

Relationer, samordningar och retorik på Språkbanken

Swe-Clarin workshop

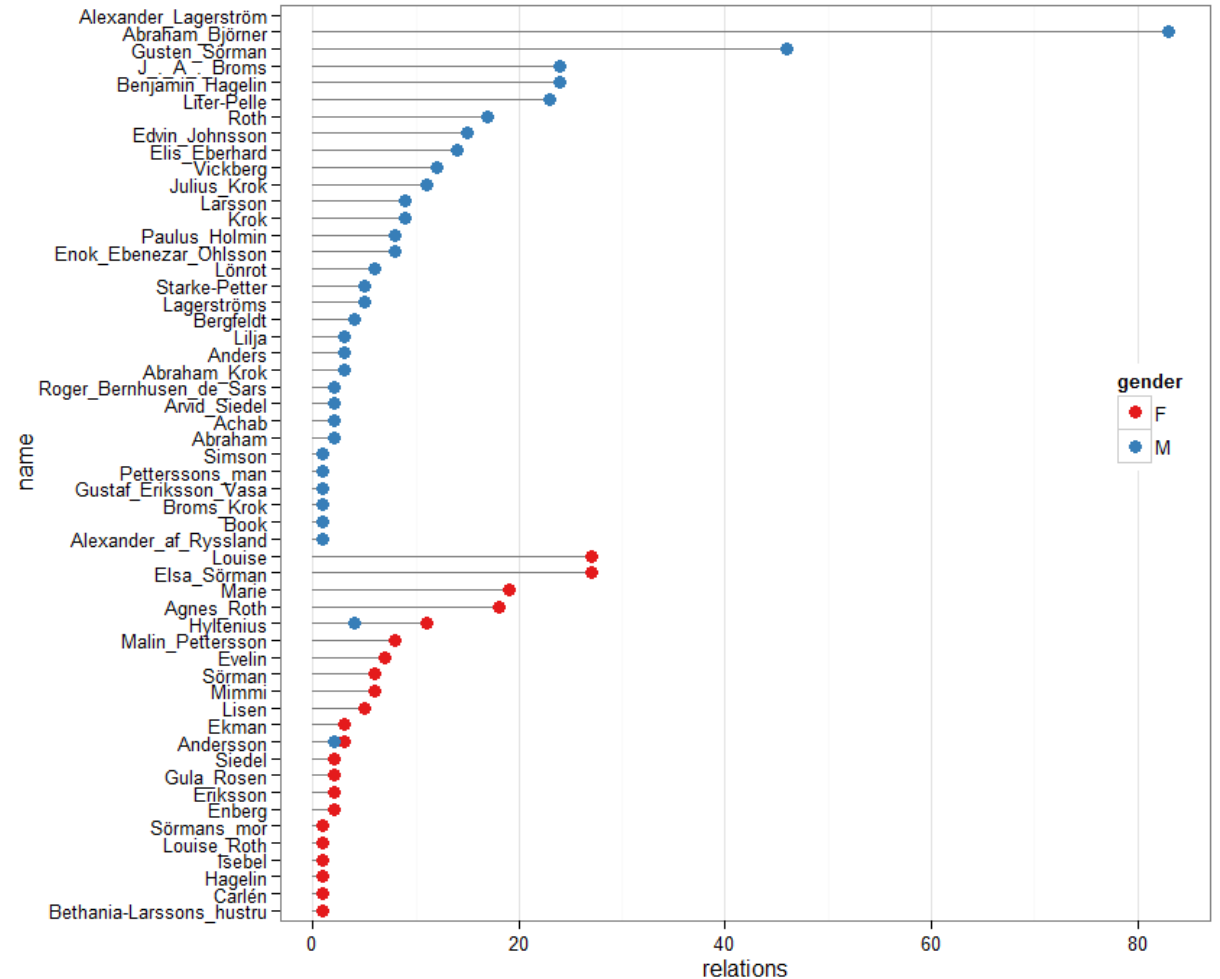
17 april 2015

Nina Tahmasebi

Semantik i historieberättande i Svensk Prosafiktion

Relationsextraktion från Svensk Prosafiktion (SPF) 1800–1900

