

Stämmer S-kurvan?

Om hållbarheten i en modell för språklig förändring

Sanna Skärhund
Lunds universitet

1 Inledning

Inom sociolingvistisk forskning har det länge varit en gängse uppfattning att språkförändringar följer vad som brukar beskrivas som en S-formad utveckling: först går förändringen långsamt, därefter snabbt och avslutningsvis åter långsamt. Såväl i vetenskaplig litteratur som i kurslitteratur framställs detta inte sällan som en vedertagen sanning (se t.ex. Aitchison, 2013; Einarsson, 2009; Labov, 1994; Nevalainen & Raumolin-Brunberg, 2003; Sundgren, 2013). Winter-Froemel (2014, s. 124) nämner ”the enormous success of the S-curve” (vilken hon dock är kritisk emot) och Nevalainen (2015 opag.) beskriver hur S-kurvan ”has stood the test of time well enough to merit its current status as a fitting abstraction in describing the unfolding of the process of language change over time”. En närmare genomgång av forskningsläget visar dock att uppfattningen om att språkliga frekvensförändringar generellt är S-formade tycks ha svagt empiriskt stöd; de undersökningar som gjorts är få och resultaten har ibland generaliserats på sätt som inte alltid stämmer med vad data faktiskt visar. På senare tid har det visserligen gjorts ett par större studier av S-kurvans möjligheter att korrekt beskriva språkliga förändringsprocesser (Feltgen et al., 2017; Ghanbarnejad et al., 2014), men resultaten av dessa två undersökningar pekar i rakt motsatta riktningar.

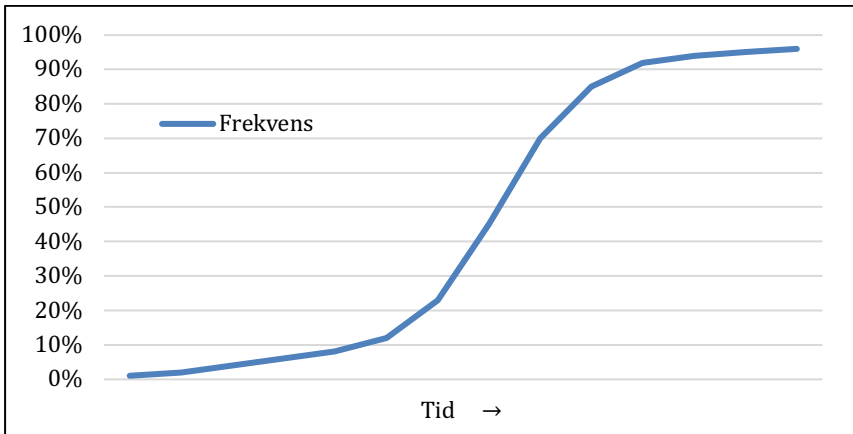
I den här artikeln vill jag ge en empirisk grund till diskussionerna om S-kurvans förmåga att på ett hållbart sätt beskriva språkförändringsprocesser, genom att visa hur väl modellen stämmer med ett antal svenska språkförändringar efter 1850. *Hållbar* ska här tolkas utifrån ordets bildliga

betydelse som ”(logiskt) korrekt eller tillfredsställande” (*Svensk ordbok*, 2021).

I det följande ger jag först i avsnitt 2 en kortfattad bakgrund till hur S-kurvan kom att bli en vedertagen modell för språkförändring samt presenterar rådande forskningsläge. Därefter beskriver jag metod och material för min undersökning i avsnitt 3. Vidare presenteras resultatet i avsnitt 4, som följs av en diskussion i avsnitt 5.

2 Bakgrund

Språkliga förändringar har ofta beskrivits som följande mönstret ”slow, slow, quick, quick, slow” (Denison, 2003, s. 56) – de går inledningsvis långsamt, snabbar sedan upp, och avslutas återigen långsamt. Schematiskt brukar en frekvensförändring som följer ett sådant mönster återges med en S-formad graf som den i figur 1.

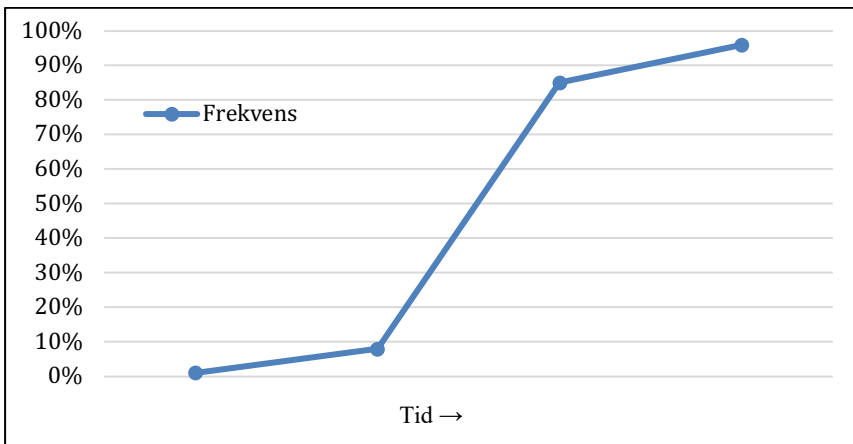


Figur 1. Språklig förändring som en S-formad frekvensökning.

Exakt vad den S-formade kurvan visar råder det dock olika idéer om (jfr Einarsson, 2009, s. 242). Illustrerar den hur en språklig variant sprids från en språkbrukare till en annan, tills slutligen nästan hela språksamhället har anammat den? Eller är det snarare en beskrivning av spridning genom själva språket som ett abstrakt system, t.ex. hur ett nytt uttal sprids från ett ord till ett annat (jfr Chen, 1972)? Den undersökning som presenteras här utgår från tidningstext, och därmed kan frekvensökningar antas motsvara en spridning bland journalister eller en ökad användning av en språklig variant som (vissa) journalister redan använder.

