finnes det et nivå 41 - verb i segmentet (eller resterende segment)? Nei: gå til trinn

Q34, Ja: gå til trinn Q21.

Q21. For det første (neste) nivå 41-verb (x) i segmentet, er x utenfor den minimale radverdien (>MIN)? Nei: gå tilbake til trinn Q20 og sjekk segmentet for et nivå 41-verb fjerntliggende fra x, Ja: gå til trinn Q22.

5 Q22. Er x en "e"-form? Nei: gå til trinn Q22A, Ja: gå til trinn Q33.

Q22A. Er x en "c"-form? Nei: gå til trinn Q24, Ja: gå til trinn Q23.

Q23. Er (x-1) et nivå 40-verb? Nei: gå til trinn Q33, Ja: gå til trinn Q32.

Q24. Er x en "f"-form? Nei: gå til trinn Q26, Ja: gå til trinn Q25.

Q25. Er (x-1) et verb på nivå 40, 41 eller 42? Nei: gå til trinn Q33, Ja: gå til trinn Q32.

10 Q26. Er x en av følgende: d41, d42, d43 eller d45? Nei: gå til trinn Q28, Ja: gå til trinn Q27.

Q27. Er (x-1) et nivå 41 eller 42 -verb? Nei: gå til trinn Q33, Ja: gå til trinn Q32.

Q28. Er x d44 eller d46? Nei: gå til trinn Q30, Ja: gå til trinn Q29.

Q29. Er (x-1) et nivå 41-verb? Nei: gå til trinn Q33, Ja: gå til trinn Q32.

15 Q30. Er x en "g"-form? Nei: gå til trinn Q32, Ja: gå til trinn Q31.

Q31. Er (x-1) et nivå 42-verb eller et verbalsubstantiv hjelpeverb (w99)? Nei: gå til trinn Q33, Ja: gå til trinn Q32.

Q32. ["ja" i Q32, 25, 27, 29 eller 31]. Gå til trinn Q32A.

20 Q32A. Er x et nivå 45 eller 46 -verb? Nei: IKKE bryt ved rommet før x. Gå i stedet tilbake til trinn Q20 og se etter andre verb i segmentet fjerntliggende fra x. (Dersom det ikke finnes flere nivå 41-verb i segmentet vil trinn Q20 direkte spørre etter trinn Q34 for nivå 42-verb. Tilsvarende vil det rettes en forespørsel fra trinn Q34 til trinn Q36 om nivå 43/44-verb, og fra trinn Q36 til trinn Q38 for nivå 45/46-verb). Ja: gå til trinn Q40.

Q33. ["ja" i Q22 etter "nei" i Q23, 25, 27, 29, 30 eller 31]. Gå til trinn Q33A.

25 Q33A. Er x et nivå 45 eller 46 -verb? Nei: vurder mulig brytning ved rommet før x men sjekk (x+n) først. Gå tilbake til trinn Q6. Ja: gå til trinn Q42.

Q34. Finnes det noen nivå 42-verb i segmentet (eller resterende segment)? Nei: gå til trinn Q36, Ja: gå til trinn Q35.

30 Q35. For det første (neste) nivå 42-verb x i segmentet, er x utenfor den minimale radverdi (>MIN)? Nei: gå tilbake til trinn Q34 og se etter det neste nivå 42-verbet i segmentet fjerntliggende fra x, Ja: gå til trinn Q22.

Q36. Finnes det noen nivå 43-verb ELLER noen nivå 44-verb i segmentet (eller resterende segment)? [Merk at verb ved nivå 43 eller 44 sjekkes i dette trinnet]? Nei: gå til trinn Q38, Ja: gå til trinn Q37.

35 Q37. For det første (neste) nivå 43/44-verbet x i segmentet, er x utenfor den minimale radverdien (>MIN)? Nei: gå tilbake til trinn Q36 og se etter det neste nivå 43/44-verbet i

segmentet fjerntliggende fra x, Ja: gå til trinn Q22.

Q38. Finnes det noen nivå 45-verb eller noen nivå 46-verb i segmentet (eller resterende segment)? [Merk at verb ved nivå 45 eller 46 sjekkes i dette trinnet]? Nei: gå tilbake til de vanlige algoritmene ved trinn 18 og gjenoppta spørring ved bruk av ordnivå b50, Ja: gå til trinn Q39.

Q39. For det første (neste) nivå 45/46-verbet x i segmentet, er x utenfor den minimale radverdien ($>MIN$)? Nei: gå tilbake til trinn Q38 og se etter det neste nivå 45/46-verbet i segmentet fjerntliggende fra x, Ja: gå til trinn Q22.

Q40. ["ja" fra trinn Q32A] Er x et nivå 45-verb? Nei: [x er derfor et nivå 46-verb]. Bryt ved rommet ETTER x. Gå deretter til trinn Q19, Ja: gå til trinn Q41.

Q41. [x er et nivå 45-verb]. Er (x+1) et verb? [Merk at dersom (x+1) nå virkelig er et verb så bør (x+1) enten være et nivå 45 eller nivå 46 -verb]. Nei: IKKE foreta brytning ved rommet før x. Gå i stedet tilbake til trinn Q38 og se etter det neste nivå 45/46-verbet i segmentet fjerntliggende fra x. Ja: IKKE bryt ved rommet før x. Bryt i stedet ved rommet før (x+1), (som også er posisjonen ETTER x). Plasser det nydannete segmentet i spor X (relativt til de andre forutgående generasjonene med brytning fra det opprinnelige segmentet/setningen). Spor D. Gå deretter tilbake til trinn Q19.

Q42. ["ja" fra trinn Q33A] Er (x+1) et verb? Nei: gå til trinn Q43. Ja: bryt ved rommet før (x+1), som også er rommet etter x. Plasser det nye segmentet i spor X (relativt til de andre forutgående generasjonene med brytning fra det opprinnelige segmentet/setningen), spor D. Gå deretter til trinn Q19.

Q43. Er x en "c"-form eller en "f"-form? Nei: gå til trinn Q45, Ja: gå til trinn Q44.

Q44. Er (x-1) ordet "to"? Nei: gå til trinn Q45, Ja: bryt ved rommet før "to". Dersom ordet umiddelbart før "to" er "og" eller "eller", bryt før "og" eller "eller" i stedet for før "to". Dersom ordet "og" eller ordet "eller" og ordet "to" alle er større enn den minimale radverdi ($>MIN$) skal det brytes før ordet "og" eller ordet "eller" eller ordet "to", hvilket av disse som måtte komme først. Dersom ordet "og" eller ordet "eller" eller ordet "to" foran x er mindre enn den minimale radverdien ($<MIN$) skal det ikke brytes, gå i stedet tilbake til trinn 38 og se etter nivå 45/46-verb i segmentet fjerntliggende fra x. Plasser det nye segmentet i spor X (relativt til de andre forutgående generasjonene med brytning fra det opprinnelige segmentet/setningen), spor C. Gå deretter til trinn Q19.

Q45. Er (x-1) et ord som finnes i den endelige subjektive pronomen-lista [liste x99; disse ordene er jeg, vi, han, henne, de, hvem osv.]? Nei: gå til trinn Q46, Ja: bryt ved rommet ETTER x. Plasser det nye segmentet i spor X (relativt til de andre forutgående generasjonene med brytning fra det opprinnelige segmentet/setningen), spor D. Gå deretter til trinn Q19.

Q46. Er (x-1) et ord i det innledende settet [y99]? Nei: gå til trinn Q46A. Ja: IKKE bryt ved rommet før x og gå tilbake til trinn Q38 for å sjekke nivå 45/46-verb i segmentet fjerntliggende fra x.

5 Q46A. Er x en "f"-form, en "d"-form eller en "g"-form? Nei: gå til trinn Q47, Ja: bryt ved rommet før x. Dersom ordet "og" eller ordet "eller" er ordet før x, (x-1), ikke foreta brytning før x. Dersom ordet "og" eller ordet "eller" er større enn den minimale radverdien (>MIN) bryt i stedet før ordet "og" eller ordet "eller". Dersom ordet "og" eller ordet "eller" foran x er mindre enn den minimale radverdien (<MIN), ikke foreta brytning - gå i stedet tilbake til trinn 38 og se etter nivå 45/46-verb i segmentet fjerntliggende fra x.

10 Q47. Er x et nivå 45-verb? Nei: IKKE bryt ved rommet før x og gå tilbake til trinn Q38 for å sjekke nivå 45/46-verb i segmentet fjerntliggende fra x. Ja: gå til trinn Q48.

15 Q48. Er (x+1) et ord i det avsluttende settet [z99]? Nei: IKKE bryt ved rommet før x og gå tilbake til trinn Q38 for å sjekke nivå 45/46-verb i segmentet fjerntliggende fra x. Ja: BRYT ved rommet FØR x. (Dersom ordet "og" eller ordet "eller" er ordet før x skal det ikke brytes før x. Dersom ordet "og" eller ordet "eller" er større enn den minimale radverdien (>MIN) skal det i stedet brytes før ordet "og" eller ordet "eller". Dersom ordet "og" eller ordet "eller" foran x er mindre enn den minimale radverdien (<MIN) skal det ikke brytes - gå i stedet tilbake til trinn Q38 og sjekk nivå 45/46-verb i segmentet fjerntliggende fra x). Plasser det nye segmentet i spor X (relativt til de andre forutgående generasjonene med brytning fra 20 det opprinnelige segmentet/setningen), spor C. Gå deretter til trinn Q19. (Slutt på verbløkke).

Tekstforsterkning ved bruk av leseutøver-aktivitet

25 Tekstpresentasjon kan forbedres ved bruk av aktiviteter eller aksjoner fra en leser eller leseutøver i form av inngangsverdier og et utvalg av audiovisuelle effekter som utgangsverdier. I en utførelse vil leseren eller leseutøveren lese teksten høyt med varierende stemmehøyde, tone, volum og pauser. Betegnelsen "leseutøver" brukes for å skille fra en leser av den forsterkede teksten og bemerke at høytlesing kan omfatte mange aktiviteter i tillegg til tale. Den uttalte tekst registreres for bruk i visning av teksten. I en 30 utførelse registreres den uttalte teksten elektronisk og analyseres med hensyn til uttalte aksjonsattributter slik som høyde, tone, volum og pauser. I en foretrukket utførelse analyseres uttalte tekstattributter og de analyserte attributtene lagres i tillegg til selve den uttalte teksten. Attributtene slik som høyde, tone og volum lagres og kartlegges for å motsvare teksttegnene, stavelsene og ordene som også er lagret, slik at attributtene kan 35 ekstraheres og reproduseres i en eller annen form i samsvar med den opprinnelige teksten. Leseutøver-aksjonsattributter kan presenteres som visningsattributter slik som

skrifttype, intensitet, skriftlinjetykkelse, skrifttegnhøyde, tegnfarge, bakgrunnsfarge, tegnnavstand, ordavstand, linjeavstand og setningsavstand.

I ett eksempel ifølge oppfinnelsen reproduseres volumet hos leseren som skriftintensitet og høyden reproduseres som skrifttykkelse, der større høyde reproduseres som smalere tegn og mindre høyde reproduseres som bredere tegn. Skriftintensitet kan vise til antall punkter satt per enhetsareal til en annen farge enn de lokale bakgrunns punktene. I ett eksempel er intens tekst svart tekst på hvit bakgrunn mens mindre intens tekst er grå tekst på hvit bakgrunn. I et annet eksempel ifølge oppfinnelsen varieres tegnfargen i henhold til tonehøyde. For eksempel reproduseres ord med større tonehøyde lengre mot den røde enden av ROYGBIV-spektret mens ord med lavere tonehøyde reproduseres mer mot den fiolette del av spektret. I et annet eksempel vises teksten forskjøvet fra den normale basislinja for teksten i henhold til tonehøyden for leseutøverens. En utførelsesform forskyver tekst oppover for over middels tonehøyde og nedover for under middels tonehøyde.

I et annet eksempel ifølge oppfinnelsen reproduseres leseutøverens pauser og uttalelengde for spesifikke ord i forsterket tekst med avstandsjusterte bokstaver, ord og setninger. Leserens uttaletid og pauser mellom setninger og ord kan registreres for å tillate måling av normale pauser mellom ord og setninger. Ord som uttales langsommere enn normalt kan reproduseres med ekstra blanke mellom bokstavene i ordet. Tidsmellomrom mellom ord som er lengre enn normalt kan reproduseres i den viste teksten ved bruk av blanke som er større enn den normale avstand mellom ordene. Enda lengre tidsmellomrom kan reproduseres ved å vise de atskilte ordene eller uttrykkene på nye linjer. Ekstremt lange tidsintervall kan vises som de atskilte ordene eller uttrykkene på nye linjer som er vertikalt atskilt med en større avstand enn vanlig.

25 Håndbevegelser

I et annet aspekt ved oppfinnelsen omfatter registrering av leseutøverens aktiviteter under lesing hånd- og fingerbevegelser. Hendenes posisjon og orientering blir fortrinnsvis registrert og lagret for å tilsvare høytlesning av hvert ord. Orientering viser til retningen som hendene, fingrene eller håndflatene peker i. I et todimensjonalt plan kan orienteringen registreres som en todimensjonal vektor. I en tredimensjonal modell kan hendene registreres som en tredimensjonal vektor. Posisjonene kan registreres som posisjonen relativt til leseutøverens skuldre eller andre kroppsreferansepunkter.

I valgte egnede tekstseksjoner kan den vertikale og horisontale forskyvningen av tekst varieres i henhold til tekstinnhold. Et eksempel på dette er visning av setningene "På den ene siden..." og "På den andre siden...". Det første uttrykket kan forskyves horisontalt mot venstre mens den andre motsatte setningen kan forskyves horisontalt mot høyre.

Tilsvarende kan det brukes vertikal forskyvning for å gjenspeile håndbevegelser slik som "Han lovet dette, men leverte bare dette".

Håndstilling og -orientering kan registreres ved bruk av kjent teknologi slik som plassering av infrarøde reflekterende punkter ved strategiske posisjoner på hendene.

- 5 Refleksjonen av projiserte infrarøde lys kan fanges og lagres og tillater på et senere tidspunkt reprodusering av håndbevegelser. I en annen framgangsmåte bruker leseutøveren hansker og håndstilling registreres direkte med et digitalt kamera ved bruk av kontrasten av hendene mot en mørkere bakgrunn.

10 Ansiktsuttrykk

Et annet aspekt ved oppfinnelsen omfatter registrering av leseutøverens ansiktsuttrykk under lesing og lagring i samsvar med ordene som leses. Ën utførelse registrerer øyeform eller øyelokk-form, øyeåpning, øyebrynstilling, hodestilling, kjevestilling og munnform.

- 15 Ansiktsuttrykk-aktivitetene hos leseutøveren kan registreres på ordbasis og vises tilbake på linjebasis. I Ën utførelse brukes ekstremstillingene for å representere posisjonen for hele linja. I en annen utførelse anvendes en gjennomsnittsposisjon for å representere posisjonen for hele linja. I en annen utførelse anvendes ansiktsuttrykkene til å endre presentasjonen av teksten i seg selv. I ett eksempel anvendes graden av åpenhet i munnen til å variere graden av åpenhet av visse lukkede vokaler, slik som "O". Dess mer 20 åpen leseutøverens munn er, dess større blir O'en og åpningen i samme. I et annet eksempel varieres høyden av visse tegn slik som "W" eller "T" avhengig av åpningsgraden i øyelokkene. Ansiktsuttrykk-representasjonen beskrevet foran kombineres med tekstsveiping beskrevet foran i ett aspekt ved oppfinnelsen. I Ën utførelsesform reproduseres ei streksskisse av et menneskeansikt over ordet som sveipes eller uthèves.
- 25 Ansiktstrekkene i det animerte ansiktet endres dynamisk. I en annen utførelse vises et ansikt ved siden av hver tekstlinje, der ansiktet animeres mens linja sveipes. I nok en utførelsesform presenteres et ansikt nær uttrykket som skal framheves, og et statisk ansiktsuttrykk vises i ansiktet.

