



**LUNDS**  
UNIVERSITET

INSTITUTIONEN FÖR PSYKOLOGI

En jämförelse av elimineringskonfrontation och sekventiell  
konfrontation beträffande öronvittnen, avseende konfidens och  
tillförlitlighet

**Mattias Sjöberg**

Kandidatuppsats HT 2011

Handledare: Mats Dahl

## Abstract

Syftet med studien var att jämföra två olika vittneskonfrontationsmetoder för öronvittnen; elimineringskonfrontation, som går ut på att eliminera de röster som låter minst lik målrösten tills det bara återstår en röst, samt sekventiell vittneskonfrontation där ett eventuellt utpekande görs direkt utan att eliminera några röster. Försökspersonerna testades i ett experiment som bestod av en instuderingsfas, där de fick höra en röst, en filler uppgift och slutligen en identifikationsuppgift där antingen en elimineringskonfrontation eller en sekventiell konfrontation användes. Elimineringskonfrontationen förväntades leda till färre felaktiga identifikationer jämfört med den sekventiella konfrontationen. Resultatet visade att elimineringskonfrontationen inte gav upphov till färre felaktiga identifikationer, samt att de båda konfrontationsmetoderna framkallade liknande utfall gällande korrekta responser och konfidens. Resultatet kan bero på ett för litet och snedvridet urval av försökspersoner. Undersökningen visar därmed att det inte förefaller någon större skillnad mellan de två konfrontationsmetoderna och att vårt minne för röster är mycket otillförlitligt oavsett konfrontationsmetod.

Nyckelord: Öronvittne, Röstidentifikation, Sekventiell vittneskonfrontation, Elimineringskonfrontation, Konfidens.

## Introduktion

Söndagen den 4 september 2011 tvingas en 23-årig man under pistolhot in i en ljus silverfärgad bil av tre maskerade rånare som kör honom till en bankomat i närheten där 23-åringen tvingas ta ut pengar till rånarna. Efter att ha tagit hans plånbok, pengar och mobiltelefon lämnar de slutligen 23-åringen. Det tar flera dagar innan mannen vågar anmäla händelsen till polisen (Andersson, 2011).

I episoden som utspelade sig ovan var samtliga rånare maskerade vilket gör att det inte finns någon möjlighet för 23-åringen att identifiera rånarnas ansikten, det vill säga 23-åringen kan inte fungera som ett ögonvittne. Det som 23-åringen däremot kan ha uppgifter om är hur de tre rånarnas röster lät. Den här informationen kan användas av polisen för att genomföra en så kallad öronvittneskonfrontation där 23-åringen ombeds att försöka identifiera någon av de tre rånarnas röster. När en sådan konfrontation genomförs är det dock viktigt att kännedom finns om vilka metoder som fungerar, och vilka som faktiskt inte fungerar för att så många skyldiga som möjligt ska gripas och att inga oskyldiga individer ska dömas för brott de inte har begått (Brewer & Wells, 2011; Hollien 2002).

I fängelser runt om i världen sitter många personer som egentligen inte borde sitta där. Personer som har blivit oskyldigt dömda till brott de inte har begått. Orsaken till att oskyldiga döms är många men en vanlig orsak är att ett vittne felaktigt har identifierat en oskyldig individ som skyldig (Wells et al., 1998). Eftersom vittnen trots allt spelar en mycket viktig roll i dagens rättsväsende är det inte rimligt att minska deras betydelse. Däremot behöver kunskapen om när man kan lita på ett vittne och när man inte kan bli bättre, samt vilka metoder som fungerar och vilka som inte fungerar för att få fram korrekta och tillförlitliga erkännanden (Brewer & Wells, 2011). Syftet med den aktuella studien är därför att jämföra skillnaden mellan två olika konfrontationsmetoder som kan användas för identifiering av misstänkta individer med hjälp av öronvittnen. Konfrontationsmetoderna som kommer att jämföras är sekventiell vittneskonfrontation och elimineringskonfrontation.

### *Sekventiell och simultan vittneskonfrontation*

Vid en sekventiell ögonvittneskonfrontation får vittnet se personerna som ingår i konfrontationen en efter en. Ett ställningstagande görs efter varje person, där vittnet får bedöma om den aktuella personen verkligen är den skyldige eller inte (Lindsay & Wells, 1985). Denna typ av konfrontation är den enda som idag också används när det gäller öronvittnen på grund av omöjligheten i att presentera alla figuranter samtidigt vid en sådan

konfrontation. En vanligare typ av vittneskonfrontationsmetod som är ofta förekommande bland ögonvittnen är en så kallad *simultan vittneskonfrontation* (Wells, 1993). Vid en *simultan vittneskonfrontation* blir vittnet presenterat för alla figuranter samtidigt och blir därefter ombedd att peka ut gärningsmannen bland dessa. En sådan bedömning kallas relativ bedömning. Ett stort problem med en sådan typ av metod är att den i många fall har visat sig leda till att vittnet pekar ut oskyldiga individer (Pozzulo et al., 2008; Pozzulo & Lindsay, 1999; Wells, 1993). En förklaring till detta fenomen tros vara att den simultana vittneskonfrontationen uppmuntrar vittnet att jämföra de personer som ingår i konfrontationsgruppen med varandra, och därefter peka ut den person som mest liknar den minnesbild som vittnet har av gärningsmannen (Wells, 1984, 1993), därav namnet relativ bedömning. Problemet med denna metod är alltså att det i många fall är den person som är mest lik gärningsmannen som bli utpekad och inte den verkliga gärningsmannen som kanske saknas i konfrontationen.

Fördelen med den *sekventiella vittneskonfrontationen* är att den uppmuntrar vittnet att jämföra sin minnesbild av gärningsmannen med varje enskild individ i konfrontationsgruppen (Lindsay & Wells, 1985). I en sådan konfrontation jämförs vittnets minnesbild först med figurant 1, därefter figurant 2 och så vidare. En sådan bedömning kallas absolut bedömning på grund av att vittnet bara har en individ att fokusera på och därmed följer också ett strängare kriterium för att göra ett utpekande vilket leder till att vittnet blir mer restriktiv i sina beslut (Granhag & Christianson, 2008; Lindsay & Wells, 1985). Den aktuella studien har använt sig av en *sekventiell vittneskonfrontation* som en följd av att studien behandlar öronvittnen, och jämfört denna konfrontationsmetod mot en elimineringskonfrontation.

### *Elimineringskonfrontation*

Elimineringskonfrontation eller *elimination-lineup*, utvecklades av Pozzulo och Lindsay (1999) som en metod för att komma tillrätta med barns höga andel av felaktiga utpekanden av oskyldiga individer när de agerar som ögonvittnen. Målet var också att metoden skulle kunna prestera samma andel korrekta utpekanden som med en simultan eller en sekventiell konfrontation som behandlas ovan. En elimineringskonfrontation kan ha lite olika utföranden men den som har valts för den aktuella studien brukar kallas för en långsam elimineringskonfrontation (*slow elimination-lineup*) (Pozzulo & Lindsay, 1999). Vid en långsam elimineringskonfrontation uppmanas vittnet att, en efter en ta bort de individer som han eller hon tycker minst liknar den skyldige tills det bara återstår en individ. Vittnet

uppmannas då att fastställa om den kvarvarande individen verkligen är den skyldige eller inte (Pozzulo & Lindsay, 1999). Fördelen med en sådan metod jämfört med till exempel en simultan vittneskonfrontation är att vittnet inte leds in att välja den person som mest liknar den skyldige, utan istället får möjlighet att jämföra sitt minne av gärningsmannen med den individ som blev sist kvar i elimineringen. Uppgiften blir således uppdelad i två olika delar, en där vittnet kommer fram till den individ som är mest lik gärningsmannen, och en del där vittnet får bestämma slutgiltigt om den kvarvarande individen är den skyldige (Pozzulo & Lindsay, 1999). En sådan metod har hos barn visat sig leda till motsvarande andel korrekta utpekanden som en vanlig simultan vittneskonfrontation när den skyldige finns inkluderad i konfrontationen (Pozzulo & Lindsay, 1999). Vad som emellertid är mer intressant är att en elimineringskonfrontationsmetod vars instruktioner modifierats för att passa barn har visat sig leda till färre falska utpekanden jämfört med en vanlig simultan vittneskonfrontation (Pozzulo & Lindsay, 1999). Senare forskning har också visat att elimineringskonfrontationen ger upphov till färre falska utpekanden jämfört med en simultan vittneskonfrontation även hos vuxna personer (Pozzulo et al., 2008).

### *Öronvittne*

Öronvittnen förekommer i många olika situationer ute i samhället; en rånare som är maskerad eller en våldtäkt som sker i mörkret är exempel på brott där ett offer inte har någon möjlighet att se förövarens ansikte och således inte är kapabel till att identifiera det vid en senare identifikation.