Beskrivningar av S-kurvans giltighet för språklig förändring brukar vanligen föras tillbaka på Rogers sociologiska översikt *Diffusion of innovations* (1962), i vilken det beskrivs hur olika nymodigheter i ett samhälle sprids enligt S-formade mönster. Redan 1954 föreslogs dock att språkförändringar följer en S-formad utveckling: "Language change in a community will be gradual and cumulative [...] The process of change in the community would most probably be represented by an S-curve" (Osgood & Sebeok, 1954, s. 155). Diskussioner om S-kurvan som modell för språklig förändring fördes också tidigt av Weinreich et al. (1968, s. 113f). Den studie det oftast hänvisas till är dock Kroch (1989), som var en av de första som presenterade data för att underbygga påståendena. Kroch (1989) jämförde fyra studier av frekvensförändringar med en logistisk funktion (vilken genererar en kurva lik den i figur 1) och visade att dessa följde S-formen.

Hur mycket en studerad frekvensförändring får skilja sig från den schematiska kurvan i figur 1 och ändå anses följa en S-form kan diskuteras. Nevalainen (2015) presenterar exempelvis förändringar som avviker avsevärt från denna form, men konstaterar trots detta att de följer S-mönstret. På liknande sätt är det relevant att fråga sig hur många datapunkter som krävs för att en likhet med S-kurvan ska kunna fastställas. För några av de studier som Nevalainen (2015) beskriver redovisas endast data från tre tidpunkter. Minst fyra nedslag bör emellertid rimligen vara nödvändiga för att alls kunna säga någonting om huruvida ökningen går snabbare under en viss period än före och efter denna period. Se figur 2.



Figur 2. S-formad frekvensökning med fyra mätpunkter.

Liknande problem med att data ansetts stämma överens med S-formen trots svagt empiriskt underlag finns i en artikel av Blythe och Croft (2012), som presenterar en metastudie av 39 språkförändringar i relation till S-kurvan. Dessa forskare beskriver hur 22 av förändringarna följer S-formen genom att först gå långsamt, därefter snabbt och sedan långsamt igen, medan 13 förändringar kan tolkas som att de utgör början eller slutet på en S-kurva. Av de senare är 8 hämtade från Nevalainen och Raumolin-Brunberg (2003), och en närmare genomgång av deras resultat ger knappast några indikationer på att det är just delar av en S-formad utveckling det handlar om. Ett särskilt flagrant exempel är hur en ökning som endast återges med två datapunkter (från ca 5 % till 35 %) av Blythe och Croft (2012) räknas in bland de förändringar som kan tolkas som början på en S-kurva. För förändringar som inte följer S-formen utan stannar av och går i motsatt riktning anger dessutom Blythe och Croft (2012, s. 279) att dessa "may be interpreted as changes following an S-curve trajectory that are then interrupted". I mina ögon ser detta ut som försök att få data att passa med den föreslagna modellen, oavsett vad analysen faktiskt visar.

Ytterligare en oklarhet i tidigare forskning är vilken matematisk funktion som bäst representerar S-formen. S-kurvan har liknats vid kumulativ frekvens av binominal distribution (Labov, 1994), "bass diffusion model" (Ghanbarnejad et al., 2014), den logistiska funktionen (Kroch, 1989, Feltgen et al., 2017) och kumulativ normalfördelning (Kroch, 1989). Dessa olika funktioner genererar visserligen liknande kurvor, men att olika modeller används innebär ändå att det inte är självklart att olika forskningsresultat är direkt jämförbara med varandra.

En del kritiska röster har också förekommit i tidigare forskning. Denison (2003, s. 68) menar att S-kurvan ibland ger en förenklad bild av den språkliga variation som finns och anser att "the S-curve should not be seized on too readily as the general shape of language change". Dessutom pekar Winter-Froemel (2014) på flera olika problem och oklarheter med S-kurvan, varav den viktigaste är avsaknaden av empiriska data för att underbygga påståendena om att språkförändringar följer detta mönster.

På senare tid har det emellertid gjorts ett par större undersökningar av S-formens möjligheter att korrekt beskriva språkförändring. I en artikel i matematisk fysik undersöker Ghanbarnejad et al. (2014) tre övergripande språkförändringar i stora datamängder: alla ord påverkade av de tyska stavningsreformerna 1901 och 1996, 50 ryska namn stavade på engelska och tyska samt regelbunden böjning för 10 engelska verb. Dessa forskares slutsatser är dels att förändringar inte kan sägas följa universella S-kurvor,

dels att det går att se vilka mekanismer som ligger bakom en förändring, eftersom språkliga reformer uppifrån skapar asymmetriska kurvor.

I en studie i fysik dras dock en motsatt slutsats. Utifrån en analys av 408 franska uttryck som genomgått semantisk utvidgning efter 1321 (t.ex. *à condition que* och *par hasard*) anger Feltgen et al. (2017) att 295 (dvs. 72 %) av dessa uttryck genomgått en S-formad ökning som är statistisk signifikant i relation till en slumpmässig ökning. De konstaterar att undersökningens resultat utgör ”a robust statistical validation of the sigmoidal pattern [dvs. S-mönstret], confirming the general claim made in the literature” (Feltgen et al., 2017, s. 3).