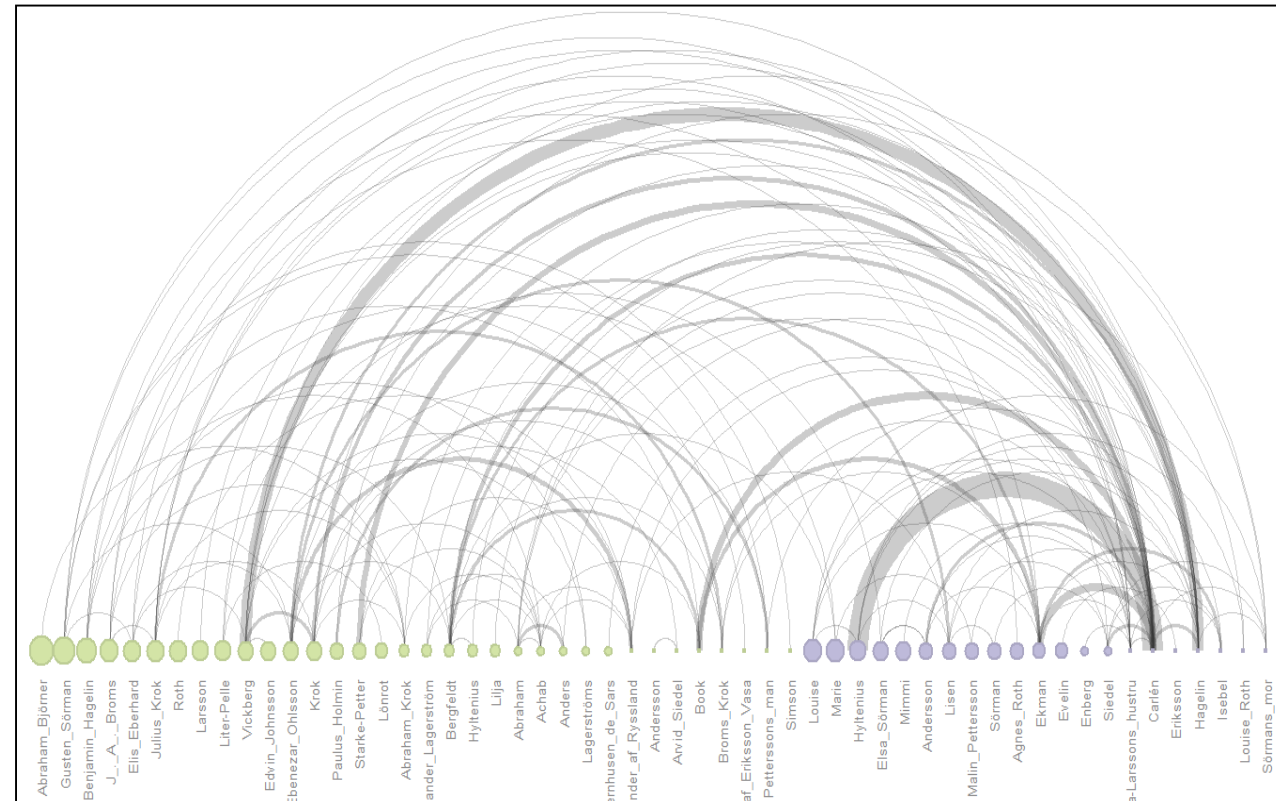
- Lista av **relationer**
- NEE för att hitta **namn** och **alias**
document-centered approach för att länka alias → namn
- Hitta meningar med minst 2 namn
- Upptäck relationer
→ Automatisk upptäckt av relationer ger många fler!



