

# Röstens teknologi

Johan Landgren

Upplevelsen av en röst. En sjungande röst. En klump, en massa. Som en muskel, ett pulserande hjärta. Eller ett svart hål, som utan att uppenbara sig för våra sinnen ändå utövar sin dragningskraft på oss. Som kittlar oss under huden och drar i något inom oss som vi inte visste var magnetiskt.

Rösten är en flicka, en varelse, en man, ett väsen, ett vidunder, en kvinna, en pojke. Naturlig, onaturlig eller övernaturlig. Men den är *ett*. En enhet, en identitet. Ett.

## Den sammansatta rösten

Låt oss stiga in i inspelningsstudion. Här produceras den identitet vi kallar *popsångare*. Det är denna drömfabrik som gör att den röst du hör på radion är något helt annat än bara *en människa som sjunger*. De som jobbar här använder en lång rad grepp för att bygga sångaren, popstjärnan, den där mänskliga men ändå ouppnåeliga rösten.

*Sångare -> Sångbås -> Puffskydd -> Mikrofon -> Förstärkare -> Equalizer -> Auto-tuner -> Kompressor -> Reverb -> ...*

Redan i sångbåset, där rösten spelas in, finns ett av de mest kritiska momenten i denna kedja. Detta specialfall av rum tillåter oss att fånga den sjungande rösten i en miljö där vi annars i princip aldrig hör den: i ett nästan helt dämpat utrymme. (Ironiskt nog närmast raka motsatsen till de lokaler som klassisk sång föredrar: där en rik efterklang ger stöd åt vokalistens ton.) Men varför detta dämpade rum? Ett svar kan heta *intimitet*. Medan rumsklang direkt ger känslan av avstånd och rymd, tillåter sångbåsets torra akustik oss att höra sångaren som om hen vore så nära som tjugo centimeter från vårt öra.

Vilket leder oss till nästa steg i kedjan: mikrofonen, och dess väktare puffskyddet. Trots allt tal om high fidelity, trogen återgivning, är denna metalliska klump, som agerar lyssnarens öra, på intet sätt ärlig. Likt *spegeln, spegeln på väggen där*, erbjuder den en reflektion som är långt ifrån objektiv. Känsligheten i mikrofonens tunna membran är enorm, och den överdriver lätt dynamiska variationer. Dess frekvensåtergivning är heller knappast oförvanskad; hos det vi kallar "högkvalitativ studioutrustning" hittar vi ofta en närmast pornografisk förstärkning av diskanten. Precis som avsaknaden av efterklang, signalerar dessa höga frekvenser *närhet* (våra öron vet mycket väl att diskanten är det första register som försvinner när avståndet till ljudet ökar). Särskilt konsonanter – dessa ljud rika på brusartade, högfrekventa komponenter – blir genom denna förmedling vässade till skirt krispiga skatter.

(Är det känslan av att Thomas Öberg viskar mig i örat som får mig att gripas av åtrå efter klickande konsonanter i bob hunds "Blommor på brinnande fartyg"? ([http://www.youtube.com/watch?v=aFQx6J-i\\_rI](http://www.youtube.com/watch?v=aFQx6J-i_rI)))

Och så vidare genom flödet: *mikrofonförstärkaren* färgar ljudet med sina egenheter, *equalizern* spetsar till klangen ytterligare och *auto-tunern* rättar till eventuella falska toner, för att ge sångaren den där överjordiskt perfekta kvaliteten. *Kompressorn* justerar signalens dynamiska nivåer (delvis för att kompensera för mikrofonens extrema känslighet) och skapar den där spöklika *närvaron*: signalens styrka jämnas ut, så att starkt blir "starkt" och svagt blir "svagt" – den faktiska ljudnivån är konstant, men skiftningar i röstens beteende och klangfärg ger oss fortfarande illusionen av toppar och dalar. *Reverbet* ger oss den känsla av rum som sångbåsets avsaknad av efterklang förnekar oss. Inget verkligt, fysiskt rum, dock, utan ett perfekt, virtuellt utrymme. (Är det inte en intressant bild: inspelningsstudios preferens för vissa reverb-enheter placerar alla sångare i samma rum. Inte bara på samma estetiska koordinater, genom att en standardmodell för röstbearbetning homogeniserar soundet. Utan också i identiska akustiska utrymmen: likriktningen i användandet av artificiell efterklang får det att låta som om de alla bebodde samma fysiska lokal.)