30 Tekstpresentasjon langs ei krummet linje.

I et annet aspekt ved oppfinnelsen vises tekst langs ei krummet basislinje i stedet for langs ei ordinær horisontal rett linje som en funksjon av leseutøverens aktiviteter. I Ën utførelse er den krummete linja enten usynlig eller hovedsakelig mindre synlig enn teksten som vises. Linja varieres i Ën utførelse i henhold til en viss aktivitet hos leseutøveren. I ett

- 35 eksempel er den horisontale forskyvningen av linja større for høyere tonehøyder og mindre for lavere tonehøyder. Ordhøyden kan framskaffes fra et fonetisk alfabet, f.eks. kan

tonehøyden for et ord slås opp i stedet for å framskaffes fra leserens stemme. I et annet eksempel varieres den horisontale forskyvningen i henhold til volumet i leseutøverens stemme. I ett eksempel blir leseutøverens aktivitet dempet tilstrekkelig til at det opprettholdes ei jevnflytende krummet linje.

5

Tekstbøyning

I det talte språk danner endringen i tone, øyde og volum fra ett ord til det neste et signal som kan betraktes som ei kurve eller bane. Allerede etter uttale av noen få ord i ei setning "trekkes" begynnelsen av ei slik kurve på ei "aural" tavle. Lytteren bruker denne initiale kurva for å bygge og forvente, i hodet, strukturen for uttrykket eller setningen som helhet, selv før den er fullstendig uttrykt av taleren.

Denne kraftfulle informasjonen gjør talelytting svært effektiv. Dagens tekstpresentasjonsteknologi gjør det imidlertid ikke enkelt for et system for å utføre en konsistent og meningsfylt skissering av disse aurale kurvene i en grafisk form. I stedet må leseren gjenopplive den deklamatoriske kraft i setningen ved å dekode symbolene i teksten og gjendanne en aural representasjon av setningen som helhet i hodet.

Innenfor en setning kan segmenter identifiseres ved bruk av innholdsanalyse som vil ha karakteristiske prosodiske konturer. Den foreliggende oppfinnelsen beskriver en framgangsmåte for dannelse av grafiske korrelasjoner for disse konturene med setningssegmenter, og integrere disse grafiske underenhetene i en større struktur for å skissere setningen som en helhet.

Prosessen

1. Setningssegmenter identifiseres ved bruk av "setningsbrytnings"-algoritmene beskrevet foran.

2. I stedet for horisontale forskyvningsregler defineres et sett med kurver og hvert setningssegment tilordnes til en av disse kurvene.

3. Kurvene arrangeres deretter på lesefeltet i et mønster som utvetydig skisserer sekvensen for segmentene (f.eks. fra toppen og nedover); segmentenes posisjon innenfor disse arrangementene kan modifiseres lett uten å endre den impliserte sekvens for å optimalisere persepsjonen av kurvene og lesingen av setningen.

Et eksempel på ei masterkurve 340 som definerer størrelsen og orienteringen av segmentkurver er skissert i Figur 13. En rekke med tegn, slik som tegnet "H" ved 344, ligger på kurva 340. Ei pil 342 ved toppen av kurva indikerer posisjonen med hvilken ei setningssegmentkurve som ble trukket fra en startbokstav (A til G) til en sluttbokstav (H til K) vil orienteres på lesefeltet og i forhold til andre setningssegmentkurver.

Startbokstavposisjonen på kurva bestemmes av omstendighetene ved begynnelsen av segmentet. Sluttegnposisjonen bestemmes av omstendighetene ved slutten av segmentet. Disse omstendighetene er oppramset i tabell 350 i Figur 14.

Etter at setningsbrytnings-algoritmene har identifisert segmenter i setningen som vil innta separate rader (i denne beskrivelsen til de innta separate kurver) måles lengden av segmentet. Det trekkes ei linje som tilsvarer denne lengden. Linja blir deretter "bøyd" slik at formen (grad og orientering) av den nydannede krummete linja er i samsvar med start-slutttegnenes posisjon spesifisert i Figur 13 og 14. "Opphengs"-punktet eller pilas posisjon ved toppen av kurva identifiseres også.

Størrelse av ei segmentkurve bestemmes av dens lengde, dets tegn; uavhengig av dens lengde vil imidlertid formen og orienteringen av segmentkurva bestemmes av prosessen beskrevet foran. Dersom et segment er svært kort, slik at det krummete segmentet vil være for lite og kurva for ekstrem, vil det defineres ei minimalkurve. Korte segmenter med en H-slutt vil deretter plasseres ved start-bokstaven spesifisert av deres innledende omstendigheter, men det resterende av segmentet vil følge krumningen spesifisert av den minimale kurvebanen og "opphengs"-punktet vil også spesifiseres av minimalkurva. Tilsvarende vil korte segmenter med en I, J eller K-slutt plasseres ved sluttegnet spesifisert av deres avsluttende omstendigheter, og det resterende av segmentet vil følge krumningen spesifisert av den minimale kurvebanen og dens "opphengs"-punkt.

Parametere for maksimal linjelengde vil mest sannsynlig forhindre at segmenter blir så lange, dvs. har så mange tegn, at kurvene deres blir svært store og forholdsvis flate. De nydannede kurvene vil dermed i dette eksemplet posisjoneres med det første segmentet øverst. I én utførelse er posisjonen for opphengspunktet en funksjon av det neste segmentet kan utvikles som var en funksjon av enden av det forutgående segmentet. Etter at alle kurvene har kommet i posisjon, posisjoneres teksttegnene for hvert segment på kurva og linjene og opphengspunktet fjernes. Et eksempel på denne prosessen er vist i Figur 15 og 16.

Nye kvaliteter tildelt av tekstbøyeprosessen

Denne tekstvisningsmetoden tillater at leseren med hjelp av ei spesiell kurve i utlegget av dens tegn gjenkjenner et individuelt segments struktur og funksjon (ikke bare sekvensiell posisjon) i hele strukturen for setningen. Dette tilbyr grafisk informasjon som letter kognitiv prosessering av setningens segmenter under lesing. Dette er tilsvarende metoden der ord-til-ord-variabilitet i tonehøyde og klang skaper en aural kurve som bistår lytteren i å forutsi hvor langt et uttrykk kommer til å bli og hvorvidt det vil være slutten på ei setning eller et avsnitt.

Konvensjonell tekst bruker bare skilletegnmerker, som ofte ikke gjenkjennes før leseren allerede har lest hele uttrykket - dette hindrer leseren fra å lage en tilsvarende fonetisk modell i hodet av det som ble lest før etter at hele segmentet er lest. Kurvene dannet med den foreliggende framgangsmåten tillater at denne typen uttrykksstruktur og funksjonsinformasjon blir kjent i forkant, slik at denne informasjonen kan brukes til ordgjenkjenning og uttrykksforståelse. Disse kurvene er også definert av innhold (dvs. leddord) forskjellig fra skilletegn. De bidrar dessuten til ytterligere differensiering av segmenter dannet i setningsbrygningsprosessen ved å skissere omstendigheter for både hodedelen og endedelen av segmentet.

10 Med henvisning til Figur 15 er det vist ei setning 380 som er den første setningen med Gettysburg-talen vist i kaskadeform dannet med algoritmer (Trinn 1 under "Prosessen"-overskriften ovenfor). Ei kolonne 386 spesifiserer en startbokstav 382 og en sluttbokstav 384 for hver linje i setningen. Dette bestemmer kurvetypen som skal tilordnes hvert segment.

15 En integrering av disse individuelle segmentkurvene, fra trinn 3, er illustrert i Figur 16. For eksempel er setning 408 bøyd over ei kurve 400 som begynner ved en sluttposisjon A og slutter ved en sluttposisjon 404, og med et opphengspunkt betegnet ved 406. Etterfølgende segmenter plasseres under forutgående segmenter ved å plassere opphengspunktet tilnærmet to tredjedeler av veien langs kurva over dem. Ekstra innholds- baserte regler kan optimalisere disse opphengspunkt-posisjonene. I en utførelsesform er hode/hale-tegnene fraværende, og bare teksttegnene er tilstede. En rekke bøyde tekstkurver dannet av eksemplet i Figur 15 er illustrert med kurver 410, 412, 414, 416 og 418.

25 Kurvene fremmer den visuelle prosodi av tekstpresentasjonen. Hver setning tilegnes en enda mer unik visuell framtreten. Kurvene bryter også opp det monotone særpreget i lineært presentert tekst, som kan redusere trettheten i øyemusklene.

Visning av teksthierarki

30 Et annet aspekt ved oppfinnelsen framskaffer et forbedret middel for visning av teksthierarkier og posisjonen for tekst som leses innenfor disse hierarkiene. Et konvensjonelt teksthierarki, opplistet fra høyere til lavere rangering, omfatter volumantall, seksjon, kapittel, delseksjon, side, avsnitt og setning. Den foreliggende oppfinnelsen anviser en framgangsmåte for å informere leseren om både antall av enhetene i hvert hierarki og posisjonen for tekst som leses innenfor dette hierarkiet.

35 Med henvisning til Figur 17 er det illustrert ei uåpnet bok 200. Boka velges for åpning av leseren, som indikert ved "X". Leserens valg kan utføres ved hjelp av enhver kjent metode

kjent innen faget slik som markørbevegelse og valg med mus. Med henvisning til Figur 18 er boka valgt, åpnet, og vises som et plan 202 som fyller en vesentlig del av en visningsskjerm 201. Åpningen av boka 200 har avdekket et flertall kapitler 206 som hver har en eksponert marg 204. I en utførelsesform er kapitelmargene forsynt med

5 kapitellnumrene innført i samme, som indikert ved 208. Leser-utvelgelse av kapittel 4 er indikert ved "X" i Figur 18, og forårsaker at kapitlene skilles ved kapittel 4, som illustrert i Figur 19.

Med henvisning til Figur 19 er resultatet av utvelgelsen av kapittel 4 illustrert. Kapitlene 206 er illustrert som atskilt for å vise kapittel 4 og eksponere sidene 210 i kapitlet. Sidene

10 210 har vertikale marger 212 eksponert. I en utførelse av oppfinnelsen er sidene nummerert og arrangert for å speile sidene i en trykket utgave av boka. I en annen utførelse av oppfinnelsen er sidene nummerert for å dra fordel av det forbedrede formatet i tekstpresentasjonen, uten noe forsøk på å speile noen trykket utgave av boka. Nærmere bestemt kan sidene formatteres og pagineres for å inneholde bare et maksimalt antall

15 avsnitt eller et maksimalt antall ord, hvilket av disse som måtte komme først. I denne utførelsen trenger ikke avsnitt brytes over sider. I denne utførelsen er antall avsnitt og ord i maksimum fortrinnsvis satt til verdier som er tilnærmet de for ei normal publisert bok. I nok en utførelse av oppfinnelsen settes antall sider i et kapittel til et maksimum, for eksempel

20 50, der avsnittene er jevnt inndelt i disse 50 sidene. I dette siste eksemplet presenteres på det meste 50 sider på visningsskjermen. Leser-utvelgelse av side 3 er indikert ved "X" i Figur 19 og forårsaker at sidene separeres ved side 3, som illustrert i Figur 20.

Med henvisning til Figur 19A er resultatet av utvelgelsen av kapittel 4 i Figur 18 illustrert i en alternativ utførelse. Kapitlene 206 er illustrert atskilt for å vise kapittel 4, og eksponere sidene 210A i kapitlet. Sidene 210A har horisontale marger 212A eksponert. I denne

25 utførelsen er nøstingen av teksthierarkiene betegnet ved at margene har en kortere lengde for lavere rangerte hierarkier. I eksemplet i Figur 19A, har sidene 210A marger som er orientert hovedsakelig parallelt med kapitlene 206, men der margene 212A for sidene 210A har en kortere lengde enn margene i kapitlene 206.

Med henvisning til Figur 20 er resultatet av utvelgelsen av side 3 illustrert. Sidene 210 er

30 atskilt ved side 3 for å vise avsnitt 214 innenfor samme. I den initiale visning vises bare avsnittets horisontale nedre marger 216. Leserutvelgelse av avsnitt 3 er indikert ved "X" i Figur 20, og forårsaker at avsnittene atskiltes ved avsnitt 3 som illustrert i Figur 21.

Med henvisning til Figur 21 er resultatet av utvelgelse av avsnitt 3 illustrert. Avsnittene 214 er separert ved avsnitt 3 for å avdekke setninger 218 med vertikale marger 220

35 innenfor samme. I en utførelse vises initialt den første setningen, som er klar for lesning, som illustrert ved første setning 222. I en annen utførelse vises i begynnelsen bare marger

220 ved åpning av avsnittet. Leser-utvelgelse av avsnitt 4 er indikert ved "X" i Figur 21, som forårsaker at setningene atskilles ved setning 4, som illustrert i Figur 22.

Med henvisning til Figur 22 er resultatet av utvelgelse av setning 4 ved 223 illustrert.

Setning 4 er setningen som leses akkurat nå. I en utførelse er avslutningen av lesingen av setning 4 indikert ved leser-aktivering av en datamaskin-perifer-anordning slik som en museknapp. I en annen utførelse indikeres slutten på lesing av en setning ved utløp av en beregnet tidsperiode i samsvar med vekten av setningen og leserens siste lesehastighet. I en foretrukket utførelse kan indikasjonen på slutten på lesing av en setning veksles mellom en tidsbasert metode og en leseraktivert metode. I en utførelse er hastighet leser-justerbar ved bruk av en inngangsverdi slik som ei rullstang. Dette tillater at leseren på en enkel måte kan øke og redusere hastigheten for å tilpasse seksjonen som leses.

Når det indikeres at lesing av ei setning er ferdig, skyves den neste setningen i posisjon. I Figur 22 glir i en utførelsesform setningen rakst mot venstre for å avdekke den underliggende neste setning 5. I en annen utførelse glir setningen 4 og alle påfølgende setninger i avsnittet mot venstre med bredden av den vertikale margen, der setningen 4 fortsetter å gli mot venstre for å ta sin plass under setningen 3. I denne siste utførelsen kan setningen som leses beholdes i samme posisjon ved å bevege setningen før teksten i setningen blir synlig for leseren. Etter at den siste setningen i avsnittet er lest åpnes det neste avsnittet for å avdekke setningene i samme, som likner mye på illustrasjonen i Figur 21 men der det neste avsnittet eksponeres eller trekkes ned.

I en foretrukket utførelse av oppfinnelsen presenteres alternative teksthierarkier med plan som har alternerende horisontale og vertikale eksponerte marger. Uttrykt på en annen måte ved hjelp av Microsoft Windows - teknologi, alternerende teksthierarkier representeres ved alternerende horisontale og vertikale flislagte plan. I en utførelse ser det ut til at de ytre høyere rangerte teksthierarkiene ligger bak de indre lavere rangerte teksthierarkiene, siden de åpnete høyere rangerte teksthierarkiene har de neste lavere rangerte teksthierarkiene vist over samme. I en annen utførelse ser det ut som at de ytre høyere rangerte teksthierarkiene er nærmere leseren, der de indre lavere rangerte hierarkiene vises lengre vekk fra leseren. I denne utførelsen blir fortrinnsvis ikke det nåværende åpnete teksthierarkiet vist, bortsett fra setningen som er under lesing. I et eksempel ved bruk av horisontalt flislagte tekstenheter og med et åpnet tekstenhetsnummer 3, framtrer tekstenhetene som en utspredd kortstokk med 1 oppå 2 på toppen av toppkanten av 3 og 4 under 3 og 5 under 4. Enhetene mot toppen av skjermen ser ut som at de er nærmere leseren enn enhetene mot bunnen av skjermen. Dette er arrangementet som anvendes i utførelsen i Figur 22 med hensyn til avsnitt, men der avsnitt 3 er åpnet.

I et annet eksempel er tekstenhetene 1 til 3 som beskrevet foran, men tekstenhet 4

ligger oppå den nedre kanten av tekstenhet 3, og tekstenhet 5 ligger på toppen av alle bortsett fra den øvre kanten av tekstenhet 4. I dette arrangementet ser det ut som at tekstenhetene vokser vekk fra leseren fra yttermarginene inntil den åpnete tekstenheten vises hvoretter tekstenhetene ser ut til å komme nærmere leseren inntil en kommer fram til den motstående yttermarginen. Dette arrangementet er tilsvarende en rekke nøstete scenetepper der de ytre teppene ligger i front, de mer sentrale teppene ligger bak de ytre teppene og de indre teppene avdekker en mindre åpning enn de ytre teppene.