Det mest kända fallet där ett öronvittne spelat en avgörande roll för en fällande dom är förmodligen kidnappningen av Charles Lindberghs 20 månader gamla son 1932. En okänd man tros ha klättrat in genom pojkens fönster med hjälp av en stege och kidnappat barnet medan Charles Lindbergh befanns sig i en annan del av huset. Efter några månader fann man kvarlevorna av Lindberghs son i en slarvigt grävd grav vilket nu gjorde kidnapparen till mördare. Efter två år greps till slut en misstänkt tysk emigrerad matläggare vid namn Bruno Richard Hauptmann. Detta trots att bevisläget mot honom var skralt. Charles Lindbergh själv menade dock att han hade hört en man tala med tysk brytning på en närbelägen kyrkogård, natten då hans son försvann. Hauptmann nekade till anklagelserna men dömdes trots sitt nekande till döden. Bruno Richard Hauptmann tänkte sina sista tankar den 2 april 1936 då han blev avrättad, endast 36 år gammal. Fallet väckte stor uppmärksamhet och många frågade sig

framför allt om det är möjligt för en person att bevara en röst i minnet så länge som 29 månader, vilket var fallet med Charles Lindbergh (Granhag, 2001; Yarmey, 2007).

En av de första forskare som började intressera sig för den här typen av frågor var Francis McGehee. Hon utförde ett antal experiment i slutet av 1930-talet för att undersöka hur tiden mellan hörandet av en röst och senare identifikation av samma röst påverkar tillförlitligheten i ett eventuellt utpekande. I en av hennes studier fick ett antal försökspersoner höra en talare som läste upp ett stycke text. Deltagarna delades sedan in i olika grupper som fick höra samma röst tillsammans med ett antal figuranter efter 1, 2 och 3 veckor samt 1, 3 och 5 månader. Träffsäkerheten låg på runt 80 procent efter 1 vecka medan den efter 5 månader endast låg på 13 procent, vilket var under slumpnivå (Eriksson, 2008; Yarmey, Yarmey & Todd, 2008). Med utgångspunkt i Francis McGehees resultat tycks antingen Charles Lindbergh ha haft ett otroligt bra minne för röster eller så har han med stor sannolikhet pekat ut fel röst och inte rösten han hörde på kyrkogården, natten då hans son försvann.

### *Öronvittnesidentifikation*

I en verklig så kallad öronvittneskonfrontation presenteras ett vittne för ett antal olika röster där den skyldiges röst kan vara inkluderad eller exkluderad. För att en sådan konfrontation ska vara rättssäker att tillämpa behövs ett tillräckligt antal röster som presenteras (figuranter) för att undvika att distinkta röster drar till sig alldeles för stor uppmärksamhet. Detta beror på att öronvittnesidentifikationer inte bara handlar om minneprocesser utan också om en rösts speciella karaktär (Philippon, Cherryman, Bull & Vrij, 2007). Många forskare är idag eniga om att ett lämpligt antal röster som presenteras (figuranter) är fem eller sex stycken (Eriksson, 2008). Vissa röster drar även till sig mer uppmärksamhet och har därmed större sannolikhet att bli utpekade som skyldiga. Det är därför viktigt att personen som konstruerar öronvittneskonfrontationen har specialkompetens inom området för att rättssäkerheten ska bli god.

Ett problem som kan uppstå i försöken att generalisera resultat från olika forskningsfynd om öronvittnen till verkliga situationer är att försökspersonerna som deltar i sådana undersökningar i regel inte utsätts för traumatiska upplevelser (Yarmey, 2007). På grund av den anledningen är det viktigt att vara medveten om begränsningarna vad gäller ekologisk validitet som uppstår i sådana undersökningars förhoppning att generalisera till

verkliga brottsituationer. Att en viss metod visar sig fungera i en experimentsituation betyder inte att den med säkerhet fungerar lika bra under mer realistiska förhållanden.

Vid en öronvittnesidentifikation är det också viktigt att vittnet instrueras att lyssna igenom alla röster, för att senare göra en bedömning. Vittnet bör kunna lyssna på materialet upprepade gånger och det är viktigt att klargöra att den misstänktes röst kan finnas med eller inte finnas med i presentationen (Eriksson, 2008). Det är också önskvärt att presentationen bör ledas av någon som inte känner till vilken röst som tillhör den misstänkte. Detta för att minska olika signaler som sänds ut omedvetet och som kan ligga till grund för en felaktigt grundad bedömning (Granhag, 2001). Vad gäller inspelningens längd är det viktigaste att så stor variation som möjligt i figuranternas röster presenteras, vilket underlättar identifikationen. Det är naturligtvis så att längre inspelningar innehåller större variation än korta inspelningar men det finns även andra faktorer som påverkar variationen, t.ex. vilken text som läses upp av figuranterna. Trots delade meningar om hur långa inspelningar som ska användas har inspelningar mellan 20 till 30 sekunder visats sig vara en bra och ganska lämplig längd (Eriksson, 2008).

Det varierar kraftigt hur ofta öronvittneskonfrontationer genomförs i brottsutredningar. I England, Tyskland och Nederländerna förekommer det några gånger per år medan metoden i Sverige inte alls har används speciellt mycket. En orsak kan vara att det kräver tämligen mycket resurser för att genomföra en öronvittnesidentifikation och att frågan inte har prioriterats i Sverige. Trots det har metoden visat sig vara framgångsrik när den har används rätt och ett kraftfullt hjälpmedel till rättsväsendet (Eriksson, 2008). Vidare menar Yarmey (1995) att polisväsendet i fråga om öronvittnen skiljer sig åt mellan mindre och större städer gällande kompetens och handlingsförmåga. Det borde även finnas information tillgänglig som är relativt enkel att ta till sig som kan användas i situationer där öronvittnen är aktuella. Misslyckade insatser inom detta område kan leda till minskad kvalitet inom rättsväsendet vilket kan få förödande konsekvenser för enskilda individer (Yarmey, 1995).

Att identifiera en röst som man kanske bara har hört en enstaka gång är svårt och osäkerheten stor vad gäller tillförlitligheten i ett eventuellt utpekande. Tidigare forskning av Philippon, Cherryman, Bull och Vrij (2007) har visat att försökspersoner som vid en öronvittnesidentifikation använder sig av en så kallad elimineringsstrategi (eliminera de figuranter som är minst lik gärningsmannen) presterar signifikant bättre än de försökspersoner som antingen inte använder någon strategi alls eller personer som använder en så kallad ”matching strategy” där försökspersonen fokuserar på en eller flera specifika karakteristiska

hos målrösten som sedan jämförs med de olika rösterna som ingår i presentationen. Dessa karakteristiska kan t.ex. vara tonhöjd eller talhastighet (Philippon, Cherryman, Bull & Vrij, 2007). Samma studie visar också att personer är bättre på att ge korrekta svar då målrösten är inkluderad i presentationen (target present) i motsats till när målrösten inte finns med i presentationen (target absent).

### *Rösten*

Grundläggande kunskaper om den mänskliga rösten är oerhört viktigt för att förstå möjligheter och begränsningar vid röstigenkänning. Rösten hos människan består enligt Eriksson (2008) av en kombination mellan en ljudkälla och ett filter som tillsammans bidrar till att påverka röstens egenskaper. Som ljudkälla brukar stämbandstenen nämnas tillsammans med olika brusljud som skapas på vägen från stämbanden till läpparna. Ett s-ljud är ett typiskt exempel på ett brusljud som skapas uppe i munhålan. För att det vi kallar rösten ska formas krävs dock att ljudkällan fortplantar sina vibrationer i något material, så kallad resonans. Detta sker med hjälp av det som inom fonetiken kallas "talröret" som representerar filtret. Talröret består av svalget och munhålan och kan vara olika stort på olika människor. Precis som musikinstrumentets storlek och form skapar olika förutsättningar för ljud skapar en människas storlek på bland annat talröret samt stämbandens vibrationstakt olika förutsättningar för en rösts karaktär (Yarmey, 1995; Hollien, 2002; Engstrand, 2007). Det är till exempel inte svårt att höra skillnad på en vuxen mans och en vuxens kvinnas röst. Denna skillnad beror på att en kvinnas stämband vibrerar ungefär dubbelt så fort som en mans. En tumregel är att en kvinnas stämband vibrerar med cirka 200 svängningar i sekunden (Hz) medan en mans stämband vibrerar med cirka 100 Hz (Kent, 1997). Skillnaden beror också på att en kvinnas talrör är kortare och har lite andra proportioner i jämförelse med en mans. Man ska ha klart för sig att det finns stora variationer mellan olika kvinno- respektive mansröster. Det kan till exempel finnas män med en mycket ljus röst och kvinnor med en markant basröst. Den akustiska överlappningen är dock liten vilket beror på att olikheterna i bland annat stämbandslängd är så pass stora mellan könen. Detta får till följd att det är relativt lätt att höra skillnad mellan en mans och en kvinnoröst (Eriksson, 2008). Det har emellertid visat sig att röstfrekvensen hos män tenderar till att höjas med åldern medan den hos kvinnor istället tycks sänkas med åldern. Detta får till följd att det kan vara relativt svårt att särskilja en äldre man med en ljus röst från en äldre kvinna med en mörk röst, vilket måste beaktas i fråga om öronvittnesidentifikationer angående äldre gärningsmän som till exempel har ringt in ett



hotsamtal. Störst verkar skillnaderna mellan mans och kvinnorösten vara mellan 20 till 40 år, för att sedan avta. Upp till cirka 12 år låter däremot pojkar och flickor mycket lika varandra, vilket betyder att en identifiering av enbart rösten i detta åldersspann blir synnerligen opålitlig (Kent, 1997).