Semantik i historieberättande i Svensk Prosafiktion

Relationsextraktion från Svensk Prosafiktion (SPF) 1800–1900

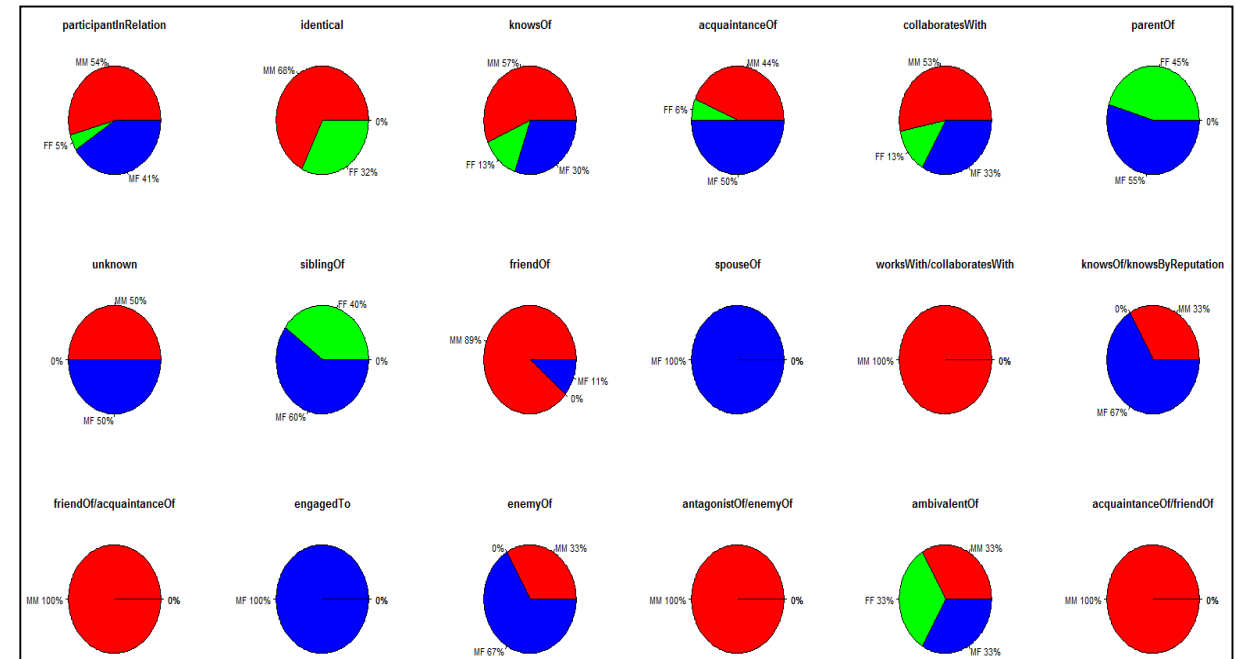
- Lista av **relationer**
- NEE för att hitta **namn** och **alias**
document-centered approach för att länka alias → namn
- Hitta meningar med minst 2 namn
- Upptäck relationer
→ Automatisk upptäckt av relationer ger många fler!



Semantik i historieberättande i Svensk Prosafiktion

Relationsextraktion från Svensk Prosafiktion (SPF) 1800–1900

- Lista av **relationer**
- NEE för att hitta **namn** och **alias**
document-centered approach för att länka alias → namn
- Hitta meningar med minst 2 namn
- Upptäck relationer
→ Automatisk upptäckt av relationer ger många fler!



Relations between 2 males = red, between 2 females = green, otherwise blue.