\* \* \*

Detta är bara en grov skiss. Produktionsbandet är långt, och kan innefatta än fler steg i byggandet av *sångaren*. Men trots dess sammansatta natur har vi knappast några problem med att höra detta aggregat, detta komplex av ljudbearbetning som en enhet, *ett* objekt, *en* varelse. Kanske framstår denna hopfogade, studioprocesserade

identitet till och med som enklare, mer greppbar och mindre komplex än den "nakna" rösten hos din kompis som nynnär några takter ur den senaste list-topparen. Där amatören utgör en levande individ, vars identitet, musikaliska talang och kulturella tillhörighet vi måste pejla, läsa av, lokalisera, erbjuda oss den mödosamt bearbetade studioproducten en bekväm paraplyterm: *popsångare*.

Oavsett vad den är förutom detta, eller inom detta, så är denna röst först och främst *popsångare*. Den har kvalat in, uppfyllt kraven, den *passerar* som popsångare, och detta faktum överskuggar alla andra egenheter och variationer.

Vi kan tryggt – utan att titta för nära på dem – bokföra eventuella skönhetsfläckar och egenheter såsom varande "inom det rimligas gränser"; vore de för stora, hade rösten ju aldrig tillåtits inträde till popsångardomänen från första början. Ett slags moment 22: sångaren avskalats sin mänsklighet just genom att *dess status som sångare* får oss att avfärda eventuella avvikelser som irrelevanta.

### Teknik och tekniker

*(Upplevelsen av en röst. En sjungande röst. En klump, en massa. Som en muskel, ett pulserande hjärta. Eller ett svart hål, som utan att uppenbara sig för våra sinnen ändå utövar sin dragningskraft på oss. Som kittlar oss under huden och drar i något inom oss som vi inte visste var magnetiskt.*

*Rösten är en flicka, en varelse, en man, ett väsen, ett vidunder, en kvinna, en pojke. Naturlig, onaturlig eller övernaturlig. Men den är ett. En enhet, en identitet. Ett.)*

Det finns olika sätt att höra. Vi kan lägga vårt rent estetiska sinne åt sidan; inte längre det korsdrag genom vidöppna fönster som karakteriserar det utomrationella upplevandet. I stället: lyssna, studera, analysera. Och så fort vi lämnar det sinnligas och det oreflekterandes domän, för att i stället tala i mer tekniska termer, förändras bilden av sångaren. Inte längre en helhet: den estetiska kroppen, den bekanta formen eller den hårda, glödande knuten av uttryck, intention, identitet. Framför oss sträcker sig nu i stället en avlång konstellation, något segmenterat och därigenom bräckligt, poröst. En kedja:

*lungor -> strupe -> stämband -> munhåla -> tunga -> tänder -> läppar -> rum -> puffskydd -> mikrofon -> förstärkare -> auto-tuner -> kompressor -> equalizer -> delay -> reverb*

Det som förut var nolldimensionellt – gränslöst, blott en närvaro, *en estetisk singularitet* – har nu antagit utsträckning: ett endimensionellt objekt, från vänster till höger. Detta är på intet sätt ett oskyldigt eller obetydligt faktum. Nej, det är att betrakta som inget mindre än ett paradigmskifte.

Möjligheterna nu svindlande: utveckling, fortskridande, hierarki, narrativ, spänning-upplösning, ett före och ett efter, ett över och ett under. Ett primärt och ett sekundärt. Delar, fack, kategorier.

Ett nytt seende skapar en ny bild av objektet. Kanske varseblir vi det först som toppar och dalar. Sedan klarnar bilden, blicken får tid att fokusera: det framstår nu som ett pärlband, som knoppar i rad eller bär på ett strå. Av-skilj-bar-a. Nu finns det en kontur, ett ställe att ta fäste, att *driva in kilen*. Att splittra. SEPARERA.

Inte så mycket att separera *lungor* från *strupe* från *stämband* från *munhåla* från *tunga* från *tänder* från *läppar*. Men mellan *läppar* och *rum* slås en kil, införs en kofot. Det som är *människokropp* skiljs av från det som är *icke-människokropp*. Rum, puffskydd, mikrofon, förstärkare, auto-tuner, kompressor, equalizer, delay, reverb – den här delen av kedjan förblir intakt, sitter fortfarande ihop. Den förblir en enhet: den är *studio*. Den är *icke-människokropp*.

(Inom det vi kallar *studio* går visserligen också underdelningar att göra: *puffskydd* buntas ihop med *mikrofon* med *förstärkare* till det vi kallar *inspelningsutrustning*. Följande steg sammanfattar vi med benämningen *effektkedja*. Men dessa avskiljningar är underordnade och av obetydlig rang jämfört med den avgrund, den ravin som slits upp mellan *människa* och *icke-människa*.)

\* \* \*

Språket understödjer alltså denna splittring, även om det också visar på hur intimt förbundna de två hälfterna är. Ett exempel: genom att sjunga i falsett kan vi lägga vår röst i ett högre register än det till vardags talar i.