Människor är ofta förvånansvärt bra på att identifiera könet hos den som talar. Det har visat sig att till och med icke språkliga ljud som hostningar och skratt kan vara nog med information för en i det närmaste helt säker identifikation av talarens kön (Eriksson, 2008). Det är en vanlig uppfattning att man kan imitera en kvinnoröst genom att tala i falsett, men olikheter i klang vilket påverkas av dimensionerna i talröret avslöjar dock oftast bluffaren som man (Eriksson, 2008). Något som också verkar vara en vanlig föreställning är att man kan bedöma en människas kroppsstorlek genom att enbart höra rösten. Många tror till exempel att en grov röst tyder på en storväxt person medan en ljus röst förmodas tillhöra en person med mindre kroppshydda. Tillförlitligheten i sådana uttalande är trots allt så pass låg att uppskattningarna inte har några större användningsområden inom rättsväsendet. En anledning till att det är på det viset är förmodligen att det som ger rösten dess grundegenskaper – stämbandslängden och talrörsdimensioner – inte är speciellt starkt korrelerat med kroppslängd eller kroppsvikt (Eriksson, 2008). Vad gäller kroppsvikt är det svårt att föreställa sig att en person skulle förändra sin röst genom att gå ner eller upp i vikt, vilket kan ge en antydning om orsaken till den låga korrelationen mellan kroppsstorlek och röstegenskaper.

### *Språk och accent*

Människans tal påverkas inte enbart av röstens karaktär och förutsättningar utan även av språk och accent. Den information som ett öronvittne har tillgång till består således av två delar; röstens karaktär och hur denna röst används, det vill säga språket. Olika språk världen över består av speciella rytmer och betoningar som gör att ett visst språk får sin specifika karaktär. Rytmen i ett språk handlar mycket om hur olika ord och stavelser betonas. Denna kan skilja sig ganska mycket mellan olika språk vilket gör att det är lätt att höra skillnad mellan till exempel tyska och spanska (Engstrand, 2007). Vad gäller vårt eget språk eller dialekt har vi ofta en bestämd uppfattning om hur det ”ska låta” och är närmast överkänsliga för skiftningar i människors uttal som inte matchar det egna uttalet. Detta medför att regionala skillnader från den egna normen gällande språket registreras omgående trots att talförståelsen är opåverkad (Engstrand, 2007). Vår dialekt kan till exempel bero på var vi kommer ifrån, vilken utbildningsnivå vi har, var våra föräldrar kommer ifrån eller vilka förebilder vi har. Ett

känt fenomen gällande ögonvittnen och ansiktsgenkänning är den så kallade ”other race effect” (Wells & Olson, 2001). Denna effekt syftar på att människor har svårigheter att identifiera ansikten från personer som kommer från en annan grupp än den egna. I ett experiment av Kerstholt, Jansen, Van Amelsvoort och Broeders (2006) undersöktes effekten av dialekt på röstigenkänning. Hypotesen var utifrån ”the other race effect” att deltagarna skulle uppleva det svårare att identifiera röster med dialekter som var okända jämfört med dialekter som var familjära. Resultatet visade en tendens till att försöksdeltagarna hade svårare att identifiera röster med en dialekt de inte var familjära med i jämförelse med en dialekt som de kände till. Dock var denna effekt inte signifikant vilket höjer ett varningens finger i fråga om dialektens verkliga betydelse för röstidentifikation.

### *Minnet*

Vid identifikation av misstänkta individer med hjälp av öronvittnen är det inte bara den misstänktes röst som har betydelse för utfallet utan även tillförlitligheten hos vittnets minne. Människors minnen kan innehålla en kolossal mängd information och man brukar för att förenkla det hela dela in människans långtidsminne i fyra olika system: det *procedurala* (det vi kan), det *semantiska* (det vi vet), det *perceptuella* (det vi känner igen) och det *episodiska* (det vi upplevt) (Schacter & Tulving, 1994). Det *procedurala* minnet lagrar det vi kan utföra motoriskt i handling. sådana saker kan vara att gå, cykla, skriva, simma etc. Procedurminnet kännetecknas av att det innebär en viss automatik. Vi tänker inte på hur vi kör bilen medan vi kör den utan kan istället koncentrera oss på andra saker som mobiltelefonen i handen eller samtalet med en medpassagerare. Saker som har lärts in i procedurminnet är väldigt svåra att bli av med. Det *perceptuella* representationssystemet byggs upp under barnets uppväxt och innefattar en grundläggande kunskap om världens former och koncept (Schacter & Tulving, 1994). Med hjälp av detta minnesystem kan vi navigera i omgivningen. När vi går omkring i ett rum och observerar en stol eller ett bord reflekterar vi inte över dessa objekt, utan tar automatiskt fram dem ur vårt perceptuella representationsminne. Det *semantiska* minnet lagrar kunskaper som att Norge är en del av Skandinavien, att August Strindberg skrev *Hemsöborna*, att den kemiska beteckningen för syre är O etc. (Schacter & Tulving, 1994). Det semantiska minnet lagrar alltså information som går att hämta utan större problem till skillnad från procedurminnet, och som i princip endast kan uttryckas med hjälp av ord. Det semantiska minnet lagrar information som inte nödvändigtvis har någon personlig relevans utan kan vara rena faktauppgifter (Schacter &

Tulving, 1994). För att inte det semantiska minnet ska falla i glömska krävs dock att man underhåller det kontinuerligt. Det *episodiska minnet* rör sådant som vi själva har upplevt. Med andra ord är det ett självbiografiskt minnesystem. Detta system är det mest komplexa och nyaste av våra minnessystem och därför det mest sårbara. Det episodiska minnet kräver en aktiv sökprocess för att finna ett visst minne (Schacter & Tulving, 1994). När man försöker komma ihåg en röst man har hört använder man sig av sitt episodiska minnesystem. Den aktuella studien kommer alltså att använda sig av försökspersonernas episodiska minne för röster.

För att man ska kunna minnas något över en längre tid behöver man först ha lärt in det, därefter lagrat denna information så att man vid ett senare tillfälle kan plocka fram samma information. Denna process brukar delas in i tre steg vilket består av *inkodning, lagring och framlockning* av information (Lundh, Montgomery & Waern, 1992). Inkodningen av information är i hög utsträckning beroende av vilka erfarenheter vi har och vad vi är uppmärksamma på. Personer som är intresserade av cyklar har till exempel lättare att minnas olika cykelmärken jämfört med en person som inte alls är intresserad av cyklar utan istället har ett intresse för växter och blommor. Denna person kan likväl vara lika bra på att komma ihåg olika blomsorter som personen som är intresserad av cyklar är på att komma ihåg cykelmärken (Granhag & Christianson, 2008). Minnesförmågan är med andra ord inte direkt generell utan är i hög grad beroende av i vilken kontext ett minne lärs in i och vad denna kontext har för speciell anknytning till personen som lagrar minnet. Lagring av information sker genom att tidigare inkodad minnesinformation kopplas samman med ny information som vi hämtar från våra tankar och upplevelser. En sådan process brukar kallas *elaborering* (Lundh, Montgomery & Waern, 1992).

Elaborering betyder att slutsatser gällande ny information dras utifrån tidigare erfarenheter och kopplas samman med vårt förflutna på ett flexibelt sätt. Det innebär att nya intryck transformeras för att passa redan befintliga minnesstrukturer. Detta får till konsekvens att minnet av upplevda händelser förbättras samtidigt som korrektheten rörande samma händelse får ge vika en aning. Det beror på att det ofta är det transformerade minnet som man kommer ihåg och inte den precisa händelsen. Ett exempel kan vara att du är ute och går i en stadsmiljö sent på kvällen. Plötsligt stiger en storvuxen man med rånarluva för ansiktet fram och ropar ”hit med plånboken”. Du ger snabbt den maskerade mannen din plånbok vilket följs av att mannen skyndar iväg från platsen. Du beger dig djupt chockad iväg från brottsplatsen. Du funderar över vem som kan ha varit förövaren. Lät inte rösten väldigt bekant? Efter en

stunds funderande kommer du fram till att du har hört en liknande röst på bussen samma dag. Du börjar jämföra ditt minne av personen på bussen med gärningsmannen. Hade de inte samma klocka? Plötsligt blir det väldigt svårt att separera de båda personerna. Ju mer du funderar över det, desto säkrare blir du på din sak. Gärningsmannen och personen på bussen är samma person! Slutsatserna som kan dras av exemplet ovan är att när vi i efterhand ska försöka erinra oss den aktuella händelsen (t.ex. i domstol), kommer vi inte bara ihåg gärningsmannens röst eller utseende utan också en del av våra egna slutsatser, sammankopplingar, och fantasier i samband med den aktuella händelsen (Lundh, Montgomery & Waern, 1992). Genom en sådan process kan vårt minne modifieras och till och med leda till falska minnen.

Fraplockning av mer komplex information kräver ofta att man har tillgång till adekvata effektiva ledtrådar. Effektiva ledtrådar kan ses som information som har olika kopplingar till ett visst minne vilket gör fraplockningen av detta minne betydligt mindre komplicerad (Granhag & Christianson, 2008). I ett experiment av Tulving och Pearlstone (1966) fick en grupp studenter uppgiften att lära in en lista med 48 ord. Orden tillhörde fyra olika kategorier, till exempel mus, hund, katt och räv i kategorin däggdjur. De försökspersoner som vid minnestestet fick kategorinamnet som ledtråd kunde återge fler ord än de försökspersoner som testades utan någon ledtråd. Det vill säga de försökspersoner som fick ledtråden att orden i listan bestod av namn inom kategorin ”däggdjur” i samband med minnestestet av ordlistan kunde återge fler ord än de deltagare som inte fick denna information (Tulving & Pearlstone, 1966). Denna studie visar på styrkan hos effektiva ledtrådar för att öka minnesförmågan hos människor.