Att Ghanbarnejad et al. (2014) och Feltgen et al. (2017) kommer till rakt motsatta slutsatser kan antagligen delvis förklaras av att de analyserar olika material. Det är kanske inte så oväntat att språkvårdande påbud uppifrån, som de tyska stavningsreformerna, ger förändringar som ser annorlunda ut än semantiska förändringar, vilka kan tänkas vara mer omedvetna för språkbrukaren. I min studie finns både förändringar som språkbrukaren sannolikt varit medveten om och förändringar som troligen är mer omedvetna för språkbrukaren inkluderade i materialet, vilket jag nu går vidare till att beskriva i nästa avsnitt.

3 Metod och material

Materialet för undersökningen utgörs av data från två stora pressdatabaser: Retrievers Mediearkivet (Retriever, 2024) för språkförändringar som inletts efter 1990 och Svenska tidningar från Kungliga biblioteket (KB, 2024) för språkförändringar längre tillbaka i tiden. Mitt urval från Mediearkivet har utgjorts av svensk presstext i formen print och webb (inte TV, radio, podd) från riksmmedia, regionmedia och lokalmedia. Antalet medier i databasen är föränderligt, men 27 december 2024 fanns det totalt 1 172 källor, med de nyss angivna begränsningarna (olika år innehåller dock olika många källor). KB:s Svenska tidningar innehåller tidningstext från 1645 till idag: hela KB:s samling av svensk tidningsutgivning till och med 1906 och efter 2013 samt ett ”mindre urval” av tidningar däremellan (se Ahlberg, 2024). Detta mindre urval är dock knappast obetydligt i storlek: en sökning på det frekventa ordet *att* för perioden 1910–1930 genererar belägg från 34 tidningar (t.ex. *DN*, *SvD* och *Aftonbladet*).

Sökningar i Mediearkivet ger uppgifter om antal artiklar som innehåller söksträngen. I KB:s material ger en sökning istället svar på hur många tidningssidor som innehåller den. Att materialen inte är helt jämförbara

med varandra spelar dock inte någon roll i det här fallet, eftersom jag inte blandar data från olika databaser.

För att finna exempel på språkförändringar har jag analyserat 150 lexikala, morfologiska och ortografiska variabler för vilka jag misstänkt att en variant har bytts ut mot en annan i svensk dagspress. (I några fall rör det sig istället om att två varianter bytts ut mot en eller tvärtom, t.ex. *icke/ej* mot *inte*.) Några exempel på variabler jag har analyserat är följande (nyare/äldre variant):

Lexikon

- tjej/flicka
- quiz/frågesport
- fysioterapeut/sjukgymnast

Morfologi och konstruktioner

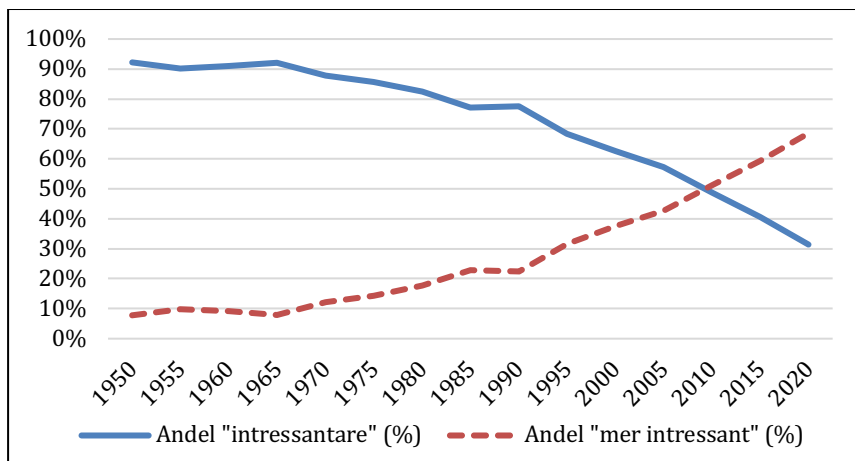
- beslutade/beslöt
- har växt/har vuxit
- kommer vara/kommer att vara

Ortografi

- av/af
- tar/tager
- skanna/scanna

De variabler jag studerat har dels varit sådana som tidigare nämnts i forskning om språkförändring (se Mårtensson, 1988), dels valts utifrån egna idéer om möjliga pågående språkförändringar (det senare gäller nutida utvecklingar som t.ex. skiftet från *frågesport* till *quiz*).

Winter-Froemel (2014) påpekar i en kritisk artikel att S-kurvan förutsätter en övre gräns (dvs. 100 %, jfr figur 1 och 2), men att det är oklart vad som ska anses vara en övre gräns för frekvens för språkliga former. I min studie ställer jag genomgående två varianter mot varandra, och därmed utgör 100 % ett scenario där en av dessa varianter helt har tagit över i presstext. Ett exempel presenteras i figur 3, som illustrerar hur *mer intressant* har ökat i relation till *intressantare* över tid i KB:s material.



Figur 3. Exempel på frekvensförändring: *mer intressant* ökar i relation till *intressantare* i KB:s tidningar.

Den förändring som syns i figur 3 är dock inte medtagen i min studie, eftersom jag ställt upp ganska rigorösa kriterier för att en frekvensökning ska få inkluderas i denna. För det första krävs att antalet belägg på den gamla och nya varianten tillsammans uppgår till minst 500 belägg per år (år 1950 finns i KB:s material endast 414 exempel på *intressantare* och 35 på *mer intressant*, dvs. totalt något färre än 500 belägg).¹ För det andra ska den nya varianten uppgå till mindre än 15 % när frekvensökningen börjar, och utgöra minst 80 % i slutet av den analyserade perioden (som framgår av figur 3 utgör exemplen på *mer intressant* knappt 70 % år 2020; frekvensen uppgår därmed (ännu?) inte till 80 %). Dessa gränsvärden är valda utifrån Nevalainen och Raumolin-Brunberg (2003, s. 55), som beskriver pågående förändringar som täcker olika delar av lutningen på S-kurvan i olika steg med "Incipient: below 15 per cent" och "Nearing completion: between 66 and 85 per cent". Att jag valt 80 % (slutfasen av "Nearing completion") och inte 85 % (vilket i deras modell beskrivs som "Completed") som övre gräns beror på att det har varit mycket svårt att hitta förändringar som uppfyller de krav jag ställt upp: trots att jag totalt undersökt 150 variabler har jag endast kunnat finna 32 som mött de angivna kriterierna.