I det generelle tilfellet varieres orienteringen av margene avhengig av teksthierarki-nivået som presenteres. Mens de illustrerte utførelsene har margene orientert alternerende horisontalt og vertikalt, 90 grader forskjøvet, kan i det generelle tilfellet forskyvningen være mindre. I ett eksempel har hierarkiet på det høyeste nivået horisontalmarger (som rager fra kl. 9 til kl. 3), og det neste lavere nivået er forskjøvet med 30 grader ragende fra kl. 10 til kl. 4.

Med nok en gang å henvise til Figur 22 kan visningen av den nåværende "boka" som leses forklares uten henvisning til hvordan leseren kom fram til lesing av ei spesiell setning. Det neste lavere teksthierarkiet er indikert med kapitler 206 som er flislagt horisontalt med horisontale nedre marger 204 eksponert. Det neste lavere teksthierarkiet, sidene 210, vises lagt oppå kapitlene 206. Sidene 210 er flislagt vertikalt med vertikale marger 212 eksponert. Det neste lavere teksthierarkiet, avsnittene 214, vises over sidene 210. Avsnitt 214 er flislagt horisontalt med horisontale marger 216 eksponert. Det neste lavere teksthierarkiet, setninger 218, er flislagt vertikalt med vertikale marger 220 eksponert.

Leseren er i stand til å visualisere posisjonen av setningen som leses innenfor avsnittet, avsnittet innenfor sida, avsnittet innenfor kapitlet og avsnittet innenfor boka. Ved bruk av kjent tredimensjonal skyggelegging og avfasing kan nøstete planare teksthierarkier gi leseren en tredimensjonal følelse for posisjonen av teksten som leses.

I en utførelsesform vises setningene i horisontalt glidende plan. I en annen utførelse vises setningene i en rekke horisontalt flislagte plan som glir horisontalt, tilsvarende avsnittene 214 i Figur 22. I en utførelse settes alltid setningene som skal vises som enten horisontalt eller vertikalt flislagte plan, med de høyere rangerte teksthierarkiene alternerende som horisontalt og vertikalt flislagte plan.

I noen tekstblokker kan det være et stort antall enheter, for eksempel et svært høyt antall sider i et kapittel, eller ei bok uten kapitler. For å håndtere slike situasjoner kan det skapes et kunstig hierarki. I en utførelse kan det dannes et ti-trinns hierarki. For eksempel dersom antall kapitler er større høyere enn 50 kan det dannes en kunstig enhet med 10 kapitler.

Denne kunstige enheten kan dannes enten som et høyere rangert teksthierarki som skal vises som det er eller ganske enkelt som en enhet med tilsvarende rangering som skal

flislegges likeledes men vises forskjellig. I ett eksempel er enheten med 10 kapitler flislagt vertikalt mens kapitlene er flislagt horisontalt. I et annet eksempel er både enhetene med 10 kapitler og selve kapitlene flislagt horisontalt, men der 10-kapittel-enheten har en mørkere farge eller på annen måte forskjellig utseende.

5

Visning av tekst over motstående sider

I et annet aspekt ved oppfinnelsen er det framskaffet en framgangsmåte for visning av tekst på etterfølgende sider. Med henvisning til Figur 23 er det illustrert ei venstreside 250 med en høyremarg 252 og ei høyreside 254 med en venstremarg 255, satt i skyggen av høyremargen 252. Side 3 vises som venstreside 250 og side nummer 4 vises som høyreside 254. Side 3 er indikert for å være under lesing med "X". Når lesing av side 3 er ferdig kan dette indikeres til lesesystemet med leseraktivitet eller ved utløp av et tidsintervall. Når lesing av side 3 er ferdig, som indikert ved pila 251, vises side 4 "foran" side 3, med høyresidens 254 venstremarg overskyggende venstresidens 250 høyremarg 252. Når lesing av side 6 er ferdig, indikert ved pila 253, vises nye sider der side 5 vises som venstreside 250 og side 6 vises som høyreside 254. På denne måten alterneres sidene mellom venstre og høyre for å etterlikne det tilsvarende mønstret for papirbøker som leseren er vant til.

Med henvisning til Figur 24 er det illustrert en annen utførelsesform av oppfinnelsen. Ei venstreside 260 med en høyremarg 262 er illustrert mens ei høyreside 264 med en venstremarg 265 overskygget av høyremargen 262. Side 3 vises som venstreside 260 og side 4 vises som høyreside 264. Side 3 er indikert ved "X". Når lesing av side 3 er ferdig, indikert ved pila 261, bringes side 4 til front slik at venstremargen 265 overskygger høyremargen 262. Når lesing av side 3 er ferdig fjernes i tillegg side 3 og erstattes med side 5. Denne konfigurasjonen er ikke mulig ved bruk av tradisjonell papirinnbinding. Når lesing av side 4 er ferdig, indikert ved pila 263, bringes side 5 til front og side 4 fjernes og erstattes med side 6.

Utførelsen illustrert i Figur 24 har den fordel at den tillater at tekstsegmenter kan brytes over sider mens en alltid tilbyr neste side for lesing på skjermen. Dette tillater at leseren kan gå framover fra tidligere illustrert tekst uten å måtte vente på ny visning av tekst. For eksempel kan et tekstsegment brutt over side 3 og 4 leses uten å kreve pause for at innholdet på side 4 skal vises. Et tekstsegment som er brutt over side 4 og 5 kan leses uten å kreve pause for at innholdet på side 5 skal vises. Følgelig er leseren i stand til å lese elektronisk presentert tekst uten å måtte vente på sideskift mens en på samme tid har den kjente presentasjonen av tekst på motstående sider, et slag tilbake mot innbundne papirbøker. Den midtre konfigurasjonen i Figur 24 er ikke bundet av innbinding.

Tekstvisning for lesing fra front til bakside.

Med henvisning til Figur 25 er det illustrert et annet aspekt ved oppfinnelsen i en skjermvisning av tekst 280 med ei første tekstlinje 282, ei andre tekstlinje 284, ei tredje tekstlinje 286 og ei fjerde tekstlinje 288. Ved bruk av perspektivillustrasjons-teknikker kan den første linja 282 bringes til å opptre nærmere leseren enn den andre linja 284, der den andre linja 284 er nærmere enn den tredje linja 286 og den tredje linja 286 nærmere enn den fjerde linja 288. I én utførelse av oppfinnelsen vises tekst på én skjerm eller visningsside om gangen med alle linjene på skjermen, og selv de mest fjerntliggende linjene er lesbare. I én utførelse er linja som leses aktivt indikert mer lyssterk eller med større kontrast enn linjene med tekst som ennå ikke er lest. Linja som er under aktiv lesing kan estimeres ved å mate fram en opplysende linje-bred markør med en hastighet lik leserens estimerte lesehastighet. I en variant av denne utførelsen frammates et linje-bredt gjennomskinnelig filter over linja som er under lesing slik at teksten som ikke befinner seg på linja som leses framtrer noe uklar og mindre lesbar enn teksten som leses.

Med henvisning til Figur 26 er visningen for Figur 25 illustrert som ei modifisert visningsside 290 der teksten er frammatet ei linje framover. Den første linja 282 har blitt tvunget framover vekk fra visningsorganet mens den andre linja 284 nå ligger i forgrunnen. I denne utførelsen av oppfinnelsen kan tekst frammates ei linje om gangen mot leseren. Med dagens teknologi kan teksten bringes til å rulle jevnt framover mot leseren. En perspektiv-effekt kan oppnås ved å lage de nedre tekstlinjene med større skriftstørrelse og større linjeavstand. Det kan også brukes tredimensjonale skrifttyper. I en foretrukket utførelse vises teksten som forsterket tekst etter prosessering ved bruk av brytningsmetoder beskrevet foran.

I en alternativ utførelse presenteres teksten fra front til bakside når den leses fra bunnen til toppen, som beskrevet foran. I denne utførelsen vises teksten statisk en skjerm om gangen, der teksten sideinndeles med et flertall linjeblokker i stedet for å gjennomrulles jevnt.

Utførelsene illustrert i Figur 25 og 26 drar fordel av den naturlige bildeprosesseringen som er valgt for den menneskelige utvikling. Nærmere bestemt prosesserte mennesker, før de leste, informasjon under gange eller på annen måte fremoverrettet bevegelse, der nyere informasjon framtrådte først lengst vekke og beveget seg nærmere med tiden og bevegelsen gjennom landskapet. Det norske språk skrives fra topp til bunn for å tillate at skriveren kan lese hva som er skrevet på den forrige linja. Skrivning fra topp til bunn vil dekke over disse linjene som akkurat er skrevet med skriverens hånd. Dette aspektet er velkjent for venstrehendte skrivere i en noe ulik form. Norsk skrives fra venstre mot høyre,

og tillater at høyrehendte skrivere kan se de ordene som nettopp er skrevet. Venstrehendte skrivere dekker over det de akkurat har skrevet, og antar ofte en ubekvem eller krampaktig skrivestilling i et forsøk på å kompensere for denne ulempen. Publisering har hittil tilpasset topp-til-bunn-stilen som er brukt for håndskrevet tekst uten egentlig å være nødt til det. Den foreliggende oppfinnelsen frigjør tekstvisning fra topp-til-bunn-konvensjonen og tillater visning av tekst i et mer naturlig skjønt ukonvensjonell visningsformat.

Variasjon av overgangsmønstre med tekstnivå

I et annet aspekt ved oppfinnelsen varieres overgangsmønstre mellom tekstvisninger som en funksjon av nivået eller rangeringshierarkiet for teksten. I en utførelse er overgangen mellom setninger innenfor et avsnitt forskjellig fra en overgang mellom avsnitt. Overgangen mellom setninger kan varieres som en funksjon av setningens posisjon innenfor avsnittet. Overgangen mellom avsnitt kan varieres som en funksjon av avsnittets posisjon innenfor kapitlet eller andre teksthierarkier som det ligger innenfor.

I ett eksempel ifølge oppfinnelsen utføres overgangen mellom avsnitt ved oppløsning av den nylig viste setningen før visning av den neste. Oppløsningen kan være en finkornet oppløsning av det første avsnittet i et kapittel, og en rekke granulariteter mellom samme for umiddelbart posisjonerte avsnitt. I et annet eksempel ifølge oppfinnelsen kan hver påfølgende setning i et avsnitt vises ved å begynne enten ved en mer høyrerettet posisjon, en mer nedadrettet posisjon, eller begge. På denne måten vil påfølgende setninger i et avsnitt marsjere enten mot høyre, nedover, eller begge sammen med progresjonen av avsnittet.

I en annen utførelse av oppfinnelsen er en musikalsk nøkkel, spilt av et instrument, forbundet med et avsnitt som skal vises. I ett eksempel er den samme nøkkel og instrument forbundet med hvert avsnitt. Et antall toner tilsvarende antallet setninger i avsnittet spilles nedover skalaen med denne nøkkelen ved begynnelsen av avsnittets presentasjon. Ved begynnelsen av visningen av hver setning spilles den tilsvarende tonen i skalaen opp eller ned skalaen. I et tilsvarende eksempel spilles et antall toner tilsvarende antallet eller posisjonen for den aktuelle setningen. I en utførelse spilles en akkord dersom setningen inneholder mer enn ett uttrykk. I en utførelse kan det spilles en mollakkord eller en disharmonisk akkord dersom setningen inneholder kontrastgivende uttrykk slik som "på tross av" eller "imidlertid". I en utførelse anvendes et annet instrument og/eller nøkkel for hvert avsnitt.

35 Tidsmessig framskyting

I et annet aspekt ved oppfinnelsen velges visse ord for å rage midlertidig fram fra de

andre ordene i setningen. De valgte ordene i ett eksempel vises før de resterende ordene i setningen blir vist. I et annet eksempel vises de valgte ordene etter at de resterende ordene ikke lenger er synlige. I nok et eksempel vises de valgte ordene både før og etter at de resterende ordene er vist. Visningen før og etter visningen av de resterende ordene er antatt å motta mere oppmerksomhet fra leseren. Ord som vises før de resterende ordene antas å bli registrert og delvis analysert av leseren før det gjenværende av setningen leses, dette antas å gjøre leseren oppmerksom på ordet slik at leseren ikke blir usikker og må gå tilbake.

I en utførelse velges ordene basert på den vanskelighetsgrad som en forventer at leseren vil ha med de valgte ordene. I ett eksempel velges ord over et visst utdanningsnivå for midlertidig framskyting. I et annet eksempel velges ord med en lyd-stavende misstilpasning for midlertidig framskyting. En lyd-stavende misstilpasning viser til ord hvis uttale ikke kan avledes fra stavingen. I nok et eksempel brukes ei ordbok med ordfrekvens, og ord som brukes med lav hyppighet velges for midlertidig framskyting.

Prosa-maling

I et annet aspekt ved oppfinnelsen anvendes farge sammen med tekstanalyse for å vise det innbyrdes forhold mellom deler av ei setning med en annen. En utførelse av denne framgangsmåten kan omfatte trinnene angitt nedenfor.

a. Analysere setningen, inndele setningen i et flertall tekstsegmenter basert på de innbyrdes forhold mellom tekstsegmentene.

b. Organisere disse tekstsegmentene i et tre med en rotnode og et flertall foreldrenoder og barnenoder.

c. Tilordne hver barnenode en attributt basert på forholdet med foreldrenoden.

d. Tilordne rotnoden en farge.

e. For hver barnenode under rotnoden, tilordne en farge til hver barnenode basert på fargen på foreldrenoden modifisert av forholdet mellom barnenoden og foreldrenoden.

f. Gjenta trinn e inntil alle noder er tilordnet en farge.

g. Vise teksten i linjer i hovedsakelig samme rekkefølge som spesifisert av forfatteren ved bruk av fargetilordningen i treet.

Med henvisning til Figur 27 velges, i en utførelsesform av algoritmen skissert foran, foreldre-barn-forholdene fra ei gruppe omfattende entitet, aksjon og modifikasjon. Ei opprinnelig setning 300 er underinndelt i tre tekstsegmenter 302, 304 og 306, som indikert ved et tre 301. Tekstsegmentet 302 har et entitets-forhold til den overordnede 300,

segmentet 304 har et aksjonsforhold til den overordnede 300, og segment 306 har et aksjonsforhold til den overordnede 300. Den overordnede 300 har i en generasjon 0 ved 308 en initial fargetrippel (Rød Grønn Blå) lik R00-G00-B00, som indikert i segment 300. Mengden farge addert til hver generasjon av barnesegmenter kan være en funksjon av selve generasjonstallet. I Figur 27 er 100 fargeenheter addert til en første barnegenerasjon ved 310, 50 fargeenheter er addert til en andre generasjon 312 og 25 fargeenheter er addert til en tredje generasjon 314. Selve fargen adderes som en funksjon av foreldre-barn-forholdet. I Figur 27 er blått addert for entiteter, rødt for aksjoner og grønt for modifikatorer.

10 I Figur 27 har segment 302 et entitetsforhold med den overordnede 300 og er ved det første generasjonsnivå, så 100 enheter med blått adderes. Mengden farge addert er mindre ved hver generasjon. Funksjonen for hvert ord i setningen er slik sett på det meste influert av funksjonen for det største segmentet som den er en del av.

15 Nye kvaliteter tildelt av prosa-malingsprosessen

Tale er et kommunikasjonssystem der individuelle uttalte enheter, slik som ord, har bestemte meninger, og en større betydning konstrueres av forholdet mellom ordgrupper. Ordsekvensen er en viktig strukturell informasjon som brukes til å bygge betydningen av forholdene mellom ordene. Innenfor ordsekvenser har mindre ordsekvenser spesielle forhold seg i mellom. På denne måten er også rekkefølgen av to eller flere mindre ordsekvenser innenfor en større ordsekvens en form for strukturell informasjon som anvendes til bygging av betydningen av forholdet mellom alle ordene.

20 I tillegg til å velge rekkefølgen som ordene uttales i uttaler også en taler hvert ord eller ordgruppe med visse lydattributter som kommer i tillegg til den enkle uttalen av fonemene som indikerer hva ordet er. Disse ekstra lydattributtene omfatter tonehøyden, volum og klangfarge som ordene uttales med. Disse lydattributtene omgjør og kombinerer ved at et spesielt ord i et uttrykk vil modifiseres i forhold til hva disse ekstra lydattributtene er for uttrykket som helhet.