#### *Faktorer som påverkar ett öronvittne*

För att förenkla beskrivningen av öronvittnen kan man enligt Granhag (2001) dela upp de faktorer som påverkar öronvittnen i tre huvudkategorier. Dessa är situationsspecifika faktorer, röstspecifika faktorer och vittnesspecifika faktorer.

#### *Situationsspecifika faktorer*

Situationsspecifika faktorer avser de faktorer som är relaterade till i vilket sammanhang vittnet hör den röst som han eller hon senare ska försöka identifiera. En viktig situationsspecifik faktor är initial exponeringstid. Tiden som ett vittne hör en röst har visat sig vara av betydelse för hur tillförlitlig en senare identifikation är, där längre exponeringstider

leder till fler korrekta identifikationer (Orchard & Yarmey, 1995). Med ökad exponeringstid följer dock också fler felaktiga identifikationer (Granhag, 2001). Trots att längre exponeringstider ofta leder till fler korrekta identifikationer verkar det vara så att många vittnen är dåliga på att uppskatta hur länge de har blivit exponerade för en viss röst. I ett experiment av Orchard och Yarmey (1995) fick försökspersonerna som ingick i studien höra en röst i antingen 30 sekunder eller 8 minuter med uppgiften att uppge hur länge de hade blivit exponerade för rösten. Resultatet visade att det i båda grupperna förekom en överskattning gällande initial exponeringstid och att denna var störst i gruppen som hade hört rösten i 30 sekunder (Orchard & Yarmey, 1995). Det verkar även vara så att identifikationen av en viss röst gynnas av att ha hört rösten flera gånger än enbart en. Detta gäller även om den sammanlagda tidsperioden för hörandet av rösten är densamma (Granhag, 2001). Tonfallet hos en eventuell gärningsman påverkar också hur tillförlitligt ett utpekande är. Inte sällan blir vittnet till ett brott exponerat för röster som till exempel skriker eller talar väldigt fort (Yarmey, 2007). Detta har visat sig försvåra identifikationen och är något som kan vara ett stort problem vid öronvittneskonfrontationer. Samma sak gäller för planerad vs. oplanerad inkodning. I de allra flesta fall där man använder sig av öronvittnen i rättssammanhang har vittnet inte förberett sig på att aktivt försöka inkoda gärningsmannens röst (Granhag, 2001). Det motsatta förhållandet råder dock i de flesta experimentsammanhang vilket kan skapa problem gällande ekologisk validitet. Något som också kan vara av betydelse vid öronvittnesidentifikationer är vad som sägs av den misstänkte och av eventuella figuranter vid konfrontationstillfället, där vokaler anses tillföra mer information till en rösts karaktär än konsonanter (Cook & Wilding, 1997; Engstrand, 2007). Det finns även forskning som visar på en tendens till att personer är bättre på att känna igen röster av samma kön som dem själva i jämförelse med röster av det motsatta könet (i.e. män känner lättare igen mansröster och kvinnor känner lättare igen kvinnoröster) (Cook & Wilding, 1997).

### *Röstspecifika faktorer*

Röstspecifika faktorer är relaterade till den specifika röst som öronvittnet tagit del av. En röstspecifik faktor är hur vida vittnet är bekant med gärningsmannen eller inte. Forskning har visat att det är lättare att identifiera en bekant röst jämfört med en obekant och även att antalet felaktiga identifieringar ökar när försökspersonen inte är bekant med en viss röst i motsats till när denna röst istället är bekant (Yarmey, A.D., Yarmey, A.L., Yarmey, M.J. & Parliament, 2001). Det kan dock ställa till problem när en bekant röst hörs i en kontext som

den inte är förknippad med. I ett sådant fall blir identifieringen mer komplicerad (Granhag, 2001). En del gärningsmän är medvetna om risken med att bli avslöjad genom sin röst och vidtar därför vissa åtgärder som försvårar identifikationen avsevärt. Ett sådant är genom att viska istället för att tala med normal ton (Yarmey, 2007). Det är svårt att säga hur vanligt förekommande en sådan förställning av rösten är men det sker säkerligen oftare vid planerade brott än oplanerade. En förklaring till varför det är svårare att identifiera en röst som viskar är att en sådan röst är rensad på tonfallsvariationer som annars uppstår vid naturlig verbal kommunikation (Granhag, 2001). Forskning har visat att det är betydligt svårare att identifiera en röst när denna har blivit förvrängd genom till exempel viskning jämfört med en röst som inte har blivit förvrängd (Clifford, 1980). Accenten eller vilket språk (se s.9) en gärningsman pratar har också visat sig ha betydelse för hur säker identifikation som kan tillämpas av ett vittne, där accent eller språk som ligger väldigt nära offrets har visat sig leda till fler korrekta utpekanden jämfört med okänd accent eller okänt språk hos gärningsmannen (Philippon, Cherryman, Bull & Vrij, 2007).

#### *Vitnessspecifika faktorer*

Något som man också måste ta hänsyn till är vilka egenskaper ett vittne besitter som har betydelse för dennes förmåga att åstadkomma korrekta utpekanden. Sådana faktorer brukar gå under benämningen vittnesspecifika faktorer. En sådan viktig faktor är öronvittnets ålder. Rent generellt kan man säga att barn upp till 10 till 15 år och äldre vuxna (över 60 år) är sämre öronvittnen jämfört med personer som befinner sig i medelåldern (20-40 år) (Clifford, 1980). Något som verkar vara en vanlig uppfattning är att personer som har tränat sig på att känna igen röster skulle vara bättre öronvittnen än vanliga lekmän. Denna uppfattning har dock begränsat stöd i forskningen som istället menar att det är mycket svårt att genom träning bli ett bättre öronvittne (Granhag, 2001). Också frågan om i vilken grad olika människor är olika bra på att känna igen röster har fascinerat många forskare. I ett experiment av Cook och Wilding (1997) undersöktes om försöksdeltagare som visade sig ha en bra förmåga att känna igen okända röster också var bättre på att namnge berömda röster än deltagare som presterade mindre bra för okända röster. Resultatet bekräftade hypotesen, nämligen att det verkar finnas en generell förmåga att känna igen röster hos vissa personer vilket avspeglar sig i igenkänningen för både okända röster och för berömda röster. Förmågan att känna igen berömda röster kan alltså vara en indikator på tillförlitligheten i förmågan att känna igen okända röster (Cook & Wilding, 1997). Hur rättssäker denna bedömning blir är dock ovisst.

Tilltron till att blinda personer skulle vara bättre på att identifiera röster kan verka logisk med tanke på att dessa individer måste kompensera sin oförmåga att se med en förbättrad kapacitet att höra och uppfatta olika ljud. Denna förmåga har dock inte visats sig leda till en bättre prestation gällande röstidentifikation (Yarmey, 2007). I ett experiment av Elaad, Segev och Tobin (1998) testades träffsäkerheten hos röstidentifikations-experty, blinda individer och en grupp av försökspersoner med normal syn. Det visade sig att experterna var signifikant bättre än de två andra grupperna beträffande röstidentifikation. Ingen skillnad hittades för prestation mellan gruppen blinda och gruppen seende. Dock visade det sig att de blinda individerna var mer realistiska i fråga om tillförlitligheten i deras beslut jämfört med gruppen som hade normal syn (Elaad, Segev & Tobin, 1998).

### *Konfidensbedömning*

När ett öronvittne har pekat ut en misstänkt röst är det intressant att avgöra om identifikationen är korrekt eller felaktig. Det finns dock inget säkert sätt att bedöma säkerheten i ett sådant utpekande utan man får hålla till godo med olika typer av indikationer. En sådan indikation är graden av konfidens som vittnet kopplar till sitt utpekande. Det tycks logiskt att vittnen som är mer säkra på sin sak oftare har rätt än vittnen som tvekar inför sitt beslut (Granhag & Christianson, 2008). Forskning visar dock att det inte är självklart att det finns ett samband mellan konfidensbedömningar och korrektheten i ett utpekande vilket är viktigt att påpeka. Till exempel visar en studie av Philippon, Cherryman, Bull och Vrij (2007) att konfidensbedömningar är korrelerat med antalet korrekta beslut, men bara när målrösten inte finns inkluderad i presentationen (target absent). Någon korrelation mellan konfidensbedömning och korrekt beslut hittades alltså inte när målrösten var inkluderad i presentationen (target present) (Philippon, Cherryman, Bull & Vrij, 2007). Det verkar följaktligen vara så att korrektheten i ett vittnes beslut stämmer någorlunda överens med dennes konfidensbedömning när målrösten saknas i presentationen men däremot inte när målrösten är inkluderad i presentationen.