Genomgående har jag först studerat frekvenser för de konkurrerande varianterna i Medicarkivet. I de fall data visat att förändringen sannolikt

¹ Detta kriterium har jag i några få fall gjort undantag ifrån om det under något enstaka år (men inte inledningsvis) funnits något färre än 500 belägg.

påbörjats före 1990 har jag därefter gått vidare till KB:s material. I KB:s material har jag noterat antal belägg vart femte år istället för varje år, eftersom förändringarna spänner över längre perioder och frekvensdata inte går att få ut direkt som Excel-filer (vilket är fallet med data från Mediearkivet), vilket gör sökningarna betydligt mer tidskrävande i och med att antalet belägg per år måste noteras i Excel manuellt.

För vissa av varianterna förekommer belägg som egentligen inte är relevanta för undersökningen. Till exempel används *af* idag i (adliga) efternamn snarare än som stavningsvariant av *av*. Dessutom förekommer det en del felaktigheter i materialet från KB. Denna typ av brus i data har jag dock varit tvungen att bortse från i en så omfattande kvantitativ undersökning som den här. Jag har dock genomgående studerat avvikelser från förväntade resultat för att utesluta att andelen felaktiga belägg är så stor att det påverkar resultatet i någon avgörande riktning. Generellt gäller att det är lättare att studera längre söksträngar än kortare, eftersom andelen felaktiga belägg då blir mindre (av den anledningen har t.ex. skiftet från *taga* till *ta* undersökts utifrån söksträngarna *att taga/att ta*). I en del fall har felaktiga träffar gjort att en variant inte kunnat undersökas alls. Ett exempel är *hen*, som ofta förekommer i avstavningen *hen-nes* i materialet från Mediearkivet. En sökning på detta ord ger därmed alltför många felträffar för att en förändring från t.ex. *han* och *hon* till *hen* ska kunna studeras med de metoder jag har använt här. Ett annat exempel är *toalett* i relation till *WC* och *vattenklosett* – i det äldre materialet ger *toalett* alltför många träffar med annan betydelse än *WC* för att det ska vara möjligt att studera en eventuell förändring i språkbruket. Det går naturligtvis att genomföra en mer exakt studie än den som presenteras här, men med en sådan skulle det inte vara möjligt att analysera ett lika stort material. Eftersom jag anser att långtgående slutsatser utifrån begränsade datamängder varit en svaghet i tidigare forskning har det varit viktigt för mig att analysen bygger på stora mängder data – som framgick ovan har gränsvärdet här satts till minst 500 belägg/år.

Om en förändring visat sig uppfylla de kriterier jag ställt upp har jag räknat ut hur många procent av det totala antalet belägg som utgörs av den nya varianten vid olika tidpunkter och ritat grafer över utvecklingen.² Därefter har jag studerat dessa grafer noga för att se huruvida frekvensökningen kan anses likna en S-form eller snarare följer ett annat mönster.

² De frekvensdata från pressdatabaserna som analysen bygger på finns för *riksdagsledamot/riksdagsman*, *idrottare/idrottsman-män*, *talesperson/talesman* och *presstalesperson/presstalesman* publicerade som bilaga i Skärlund (2023); övriga data delas gärna – kontakta mig för att få tillgång till Excel-filerna.

De olika typer av frekvensförändringar jag funnit i materialet presenteras i nästa avsnitt, i vilket undersökningens resultat redovisas.

4 Resultat

Min undersökning av 150 språkliga variabler har resulterat i en lista med 32 frekvensökningar vilka uppfyller de kriterier som ställdes upp i föregående avsnitt. Av dessa sker 11 förändringar från och med 1993 och framåt (studerade via Mediearkivet), medan 21 påbörjas längre tillbaka i tiden (analyserade via KB:s tidningsdatabas). Förutom S-formade frekvensökningar har jag funnit dels linjära och delvis linjära ökningar, dels ojämna ökningar för vilka det är svårt att se något tydligt mönster. En översikt presenteras i tabell 1 nedan.

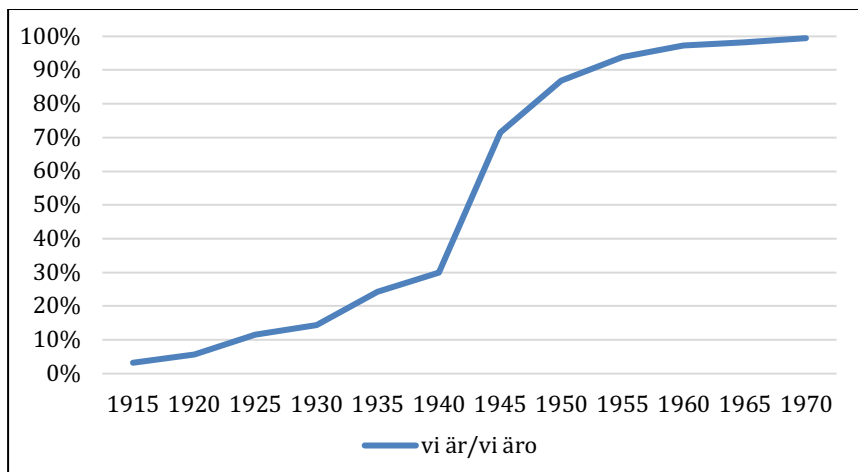
Startår för frekvensökningen är i tabellen satt till det år efter vilket ökningen överstiger en procentenhet till efterföljande nedslag och sedan fortsätter uppåt, och slutvärde till då varianten överstigit 80 % av beläggen och efterföljande ökning är mindre än en procentenhet, så att ökningen kan anses ha avstannat. Mediearkivet analyseras varje år med sista nedslag 2024, KB:s material vart 5:e år med sista nedslag 2020.

