

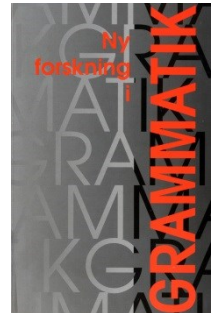
Ny Forskning i Grammatik

Titel: Sprogtypologi og anaforer
Om anaforrelationer i endo- og exocentriske sprog

Forfatter: Iørn Korzen

Kilde: Ny Forskning i Grammatik 22, 2015, s. 149-166

URL: <http://ojs.statsbiblioteket.dk/index.php/nfg/issue/archive>



© Forfatterne og Dansk Sprognævn, 2015

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Søgbarhed

Artiklerne i de ældre numre af Ny Forskning i Grammatik (1993-2012) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

Sprogtypologi og anaforer

Om anaforrelationer i endo- og exocentriske sprog

Iørn Korzen

1. Indledning

Det er alment kendt at anaforrelationer er særdeles centrale i teksters struktur, og det er også alment kendt at tekststruktur varierer fra sprog til sprog, og især fra sprogfamilie til sprogfamilie. Men det er nok mindre kendt om – og i givet fald præcis hvordan – typologiske forskelle i tekststruktur manifesterer sig i anaforrelationerne. Formålet med dette indlæg er at beskrive nogle fundamentale forskelle i anaformanifestationer i de såkaldte “endo-” og “exocentriske” sprog og relatere disse forskelle til generelle og typologiske forskelle i tekstuelle og leksikalske strukturer i de to sprogtyper.

Mit indlæg har følgende disposition: I afsnit 2 beskriver jeg kort forskellene mellem endo- og exocentriske sprog samt den empiriske basis for min undersøgelse. I afsnit 3 dokumenterer jeg nogle væsentlige kvantitative forskelle i forekomsterne af de forskellige nominale anafortyper i mit korpus, og i afsnit 4 og underafsnit giver jeg en række eksempler på anaforforekomsterne, som jeg i afsnit 5 sammenholder med leksikalske strukturer i endo- og exocentriske sprog. I afsnit 6 relaterer jeg den sproglige markering af anaforer til specifikke fænomener i tekststruktur.

2. Endo- og exocentriske sprog og den empiriske basis

“Endocentriske” og “exocentriske” sprog er CBS-forskergruppen TYPOLex’s termer for hhv. germanske og romanske sprog, og termerne er baseret på grundlæggende forskelle i leksikaliseringmønstre.¹ Forskellene kan illustreres ved sætningerne i tabel 1.

1 TYPOLex-gruppen ledes af Michael Herslund og består i øvrigt af Irene Baron, Hanne Korzen, Iørn Korzen, Lita Lundquist, Henrik Høeg Müller og Viktor Smith. En oversigt over gruppens fællespublikationer findes i Herslund (2009). Herefter er udkommet Cresti & Korzen (red.) (2010) og Korzen, Ferrari & De Cesare (red.) (2014).

		Periferi	Centrum	Periferi
Endocentriske (germanske) sprog	Dansk Engelsk	[FUNKTION] <i>Bilen</i> <i>The car</i>	[+ MÅDE] kører ind i drives into	[FUNKTION] <i>gården</i> <i>the yard</i>
Exocentriske (romanske) sprog	Italiensk Fransk	[KONFIGURATION] <i>L'automobile</i> <i>La voiture</i>	[- MÅDE] <i>entra ne</i> <i>entre dans</i>	[KONFIGURATION] <i>(i)l cortile</i> <i>la cour</i>

Tabel 1. Leksikaliserede betydningskomponenter i verber og substantiver i endo- og exocentriske sprog

De germanske sprog har stor leksikalsk præcision eller “vægt” i deres verber pga. den typiske leksikalisering af betydningskomponenten MÅDE. I tabel 1 har verberne *køre* og *drive* en mere præcis betydning end *entrare*, *entrer*, idet de specificerer MÅDEN hvorpå verbalhandlingen udføres. Det betyder at der i germanske sprog ligger en stærk betydningsvægt i centrum af sætningen, og da verbalet opfattes som dette centrum, er termen **endocentrisk** valgt for disse sprog.