Svenska Pseudosamordningar – Upptäckt och förändring

- Verbpar där första verbet är betydelsesvagt
- "åka och handla", "gå och gifta sig", "ringa och berätta"
- Typiska egenskaper:
 - "Både" är inte möjligt: "jag både satt och läste"
 - Omformuleringar är orimliga: "Mona satt och hon läste".
- Försök att skilja SPCer från icke-SPCer genom att använda dessa egenskaper
- Bloggmixen
 - 505 miljoner tokens
 - 1998-2014
- Bonniers romaner I & II
 - 11 miljoner tokens
 - 1976-1981

Svenska Pseudosamordningar – Upptäckt och förändring

| Precision | Recall | Class |
|-----------|--------|--------------|
| 0.874 | 0.899 | SPC |
| 0.916 | 0.894 | Non-SPC |
| 0.897 | 0.897 | Weighted avg |

Precision and recall for Blogmixen

De egenskaper vi har definierat kan separera SPCer från icke-SPCer med ca 90% säkerhet

→ Hur hittar vi nya SPCer?

| Första verb | Konf. | #par | Par |
|---------------|-------|------|--|
| åka | 1.0 | 120 | [campa, beställa, bidra,...] |
| vakna* | 1.0 | 63 | [se, leva, ligga, ...] |
| stanna | 1.0 | 65 | [käka, städa, säga,...] |
| packa | 1.0 | 19 | [förbereda, vänta, åka, ...] |
| äta | 1.0 | 35 | [lyssna, fixa, titta, ...] |
| ramla* | 1.0 | 9 | [bryta, slå, skrapa, skada, skylla, ...] |
| dra | 1.0 | 47 | [kolla, storhandla, gymma, ...] |
| nicka* | 1.0 | 4 | [säga, se, komma, ...] |
| fara | 1.0 | 30 | [käka, fixa, fika, ...] |
| googla | 1.0 | 9 | [titta, hitta, upptäcka, ...] |
| maila | 1.0 | 13 | [vilja, tipsa, fråga, ...] |
| stog | 1.0 | 7 | [vänta, titta, kolla ...] |
| kommentera* | 0 | 14 | [dela, motivera, fråga, ...] |
| mejla | 1.0 | 7 | [berätta, säga, kolla, ...] |
| messa | 0.9 | 6 | [vilja, undra, säga, ...] |
| prata | 0.9 | 25 | [bestämna, räkna, låtsas,] |

Retorik

- Hur har begrepp retorik ändrats över tid?
- I *vilka sammanhang* talades det om t.ex. retorik och vältalighet förr och hur har detta ändrats över tid?
- Använder oss av begreppskluster:
 1. Vältalighet, wältalighet, eloquentia, vältalare ...
 2. Deklamatorisk, deklamation, declamation, föredrag ...
 3. Retorik, rhetorik, retorica, ...
 4. Sofist, sophist, mångordig, vidlyftig, ordprål,...
- Kubhist-materialet
- 1750-1925

Tre metoder – **kontext**, attityd och ordbild

Kontext genom att kolla på närliggande ord.

Ett exempel: Stockholmsposten 1830-talet

”Profvet på afsöndring af **den franska retorikens betydelselösa ordmassa** och dess af en falsk smak tillskapade prydnader...”

Översikt ger mer!

tjusningskraft

betydelselös

oemotståndlig

förnuftig

mäktig

→ Dela upp texten i 10 års-perioder och jämför dessa för att hitta förändringar i kontext.