Som framgår av tabell 1 kan 14 av de 32 funna frekvensökningarna anses följa det S-kurvemönster som illustrerades i figur 1 och figur 2 ovan: den nya varianten ökar först långsamt, därefter snabbt, och slutligen långsamt igen. Ett exempel presenteras i figur 4 nedan, som visar hur *vi är* ökar i relation till *vi äro* i KB:s tidningar 1915–1970.

Tabell 1. De 32 frekvensökningar som inkluderats i undersökningen.

Analyserade variabler (nyare/äldre variant)	Period för frekvensökning av ny variant	Typ av frekvensökning
<i>bara/blott</i>	1855–1980	S-form
<i>inte/icke-ej</i>	1855–2020	linjär
<i>buss/omnibus-s</i>	1865–1940	ojämn
<i>forskare/vetenskapsman-män</i>	1865–2020	ojämn
<i>att ta/att taga</i>	1870–1960	ojämn
<i>ta emot/emottaga</i>	1880–1955	ojämn
<i>att dra/att draga</i>	1880–1960	S-form
<i>ska/skall</i>	1880–2010	delvis linjär
<i>cykel-bicykel/velociped</i>	1885–1945	S-form
<i>bio/biograf</i>	1890–1960	delvis linjär
<i>rött/rödt</i>	1895–1930	S-form
<i>gett/givit-gifvit</i>	1895–2015	ojämn
<i>av/af</i>	1900–1935	S-form
<i>flygplan/flygmaskin</i>	1915–1960	ojämn
<i>vi är/vi äro</i>	1915–1970	S-form
<i>USA/Förenta staterna</i>	1915–1980	S-form
<i>döms/dömes</i>	1935–1990	S-form
<i>kompis/kamrat</i>	1940–2015	ojämn
<i>riksdagsledamot/riksdagsman</i>	1965–2020	S-form
<i>idrottare/idrottsman-män</i>	1970–2020	delvis linjär
<i>loppis/loppmarknad</i>	1975–2015	S-form
<i>förskola/dagis</i>	1993–2023	linjär
<i>dopning/doping</i>	1998–2004	linjär
<i>e-mail-mail-e-mejl-mejl/e-post</i>	2001–2017	ojämn
<i>funktionsnedsättning/handikapp</i>	2002–2020	S-form
<i>talesperson/talesman</i>	2004–2019	S-form
<i>Liberalerna/Folkpartiet</i>	2006–2020	linjär
<i>transperson/transsexuell</i>	2006–2022	ojämn
<i>presstalesperson/presstalesman</i>	2008–2020	S-form
<i>podd/pod</i>	2010–2020	S-form
<i>Myanmar/Burma</i>	2011–2021	ojämn
<i>regioner/landsting³</i>	2014–2021	linjär

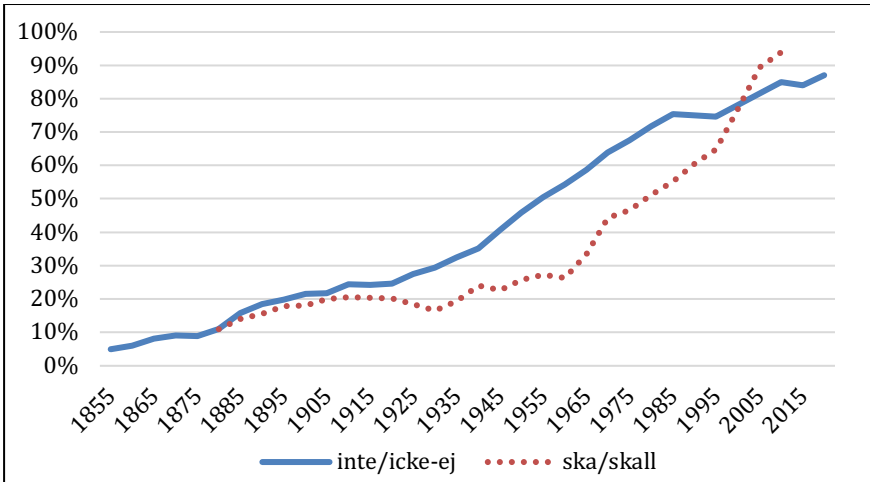
³ Eftersom *regioner* också förekommer i andra sammanhang än den här avsedda användes söksträngarna "landsting och kommuner" OR "kommuner och landsting" respektive "kommuner och regioner" OR "regioner och kommuner".