I de flesta typer av öronvittnesidentifikationer gör det aktuella vittnet bara ett utpekande. Vad som däremot kan skilja sig åt är hur många figuranter som presenteras i anslutning till den misstänkta rösten. I en studie som hade för avsikt att undersöka tillförlitligheten i konfidensbedömningar gjorda av försökspersoner som försökte identifiera en misstänkt röst användes två olika typer av vittneskonfrontationer av dessa röster. Den ena var en så kallad ”one-person lineup” där försökspersonerna blev presenterade för en målröst

och fick efter 5 minuter höra en röst som antingen kunde vara målrösten eller ej. Den andra typen av vittneskonfrontation som användes var en så kallad "six-person lineup" där försökspersonerna blev presenterade för en målröst i likhet med föregående konfrontation men blev efter 5 minuter istället presenterade för sex stycken röster där den misstänktes röst antingen kunde vara inkluderad eller ej. Resultaten visade att konfidensbedömningar som rapporterades av försökspersonerna efter varje taget beslut, stämde dåligt överens med det verkliga utfallet i de båda konfrontationsmetoderna. Sämst stämde konfidensbedömningarna för konfrontationen med sex stycken figuranter (six-person lineup) jämfört med konfrontationen med enbart en röst (one-person showup) (Yarmey, 1995).

Hur säker man som vittne är på sitt utpekande beror självklart på en mängd olika faktorer; hur länge man har blivit exponerad för gärningsmannen; hur bra man har sett/hört gärningsmannen; hur bra minne man upplever sig ha etc. Det är dock inte säkert att konfidensbedömningar av vittnen stämmer väl överens med korrektheten i deras bedömning. Detta gäller för både öron och ögonvittnen. Normalt uppgår korrelationen mellan konfidensbedömningar och korrekthet inte till mer än 0.30 (Olsson, 2000), vilket kan förefalla mycket men betyder i praktiken att konfidensbedömningar är en svag indikator på tillförlitligheten i ett eventuellt utpekande.

### *Syfte och hypoteser*

I ljuset av tidigare forskning ter sig kunskapen om öronvittnen i många avseenden bristfällig. Därför föddes idén att jämföra två olika konfrontationsmetoder som sedan tidigare har visat sig framkalla olika resultat gällande ögonvittneskonfrontationer. Dessa båda konfrontationsmetoder är sekventiell vittneskonfrontation och elimineringskonfrontation. Tidigare forskning har bland annat visat att elimineringskonfrontationen leder till en lägre andel felaktiga utpekanden jämför med en simultan vittneskonfrontation vilken är vanligt förekommande bland ögonvittnen (Pozzulo et al., 2008). Frågan som ställdes i den aktuella studien var huruvida sådana effekter även skulle förekomma gällande öronvittnen. Forskning som undersöker skillnaden mellan olika metoders tillförlitlighet när det gäller vittneskonfrontationer med öronvittnen saknas just nu vilket betyder att mer kunskap behövs för att med säkerhet uttala sig om skillnaderna mellan olika vittneskonfrontationsmetoder. Sådan forskning är relevant om man beaktar alla de brott som sker i mörker eller där gärningsmannen är ordentligt maskerad. Dessa brott har den gemensamma nämnaren att ögonvittnen ofta inte tillför den aktuella brottsutredningen någon ytterligare information



eftersom gärningsmannen inte har varit fullt synlig. Däremot kan identifikation av den skyldige med hjälp av öronvittnen utgöra en viktig pusselbit i ett uppklärt rättsfall om detta sker på ett rättssäkert och pålitligt sätt (Granhag, 2001; Hollien, 2002). Förhoppningen är precis som nämndes i inledningen att den aktuella studie ska ge en första beskrivning av hur sekventiell konfrontation och elimineringskonfrontation skiljer sig åt beträffande tillförlitlighet och konfidens när det gäller öronvittnen och att den därmed kan bidra med ökad kunskap inom ämnet.

Studien liknar undersökningen av Pozzulo et al. (2008), som studerade huruvida simultan, sekventiell och elimineringskonfrontation hade någon påverkan på hur många korrekta utpekanden som gjordes samt hur många felaktiga utpekanden som utfördes av ögonvittnen. Skillnaden är att den aktuella studien, vilket har nämnts tidigare har som mål att undersöka ifall de skillnader som rapporterades av Pozzulo et al. (2008) för ögonvittnen även gäller för öronvittnen, då det finns belägg för att hävda att en identifikation av någons röst är betydligt svårare än att känna igen någons ansikte (Yarmey, 2007). Med utgångspunkt i denna tidigare forskning om ögonvittnen avseende elimineringskonfrontation och sekventiell vittneskonfrontation som därmed existerar, väcktes tanken att undersöka ifall det även skulle finnas en skillnad mellan dessa båda metoder med avseende på öronvittnen. Utifrån det föregående resonemanget formulerades tre hypoteser som den aktuella studien har för avsikt att besvara. Dessa hypoteser är följande:

Hypotes 1: Elimineringskonfrontationen kommer att framkalla motsvarande andel korrekta responser jämfört med den sekventiella konfrontationen.

Hypotes 2: Elimineringskonfrontationen kommer att framkalla en lägre andel felaktiga identifikationer jämfört med den sekventiella konfrontationen.

Hypotes 3: Andelen korrekta responser är högre när målrösten finns inkluderad i konfrontationen jämfört med när målrösten saknas, oberoende av konfrontationsmetod.

## Metod

### *Deltagare*

32 studenter (12 män och 20 kvinnor) från Lunds Universitet rekryterades via en egeninsamlad deltagar-pool och erhöll en trisslott som kompensation för deltagandet i studien.

Deltagarna var mellan 19 och 36 år ( $M = 23,6$  år,  $SD = 3,8$ ). Bara deltagare som hade svenska som modersmål rekryterades.

### *Material*

*Röster.* 28 röster spelades in (14 män och 14 kvinnor) med hjälp av en Logitech USB-mikrofon av bra kvalitet. Upphovsmännen till rösterna var mellan 18 och 46 år ( $M = 27,8$  år,  $SD = 7,4$ ). Rösterna spelades in i en tyst miljö utan störande ljud. Alla upphovsmän till rösterna hade svenska som modersmål vilket var viktigt för att stärka studiens validitet och var införstådda i hur deras röster var tänkta att användas. Rekryteringen av röster skedde vid psykologiska institutionen vid Lunds universitet under tre olika dagar under en vanlig arbetsvecka. Således har ett bekvämlighetsurval tillämpats. Detta på grund av knappa resurser men även svårigheten med att genomföra ett totalt randomiserat urval. Deltagandet var frivilligt. Ingen kompensation utgick. Rösterna spelades in genom att upphovsmännen fick högläsa varsitt textstycke framför en mikrofon två gånger efter varandra. Detta gjordes för att försäkra sig om att upprepningar och stakningar som kan vara vanligt förekommande hos personer som läser högt, inte skulle förenkla identifikationen i studiens huvudsakliga experiment. Totalt användes fyra olika textstycken (se bilaga 1).

Textstycke nummer 1 och 3 höglästes enbart av manliga röster och textstycke nummer 2 och 4 höglästes enbart av kvinnliga röster. Detta gjordes för att göra experimentet mer verklighetstroget och relevant samt att öka resultatets generaliserbarhet.

### *Design och procedur*

En 2 (elimineringskonfrontation eller sekventiell konfrontation) \* 2 (målröst inkluderad eller målröst exkluderad) inom-grupps design användes. De beroende variablerna var prestation avseende igenkänning av röster samt konfidensbedömningar som rapporterades i samband med varje taget beslut. Dessa meddelades på en skala från slumpnivå till 100 %, där 100 % indikerade fullständig säkerhet beträffande sitt beslut. Slumpnivån skiljde sig åt beroende på vilken typ av konfrontation som användes, där den sekventiella konfrontationen medförde sex stycken valmöjligheter och således cirka 17 % slumpchans att välja rätt röst medan deltagarna i elimineringskonfrontationen endast blev ombudda att rapportera sin konfidens beträffande den kvarvarande rösten vilket gjorde att slumpnivån var 50 %.

Med hjälp av de 28 rösterna konstruerades fyra röstkonfrontationer (lineups) med sex röster i varje, två med enbart mansröster och två med enbart kvinnoröster. För att kunna

konstruera en röstkonfrontation (lineup) där målrösten saknades, spelades sju röster in för varje eventuell röstkonfrontation för att de olika betingelserna skulle kunna roteras fritt mellan de fyra textstyckena. Alla textstycken förekom sålunda i alla olika betingelser (elimineringskonfrontation eller sekventiell konfrontation samt målröst inkluderad eller målröst exkluderad). För att eliminera eventuella övningseffekter togs beslutet att använda en så kallad ABBA-design som i praktiken betyder att man låter försökspersonerna använda sig av de olika metoderna i en slumpmässigt vald ordning följt av samma betingelser i motsatt ordning. De olika konfrontationsmetoderna (lineups) roterades därmed så att alla textstycken förekom lika många gånger med varje enskild röstkonfrontationsmetod. För att balansera de betingelser då målrösten fanns inkluderad i röstkonfrontationen slumpades ordningen av ”målröst inkluderad” och ”målröst exkluderad” på alla försökspersoner och alla betingelser. Det innebar att det sammanlagt fanns 64 röstkonfrontationer där målrösten var inkluderad och 64 röstkonfrontationer då målrösten var exkluderad. Varje block i experimentet bestod av en instuderingsfas, där deltagarna fick höra en målröst, fem stycken utfyllnadsuppgifter som sammanlagt tog cirka fem minuter och slutligen en testfas, där försöksdeltagarna skulle identifiera målrösten de hört tidigare. Experimentet konstruerades i programmet Power-Point.