Men en liknande förflyttning kan också åstadkommas genom en *pitch shifter* – en elektronisk modul som används för att höja eller sänka ett ljuds tonhöjd. Det förstnämnda (falsett) kallar vi ”en teknik” (närmare bestämt ”en sångteknik”). Det senare (pitch shifter) faller däremot under benämningen ”teknik”.

När teknik blir utomkroppslig blir den alltså grammatiskt *öräknebar*. Precis som vi inte kan säga ”en luft” eller ”flera luft” – vi måste säga ”mycket luft” eller ”lite luft” – kan vi inte sätta elektronisk teknik i singular eller plural. De inomkroppsliga teknikerna, däremot, behandlar vårt språk på annat sätt: vi talar om ”en sångteknik” eller ”flera sångtekniker” då vi talar om olika sätt att använda vår röst.

(Engelskans uppdelning i *technique* respektive *technology* visar på ett kanske ännu tydligare sätt på denna relation, som på samma gång är en klyvning och en samhörighet.)

Genom att utgå från samma ord – teknik – erkänner språket alltså att de två sätten att forma rösten (inomkroppsligt respektive utomkroppsligt) är sinsemellan relaterade. Vår grammatik ser dock samtidigt till att på ett aldrig så subtilt sätt markera skillnad mellan begreppen, för att försäkra sig om att en grundläggande regel följs: det mänskliga får inte blandas ihop med det icke-mänskliga.

\* \* \*

Denna uppdelning må vara hållbar så länge vi håller oss på ett ganska enkelt plan: tekniken är något som *appliceras på* människan. Det artificiella, onaturligt tekniska *förvränger* det naturligt mänskliga. En sopran flyttas ner i basregister för komisk effekt. En mansröst spelas upp i dubbel hastighet, varpå alla maner blir överdrivna, vibrato dubbelt så dallrigt. Människan är utgångspunkten, maskinen något sekundärt, en inkräktare.

Men verkligheten har en tendens att inte vara så enkel. Hierarkier inte enkelriktade; över inte överordnat under. Ett nyckelbegrepp är *växelverkan*. Så också när det gäller ljudproduktion. Röst spelas medvetet in i en lägre oktav, för att sedan flyttas upp på elektronisk väg. Detta som ett sätt att överföra ett avslappnat lågt registers speciella kännetecken till en ny domän. Genom diskanten, där vi väntar oss en spänd röst som kämpar för att nå de höga tonerna, svävar plötsligt en obekymrad, lättsamt hummande stämma.

Eller lyssna på T-Pain (<http://www.youtube.com/watch?v=rxRvDpF2FDA>). Artisten som blivit synonym med *auto-tune*, 00-talets mest åtrådda och avskydda röstbearbetningsmjukvara. Som drivit sin estetiska spjutspets så långt att han anklagats för att ha startat en epidemi. Se bara hur Jay-Z producerade en låt – D.O.A - Death of Autotune ([http://www.youtube.com/watch?v=aMuf\\_ekJhOs](http://www.youtube.com/watch?v=aMuf_ekJhOs)) – enbart med syftet att lyfta ett bannande finger åt sina kollegor som influerats av det förföriska soundet: ”Get back to rap, you T-Pain-in’ too much”.

Hur låter T-Pain? T-Pain låter *inte* som Faheem Rasheed Najm, personen som ligger bakom artistnamnet. T-Pain låter som de synergetiska kopplingarna mellan Najms kropp, sångbåset, mikrofonen, kompressorn, auto-tunern ... Utan dessa komponenters verkan och växelverkan finns inte den röst, den identitet vi kallar T-Pain. Det handlar om en samverkan mellan organ och kretskort, där ingen av parterna kan sägas vara primär eller mer central.

Och så går relationerna, från höger till vänster, vänster till höger, nerifrån och upp, utifrån och in. Kompressorns möjlighet att tona ner dynamiska variationer öppnar möjligheten att använda den känsliga stormembransmikrofonen. Mikrofonens överdrift av diskanten lyfter fram det sensuella läckage av luft som avkrävs av envar som vill sjunga pop. Samtidigt ser en *de-esser* till att redan högfrekventa s- och f-ljud inte skenar iväg och rubbar ljudbildens balans. Sångaren behöver heller inte bekymra sig om att jämna ut styrkeförhållandet mellan höga och låga toner, då detta kan balanseras på annan plats i effektkedjan. Sångbåsets stumma akustik ger inspelningen en känsla av total närhet, samtidigt som reverb-effekten förser rösten med ett diskret men ack så väsentligt hav av rumsklang att glida på. För att inte tala om det faktum att studioformatets möjlighet till pålägg och omtagningar tillåter vokalisten att ständigt prestera på toppen av sin kapacitet. Mellan en svår ton och nästa kan stämbanden få timmar av vila, något som den klassiska sångaren inte ens kan drömma om.