30 Disse ekstra uttaleattributtene, slik som tonehøyde og klangfarge, (heretter omtalt som intonasjonsattributter) representerer derfor et flerdimensjonalt system med verdier som kan kombineres og som veksler mellom suksessive nivå med differensiering av ordgrupper til mindre delenheter innenfor de større ordgruppene. Disse verdiene representerer forholdet mellom ordet (eller ordene) og den større enheten av ordgruppe som det var en del av.

35 I tillegg holdes hver delenheter innenfor et område med intonasjonsverdier som er egnet for dette umiddelbare tekstinnholdet. En større enhet, som kan inneholde flere delenheter i seg selv, vil tilsvarende ha et forhold til den enda større enheten som inneholder

delenheter lik seg selv, med en endring av en av disse samme intonasjonsverdiene som er egnet for dette større tekstinnholdet.

Eksempelvis en grundig lytting på setningen: "Den rosa hatten for den rosa klovnen ble lagt i ei rosa eske" - vil avsløre at intonasjonsverdiene for hvert av ordene "rosa" skiller ut ordet "rosa" tonalt fra de større enhetene ordet "rosa" modifierer i hvert tilfelle. I det første tilfellet er "den rosa hatten" grunnentiteten i setningen som en helhet: intonasjonsverdien for "rosa" i dette tilfellet er bygget opp fra et basislinjesett med verdier brukt til å representere grunnentiteten.

I det andre tilfellet er uttrykket "for den (rosa) klovnen" i seg selv et segment som har et modifierende forhold til hele uttrykket. Denne andre "rosa" har derfor et lavere volum enn den første "rosa" men har fremdeles en tonehøyde som skiller den ut på en tonal måte fra "klovn", på samme måte som den første "rosa" skilte seg selv ut fra "hatt".

I det tredje tilfellet varieres hele uttrykket "ble lagt i en rosa boks" i en annen tonal dimensjon, (klangfargen er skarpere) for å skille den ut fra den første halvdel av setningen, som en måte å skildre at dette segmentet er i en "aksjons"-dimensjon i forhold til setningen. Del-enhetene i dette uttrykket beholder denne basislinje-klangfarge dimensjonsverdi og modifierer den deretter ved å endre intonasjonsverdiene for hver del-enhet basert på forholdet mellom del-enheten og det større segmentet de var en del av. På denne måten har fremdeles den tredje "rosa", i den mindre tekstsammenhengen for dens forhold til "den (rosa) eska", en tonehøyde som skiller den tonalt ut fra "eske", men den bærer også en litt skarpere klangfarge enn de første to "rosa"-ene fordi den også er i "aksjons"-dimensjonen som en del av "ble lagt i ei (rosa) eske".

For å gi en ytterligere illustrasjon på denne intonasjonsstrukturen skal en legge merke til hvordan ordet "hatt" har to ulike sett med intonasjonsverdier i den etterfølgende setningen: "Den rosa hatten for den rosa klovnen ble lagt i hatteeska". Den andre "hatt" i denne setningen har et sett med intonasjonsverdier som betegner dens modifierende egenskaper for en entitet ("eska", objektet) i aksjonssegmentet ("ble lagt i") i setningen. I motsetning til dette har den første "hatt" et sett med intonasjonsverdier som plasserer den som grunnentiteten i setningen som helhet. Den andre "hatt" i dette andre eksemplet har faktisk et sett med tonehøyde, tone og volum -verdier som er ekvivalent med den tredje "rosa" i den første eksempelsetningen.

Disse dimensjonene med intonasjon i tale, som angitt foran, har kombinerbare og utbyttbare verdier. Disse dimensjonene med intonasjon er også transparente ved at de oppfattes samtidig med ordenheten som bærer den individuelle ord-givende informasjon: hvert uttalt ord bærer derfor en pakke med både semantiske og syntaktisk informasjon. På denne måten er dimensjonene med intonasjon ikke-kommenterbare; det vil si at de ikke

krever tilleggsytring. En person kommer for eksempel ikke med et utsagn foran hvert uttrykk i ei setning: "Subjekt-den rosa hatten-modifikator for subjekt-den rosa klovn-verb-ble lagt-i-objekt-den rosa eska".

5 De eksisterende framgangsmåtene for å representere forholdet mellom ordgrupper i ei setning, slik som analysetrær eller grammatikalske diagrammer, er ikke transparente men er i stedet avhengig av ekstra kommentarer og symboler for å merke ordgruppene. Disse merkene er igjen fikserte og verken kombinerer eller veksler verdier over påfølgende nivå for differensiering av del-enheter, eller begrenser slike verdier i henhold til deres tekstinnhold.

10 Det er derfor ønskelig at forsterket tekstpresentasjon bør berike framturen av de aktuelle ordene i teksten med visuelle attributter som likner på intonasjonsattributtene i talen. Disse visuelle attributtene bør skape et flerdimensjonalt system av kombinerbare og vekslbare verdier som representerer attributter for ordene i tillegg til deres ordspesifikke betydninger. Disse visuelle attributtene bør også være transparente, dvs. merkbare
15 samtidig med persepsjonen av selve ordene.

Med mindre en tekstpresentasjonsmetode systematisk sikrer at den forfatter-spesifiserte sekvensen av ord representeres utvetydig, er det dessuten ikke tekstpresentasjon i betydningen framsatt i foreliggende beskrivelse, men ganske enkelt en framgangsmåte for å demonstrere resultatene fra noen smale aspekter ved tekstanalyse som underordner
20 integriteten i forfatterens beskjeder til formålene med en slik analyse.

"Prosa-malings"-prosessen beskrevet foran representerer på en unik måte disse intonasjonstilstandene i tale med en utvetydig tekstpresentasjon. I tillegg er "prosa-male"-prosessen en ny tilnærming til syntaktisk analyse i seg selv ved at egenskapene ved entitet, aksjon og modifikasjon, som representerer forholdet mellom barne- og foreldre-
25 segmenter, også kombinerer og veksler mellom suksessive nivå med differensiering av ei setning.

Det er med andre ord ikke bare det at forsterket visuell presentasjon av tekst bør streve etter tilstandene i talen men at syntaktisk analyse i seg selv bør etterstreve tilstandene i intonalitet. En grunnleggende øvelse for testing av grammatikalske regler i språkanalyse er
30 å lage en rekke ord ved bruk av grammatikalske regler og deretter lytte til den foreslåtte setningen for å bestemme om den "lyder riktig". Som illustrert i diskusjonen foran rundt intonasjonsverdier, kan det indre "resonnementet" med hvilket "komplette setninger" intuitivt gjenkjennes som gyldige forslag formidles av et flerdimensjonalt system med attributter for forholdet mellom del-enheter og større enheter hvis eksterne uttrykk er
35 intonalitet. Matematikken i tanken kan være musikalsk.

Framgangsmåten kan anvendes til å vise tekst ved enten å bruke tekstbrytning som

beskrevet foran eller ved bruk av vanlig tekstvisning, men i farger. Framgangsmåten kan også stoppe begrensningen med segmentering av ei setning i ord. For eksempel kan algoritmen stoppe etter et fast antall nivå eller etterlate uttrykk intakte, og etterlate alle ordene i uttrykket med samme farge.

5

Appendiks 1

b10 etter

b10 selv om

b10 så mye som

10

b10 fordi

b10 hvordan

b10 dersom

b10 om bare

b10 ettersom

15

b10 i tilfelle

b10 for at

b10 bortsett fra

b10 i mangel av

b10 på tross av

20

b20 men

b20 for hvem

b20 i hvilken

b20 for hvem

b20 over hvilken

25

b30 jeg

b30 han

e40 kan

e40 kan ikke

e40 kunne

30

e40 kunne ikke

e40 jeg vil

e40 han vil

e40 det vil

e40 jeg gjorde

35

e40 han gjorde

c40 må

- e40 måtte
c40 gjør
c40 ikke gjøre
40 gjør ikke
- 5 e40 gjorde
e40 gjorde ikke
c41 har
e41 jeg har
e41 har ikke
- 10 d41 hadde
d41 hadde ikke
g41 har
c42 være
d42 vært
- 15 g42 er
e42 jeg er
e42 han er
e42 det er
e42 er ikke
- 20 e42 jeg er ikke
e42 han er ikke
c45 våken
d45 våknet
g45 våkner
- 25 e43 vekker
e43 var våken
e43 åt
c45 forlate
d45 forlatt
- 30 g45 forlater
c44 avholde
d44 avholdt
g46 avholder
c43 akseptere
- 35 d45 aksepterte
g45 aksepterer

- c43 ledsage
- d45 ledsaget
- g45 ledsages (av)
- c43 ledsager (noe)
- 5 c43 regne med
- d45 regnet med
- g45 regner med
- c43 akkumulere
- d45 akkumulert
- 10 g45 akkumulerer
- c43 anklage
- d45 anklaget
- g45 anklager
- c45 opptre
- 15 d45 opptrådte
- g45 opptrer
- c43 tillegge
- d45 tilla
- g45 tillegger
- 20 c43 justere
- d34 justere
- g45 justering (under --)
- e43 justerer
- c43 tilpasse
- 25 d45 tilpasset
- g45 tilpasning (under --)
- c43 tilpasser
- c43 tilrå
- d45 tilrådt
- 30 g45 tilråing
- e43 tilrår
- c43 bekrefte
- d45 bekreftet
- g45 bekrefter
- 35 e43 bekrefter
- c46 aggregere

- d46 aggregerte
- g46 aggregering (under --)
- e46 aggregerer
- c44 være enig
- 5 d44 ble enig
- g46 blir enig
- c43 gå med på
- d45 gikk med på
- c43 tillate
- 10 d45 tillot
- g45 tillater
- c43 ta hensyn til
- c45 tok hensyn til
- g45 tar hensyn til
- 15 c43 endre
- d45 endret
- g45 endrer
- c45 svar
- d45 svarte
- 20 g45 svarer
- c45 svarer
- c44 opptre
- d44 opptrådte
- g46 opptrer
- 25 c43 anvende
- d45 anvendte
- g45 anvender
- c44 ankomme
- d44 ankom
- 30 g46 ankommer
- c43 kommer til
- d45 kom til
- g45 kommer til
- c43 spørre
- 35 d45 spurte
- g45 spør

	c43 spørre etter
	d45 spurte etter
	g45 spør etter
	c43 feste
5	d45 festet
	g45 fester
	e43 fester
	c45 angripe
	d45 angrep
10	g45 angriper
	c43 unngå
	d45 unngikk
	g45 unngår
	c43 avskaffe
15	d45 avskaffet
	g45 avskaffer
	e43 avskaffer
	c43 aktivere
	d45 aktiverte
20	g46 aktiverer
	e43 aktiverer
	c43 tilpasse
	c45 tilpasset
	g45 tilpasser
25	e43 tilpasser
	c45 forsvare
	d45 forsvarte
	g45 forsvarer
	e45 forsvarer
30	c43 lindre
	d45 lindret
	g45 lindrer
	e43 lindrer
	c43 hevde
35	d45 hevdet
	g45 hevder

- e43 hevder
- c43 tilføye
- d45 tilføyde
- g45 tilføyer
- 5 e43 tilføyer
- c43 fordømme
- d45 fordømte
- g45 fordømmer
- e43 fordommer
- 10 c43 autentisere
- c45 autentiserte
- g45 autentiserer
- e43 autentiserer
- c43 tilstå
- 15 d45 tilsto
- g45 tilstår
- e43 tilstår
- b50 både
- b50 enten
- 20 b60 igjen
- b60 som
- b60 vekk
- b60 ved siden av
- b60 alltid
- 25 b60 i løpet av
- b60 ved
- b60 angående
- b60 om
- b70 over
- 30 b70 på tvers
- b70 mot
- b70 langs
- b70 midt iblant
- b70 blant
- 35 b70 rundt
- b70 under

- b70 ved siden av
- b70 inne i
- b70 mellom
- b70 under
- 5 b70 bortenfor
- b75 ved
- b70 inn i
- b75 inn
- b75 for
- 10 b75 fra
- b80 av
- b81 og
- b90 alle
- b90 enhver
- 15 b90 enhver
- b90 noen
- b90 begge
- b90 hver
- b90 enten
- 20 b90 alle
- b90 alle og enhver
- b90 fa
- b90 seg selv
- b90 henne selv
- 25 b90 ham selv
- b90 større enn
- b95 en
- w99 tilrå
- w99 tilrådde
- 30 w99 tilrår
- w99 innrømme
- w99 innrømte
- w99 innrømmer
- w99 foregripe
- 35 w99 foregrep
- w99 foregriper

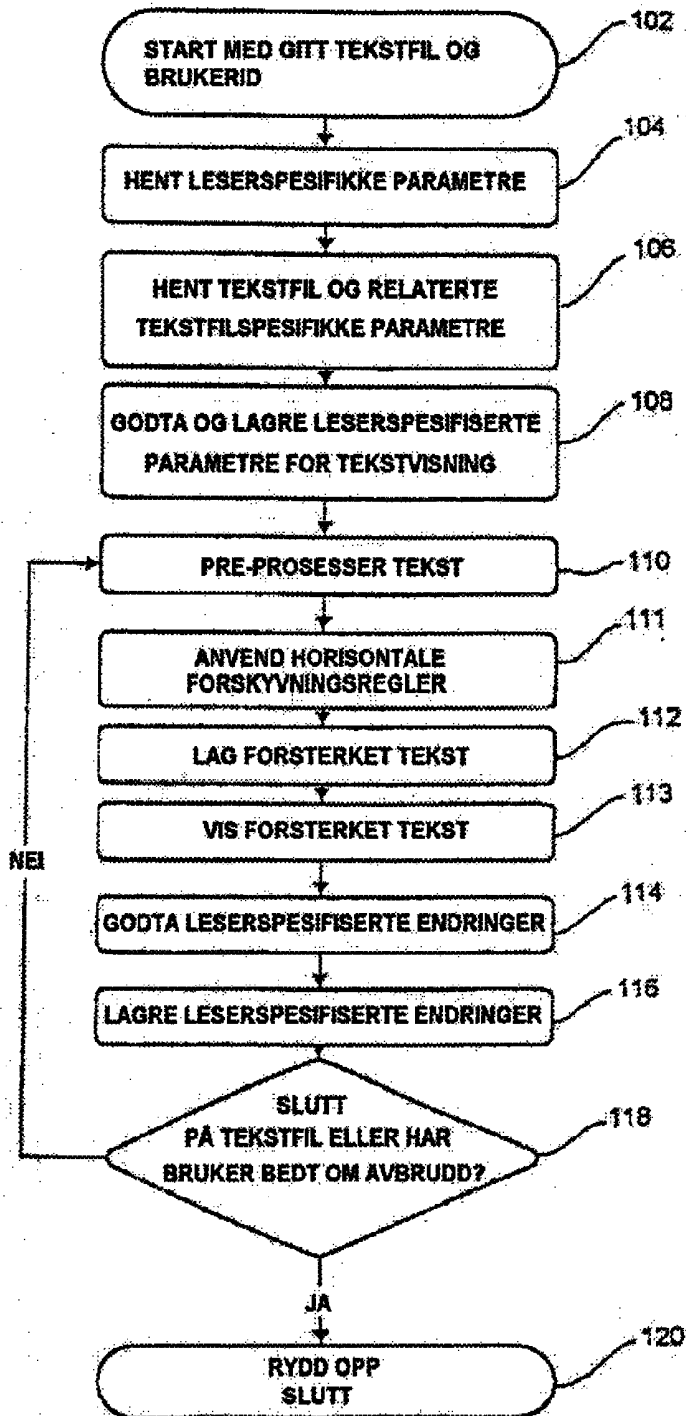
- w99 unngå
- w99 unngikk
- w99 unngår
- x99 jeg
- 5 x99 han
- y99 en/et
- y99 hans
- y99 hennes
- y99 dets
- 10 z99 en/et
- z99 hans
- z99 ham
- z99 henne
- z99 hennes
- 15 z99 det
- z99 dets
- z99 alle
- z99 begge
- z99 alle
- 20 z99 eventuelle
- z99 hver
- z99 alt
- z99 hvem som helst
- z99 eller

Patentkrav:

1. Datamaskinkjørbar leseproduktfremstillende framgangsmåte for å danne en leseproduktvisning av en tekst på naturlig språk som har et forfatterspesifisert karakterinnhold, og en forfatterspesifisert ordrekkefølge inne i karakterinnholdet, hvori leseproduktvisningen omfatter bare et forfatterspesifisert karakterinnhold, **karakterisert ved** at nevnte framgangsmåte omfatter:
 - a) ekstrahere tekstspesifikke attributter fra nevnte tekst på naturlig språk, hvor nevnte tekstspesifikke attributter omfatter et mål for tekstens vanskelighet, og
 - b) variere vertikal og horisontal plassering for tekstdeler av karakterer av karakterinnholdet for nevnte tekst på en visningsflate i henhold til nevnte attributter og samtidig opprettholde den nevnte forfatterspesifiserte ordrekkefølge i visuelle visningsmodeller for nevnte karakterposisjoner, hvor nevnte tekstpresentasjon omfatter en automatisk tekstforflyttingshastighet og variere en hastighet for nevnte tekstpresentasjon i samsvar med nevnte mål for tekstens vanskelighet.
2. Framgangsmåte ifølge patentkrav 1, **karakterisert ved** at målet for tekstens vanskelighetsgrad omfatter en estimert uttaletid for nevnte tekst.
3. Framgangsmåte ifølge patentkrav 1, **karakterisert ved** at målet for tekstens vanskelighetsgrad omfatter et estimert utdanningsnivå for teksten.