Derimod har de romanske sprog en større betydningsvægt i deres substantiver, hvilket skyldes at betydningskomponenten KONFIGURATION er distinktiv i leksikaliseringsmønstret af romanske substantiver. De germanske substantiver adskiller sig derimod typisk ved betydningskomponenten FUNKTION, og da forskellige objekter godt kan have den samme funktion, men se forskellige ud, medfører dette en mere specificeret leksikalisering af romanske substantiver end af germanske. I tabel 1 er de romanske substantiver *automobile*, *cortile*, *voiture* og *cour* mere specificerede og præcise – de har større intension og dermed mindre ekstension, dvs. de denoterer færre entiteter i verden – end *bil*, *gård*, *car* og *yard*. I de romanske sprog ligger der således en større leksikalsk vægt uden for sætningens centrum, derfor termen **exocentrisk** for disse sprog.²

Forskellen i semantisk vægt kan illustreres som i tabel 2.

Endocentriske (germanske) sprog	n - V - n
Exocentriske (romanske) sprog	N - v - N

Tabel 2. Fordeling af semantisk vægt i endo- og exocentriske sprog

2 Jf. fx Herslund & Baron (2003), Baron & Herslund (2005), Korzen (2005a/b), Herslund (2010). Disse leksikaliseringsmønstre gælder for substantiver der angiver artefakter, og de beskriver generelle tendenser i de to sprogfamilier, men er ikke uden undtagelser. Også naturligt skabte genstande leksikaliseres typisk på et hyperonymniveau i de germanske sprog i forhold til de romanske. For dansk-franske sammenligninger jf. Herslund (1997: 30ff) og for dansk-italienske Korzen (2000: 186ff).

På de følgende sider vil jeg som eksempler på de to sprogtyper benytte dansk og italiensk, og som empirisk basis anvender jeg et korpus af danske og italienske L1-tekster fra det såkaldte “Europarl corpus” (Koehn 2005, <http://statmt.org/europarl/>), der består af politiske taler holdt i Europaparlamentet fra 1996 og frem. Til mine statistiske beregninger benytter jeg et lige stort dansk og italiensk underkorpus af taler, der er udvalgt således at de er indbyrdes nogenlunde lige lange og repræsenterer en så stor variation som muligt af forfattere og politiske temaer.

	Danske	Italienske
Antal politiske taler (alle L1)	50	50
Samlet antal ord	14.737	14.708

Tabel 3. Europarl-korpus anvendt til statistik

3. Anafortypologi og lidt statistik

De anafortyper og -undertyper der er specielt interessante i relation til typologiske forskelle i tekststruktur og leksikon, er nævnt i tabel 4. De koreferentielle anaforer henviser til samme referent som antecedenten, de “tro” består af NP’er der er leksikalsk identiske med antecedenten, de “utro” af NP’er der er leksikalsk forskellige fra antecedenten. De associative anaforer refererer til en størrelse der kan associeres med antecedenten, og de resumptive indkapsler (“resumerer”) en hel sætning eller tekstsegment:

Anafortyper	Eks. (antecedent ← anafor)
Koreferentielle anaforer	
• NP’er, “tro”	<i>et dokument</i> ← <i>dokumentet</i>
• NP’er, “utro”	<i>et dokument</i> ← <i>teksten</i>
• Pro-former (cvt. ital. Ø-subj.)	<i>et dokument</i> ← <i>det, dette, ital.: lo, esso, Ø</i>
Associative anaforer	<i>et dokument</i> ← <i>forfatteren</i>
Resumptive anaforer	
• NP’er	<i>Der er ikke mere vin.</i> ← <i>Det problem (må løses)</i>
• Pro-former (cvt. ital. Ø-subj.)	<i>Der er ikke mere vin.</i> ← <i>Det (er et stort problem)</i>

Tabel 4. Anafortypologi

Jeg optalte forekomsterne af disse anafortyper i mit korpus og kom til resultatet i tabel 5.