Sådana kopplingar fram och tillbaka genom röstinspelningens kedja av kroppsligheter, rum och apparater är vanliga, och ifrågasätter relevansen i att dra den obönhörliga skiljelinjen mellan början och slut, original och förvanskad kopia, människa och icke-människa. Då vi talar om rösten – den röst vi hör i den inspelade värld av poplåtar och reklamfilmer som i allt högre grad utgör vår verklighet – finns inget före och inget efter, inget ”det mänskliga” och ”det omänskliga”. Där finns bara *rösten*. Rösten som kittlar oss, attraherar oss, äcklar oss.

Som får oss att vilja konsumera, och som får oss att säga att vi minsann inte går på sådana enkla knep, att vi minsann inte tänker handla mer, vad den där glödande klumpen, den där närvaron än säger. Som får oss att åtrå och äcklas över vår åtrå och bada i vår ljuvliga åtrå.

### Ett förslag: metakropp

Jag har tänkt på begreppet *metakropp*. Har ingen myntat det tidigare? Det känns ju som öppet mål; meta-, detta eviga buzzword, detta gyllene prefix som kan förvandla vilket tråkigt ämne som helst till en kulturteoretisk guldgruva. Och kropp, betonandet av kroppslighet som en moderiktig känga åt den akademiska traditionens torra rationalism. Metakropp.

Jag använder då ordet *meta* i samma bemärkelse som i *metafilm* – en film om film, eller *metafilosofi* – filosofi om filosofi (i motsats till den ursprungliga grekiska betydelsen ”efter” eller ”bortom”).

Och *kropp* som i levande kropp, det som gör oss till *varelser*, men också som i fysikalisk eller astronomisk kropp: något som är materiellt, som därigenom utövar krafter på sin omgivning. (Den fysiska kroppen utsätter närliggande kroppar för gravitation. Vad kan vi kalla den estetiska kroppens attraktions- och repulsionskraft?)

I alla fall, om kusten verkar vara klar (och ordet *metakropp* faktiskt visar sig vara tillgängligt för min appropriering), och jag faktiskt bestämmer mig för att prova detta begrepp – sjösätta det – då borde det användas just här.

För detta aggregat, detta komplex bestående av *lungor, strupe, stämband, munhåla, tunga, tänder, läppar, rum, puffskydd, mikrofon, förstärkare, auto-tuner, kompressor, equalizer, delay, reverb*, vad är det, såsom det uppträder för våra sinnen, om inte en metakropp? *En* kropp, singular – för att det var där det började: vi upplevde en enhet, hörde den sjungande rösten som *ett*. Men en *meta-kropp*, en kropp bestående av kroppar. För när vi ändrar vårt perspektiv, förbyts upplevelsen av det som tidigare framstod som odelbart. Den musikaliska atomens knådade klump visar nu upp en yta av skiftande färger. Nyanser som vittnar om dess beståndsdelar, de olikfärgade komponenter som bygger upp det skenbart enhetliga.

(Det är där splittringens process börjar. Där vi kan hitta en skreva, punkt att driva in kilen. Separera – krossa en illusion. Vi kommer göra det igen. Men låt oss inte göra det just nu. Låt oss vila.)

En kropp bestående av kroppar. Av saker och ting som under vissa förutsättningar är begripliga för oss var för sig (”En mikrofon. En lunga. En kompressor-plugin. En tunga.”), men som, när vi hör på ett annat vis, upphör att existera som separata element. Nu hör vi inte något aggregat av kroppsdelar och elektroniska apparater. Vi hör *Cher* (som genom sitt användande av auto-tune i superhitten Believe (<http://www.youtube.com/watch?v=LbXiECmCZ94>) myntade begreppet ”Cher-effekten”). Vi hör *Karin Dreijer Andersson* (som i gruppen The Knife och soloprojektet Fever Ray ingått ett närmast symbiotisk relation med sin röstförvrängningsapparat (<http://www.youtube.com/watch?v=LHkIcZsORkU>)). Vi hör den identitet som är *T-Pain*, som *inte* är en sjungande Faheem Rasheed Najm, oavsett hur många skivkontrakt och födelseattester som påstår sig bevisa att de två är samma person.

*Självklart* är de inte samma person. T-Pain är inte en person; T-Pain är en stjärna, en röst, ett sound. En estetisk identitet. Faheem Rasheed Najm, däremot, är en person, en människa, ett liv. Vem kunde någonsin få för sig att blanda ihop de två?