1/27

Fig. 1



2/27

Nr.	Element	Type	Skille-tegn	Ordpos.	Ordpos.	Stor forb.	Tegnant.	Ant. stavelser	Uttaleid (s)	Utdanningsnivå
1	I	ord		pron.		Ja	1	1	.5	1
2		blank								
3	had	ord		verb			3	1	.25	1
4		blank								
5	a	ord		art.			1	1	.25	1
6		blank								
7	farm	ord		N-V-A	subst.		4	1	.75	1
8		blank								
9	in	ord		prep			2	1	.25	1
10		blank								
11	Africa	ord		subst.		Ja	6	3	1.0	3
12	,	skillet.	komma				1		.5	
13		blank							.5	
14	at	ord		prep			2	1	.25	1
15		blank								
16	the	ord		art.			3	1	.25	1
17		blank								
18	foot	ord		N-V-A	subst.		4	1	.5	1
19		blank								
20	of	ord		prep.		Ja	2	1	.25	1
21		blank								
22	the	ord		art.			3	1	.25	1
23		blank								
24	Ngong	ord		subst.		Ja	5	2	1.0	12
25		blank								
26	Hills	ord		subst		Ja	5	1	.75	2
27	.	skillet.	punktum						.5	

Fig. 2

Fig. 3

Den geografiske posisjonen
og høyden på landskapet,
kombinert for å skape et landskap
som ikke hadde sitt like
i hele verden

3/27

Fig. 4

Den geografiske posisjonen,
og høyden

på landskapet,

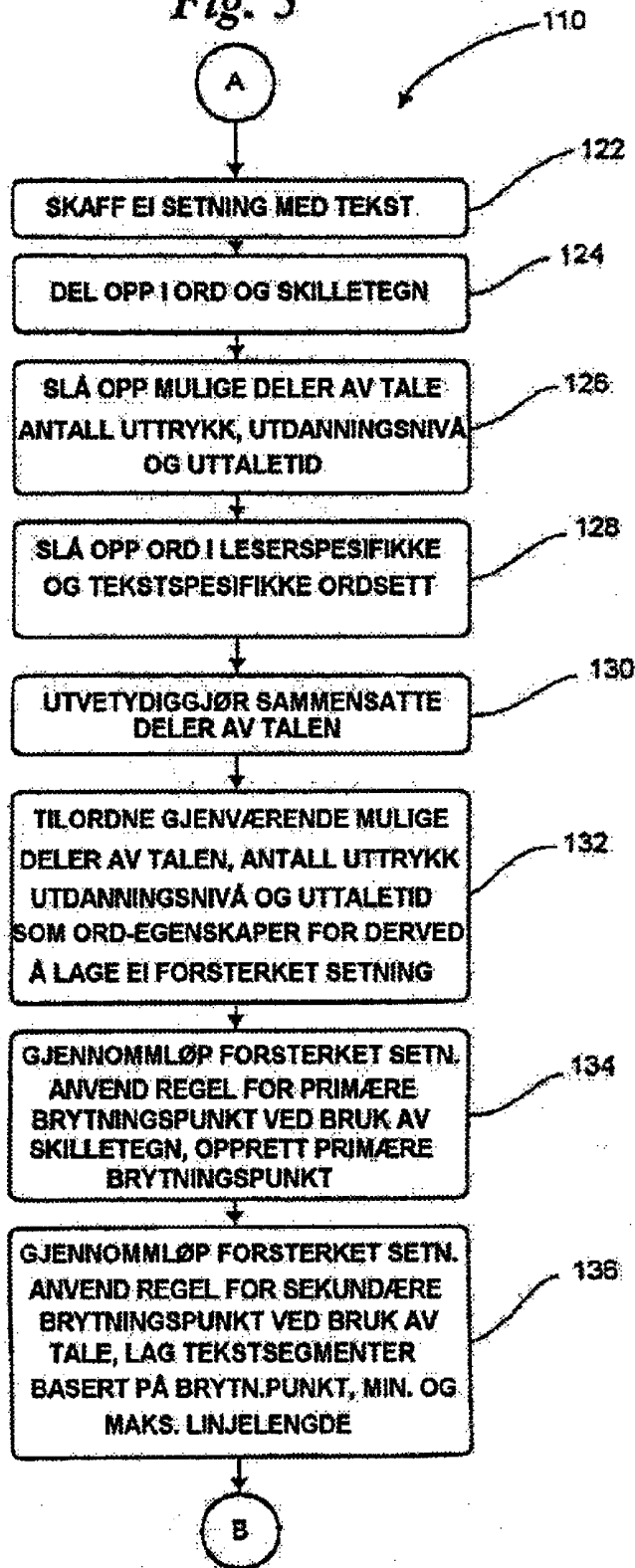
kombinert for å skape et landskap
som ikke hadde sitt like

i hele verden

4/27

5/27

Fig. 5



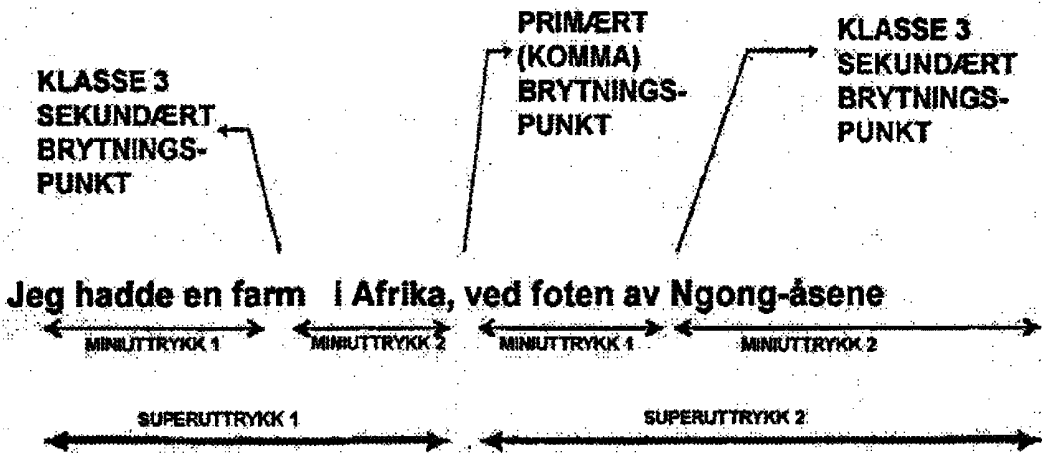
6/27

Fig. 6

Klasse 1, subklasse 1 Preposisjoner	midt iblant, dessuten, mot, blant, om, ved (med verbalsubstantiv), under, siden, helt gjennom, innen, til, langs, bortsett fra, rundt, gjennom, inntil, uten
Klasse 1, subklasse 2 Konjunksjoner	selv om, som om, med mindre, inntil, som om, slik at, for å, fordi, men
Klasse 1, subklasse 3 Relative pronomen	til den, til enhver
Klasse 1, subklasse 4 Andre	heller enn, i mangel av, i stedet for
Klasse 2, subklasse 1 Preposisjoner	over, før, nedenfor, mellom, opp, på tvers av, mot, på, etter, under, til (med verb)
Klasse 2, subklasse 2 Konjunksjoner	når, hvor, når som helst, hvor som helst, mens, og (med restriksjoner), eller (med restriksjoner), dersom
Klasse 2, subklasse 3 Relative pronomen	hva, hvem, hvilket, hvis
Klasse 2, subklasse 4 Andre	substantiv/verb-grensesnitt, adverb som følger et verb
Klasse 3, subklasse 1 Preposisjoner	for, ved, til (med ikke-verb), med, av, i, av (uten verbalsubstantiv), fra, på

7/27

Fig. 7



8/27

Fig. 8

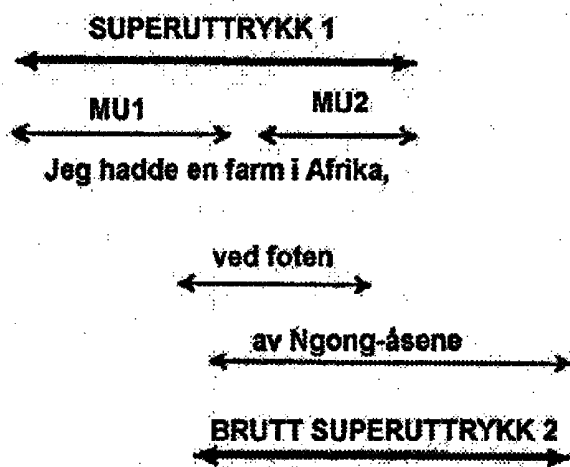
```
// SU = Super-uttrykk
// SBR = Sekundær brytningsregel
// MIN = minimal linjelengde
// MAX = maksimal linjelengde

FOR alle SU i setningen
  FOR alle SBR-nivå, fra lav til høy
    FOR alle ord i SU, fra venstre til høyre
      IF resulterende tekstsegment vil være > MIN
        anvend SBR, lag nytt tekstsegment
      ELSE
        ikke anvend SBR for derved \
        å addere til lengden av neste\
        tekstsegment
      ENDIF
    END for alle ord i SU
  END for alle nivå i SBR
END for alle SU i setning

FOR alle tekstsegmenter laget
  WHILE tekstsegment-lengde er > MAX
```