På grund av svårigheterna med att identifiera röster gjordes förhållandena för testpersonerna så gynnsamma som möjligt genom att minska störande ljud och låta deltagarna lyssna på rösterna i den aktuella konfrontationen så många gånger som önskades. Vidare användes bara mansröster i samma konfrontation och kvinnoröster i samma konfrontation. Detta gjordes för att efterlikna en verklig konfrontation som uteslutande sker med samkönade figuranter på grund av att det är en markant skillnad mellan mans och kvinnoröster (Yarmey, 1995) vilket förmodligen hade gjort uppgiften allt för simpel. Varje röstkonfrontation bestod av sex figuranter vilket har visat sig vara ett lämpligt antal (Eriksson, 2008) och har även används i många studier med besläktade frågeställningar, bland annat av Philippon, Cherryman, Bull och Vrij (2007).

Experimentet genomfördes i ett laboratorium på institutionen för psykologi vid Lunds Universitet genom att försökspersonerna satt framför en datorskärm iförda hörlurar samt blev ombedda att följa de instruktioner som redogjordes på skärmen. Experimentledaren befann sig också i rummet för att kunna besvara eventuella frågor som kunde dyka upp under experimentets gång. Alla försöksdeltagare blev testade individuellt och testet tog cirka 35 minuter att genomföra. Försökspersonerna fick reda på att volymen kunde justeras till för dem en behaglig nivå under experimentets gång. Ingen tidsbegränsning förekom. Experimentet

bestod av fyra olika block. Varje block inleddes med att deltagarna fick höra en målröst läsa något av de ovanstående nämnda textstyckena. Därefter genomfördes fem stycken utfyllnadsuppgifter som sammanlagt tog ungefär fem minuter att genomföra.

Utfyllnadsuppgifterna var av typen intelligensuppgifter som är vanliga bland IQ-tester. Dessa hade som syfte att förhindra att målrösten fanns kvar i korttidsminnet hos försöksdeltagarna vilket i sådant fall torde medföra att igenkänningen blivit betydligt mindre komplicerad och följaktligen inte vara relevant för brottsituationer, där ju ofta en längre tid passerar mellan ett brott och en senare vittneskonfrontation (Granhag & Christianson, 2008). Efter utfyllnadsuppgiften blev försökspersonerna presenterade för sex stycken röster som högläste samma text som målrösten hade läst. Identifikationen skedde med hjälp av antingen en elimineringskonfrontationsmetod eller en vanlig sekventiell konfrontationsmetod. Försökspersonerna blev efter varje ställningstagande ombedda att skatta sin konfidens gällande deras bedömningar på en skala från slumpnivå till 100 %.

## Resultat

Sammanlagt 128 röstkonfrontationer har genomförts varav 64 stycken med en så kallad elimineringskonfrontation och 64 stycken med en sekventiell vittneskonfrontation. Vidare förekom målrösten i 64 stycken vittneskonfrontationer och var exkluderad i 64 stycken. För att kunna jämföra de olika konfrontationsmetoderna summerades antalet korrekta responser för varje försöksperson och konfrontationsmetod och delades med antalet felaktiga responser samt multiplicerades med hundra. På så sätt kalkylerades ett procent-värde mellan 0-100 för varje försöksperson och varje aktuell metod. Vidare kunde försökspersonerna när målrösten var inkluderad i konfrontationen antingen gjort en korrekt identifiering (hit), felaktigt avstått från att peka ut någon röst (miss) eller felaktigt pekat ut en annan röst än målrösten (falskt alarm). När målrösten var exkluderad i konfrontationen kunde försökspersonerna antingen korrekt, avstått från att peka ut någon röst (hit) eller felaktigt pekat ut en annan röst än målrösten (falskt alarm). Andelen av de olika typerna av responser i procent för de olika konfrontationsmetoderna finns presenterade i *tabell 1*.

Tabell 1: Andel av korrekta utpekanden, felaktiga utpekanden, korrekta avvisande och felaktiga avvisanden för de två konfrontationsmetoderna i procent

	Korrekt utpekande	Felaktigt utpekande	Korrekt avvisande	Felaktigt avvisande
<b>Eli-konfrontation</b>	31.3	32.8	29.7	6.2
<b>Sek-konfrontation</b>	34.4	28.1	23.4	15.6

### *Konfrontationsmetod*

Ett paired-samples t-test genomfördes för att undersöka skillnaden mellan elimineringskonfrontationen och den sekventiella konfrontationen beträffande andelen korrekta responser, vilket alltså innefattade korrekta utpekanden samt korrekta avvisanden. Det var ingen statistisk signifikant skillnad mellan elimineringskonfrontationen ( $M = 60.9$ ,  $SD = 37.5$ ) och den sekventiella konfrontationen ( $M = 57.8$ ,  $SD = 40.4$ ),  $t(31) = 0.33$ ,  $p > .05$ . Vilket talar för att de båda konfrontationsmetoderna var jämförbara i fråga om andel korrekta responser.

För att undersöka huruvida det förelåg någon signifikant skillnad mellan de båda konfrontationsmetoderna avseende felaktiga identifikationer genomfördes ett paired-samples t-test. Det var ingen statistisk signifikant skillnad för andelen felaktiga identifikationer för elimineringskonfrontationen ( $M = 32.8$ ,  $SD = 35.0$ ) och den sekventiella konfrontationen ( $M = 28.1$ ,  $SD = 35.8$ ),  $t(31) = 0.59$ ,  $p > .05$ . Detta indikerar att de båda konfrontationsmetoderna är jämförbara även när det gäller andelen felaktiga identifikationer.

### *Målröst inkluderad/exkluderad*

Ett ytterligare paired-samples t-test genomfördes för att undersöka ifall det fanns någon skillnad i andelen korrekta responser när målrösten var inkluderad i konfrontationen i jämförelse med när målrösten istället var exkluderad, oberoende av konfrontationsmetod. Det var ingen statistisk signifikant skillnad mellan andelen korrekta responser när målrösten var inkluderad i konfrontationen ( $M = 64.6$ ,  $SD = 40.1$ ) jämfört med när målrösten var exkluderad i konfrontationen ( $M = 51.0$ ,  $SD = 43.6$ ),  $t(31) = 1.30$ ,  $p > .05$ . Resultatet talar för att

försökspersonerna inte rapporterade fler korrekta svar då målrösten fanns inkluderad i konfrontationen jämfört med när målrösten istället var exkluderad.

### *Konfidens*

Ett paired-samples t-test genomfördes också för att titta på om det fanns några skillnader mellan de båda konfrontationsmetoderna avseende konfidens. Det var ingen statistiskt signifikant skillnad i rapporterad konfidens mellan elimineringskonfrontationen ( $M = 80.56$ ,  $SD = 14.87$ ) och den sekventiella konfrontationen ( $M = 79.28$ ,  $SD = 14.55$ ),  $t(31) = 0.41$ ,  $p > .05$ . Vilket pekar på att rapporterad konfidens verkar ha varit oberoende av vilken konfrontationsmetod som användes i samband med det aktuella beslutet.

### Diskussion

Syftet med den aktuella studien var att jämföra tillförlitligheten mellan två olika vittneskonfrontationsmetoder gällande röstidentifikation; elimineringskonfrontation och sekventiell vittneskonfrontation. Elimineringskonfrontationen förväntades leda till en motsvarande andel korrekta responser som den sekventiella vittneskonfrontationen, vilket går i linje med forskning om vittneskonfrontationsmetoder beträffande ögonvittnen av Pozzulo et al. (2008) samt forskning om öronvittnen av Philippon, Cherryman, Bull och Vrij (2007). Vidare förväntades elimineringskonfrontationen leda till färre falska utpekanden jämfört med den sekventiella vittneskonfrontationen på grund av en uppdelning beträffande identifikationsuppgiften i två delar; en relativ bedömning där rösten som är mest lik målrösten blir sist kvar i elimineringen och en absolut bedömning, där ett slutgiltigt beslut tas huruvida den kvarvarande rösten verkligen är målrösten. Fler korrekta responser förväntades också inträffa när målrösten var inkluderad i konfrontationen till skillnad från när målrösten saknades. Genom att använda optimala förhållanden för identifikation och inspelning av röster förväntades skillnaderna mellan de båda konfrontationsmetoderna framkomma ännu tydligare än om så inte hade varit fallet. Resultatet visade att det inte finns några signifikanta skillnader mellan de två konfrontationsmetoderna.

