

och den 23 är inte »emerentius« men Emerentiana, den 25 Paulus apostolus, inte »apostola«. I kalendernamnen s. 44 undrar jag om inte Marie presentatio den 21 november hellre borde kallas »Marie offring« eller »frambärande i templet«.

*Tore Nyberg*

Identifying People in the Past, ed. by *E. A. Wrigley*, Edw. Arnold Ltd., London, 1973, 159 s.

Enhver historiker, der har beskæftiget sig med at undersøge levnedsløbet for personer, der ikke findes omtalt i biografiske leksika eller lignende, har stået over for det problem, at han må sammenstykke den pågældendes livsløb fra en række forskellige kilder. Ikke mindst inden for socialhistorien er dette forhold typisk, når man arbejder med persongrupper fra det 16.-19. århundredes historie.

På et nutidigt materiale vil denne sammenstyknings normalt være ret problemfri, fordi man sjældent vil være i tvivl om, hvorvidt den person, man finder et kildevidnesbyrd om, også er den, man søger. Kommer man imidlertid blot 70-80 tilbage i tiden, savnes hyppigt sikre vidnesbyrd for at kunne opstille et »generalieblad« for den pågældende person, idet det i mange tilfælde kan være tvivlsomt, om der er tale om den samme person, når man har en dåbsregistrering, en vielsesregistrering, et fæstebrev, et skifte o. l., som alle lyder på samme eller omtrent samme navn.

Når man alligevel gør sig store anstrengelser for at finde ud af, om det er den samme person, hænger det sammen med, at man, såfremt en sammenkædning kan foretages, normalt vil kunne få en række nye oplysninger frem ved at kombinere de enkelte kilders indhold, f. eks. kan man finde vielsesalderen, der før et stykke ind i 1800-tallet ikke er opgivet i kirkebøgerne, såfremt man kan finde personens dåb, eller man kan måske se, hvor længe en person har haft et fæste, og hvilke økonomiske forhold, det har ført med sig, hvis man kan kombinere fæstebrev med begravelses- og skifteoplysninger.

Det er derfor et centralt problem for mange forskellige grene af historien, der behandles i den foreliggende bog, hvor man – ud over enkelte overvejelser af mere erkendelsesteoretisk natur – viser forskellige metoder til at foretage sådanne identifikationer, metoder, hvis forskelle dels er dikteret af forskelle i kildematerialet, dels af forskelle i de formål, hvortil man ønsker at anvende den samling af oplysninger om de enkelte personer, som man kan kombinere sig frem til fra de anvendte kilder.

De forfattere, der har bidraget til artikelsamlingen, nævner især tre hovedidentifikationsproblemer, hvor man søger at finde frem til generelle metoder til sammenkædning, idet man ønsker at nå frem til procedurer, der gør det muligt at anvende maskinelle bearbejdningsmetoder for derved at kunne anvende dem på meget store datamængder, hvorved man naturligvis normalt også vil kunne få mere pålidelige resultater.

De tre problemer er, 1) at samme person kan have ændret sit navn gennem tiden, 2) at han anvender skiftende navne, og 3) at der i et materiale findes adskillige personer med samme navn, som kan være vanskelige at udskille fra hinanden.

Det første problem kendes i vid udstrækning i Amerika, hvor indvandreres navne gradvist er blevet amerikaniseret. Hovedmetoden til at foretage en identificering, når det drejer sig om denne type, er ifølge bogen udarbejdelse af en art lydskrift, hvorved de pågældende navne vil se ens ud i den originale udgave og i den amerikaniserede, således at en maskinel sortering automatisk vil kæde personerne sammen. Mindre heldigt er det dog, at en sådan lydskrift vil være specifik for det pågældende sprog og derfor må udarbejdes specielt for hvert sprog, som man ønsker at anvende den på.

Problemer af type 2 og 3, der synes at være mere relevante for danske forhold, behandles mest indgående af E. A. Wrigley og R. S. Schofield, der – med særligt hensyn til familierekonstitutionsstudier – viser, hvordan man kan opstille et katalog af forskellige godkendelseskriterier i form af algoritmer, der kan tilpasses materialet, idet man så kan undersøge, i hvilket omfang disse kriterier dels vil give fejlagtige sammenkædninger, dels betyde, at oplysninger om samme person ikke bliver sammenkædet. Som ved andre statistiske fordelinger gælder det også her, at en variation af godkendelsesbetingelserne giver flere fejl af den ene slags, men samtidig færre af den anden eller omvendt, og opgaven ved enhver undersøgelse er derfor at finde den algoritme, som ved at inddrage omtrentlig alder, evt. erhverv og kendskab til andre familiemedlemmer m. m. giver den optimale fordeling på de to typer af fejl. I praksis betyder dette naturligvis, at der kun kan gives generelle retningslinier for sammenkædningerne og ikke en endegyldig løsning på, hvorledes det kan ske, idet den optimale løsning vil afhænge af, hvilket formål det sammenkædede materiale skal anvendes til. Erfaringerne viser dog, at der ofte vil være en ganske vid margin for variation af godkendelseskriterierne, uden at det påvirker indholdet nævneværdigt, når det drejer sig om at finde typiske egenskaber ved materialet som helhed.

De erfaringer, som anmelderen har med hensyn til at anvende den maskinelle sammenkædning på dansk materiale, viser, at man her nok vil kunne lære en del metodisk af den foreliggende bog, men at man må være indstillet på, at den højeste identifikationsprocent målt i forhold til arbejdsindsatsen opnås gennem en kombination af maskinel og manuel bearbejdelse, hvor den maskinelle del står for grovsorteringen, mens den mere specifikke del sker manuelt. Hovedårsagen hertil er, at det danske navnemateriale synes at være koncentreret på færre navne end noget andet lands, der har været underkastet en maskinel bearbejdelse. Således er det normalt i danske sogne i det 18. århundrede, at de tre hyppigst anvendte fornavne dækker 30–50 % af alle tilfælde inden for hvert køn. Denne store navnelighed har givet anledning til forskellige former for anden navnebrug så som erhvervsbetegnelser og stednavne, som er stærkt svingende fra sogn til sogn med det resultat, at den mest anvendelige algoritme til finsorteringen for det enkelte sogn først vil kunne opstilles, når man gennem den manuelle bearbejdelse har fået foretaget den optimale sammenkædning.

*Hans Chr. Johansen*