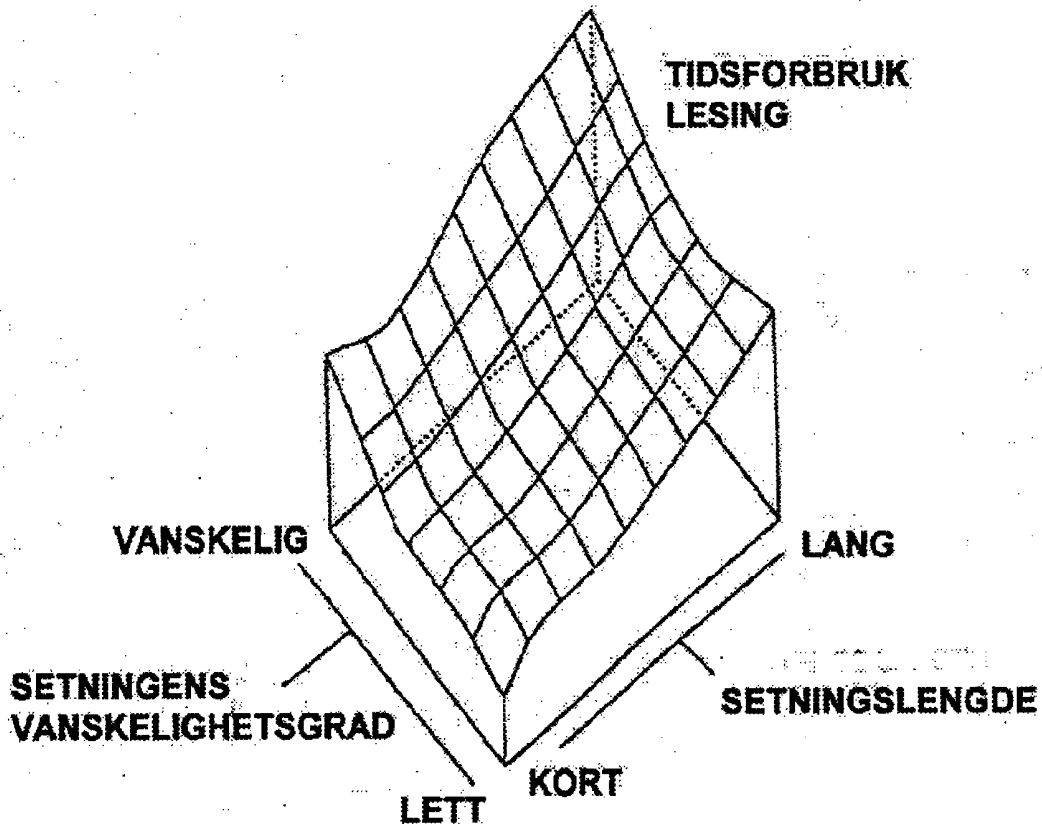
9/27

Fig. 9



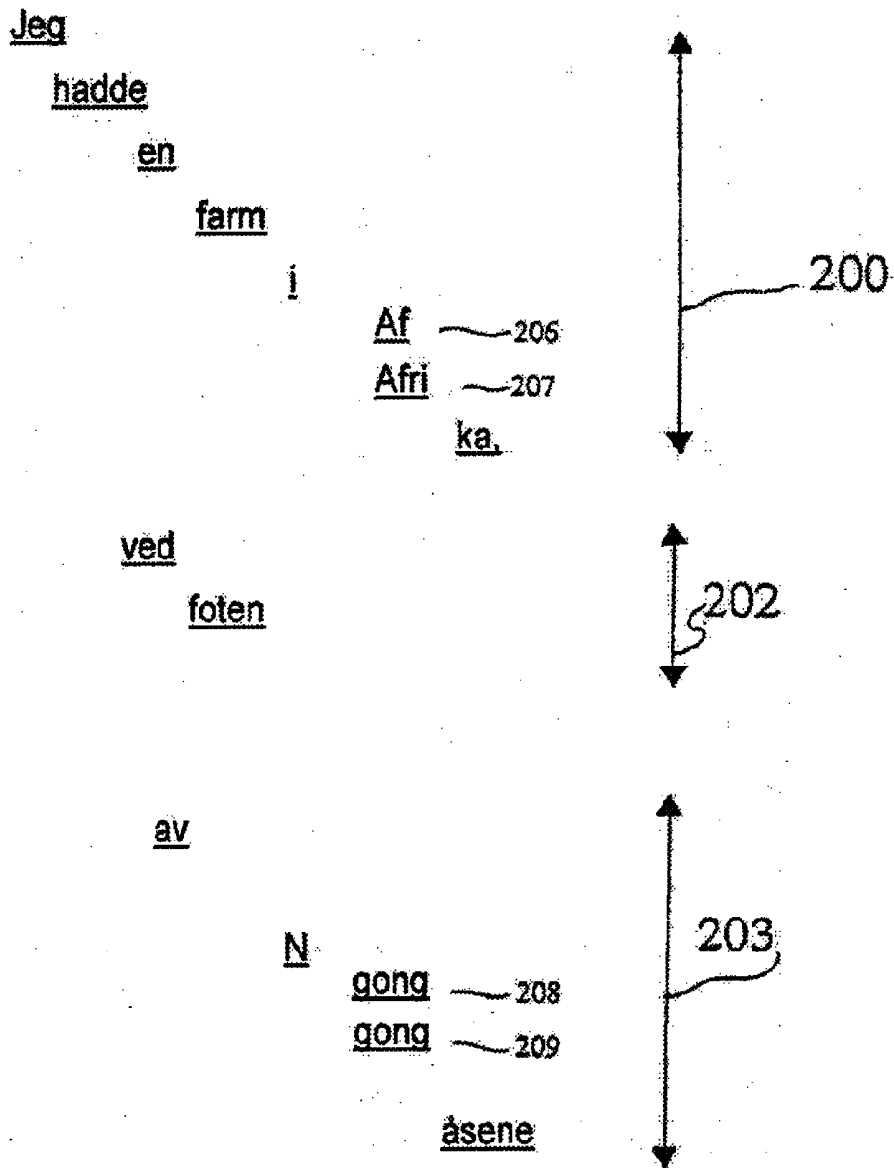
10/27

Fig. 10



11/27

Fig. 11



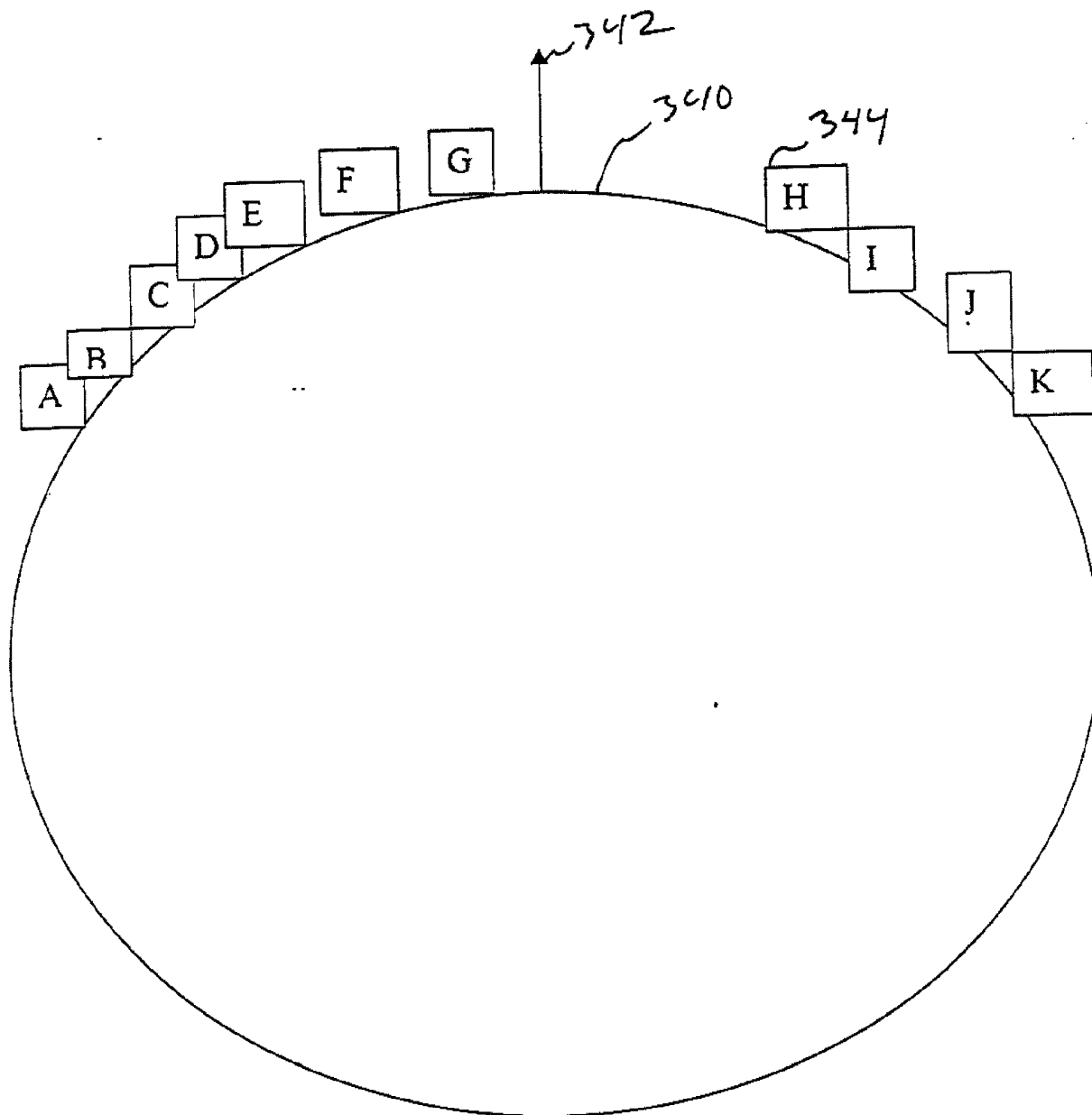
12/27

SP.	320																								
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20						
X	Gate 1					Gate 2					Gate 3					Gate 4					Gate 5				
SPor	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	
b10	Spør 1 bare brukt til skillebegynnelse																								
b20	Spør 1 bare brukt til skillebegynnelse																								
b30	Innretning, se trinn 1-15																								
b40																									
b45	Begynnelse på ord i B; tallet etter ord i ny rad i D																								
b50																									
b60																									
b70																									
b80																									
b85	Begynnelse på ord i B; tallet etter ord i ny rad i D																								
b90																									

324
330
326
328
FIG. 12

13/27

Fig. 13



14/27

Hode		Hale			
	H	I	J	K	
Beskrivelse av omstendigheter ved starten (hode) av segmentet (hale) til høyre	Splitt mellom ord	Komma, dobbelstrek, parentes eller skilletegn ved segmentslutt	Kolon, semikolon eller ellipse ved segmentslutt	Slutt på setning	
A Setningsstart	A-H	A-I	A-J	A-K	
B Kolon, semikolon eller ellipse foran segmentet	B-H	B-I	B-J	B-K	
C Komma, dobbelstrek, parentes eller skilletegn ved starten av eller foran segmentet	C-H	C-I	C-J	C-K	
D Ord b10-b39 i leddordlista	D-H	D-I	D-J	D-K	
E Ord b40-b49 i leddordlista	E-H	E-I	E-J	E-K	
F Ord b50-b70 i leddordlista (verb)	F-H	F-I	F-J	F-K	
G Ord b80-b95 i leddordlista	G-H	G-I	G-J	G-K	

Fig. 14

Fig. 15

Fire snes og sju år siden

A-I

frambrakte våre fedre

C-H

en ny nasjon på dette kontinentet:

F-J

unnfanget i frihet

B-I

og innviet med erklæringen

C-H

om at alle mennesker er skapt like.