Efter en jämförelse mellan elimineringskonfrontationen och den sekventiella vittneskonfrontationen gällande antalet korrekta responser visade det sig att elimineringskonfrontationen framkallar motsvarande andel korrekta responser jämfört med den sekventiella vittneskonfrontationen. Detta går i linje med tidigare forskning beträffande ögonvittnen som har visat att elimineringskonfrontationen framkallar motsvarande andel

korrekta responser jämfört med den sekventiella vittneskonfrontationen (Pozzulo et al., 2008), vilket också var den förevarande studiens första hypotes. En potentiell förklaring till det aktuella resultatet skulle kunna vara att de olika rösterna som användes i experimentet var så pass olika varandra att det utjämnade eventuella skillnader mellan de båda konfrontationsmetoderna. Resultatet går inte i linje med forskning av Philippon, Cherryman, Bull och Vrij (2007) som visade att försökspersoner som använder sig av en så kallad elimineringsstrategi, vilken består i att eliminera de röster som inte låter som målrösten i samband med varje taget beslut, presterade signifikant bättre jämfört med de försökspersoner som inte använde sig av en sådan minnesstrategi. Det ska framföras att i studien av Philippon, Cherryman, Bull och Vrij (2007) var rösterna som användes utvalda genom att matcha dem gällande en rad faktorer så som särprägel och ordinäritet, för att undvika att distinkta röster skulle uppta alldeles för stor uppmärksamhet. En sådan åtgärd saknas i den nuvarande studien.

Beträffande hypotes 2, visar resultatet efter en jämförelse mellan de båda konfrontationsmetoderna att elimineringskonfrontationen inte framkallar signifikant färre falska utpekanden jämfört med den sekventiella vittneskonfrontationen. Detta går emot studiens andra hypotes samt forskning av Pozzulo et al. (2008), som visar att elimineringskonfrontationen bland ögonvittnen minskar antalet felaktiga utpekanden jämfört med en simultan vittneskonfrontation. Den simultana vittneskonfrontationen i studien av Pozzulo et al. (2008) har på grund av att den aktuella undersökningen avhandlar öronvittnen, bytts ut mot en sekventiell vittneskonfrontation. En eventuell orsak till det aktuella resultatet kan vara att den förevarande studien har använt sig av en så kallad långsam elimineringskonfrontation (slow elimination-lineup) som i tidigare studier misslyckats med att producera färre falska utpekanden jämfört med en sekventiell vittneskonfrontation hos barn (Pozzulo & Lindsay, 1999). En ytterligare förklaring kan vara att den långsamma elimineringskonfrontationen helt enkelt är krångligare i jämförelse med den sekventiella vittneskonfrontationen som bara kräver ett ställningstagande från försökspersonen, vilket kan ha försämrat prestationen för elimineringskonfrontationen. Detta är något som även Pozzulo och Lindsay (1999) nämner som en möjlig orsak till det sämre resultatet för den långsamma elimineringskonfrontationen. Vidare finns det en möjlighet att elimineringskonfrontationen försvårade minnesprocessen av målrösten hos försöksdeltagarna på grund av att de var tvungna att fokusera på den röst som lät minst lik målrösten, samt upprepa denna process tills det bara återstod en röst. Detta kan ha stört minnet av den röst som försökspersonerna hade som uppgift att försöka identifiera, det vill säga målrösten.

Fortsättningsvis visar en analys av hypotes 3 att försökspersonernas korrekthet angående sina beslut inte var signifikant högre när målrösten var inkluderad i konfrontationen i jämförelse med när den istället var exkluderad, vilket alltså går emot undersökningens tredje hypotes. Detta går även emot tidigare undersökningar av bland andra Philippon, Cherryman, Bull och Vrij (2007) samt Kerstholt, Jansen, Van Amelsvoort och Broeders (2006). Något som också har diskuterats tidigare av VanWallendael, Surace, Hall Parsons och Brown (1994) är att många vittnen i riktiga vittneskonfrontationer leds in i en föreställning om att gärningsmannens röst finns med i konfrontationen, fast den i själva verket inte finns med. Som vittne är det kanske svårt att föreställa sig att rättsväsendet skulle genomföra en konfrontation där gärningsmannens röst saknas, vilket kan vara en bidragande faktor till varför en så hög andel felidentifieringar ofta rapporteras. Översatt till denna studie kan man tänka sig att försöksdeltagarna, dels på grund av att experimentledaren poängterade att målrösten inte behövde vara närvarande i konfrontationen samt att det inte rörde sig om någon verklig vittneskonfrontation inte kände samma press att peka ut någon skyldig röst som verkliga vittnen gör. Dessutom hade de aktuella rösterna som användes i studien som nämnts tidigare inte balanserats utifrån sådana effekter som dialekt, röstläge eller talhastighet, vilket är något som annars möjligen skulle kunna försvåra identifikationsuppgiften och därmed bidra till att öka skillnaderna mellan de båda betingelserna, i likhet med tidigare nämnda studier. Något som även tycks ha varit en vanlig strategi hos försöksdeltagarna för att minnas målrösten var att lyssna på stakningar eller uttal rörande vissa speciella ord hos målrösten för att på så vis underlätta identifikationen. Detta kan ha försvårat igenkänningsuppgiften, med tanke på det inte var samma röstinspelning som försökspersonerna hörde första och andra gången de hörde målrösten. Konsekvensen blir då rimligtvis att det är svårare att känna igen målrösten utan till exempel en stakning på ett visst ord när man hör den för andra gången.

Efter varje taget beslut fick deltagarna i studien skatta sin konfidens beträffande korrektheten i deras beslut. Konfidensbedömningarna rapporterades oavsett om försökspersonen hade gjort någon identifikation eller inte. Bedömningarna är således ett mått på hur säkra försöksdeltagarna var på att de hade gjort en korrekt respons, vilket kunde innefatta att ingen röst ansågs tillhöra målrösten. Efter att ha analyserat resultaten visade det sig att det inte förelåg någon statistisk signifikant skillnad i rapporterad konfidens mellan de två konfrontationsmetoderna. Vad detta beror på är svårt att svara på men eftersom de två konfrontationsmetoderna inte framkallade signifikant olika korrekta responser, vilket alltså



betyder att de är relativt likvärdiga, verkar det rimligt att även konfidensbedömningarna torde vara relativt lika.

Slutsatserna som kan dras av studiens resultat är framför allt att elimineringskonfrontationen och den sekventiella vittneskonfrontationen verkar vara relativt lika varandra i fråga om öronvittneskonfrontationer. Det verkar alltså vara så att uppdelningen i en relativ och en absolut bedömning inte gav upphov till färre felaktiga identifikationer så som Pozzulo et al. (2008) rapporterar. Vad detta beror på är svårt att svara på men några möjliga anledningar har diskuterats ovan. Eftersom de felaktiga utpekningarna var så pass stora för de båda konfrontationsmetoderna och att felaktiga utpekningar står för en stor del av alla oskyldiga individer som döms för brott de inte har begått (Wells et al., 1998), är det betydelsefullt att de båda konfrontationsmetoderna som har jämförts i den aktuella studien behandlas med största försiktighet när det gäller öronvittnen. I ett mer praktiskt avseende kan studien fungera som en indikation på att ett öronvittnes utsaga inte tycks påverkas i lika stor utsträckning gällande vittneskonfrontationens utformning som ögonvittnen gör. En sådan kännedom kan vara nog så viktigt för att på ett rimligt sätt förstå både tillgången men också begränsningarna i ett öronvittnesutpekande och vad detta kan få för konsekvenser för oskyldiga individer. Förhoppningsvis kan dessa slutsatser bidra med en ökad kunskap om öronvittnen samt en bredare förståelse för de svagheter och styrkor hos de konfrontationsmetoder som kan användas i samband med en sådan typ av vittne.

Vad gäller generaliserbarheten av resultaten till verkligheten finns det i den aktuella studien en del problem, vilka delas av i stort sett alla vetenskapliga undersökningar som har som syfte att studera vilka faktorer som påverkar olika typer av vittnen. Detta på grund av att det inte är etiskt korrekt eller försvarbart att utsätta försökspersoner för den ångest och lidande som en riktig brottsituation innebär. En förhöjd anspänning av stress kan tänkas påverka minnet av till exempel en röst, vilket innebär att det är viktigt att förstå begränsningarna med de studier som har gjorts hittills om exempelvis öronvittnen. Fortsättningsvis har även gärningsmannen många gånger en förhöjd nivå av stress vilket påverkar dennes röst och är något som sällan avspeglas i de öronvittneskonfrontationer som genomförs runt om i världen (Yarmey, 2007). En diskrepans i röstanspänningen mellan brottstillfället och vittneskonfrontationen kan leda till en sämre förmåga hos vittnet att göra en korrekt identifikation jämfört med om gärningsmannen talar med liknande eller samma anspänning i rösten vid både brottstillfället och vittneskonfrontationen, vilket också är det vanliga i experimentsituationer. Något som även är ett problem med experimentella studier angående

öronvittnen är att personer som blir utsatta för brott allt som oftast inte är förberedda på att koda in gärningsmannens röst. Följderna av detta kan bli att vittnet minns förövarens röst sämre än om han eller hon hade blivit förvarnad om att registrera rösten innan det aktuella brottet. I verkligheten förekommer det också ofta en ganska lång tid mellan hörandet av en röst och senare identifikation av samma röst. I den aktuella studien genomfördes en fördröjning på cirka fyra till fem minuter mellan inkodningstillfället och identifikationstillfället. Detta rimmar dåligt med verklighetens premisser och sänker således studiens ekologiska validitet.