	Danske	Italienske	Forskel ³
Samlet antal anaforer	798	743	7,4 %
Koreferentielle anaforer	595	535	11,2 %
• NP'er, "tro"	232	110	110,9 %
• NP'er, "utro"	44	68	54,5 %
• Pro-former	319	357	11,9 %
Associative anaforer	70	111	58,6 %
Resumptive anaforer	133	97	37,1 %
• NP'er	30	39	30,0 %
• Pro-former	103	58	77,6 %

Tabel 5. Anaforforekomster i det undersøgte korpus

Under de "tro" koreferentielle anaforer er medtalt *proprier* anvendt flere gange om samme referent (inkl. tilfælde som *Parlamentet*, *Kommissionen*, *Nice-traktaten* osv.), selvom man her kan diskutere den anaforiske værdi versus en egentlig koreference, dvs. reference udtrykt på præcis samme måde i hver enkelt NP (Korzen 1996: 133-135; 2006: 327-328). Jeg har valgt denne fremgangsmåde for bl.a. også at kunne skelne mellem gentagelser af *proprier* (der altså tæller som "tro" NP'er) versus "ægte" anaforisk reference af typen *Danmark* ← *mit land*, *Toscana* ← *min region*, *EU* ← *Unionen* osv. (der tæller som "utro" NP'er). Under pro-formerne er medtalt italienske nul-subjekter ved finitte – men ikke infinite – verbalformer.

Af tallene i tabel 5 kan udledes følgende:

1. Det samlede antal anaforer er højere i dansk end i italiensk.
2. Specielt er de koreferentielle og resumptive anaforer hyppigere i dansk end i italiensk.
3. Blandt de koreferentielle anaforer er de "tro" markant hyppigere i dansk, mens de "utro" er hyppigere i italiensk.
4. Blandt de resumptive anaforer er pro-formerne markant hyppigere i dansk, mens NP-anaforerne er hyppigere i italiensk.
5. De associative anaforer er væsentligt hyppigere i italiensk end i dansk.

3. I alle tilfælde er procentforskellen beregnet ud fra det laveste af de to relevante tal.

4. Anafortypologi. Eksempelmateriale

Da alle de her analyserede anaforer er nominale, er der tale om substantiviske ledfunktioner, dvs. subjekt, objekt og styrelse for præposition, og et højere antal valensled i dansk end i italiensk kan være tegn på en forskel i antallet af propositioner i de to korpora. Det skal jeg vende tilbage til.

4.1. Koreferentielle anaforer

Blandt de koreferentielle anaforer optræder de mest markante forskelle mellem de tro og de utro NP-anaforer. Et par typiske danske eksempler på tro anaforer:

- (1) *Denne sag* [om sporbarhed for slagtekvæg] er også et spørgsmål om troværdighed omkring implementering af fælles beslutninger i alle medlemslande. **Denne sag** handler i øvrigt om borgernes tillid til og tro på, at fødevarer kvaliteten er i orden. <ep-00-04-11.txt:271>⁴
- (2) I fordelingen af *fødevarerhjælpen* bør civilsamfundsorganisationer spille en afgørende rolle, for at undgå at **fødevarerhjælpen** misbruges politisk. <ep-02-09-04.txt:160>

I italiensk er utro anaforer derimod mere frekvente end i dansk.⁵ Til illustration af dette forhold og for at sammenligne den typiske sprogbrug ved et informationsindhold som kan formodes at have nogenlunde den samme bekendthedsgrad for et dansk og et italiensk publikum, har jeg fundet eksempel (3) og (4) i to avisers elektroniske udgaver.