D-K

15/27

16/27

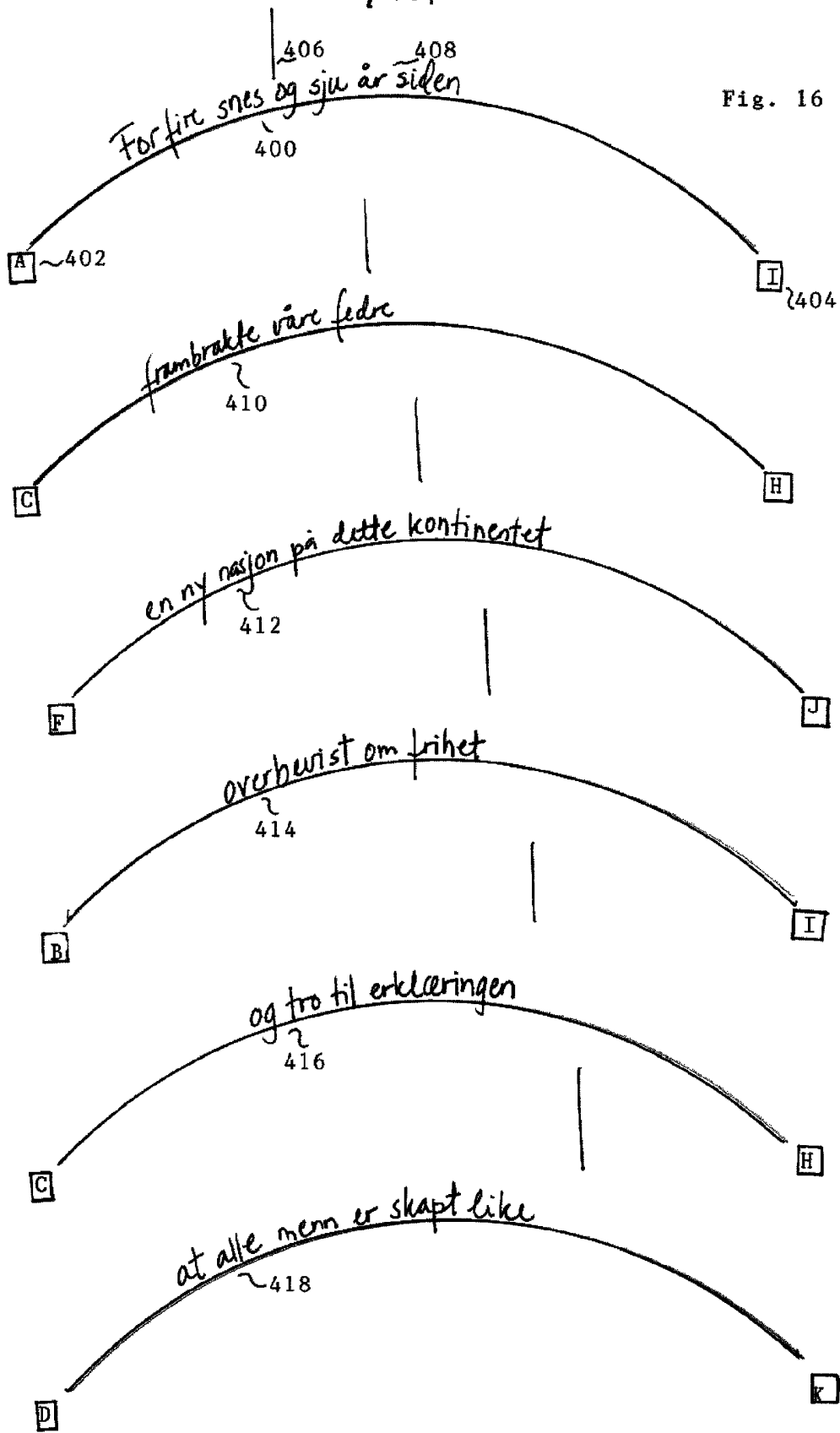


Fig. 16

17/27

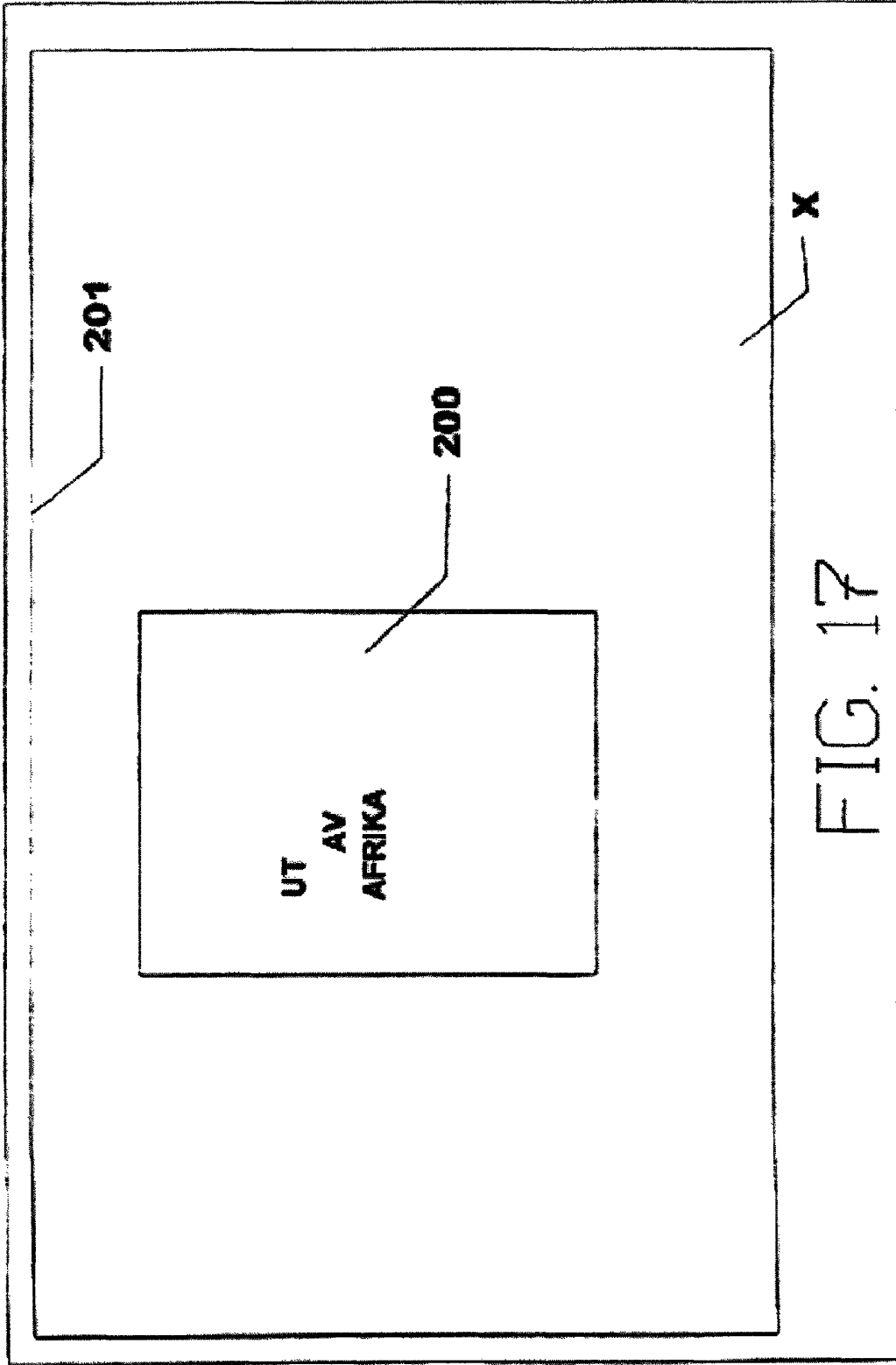


FIG. 17

18/27

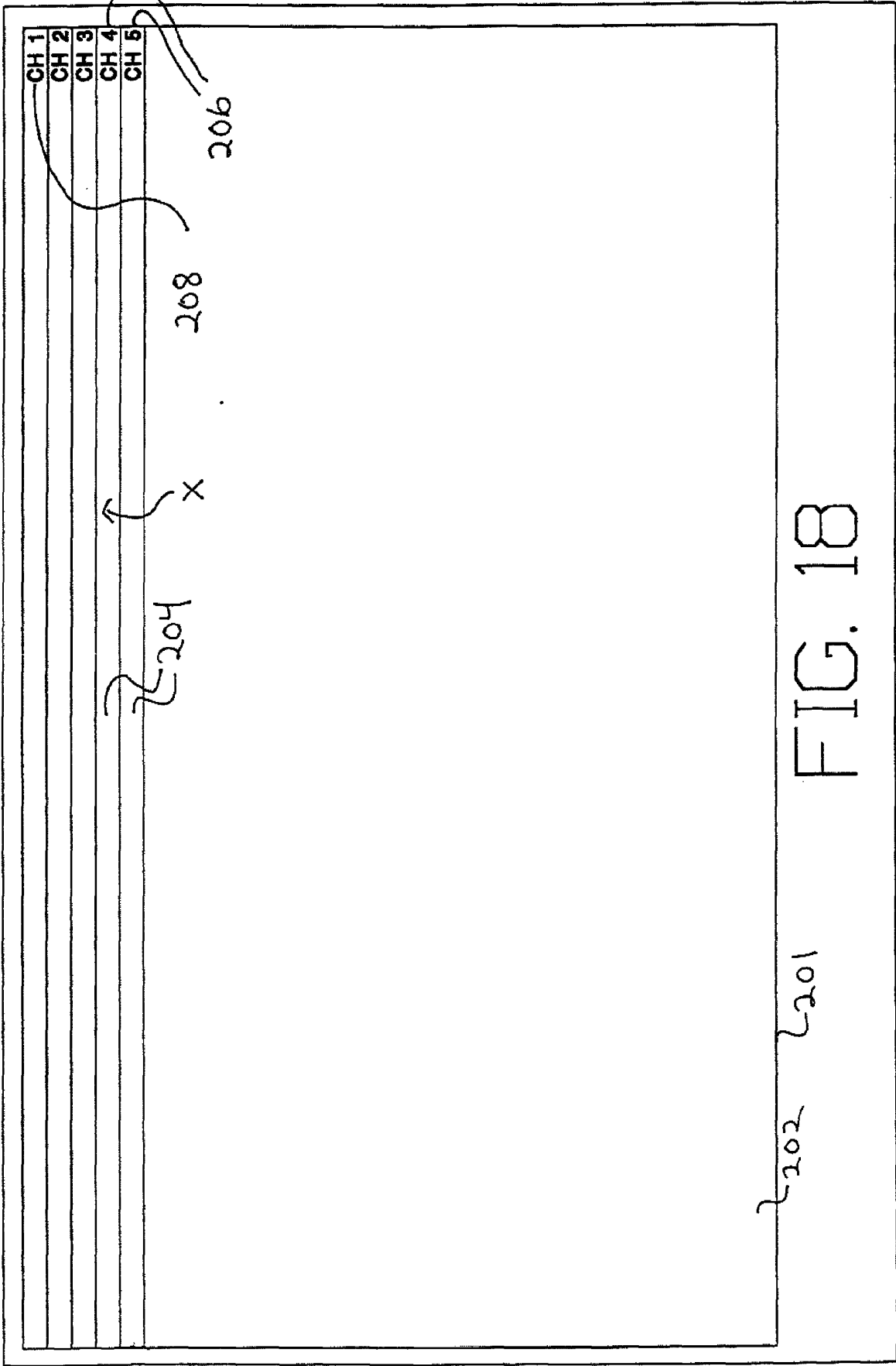


FIG. 18

19/27

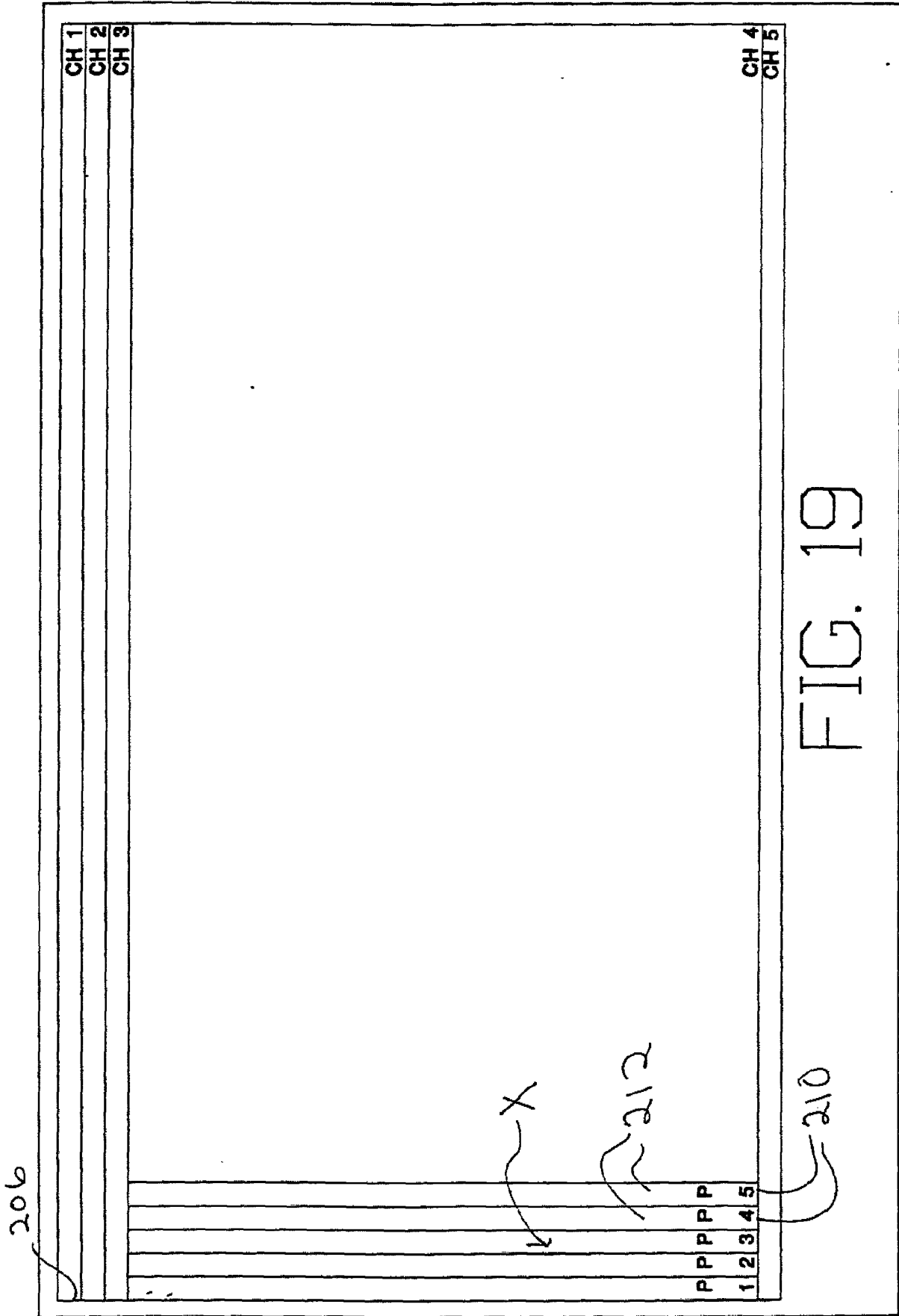


FIG. 19

20/27

206

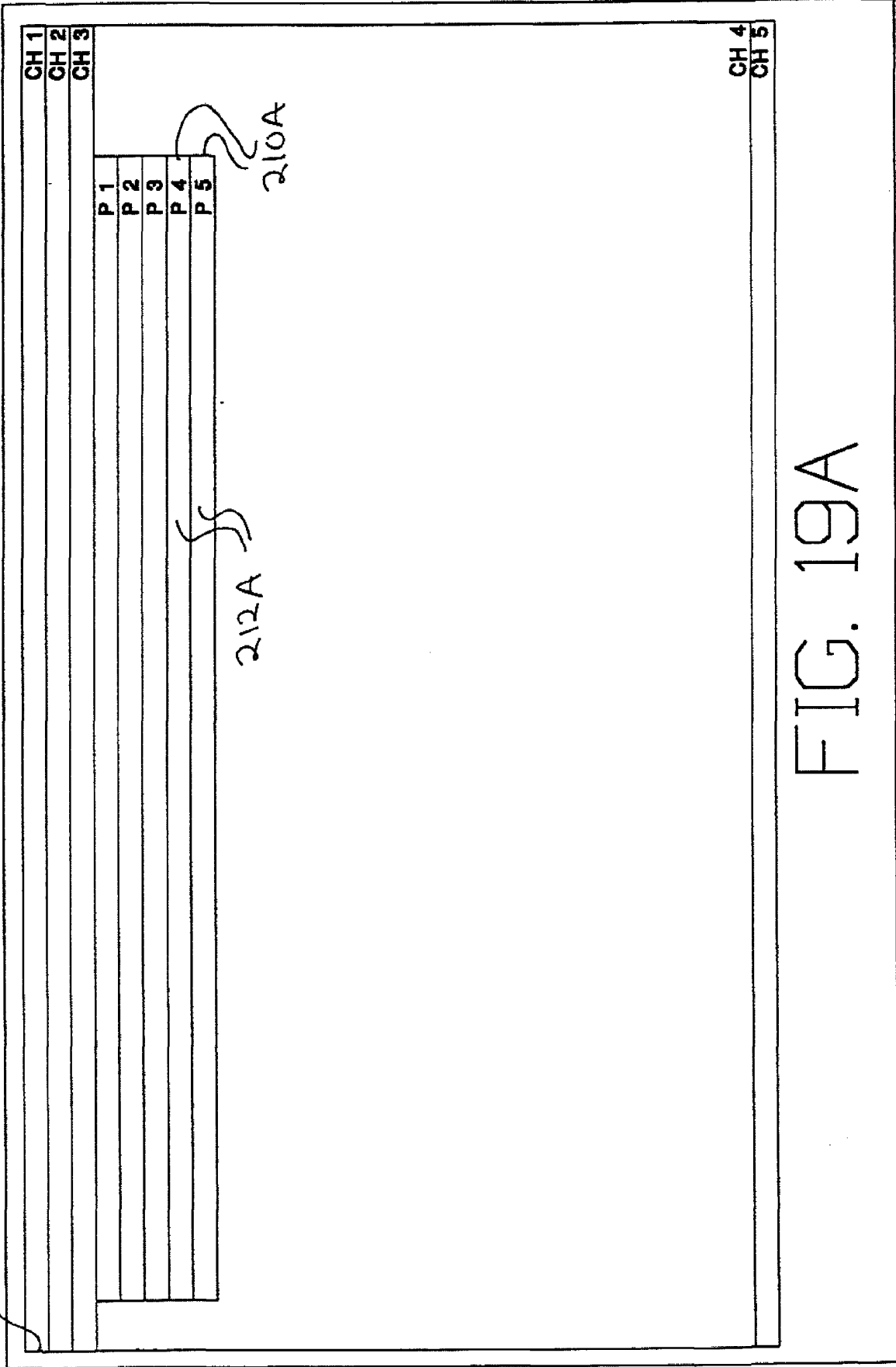


FIG. 19A

21/27

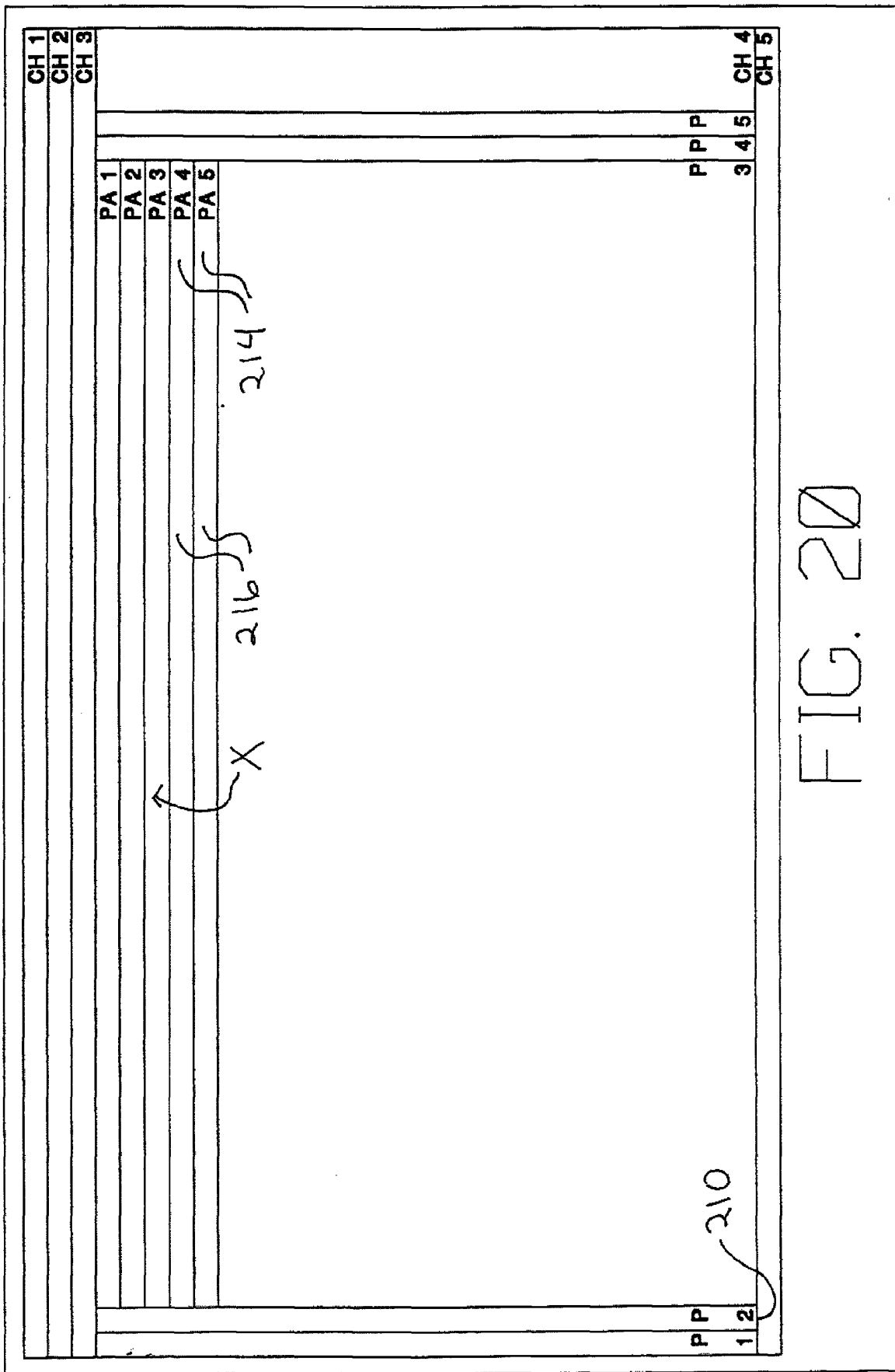
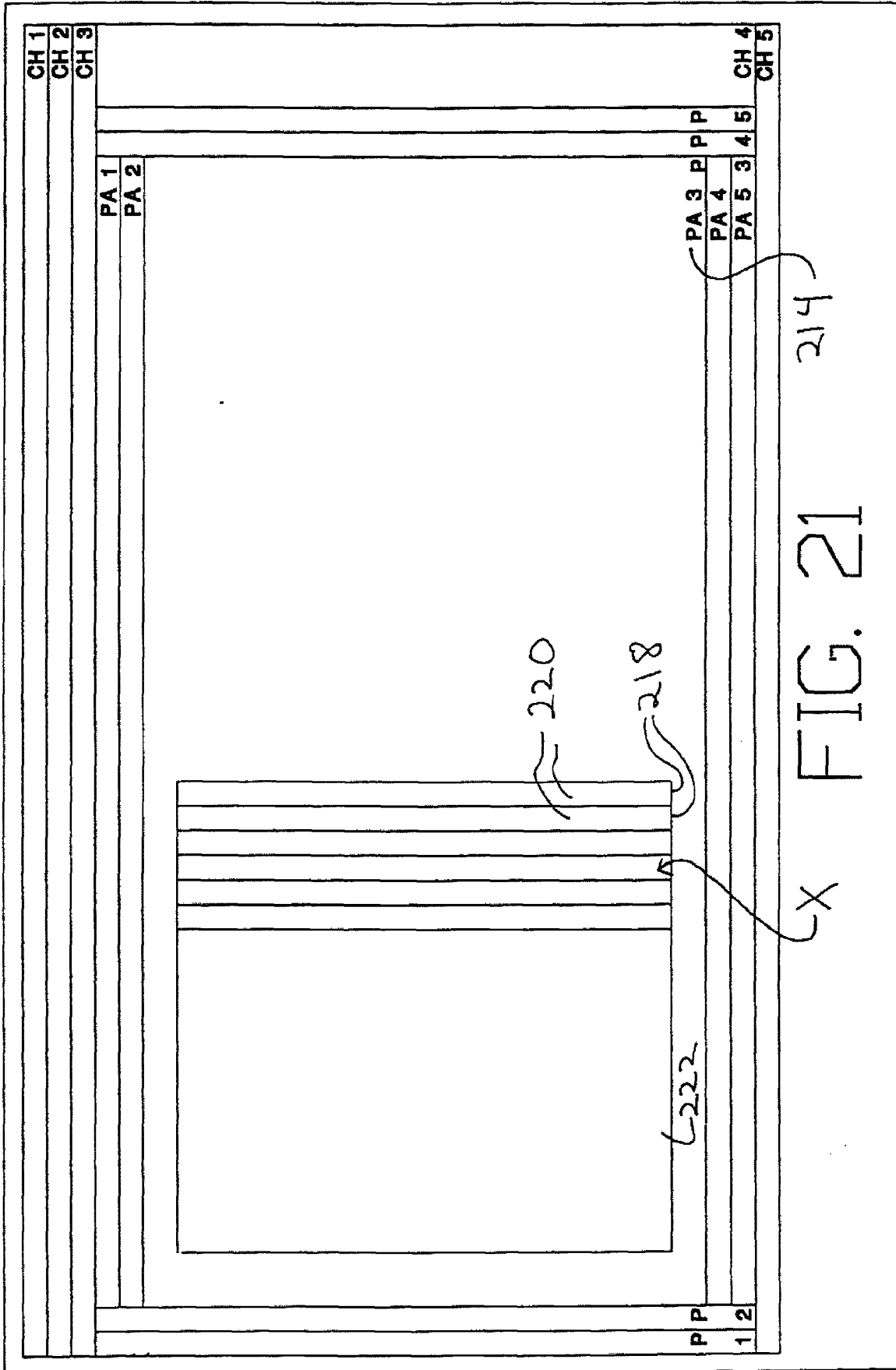