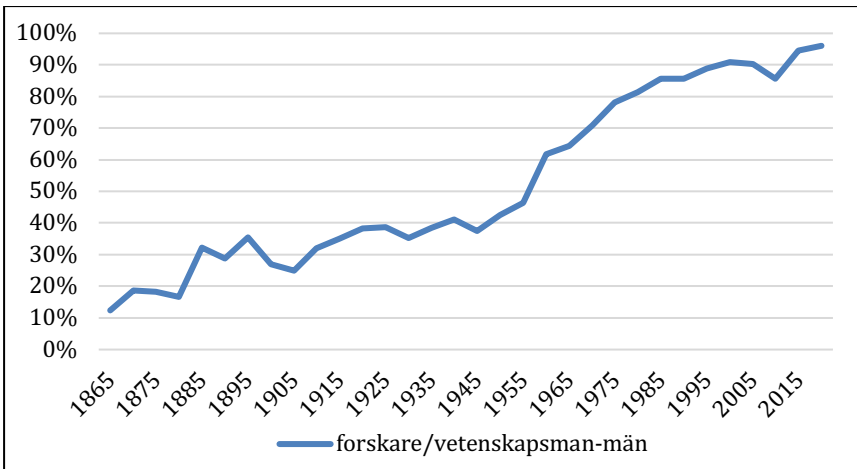
Figur 4. Exempel på S-formad frekvensökning: *vi är* ökar i relation till *vi äro* i KB:s tidningar.

Bland resterande 18 förändringar finns 5 som följer linjära utvecklingar. Som sådana har jag räknat dels ökningar som sker med ungefär samma hastighet under hela perioden (t.ex. *förskola/dagis*), dels förändringar för vilka nästan hela ökningen äger rum under endast ett eller två år (t.ex. *regioner/landsting*). Vidare kan 3 förändringar ses som delvis linjära: ökningen sker först långsamt, sedan snabbt och därefter avstannar den plötsligt utan någon nedtrappningsfas. Se figur 5, som illustrerar dessa två typer av frekvensökningar med exempel från materialet.

Slutligen finns i materialet också 10 frekvensförändringar som inte följer något av de mönster som illustrerats i figur 4 och 5. För dessa gäller att den nya varianten ökar mer ojämnt. Ett exempel syns i figur 6, som visar hur *forskare* ökar i relation till *vetenskapsman/vetenskapsmän* i KB:s tidningar 1865–2020.

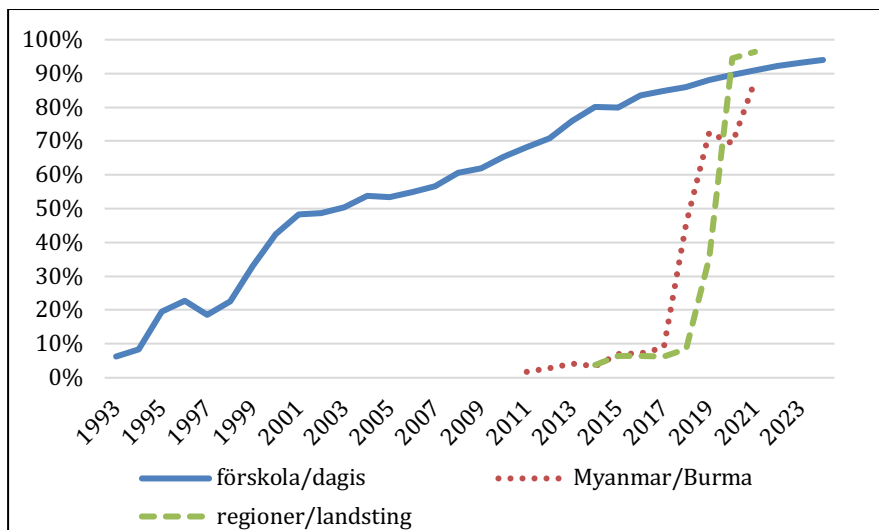


Figur 5. Exempel på linjär (*inte* i relation till *icke* och *ej*) respektive delvis linjär (*ska* i relation till *skall*) frekvensökning i KB:s tidningar.



Figur 6. Exempel på ojämn frekvensökning: *forskare* ökar i relation till *vetenskapsman/vetenskapsmän* i KB:s tidningar.

Bland de 18 variabler som inte följer S-formen finns både ökning som beror på förändrade riktlinjer uppifrån (t.ex. att *Myanmar* ersatte *Burma*, vilket beslutades av Mediespråksgruppen 2018) och sådana som troligen är mer omedvetna från journalistens sida (t.ex. ett skifte från *givit* till *gett*). Se figur 7, där tre av de frekvensförändringar som inte följer S-mönstret presenteras.

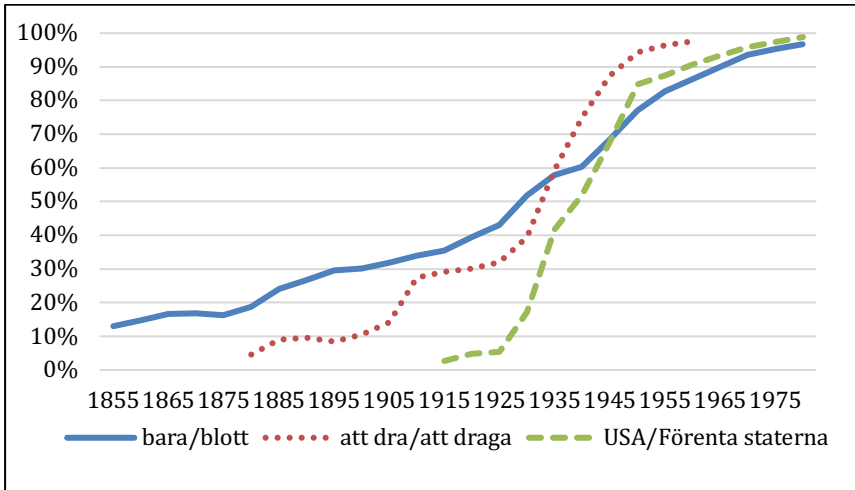


Figur 7. Tre frekvensökningar som inte följer S-formen i presstext i Mediearkivet: *förskola*, *Myanmar* och *regioner* ökar.