- (3) Det er her, USA's præsident, Barack Obama, den russiske præsident, Vladimir Putin, værten David Cameron, **der er britisk premierminister**, og adskillige andre ledere er samlet til topmøde i organisationen G8. (Børsen, 2013)
- (4) Il duello USA-Russia. [...] Tra i due litiganti, prova ad inserirsi il terzo incomodo: *Cameron. Il leader britannico* ha incontrato sia Putin che Obama, ed ora vuole giocarsi le sue carte. (Diritto di critica, 2013)

4 Henvisningerne til Europarl-teksterne ("ep") består af år-måned-dag efterfulgt af talernummer. I Europarl-eksemplerne angives antecedenterne i kursiv og anaforerne i fed kursiv. Som i eks. (1) kan en antecedent selv fungere som anafor (her: resumptiv) for en anden antecedent.

5 Lundquist (2005) er kommet til et helt parallelt resultat i en sammenligning mellem dansk og fransk.

‘Duellen USA-Rusland. [...] Mellem de to stridende parter forsøger en tredje ubekvem part at komme ind: *Cameron. Den britiske leder* har mødt både Putin og Obama, og nu vil han spille sine kort.’

Som det ses, bruges i dansk en sætning (relativ ledsætning) med den nye beskrivelse som subjektsprædikat. I italiensk er en tilsvarende prædikation udeladt, og den nye beskrivelse udtrykkes i den utro subjektsanafor.⁶

4.2. *Resumptive anaforer*

Blandt de resumptive anaforer er pro-formerne hyppigere i dansk og NP’erne hyppigere i italiensk. De følgende eksempler viser den typiske sprogbrug her:

- (5) I beretningen står der jo, *at man ikke vil samarbejde*, og jeg tror, **det** [RESUMPTIV PRO-FORM] *er et problem* vi bliver nødt til at tage op. <ep-97-11-18.txt:35>
- (6) Jeg havde personligt gerne set, at *vi havde fået alle flammehæmmere med fra starten*, men jeg synes, **det** [RESUMPTIV PRO-FORM] *er et spørgsmål*, vi må prøve at arbejde videre med. <ep-01-05-15.txt:40>

De danske strukturer består af et anaforisk pronomensubjekt med kategoriseringen af det resumerede tekstsegment (som hhv. *et problem* og *et spørgsmål*) liggende i prædikationen, som desuden indeholder en relativsætning til specificering af denne kategorisering. Hvis vi sammenligner med to tilsvarende kategoriseringer i italiensk, ser vi et andet mønster; her ses segmentkategoriseringen at finde sted i NP-anaforerne (hhv. subjekt og objekt):

- (7) *Nel cortile del nostro Parlamento è in corso – forse ancora – un’indegna manifestazione di un folto gruppo di giovani italiani, con le bandiere della Lega Nord, che sta lanciando slogan che dicono di bruciare la bandiera italiana, ...*
 Penso, quindi, che **questo problema** [RESUMPTIV NP] vada segnalato alla Presidente del Parlamento ... <ep-00-06-15.txt:100>
‘I Parlamentets gård finder der – måske stadigvæk – en forkastelig demonstration sted, hvor en stor gruppe unge italienere med Liga Nord’s flag råber sloganer om, at det italienske flag skal brændes, ...

6 Især i journalistisk stil, hvor begrænset plads kan tilskynde til nominalstil (se også afsnit 5), er utro anaforer dog langt fra ukendte i dansk, men de forekommer generelt sjældnere end i italiensk.

Jeg mener altså, at **dette problem** [RESUMPTIV NP] bør meddeles Parlamentets formand ...⁷

- (8) *I cittadini che risiedono o lavorano in Stati membri diversi dal loro paese d'origine non possono abbonarsi legalmente a piattaforme straniere, in quanto questi operatori non sono titolari dei diritti validi per gli Stati membri in questione. Il Parlamento europeo ha già sollevato **tale questione** [RESUMPTIV NP] in passato e la relatrice analizza il problema, ... <ep-00-06-13.txt:133>*