Tre metoder – kontext, **attityd** och ordbild

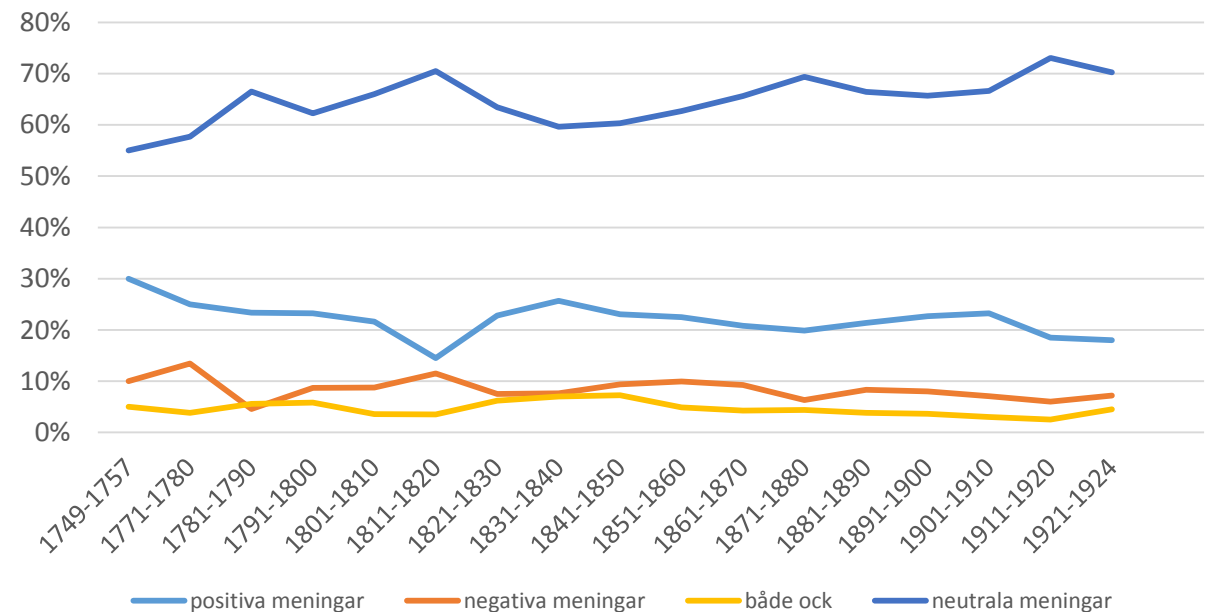
Kolla attityden genom att undersöka värdeord kring ord ur begreppsklustren.

Lista av positiva och negativa ord

Klasser av värdeord ur tesaursen *Svenskt ordförråd ordnat i begreppsklasser* (Bring 1930)

- Alla meningar kan klassas positiv/negativ/både/neutral
- Enskilda meningar & överblick

Begreppskluster Vältalighet



Tre metoder – kontext, attityd och ordbild

vältalighet (substantiv) Allt Kubhistmaterial

| Preposition | Attribut | vältalighet | Efterställt Attribut |
|------------------|----------|------------------------|----------------------|
| 1. med | 222 ☞ | 1. politisk | 30 ☞ |
| 2. genom | 39 ☞ | 2. parlamentarisk | 14 ☞ |
| 3. i sin | 5 ☞ | 3. manlig | 8 ☞ |
| 4. rörande | 4 ☞ | 4. manlig ² | 8 ☞ |
| 5. för | 31 ☞ | 5. vanlig | 13 ☞ |
| 6. öfver | 5 ☞ | 6. eldig | 6 ☞ |
| 7. i besittning | 1 ☞ | 7. stor | 31 ☞ |
| 8. af | 137 ☞ | 8. andlig | 8 ☞ |
| 9. åt | 3 ☞ | 9. romersk | 7 ☞ |
| 10. till och med | 2 ☞ | 10. verklig | 10 ☞ |
| 11. till och med | 2 ☞ | 11. sann | 8 ☞ |
| 12. trots | 2 ☞ | 12. naturlig | 8 ☞ |
| 13. efter | 3 ☞ | 13. beundransvärd | 5 ☞ |
| 14. utan | 3 ☞ | 14. stum | 5 ☞ |
| 15. i förening | 1 ☞ | 15. kraftfull | 5 ☞ |