Som framgår av figur 7 kan medvetna förändringar ge tydliga hack i kurvorna: bruket ändras snabbt från ett år till ett annat. Att nya riktlinjer för namn på länder (*Myanmar*) eller nya namn på verksamheter (*regioner*) inte följer S-formade mönster i tidningstext är inte särskilt förvånande; det kan liknas vid hur den tyska stavningsreformen tog sig uttryck (Ghanbarnejad et al., 2014). Det är emellertid också uppenbart att inte alla reformer får samma genomslag: *förskola* ersatte *daghem* (*dagis*) i och med den nya läroplanen 1998, men den nya varianten ökar istället långsamt (och närmast linjärt) under en period av 30 år.

Av de 14 S-formade frekvensökningarna är åtminstone fyra förmodligen medvetna val eller avhängiga uttalade språknormer på den tidning vid vilken journalisten arbetar, eftersom de förknippas med specifika reformer (*funktionsnedsättning* blev den rekommenderade termen från Socialstyrelsens terminologiråd 2007, *af* respektive *rödt* byttes till *av* respektive *rött* i samband med stavningsreformen 1906 och pluralformen *vi äro* byttes gradvis ut mot singularformen *vi är* under första halvan av 1900-talet – TT:s införande av singularformerna 1945 brukar ses som ett nyckelår för förändringen). Också ett ökat undvikande av de tre personbeteckningarna med *-man* (*riksdagsman*, *talesman*, *presstalesman*) kan kanske tänkas utgöra medvetna val från journalistens sida, även om dessa inte kan knytas till en enskild reform. Möjligen gäller detta också skiftet från engelsk till svensk stavning i *dopning* och *podd*. I så fall återstår endast fem S-formade frekvensökningar som journalisten eventuellt varit mer

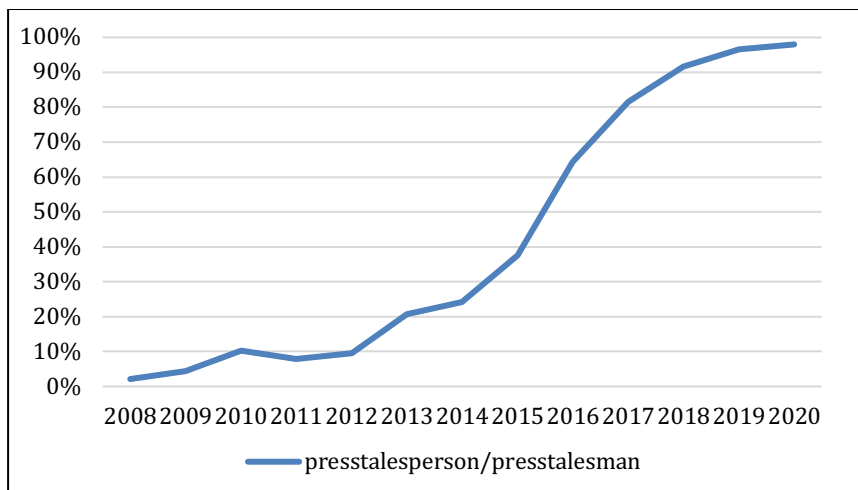
omedveten om: skiftena från *blott* till *bara*, *att draga* till *att dra*, *dömes* till *döms*, *velociped* till *cykel* och *Förenta staterna* till *USA*. Tre av dessa presenteras i figur 8.



Figur 8. Tre frekvensökningar som följer S-formen i KB:s tidningar: *bara*, *USA* och *att dra* ökar.

Figur 8 illustrerar också ett viktigt problem med undersökningen: det är i vissa fall svårt att avgöra om en förändring ska räknas som S-formad eller inte, eftersom det inte finns någon bestämd definition av hur olik den utveckling som illustrerades i figur 1 en frekvensökning får vara och ändå anses S-formad. Jag har sett ökningen av *bara* i figur 8 som S-liknande eftersom den går (något) snabbare under mitten av perioden än under början och slutet av den, men ökningen är inte långt ifrån att vara linjär. Vidare har jag ansett att ökningen av *att dra* följer S-formen trots en liten inledande puckel under vilken ökningen tillfälligt avstannar. Att utvecklingen följer en S-kurva om den är "slow, slow, quick, quick, slow" (Denison, 2003, s. 56) förefaller inte vara en tillräckligt utförlig definition.

I ett fall utgör dock frekvensökningen närmast ett skolexempel på en S-formad utveckling. Jämför figur 9 med artikelns inledande figur 1.



Figur 9. Ökning av *presstalesperson* i relation till *presstalesman* i presstext i Mediearkivet.

5 Diskussion

I den undersökning av 408 franska uttryck som genomfördes av Feltgen et al. (2017) var andelen S-formade frekvensökningar 72 %. I min studie utgjorde de mindre än hälften av de undersökta förändringarna: 14 av 32. En slutsats som kan dras av denna, förvisso mycket begränsade, studie är därmed att S-kurvan är en hållbar modell för att beskriva en del språkförändringar, men långt ifrån alla, och att kategoriska uttalanden om att förändringar generellt följer en förutsägbar utveckling i form av en S-kurva (se t.ex. Tagliamonte, 2015) förefaller behöva nyanseras. Snarare tyder den här studien på att ”the S-curve is neither as simple nor as uniform a phenomenon as is sometimes assumed”, som Denison (2003, s. 68) påpekat.

Blythe och Croft (2012, s. 180) anger att de inte känner till några språkförändringar för vilka en ny variant ökar linjärt istället för som en S-form. Min undersökning visar dock att 8 av de 32 analyserade förändringarna ökar linjärt eller delvis linjärt. Sådana förekommer alltså, även om de möjligen inte är lika vanliga som S-formen.