Ett ytterligare potentiellt problem med den aktuella studien är att ordningen i vilken de sex figuranterna presenterades i inte har randomiserats utan istället roterats vilket betyder att ordningen på de sex figuranterna är en möjlig ”confounding”. Det är omöjligt att svara på hur stor påverkan denna faktor har på det slutgiltiga resultatet men det är en effekt som måste beaktas. I praktiken är det emellertid inte troligt att denna effekt är överhängande på grund av att ordningen på figuranterna trots allt har roterats. En annan möjlig counfounding är att de olika försökspersonerna har gjort testet under olika tidpunkter på dagen och under olika dagar. Förmågan att känna igen röster påverkas troligtvis inte speciellt mycket av sådana faktorer men det är viktigt att ha i åtanke att det kan finns en rad faktorer som kan innebära möjliga counfoundings. Utmaningen ligger i att randomisera de variabler man tror har en avgörande effekt, vilket också har varit det primära målet med den aktuella studiens metodval.

Vid en större studie hade det varit aktuellt att innan utförandet av den ordinarie undersökningen genomföra analyser av de röster som var tänkta att användas i samma röstkonfrontation för att försäkra sig om att dessa inte skiljer sig allt för mycket från varandra. Detta därför att minnet av en röst tycks vara influerat av i vilken utsträckning rösten utmärker sig, där utmärkande röster blir bättre ihågkomna än mera ordinära röster. (Orchard & Yarmey, 1995). I den aktuella studien har ingen hänsyn tagits till de olika rösternas dialektala skillnader vilket kan ha medfört att identifikationsuppgiften skiljer sig i antalet korrekta respektive felaktiga responser, jämfört med om alla röster i experimentet hade haft samma dialekt. Det är därför viktigt för framtida undersökningar som berör öronvittnen att försäkra sig om att ingen av rösterna i konfrontationen utmärker sig mer än någon annan för att bättre kunna generalisera till verkliga vittneskonfrontationer hos vittnen som säger sig har hört gärningsmannen. Något som även hade varit av intresse i en mer djupgående studie är att jämföra den långsamma elimineringskonfrontationen som användes i den aktuella studien med en mer rapid elimineringskonfrontation, en så kallad fast elimination lineup, som har

utvecklats av Pozzulo & Lindsay (1999). Detta framför allt på grund av att den rapida elimineringskonfrontationen hos barn har visat sig leda till färre felaktiga identifieringar jämfört med den långsamma elimineringskonfrontationen när målrösten är exkluderad från konfrontationen. I den rapida elimineringskonfrontationen ombeds vittnet att peka ut den individ som är mest lik gärningsmannen, men som nödvändigtvis inte är gärningsmannen. Alla andra figuranter plockas sedan bort, med undantag från den person vittnet har valt. Vittnet får därefter avgöra om denna person verkligen är gärningsmannen. Uppgiften blir därmed uppdelad i två delar, en relativ och en absolut, precis som i den långsamma elimineringskonfrontationen men går här betydligt fortare vilket är något som alltså tycks minska de felaktiga identifikationer som är särskilt vanliga hos barn. En intressant frågeställning för en framtida studie hade varit att undersöka om de resultat hos den rapida elimineringskonfrontationen som funnits för barn som är ögonvittnen, också gäller för vuxna och öronvittnen.

## Referenser

- Andersson, J. (2011). Grova personrån oroar polisen. *Helsingborgs dagblad*, 7/9-2011.
- Brewer, N., & Wells, G. L. (2011). Eyewitness identification. *Current Directions in Psychological Science*, 20(1), 24-27.
- Brown, D. (2009). *Den förlorade symbolen*. Stockholm: Albert Bonniers förlag.
- Clifford, B. R. (1980). Voice identification by human listeners: On earwitness reliability. *Law and Human Behavior*, 4(4), 373-394.
- Cook, S., & Wilding, J. (1997) Earwitness testimony: Never mind the variety, hear the length. *Applied Cognitive Psychology*, 11, 95-111.
- Engstrand, O. (2007). *Fonetik light*. Lund: Studentlitteratur.
- Eriksson, A. (2008). Rättsfonetik. I P. A. Granhag, & S. Å. Christianson (Reds.), *Handbok i rättspsykologi*, (ss. 325-339). Stockholm: Liber.
- Elaad, E., Segev, S., & Tobin, Y. (1998). Long-term working memory in voice identification. *Psychology, Crime & Law*, 4, 73-88.
- Granhag, P. A., & Christianson, S. Å. (Red.). (2008). *Handbok i rättspsykologi*. Stockholm: Liber.
- Granhag, P. A. (2001). *Vittnespsykologi*. Lund: Studentlitteratur.
- Hollien, H. (2002). *Forensic voice identification*. San Diego: Academic press.
- Kent, R. D. (1997). *The speech science*. San Diego: Singular Publishing Group, Inc.
- Kerstholt, J. H., Jansen, N. J. M., Van Amelsvoort, A. G., & Broeders, A. P. A. (2006). Earwitnesses: Effects of accent, retention and telephone. *Applied Cognitive Psychology*, 20, 187-197.
- Lindsay, R. C. L., & Wells, G. L. (1985). Improved eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, 3, 556-564.
- Lundh, L. G., Montgomery, H., & Waern, Y. (1992). *Kognitiv psykologi*. Lund: Studentlitteratur.
- Olsson, N. (2000). A comparison of correlation, calibration, and diagnosticity as measures of the confidence-accuracy relationship in witness identification. *Journal of Applied Psychology*, 85(4), 504-511.
- Orchard, T. L., & Yarmey, A. D. (1995). The effects of whispers, voice-sample duration, and voice distinctiveness on criminal speaker identification. *Applied Cognitive Psychology*, 9, 249-260.
- Philippon, C. A., Cherryman, J., Bull, R., & Vrij, A. (2007). Earwitness identification performance: The effect of language, target, deliberate strategies and indirect measures. *Applied Cognitive Psychology*, 21, 539-550.

- Pozzulo, J. D., Dempsey, J., Corey, S., Girardi, A., Lawandi, A., & Aston, C. (2008). Can a lineup procedure designed for child witnesses work for adults? Comparing simultaneous, sequential, and elimination lineup procedures. *Journal of Applied Social Psychology*, 38(9), 2195-2209.
- Pozzulo, J. D., & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child witnesses. *Journal of Applied Psychology*, 84(2), 167-176.
- Schacter, D. L., & Tulving, E. (1994). *Memory systems 1994*. Massachusetts: The MIT Press.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381-391.
- vanWallendael, L. R., Surace, A., Hall Parsons, D., & Brown, M. (1994). Earwitness voice recognition: Factors affecting accuracy and impact on jurors. *Applied Cognitive Psychology*, 8, 661-677.
- Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, 14, 89-103.
- Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist*, 48, 553-571.
- Wells, G. L., & Olson, E. A. (2001). The other-race effect in eyewitness identification. *Psychology, Public Policy, and Law*, 7(1), 230-246.
- Wells, G. L., Small, M., Penrod, S., Malpass, R., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, 22, 603-647.
- Yarmey, A. D. (1995). Earwitness speaker identification. *Psychology, Public Policy, and Law*, 4(1), 792-816.
- Yarmey, A. D. (2007). The psychology of speaker identification and earwitness memory. I R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology*, Vol. 2: *Memory for people* (ss. 101-136). Mahwah, NJ, US: Lawrence Erlbaum Associates.
- Yarmey, D. A., Yarmey, A. J., & Todd, L. (2008). Frances McGehee (1912-2004): The first earwitness researcher. *Perceptual and Motor Skills*, 106, 387-394.
- Yarmey, A. D., Yarmey, A. L., Yarmey, M. J., & Parliament, L. (2001). Commonsense beliefs and the identification of familiar voices. *Applied Cognitive Psychology*, 15, 283-299.

### *Bilaga 1: Textstycken*

Textstycke 1: Upptäckten av Isaac Newtons hemliga papper 1936 hade slagit världen med häpnad genom att avslöja Newtons enorma passion för studiet av urgammal alkemi och mystisk visdom. Bland Newtons privata papper fanns ett brev till Robert Boyle där han uppmanade Boyle att iaktta "största tystnad" beträffande den mystiska kunskap som de hade fått. "Den kan inte spridas", skrev Newton, "utan att världen utsätts för ofantliga risker" (Brown, 2009).

Textstycke 2: Trots att kongressbiblioteket innehåller en sal som av många betraktas som "det vackraste rummet i världen" är det mindre känt för sin oerhörda prakt än för sina stora samlingar. Med över åttahundra tusen hyllmeter - uppställda på rad skulle de nå från Washington D.C. ända bort till Boston - kan det göra anspråk på benämningen världens största bibliotek utan att möta några invändningar (Brown, 2009).

Textstycke 3: Chefen för Capitoliums vaktstyrka Trent Anderson hade haft ansvaret för säkerheten i kongressbyggnaden i över ett årtionde. Han var en kraftig, bredbringad man med skarpa drag och rött hår som han höll kortsnaggat, vilket fick honom att utstråla militär auktoritet. Han bar sitt vapen synligt som en varning till alla som var dumma nog att ha åsikter om gränserna för hans befogenheter (Brown, 2009).

Textstycke 4: Ett villkor för att man ska bli frimurare är att man tror på en högre makt. Skillnaden mellan frimurarnas andlighet och organiserad religion är att frimurarna inte föreskriver någon särskild definition eller något namn på denna högre makt. I stället för bestämda teologiska identiteter som Gud, Allah, Buddha eller Jesus använder frimurarna mer allmängiltiga termer som "högre väsen" eller "universums store arkitekt". Tack vare det kan frimurare från olika religioner samlas (Brown, 2009).