Attributen är överlag positiva

vältalighet (substantiv) 1770-1800 talet

| Preposition | Attribut | vältalighet | Efterställt Attribut |
|-------------|----------|------------------------|----------------------|
| 1. med | 8 ☞ | 1. manlig | 4 ☞ |
| 2. rörande | 1 ☞ | 2. manlig ² | 4 ☞ |
| 3. jämte | 1 ☞ | 3. okonstlad | 1 ☞ |
| 4. för | 3 ☞ | 4. egentelig | 1 ☞ |
| 5. af | 5 ☞ | 5. män^elig | 1 ☞ |
| 6. genom | 1 ☞ | 6. oförliknelig | 1 ☞ |
| 7. i | 3 ☞ | 7. rik | 1 ☞ |
| | | 8. cill | 1 ☞ |
| | | 9. andclig | 1 ☞ |
| | | 10. latinsk | 1 ☞ |
| | | 11. owanlig | 1 ☞ |
| | | 12. mycken | 1 ☞ |
| | | 13. listig | 1 ☞ |
| | | 14. sann | 1 ☞ |
| | | 15. wanliga | 1 ☞ |

Attributen är överlag positiva

vältalighet (substantiv) 1850-talet

| Preposition | Attribut | vältalighet | Efterställt Attribut |
|-----------------|----------|-------------------|----------------------|
| 1. med | 51 ☞ | 1. politisk | 18 ☞ |
| 2. genom | 11 ☞ | 2. kraftfull | 3 ☞ |
| 3. till och med | 1 ☞ | 3. parlamentarisk | 2 ☞ |
| 4. till och med | 1 ☞ | 4. mäktig | 2 ☞ |
| 5. inför | 1 ☞ | 5. sofistisk | 1 ☞ |
| 6. liksom | 1 ☞ | 6. otänd | 1 ☞ |
| 7. i sin | 1 ☞ | 7. förrättika | 1 ☞ |
| 8. under | 1 ☞ | 8. bjerllig | 1 ☞ |
| 9. åt | 1 ☞ | 9. rgmersk | 1 ☞ |
| 10. rörande | 1 ☞ | 10. känslorik | 1 ☞ |
| 11. utan | 1 ☞ | 11. käi.ua | 1 ☞ |
| 12. för | 5 ☞ | 12. sjellbchaglig | 1 ☞ |
| 13. på | 2 ☞ | 13. diabolisk | 1 ☞ |
| 14. af | 36 ☞ | 14. ordrik | 1 ☞ |
| 15. i | 12 ☞ | 15. svulstig | 1 ☞ |

Attributen är blandade

Sammanfatta utan att tappa länkarna!

| | | | | |
|----------|----------|----|------|---|
| ordalag | vältalig | 13 | 1880 | vältalig varm och djupt allvarlig ordalag tacka landshöfding Treffenberg fara den honom och hans familj egna hyllning |
| föredrag | vältalig | 13 | 1877 | Det vältalig föredrag stöta in ed uppläsande af Etraudbergs afstedssång Tack för god wall hrvarefter ett fölhänge scen bakgrund upprulla och en tablå gaf ytterlig erinran om änniet för dag fest Allt detta känna liten Maria der hon läg timma efter timma ensam på sitt läger känna icke tänka ty rötter till verklig |
| skola | vältalig | 13 | 1875 | tanke vart jordelif ligga djupt barnahjertats känsla ehuru barn mun icke kunna uttala dem ens för dem sjelfva och ehuru somligas läpp aldrig öppna för att uttala dem för någon en tystnad af ringa vigt emedan denna tid snarare vara ämnad att vara en lärotid och många af våra mest vältalig yttrande här skola endast förefalla såsom barn joller der Hau vara icke lycklig vara den tanke om åter och åter stå för henne och hon tycka ig ännu beständigt kannan loris blyg |
| blick | vältalig | 13 | 1880 | och vördnadtfulla handtryckning ännu beständigt se hans tankfull och vältalig blick fasta på aig skymning halfdunkel |

Tack för er tid!

Delaktiga i arbete som presenterats:

Malin Ahlberg, Peter Andersson, Lars Borin, Markus Forsberg, Dimitrios Kokkinakis, Mats Malm, Nina Tahmasebi, Jon Viklund