Mest överraskande anser jag det vara att majoriteten av de 14 frekvensökningar som följer S-mönstret antagligen kan ses som utslag av medvetna val från journalistens sida. Många av de språkförändringar som kan tänkas vara mer omedvetna för journalisten – t.ex. *inte/icke-ej*, *ta emot/emottaga*, *ska/skall*, *gett/givit* – är däremot inte S-formade. Snarare hade jag förvän-

tat mig att det motsatta skulle vara fallet (jfr Ghanbarnejad et al., 2014). Omedvetna språkförändringar är kanske de mest intressanta att analysera, och för dessa tycks det alltså inte självklart att S-kurvan är den mest träffande modellen. Samtidigt saknas en tydlig definition av hur en utveckling ska se ut för att räknas som S-formad. Utan en sådan kan olika forskare inkludera vitt skilda frekvensförändringar i modellen, vilket gör det svårt att föra en diskussion om dess träffsäkerhet.

En mer omfattande studie, lik den Feltgen et al. (2017) genomfört för franskans räkning, skulle vara intressant också för svenskans del. Mina genomgångar av 150 variabler har dock visat att det är mycket svårt att finna språkliga uttryck som både genomgår stora frekvensökningar och dessutom är så vanliga i presstext att det ska vara rimligt att dra någorlunda säkra slutsatser om utvecklingen. I takt med att allt mer material digitaliseras bör det dock bli enklare att genomföra sådana undersökningar.

Litteratur

- Ahlberg, J. (19 oktober 2024). Kort om databasens innehåll. *Svenska tidningar – frågor och svar*. <https://feedback.blogg.kb.se/forums/topic/kort-om-databasens-innehall/> [2024-12-27]
- Aitchison, J. (2013). *Language change: progress or decay?* (4 uppl.). Cambridge University Press.
- Blythe, R. A., & William, C. (2012). S-curves and the mechanisms of propagation in language change. *Language*, 88(2), 269–304.
<https://www.jstor.org/stable/23251832>
- Chen, M. (1972). The Time Dimension: Contribution toward a Theory of Sound Change. *Foundations of Language*, 8(4), 457–498.
<https://www.jstor.org/stable/25000618>
- Denison, D. (2003). Log(ist)ic and simplistic S-curves. I R. Hickey (Red.), *Motives for language change* (s. 54–70). Cambridge University Press.
- Einarsson, J. (2009). *Språksociologi* (2 uppl.). Studentlitteratur.
- Feltgen, Q., Fagard, B., & Nadal, J.-P. (2017). Frequency patterns of semantic change: corpus-based evidence of a near-critical dynamics in language change. *Royal Society open science*, 4, 1–14.
<http://dx.doi.org/10.1098/rsos.170830>
- Ghanbarnejad, F., Gerlach, M., Miotto, J. M., & Altmann, E. G. (2014). Extracting information from S-curves of language change. *Journal of the Royal Society Interface*, 11, 10–44.
- KB = Kungliga biblioteket. (2024). *Svenska tidningar*. <https://tidningar.kb.se/>
- Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1(3), 199–244.
<https://doi.org/10.1017/S0954394500000168>

- Labov, W. (1994). *Principles of linguistic change Vol. 1 Internal factors*. Blackwell.
- Mårtensson, E. (1988). Förändringar i 1900-talets svenska: en litteraturgenomgång. I E. Mårtensson & J. Svensson (Red.), *Offentlighetsstruktur och språkförändring* (s. 101–217). Institutionen för nordiska språk, Lunds universitet.
- Nevalainen, T. (2015). Descriptive adequacy of the S-curve model in diachronic studies of language change. *Studies in Variation, Contacts and Change in English*, 16. <https://varieng.helsinki.fi/series/volumes/16/nevalainen/>
- Nevalainen, T., & Raumolin-Brunberg, H. (2003). *Historical sociolinguistics: Language change in Tudor and Stuart England*. Longman.
- Osgood, C. E., & Sebeok, T. A. (Red.) (1954). *Psycholinguistics: a survey of theory and research problems: report of the 1953 Summer Seminar*. Waverly Press.
- Retriever. (2024). *Mediearkivet*. <https://www.retrievergroup.com/sv/product-medicarkivet>
- Rogers, E. (1995/1962). *Diffusion of innovations* (4 uppl.). Free press.
- Skärlund, S. (2023). När mannen blev en person – Personbeteckningar med och utan efterledet -man i svensk tidningstext. *HumaNetten*, 50, 213–260. <https://doi.org/10.15626/hn.20235010>
- Sundgren, E. (2013). Språklig variation och förändring. I E. Sundgren (Red.), *Sociolingvistik* (2 uppl.) (s. 77–121). Liber.
- Svensk ordbok (2021). Ordet hållbar på svenska.se: <https://svenska.se/tre/?sok=h%C3%A5llbar&pz=1> [2024-12-27]
- Tagliamonte, S. A. (2015). Exploring the architecture of variable systems to predict language change. *Studies in Variation, Contacts and Change in English*, 16. <https://varieng.helsinki.fi/series/volumes/16/tagliamonte/>
- Weinreich, U., Labov, W., & Herzog, M. I. (1968). Empirical Foundations for a Theory of Language Change. I M. Yakov & W. P. Lehmann (Red.), *Directions for historical linguistics: a symposium [held at the University of Texas, [Austin] 1966]* (s. 96–195). Univ. of Texas Press.
- Winter-Froemel, E. (2014). What does it mean to explain language change? Usage-based perspectives on causal and intentional approaches to linguistic diachrony, or: On S-curves, invisible hands, and speaker creativity. *Energeia*, V, 123–142. <https://doi.org/10.55245/energeia.2014.008>