*‘Borgere, der bor eller arbejder i andre medlemslande end deres oprindelsesland, kan ikke lovligt abonnere på udenlandske platforme, da disse operatører ikke har de rettigheder der gælder for de pågældende medlemslande. Europa-Parlamentet har allerede rejst **dette spørgsmål** [RESUMPTIV NP] tidligere, og ordføreren analyserer problemet, ...*

4.3. Associative anaforer

De associative anaforer er betydeligt hyppigere i italiensk end i dansk; et typisk italiensk eksempel:

- (9) D'altra parte, un'azione restrittiva rispetto alle coltivazioni di *tabacco* in Europa non avrebbe alcuna rilevanza sul **consumo**, se non quella di dirottare **i consumatori europei** verso tabacchi extra-UE, di gran lunga meno pregiati e certamente più nocivi per la salute. <ep-02-03-14.txt:124>

*‘Restriktioner for tobaksdyrkingen i Europa ville i øvrigt ikke få nogen som helst indflydelse på **forbruget**, men ville blot få **de europæiske forbrugere** til i stedet at købe tobak, som er fremstillet uden for EU, og som ikke er af nær så god kvalitet og så afgjort er mere sundhedsskadelig.’*

I denne forbindelse er det interessant at sammenligne den officielle italienske og den officielle danske oversættelse af den samme engelske tale, ep-96-06-20.txt:163:

- (10) Signor Presidente, c'è un problema di traduzione per la versione inglese *dell'emendamento n° 4 dell'onorevole Pex*. Potrebbe chiedere all'**autore**

7 Med undtagelse af eks. (10) - (11) nedenfor er oversættelserne af Europarl-eksemplerne mine egne; de er holdt så tæt på Parlamentets officielle som muligt, men ligger i enkelte tilfælde lidt tættere på kildeteksten for at illustrere dennes anaforsyntaks bedst muligt.

[ASSOCIATIV ANAFOR] di chiarire cosa intende per »decisions taken on improper grounds«. <ep-96-06-20.txt:163>

‘Hr. Formand, der er et oversættelsesproblem i den engelske version af *ændringsforslag nr. 4 af hr. Pex*. Kunne De bede **forfatteren** [ASSOCIATIV ANAFOR] forklare hvad han mener med »decisions taken on improper grounds«.

- (11) Hr. formand, i *ændringsforslag nr. 4 af hr. Pex* er der et problem i den engelske oversættelse. Kan De spørge *forfatteren til ændringsforslaget* [TRO KOREFERENTIEL ANAFOR], om han kan forklare udtrykket »decisions taken on improper grounds«.

Hvor italiensk benytter en associativ anafor uden at specificere at det drejer sig om *forfatteren til ændringsforslaget*, dvs. uden at ekspliciterer en proposition af typen *ændringsforslaget har en forfatter*, ekspliciterer dansk dette semantiske indhold. Hvor tekstkohæsionen i (10) fin der sted i den associative anafor *l'autore 'forfatteren'*, ligger den i (11) i den tro koreferentielle anafor *ændringsforslaget*, og den bestemte artikels henvisning i *forfatteren (til ændringsforslaget)* er kataforisk til denne anafor.⁸

5. Tekstsemantisk “tomme” og “fulde” anaforer

En nærmere analyse af de citerede (og meget typiske) danske og italienske anaforer viser at de danske er leksikalsk-semantisk “tomme” i ko-teksten. Der er enten tale om pronominaler eller om tro koreferentielle anaforer, og ingen af disse bidrager med nyt leksikalsk indhold i teksten; det nye indhold ligger i en prædikation enten med *være* (svarende til de utro tilfælde, jf. (4), s amt ved de resumptive pronomenanaforer, jf. (5) – (6)) eller fx med *have* (svarende til de associative tilfælde, evt. – som i (11) – tekstualiseret som en præpositionsforbindelse).⁹ Derimod indeholder alle de tilsvarende italienske anaforer nyt semantisk indhold i sig selv. Dette kan beskrives som i tabel 6.