FIG. 20

22/27



23/27

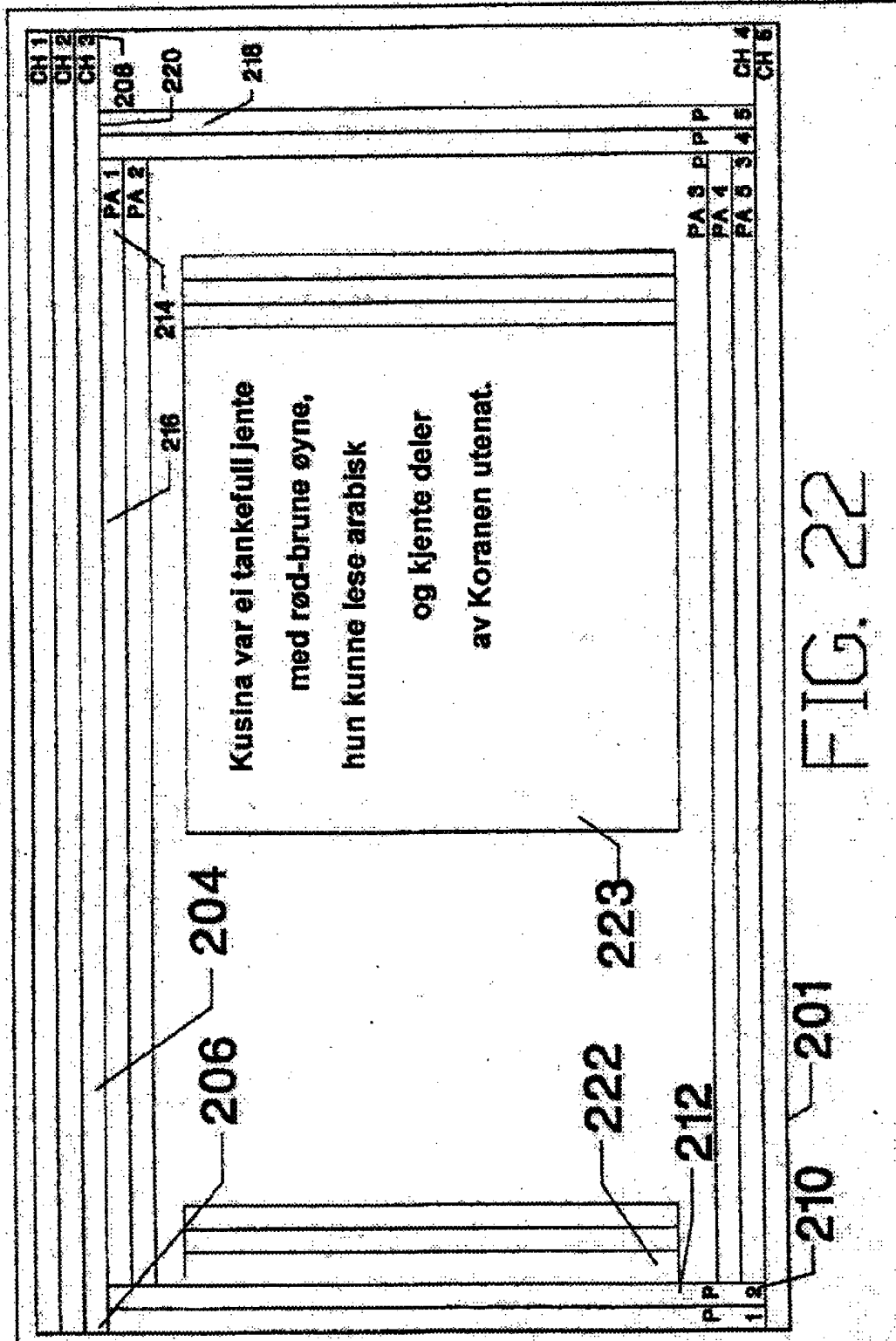


FIG. 22

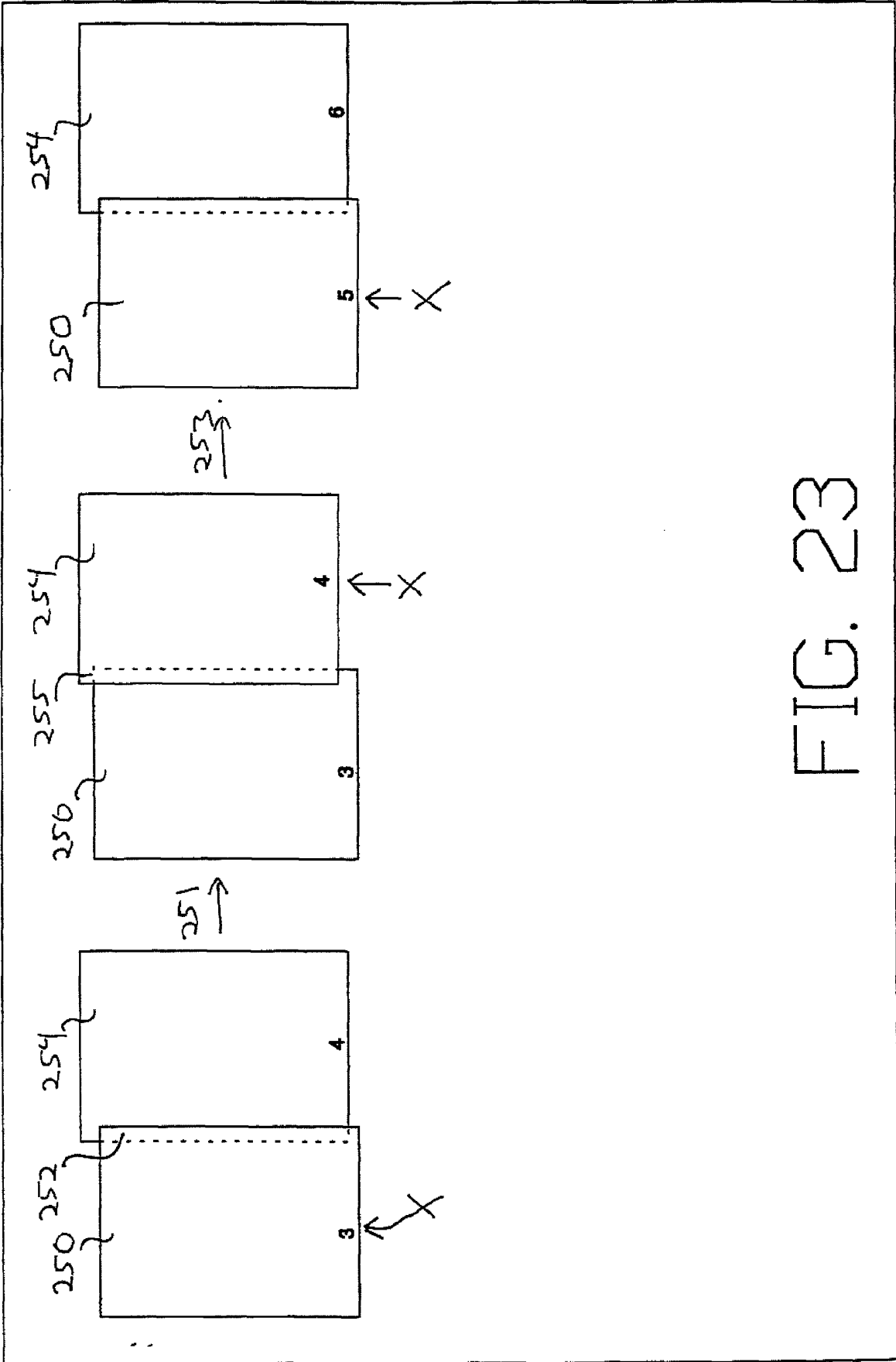


FIG. 23

25/27

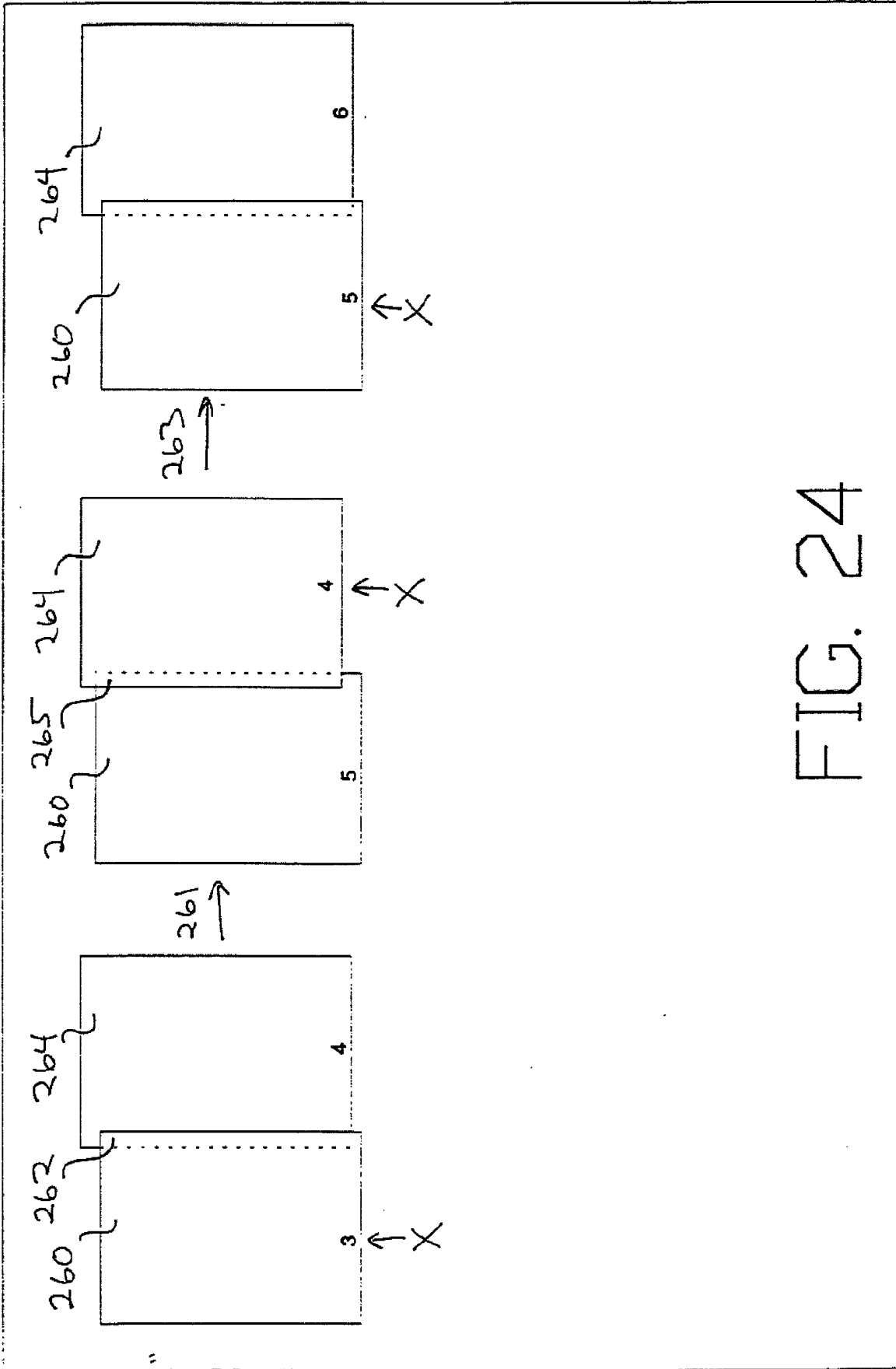


FIG. 24

26/27

280 ~
288 ~ frambrakt
286 ~ våre fedre
284 ~ for sju år siden
282 ~ Fire snes og

FIG. 25

290 ~
På dette kontinent
frambrakt ~ 288
være fedre ~ 286
for sju år siden ~ 284

FIG. 26

27/27

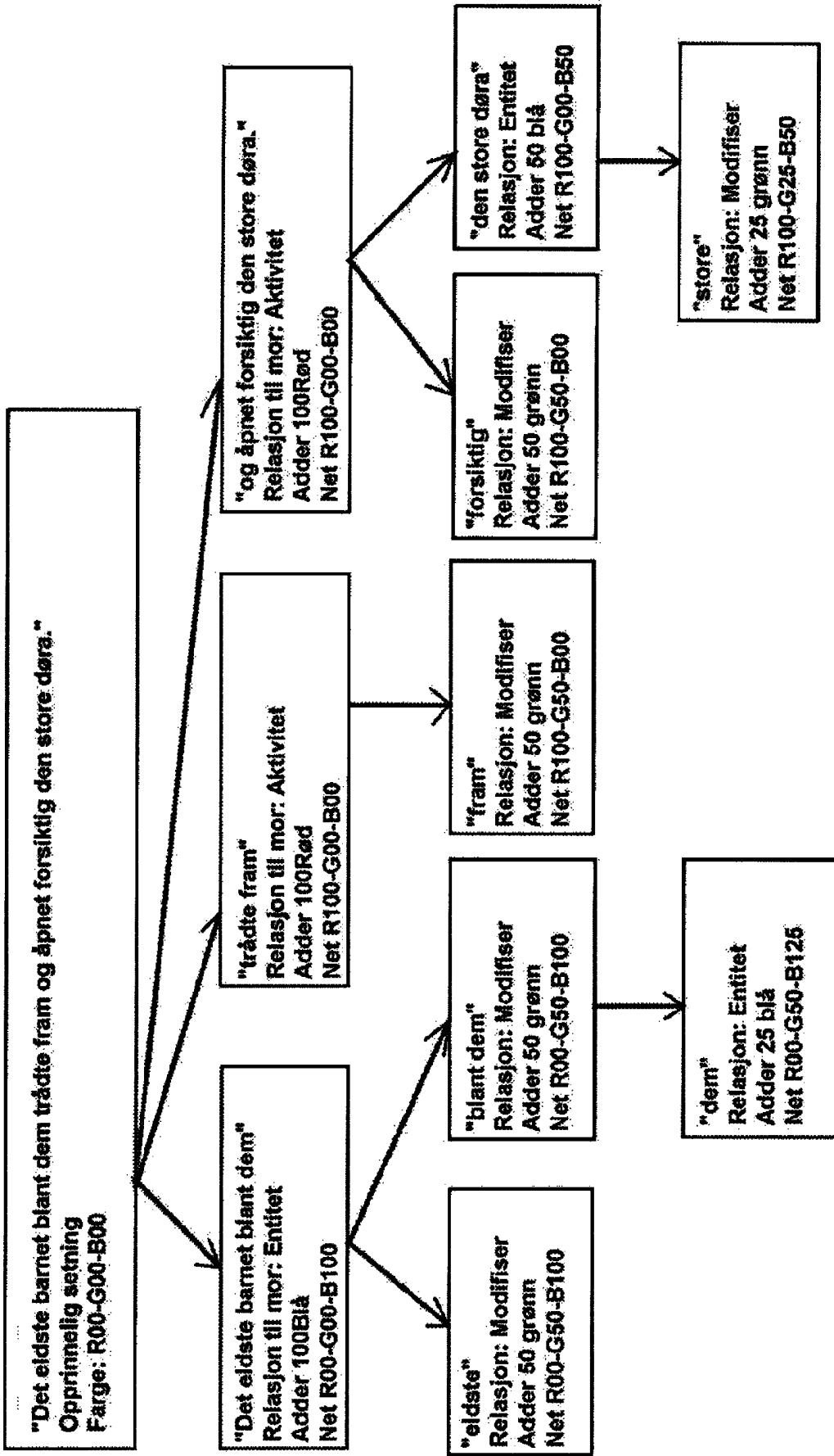


Fig. 27