8 Den engelske kildetekst lyder således: “Mr President, in *Amendment.No 4 by Mr Pex* there is a problem in the English translation. Could you ask the *author of the amendment* if he could explain the phrase ‘decisions taken on improper grounds.’” Som den danske oversættelse har den engelske kildetekst en tro koreferentiel anafor og ekspliciteret relation mellem forfatter og tekst.

9 I Korzen (2014) har jeg mere detaljeret beskrevet de semantiske relationer der kan fortolkes mellem en antecedent og en associativ anafor.

Dansk	Tom nominal anafor	–	Prædikation (fx med <i>være</i> eller <i>have</i>)
Italiensk	Fuld nominal anafor	–	Udeladt prædikation

Tabel 6. Typisk fordeling af semantisk indhold i anaforkonstruktioner i dansk og italiensk

Hvis vi sammenholder tabel 6 med tabel 2 ovenfor, ser vi en meget smuk overensstemmelse:

Endocentriske sprog	Leksikalsk struktur	n	–	V	–	n
	Anaforstruktur	tom anafor	–	prædikation	–	tom anafor
Exocentriske sprog	Leksikalsk struktur	N	–	v	–	N
	Anaforstruktur	fuld anafor	–	udeladt prædikation	–	fuld anafor

Tabel 7. Typisk fordeling af “informationsvægt” i leksikon og anaforstruktur

De endocentriske sprog er leksikalsk kendetegnet ved vage og abstrakte nominer og betydningstunge verber og anaforstrukturelt ved tomme nominalanaforer og fuld prædikation. De exocentriske sprog er leksikalsk kendetegnet ved betydningstunge nominer og abstrakte og upræcise verber og anaforstrukturelt ved fulde nominalanaforer og udeladt prædikation.¹⁰ De leksikalske strukturer og forskelle genfindes altså i anaforstrukturerne og forstærkes således på det tekstuelle niveau. Den manglende prædikation i romansk kan ses som den ekstreme manifestation af den typiske romanske nominalstil; en slags “midtvejsmanifestation” ses i de propositioner der tekstualiseres som infinitte eller nominaliserede strukturer, jf. afsnit 6.2 og tabel 12.

6. Sproglig markering af nominale anaforer

Som det fremgik af tabel 4 og 5 og de citerede eksempler, kan nominale anaforer kodes, eller markeres, sprogligt på forskellig måde. Man kan tale om fire forskellige markeringsmuligheder, hvoraf flere kan forekomme i samme anafor, som vi skal se nedenfor:

¹⁰ Det skal dog understreges at de endo- og exocentriske sprog ikke er fuldstændig indbyrdes ens. Det er fx muligt at der bør tages nogle forbehold for tysk anafor- og tekststruktur, og det er også velkendt at engelsk har været under stærk påvirkning fra fransk og latin. Det beskrevne endocentriske mønster er dog typisk for de skandinaviske sprog.

<p>A. Morfofonologisk markering:</p> <ol style="list-style-type: none"> 1. Ved brug af frie pronominer i modsætning til klitiske pronominer og nulformer. 2. Ved brug af demonstrativ i modsætning til bestemt artikel i NP'er. <p>B. Leksikalsk markering:</p> <ol style="list-style-type: none"> 1. Ved brug af NP'er i modsætning til pronominer og nulformer. 2. Ved brug af (typisk attributivt) specificerede NP'er i modsætning til ikke-specificerede NP'er. <p>C. Positionel markering:</p> <ol style="list-style-type: none"> 1. Ved brug af højre- og venstredislokering eller sætningskløvning. <p>D. Prosodisk markering:</p> <ol style="list-style-type: none"> 1. I talesprog ved særlig intonation og/eller betoning.

Tabel 8. Sproglig markering af nominalanaforer

Graden af anaformarkering afhænger af antecedentens "prominens" eller "vægt" i ko-teksten på det tidspunkt hvor den anaforiseres, dvs. genoptages ved en anafor. Jo stærkere prominens, desto svagere markering, og jo svagere prominens, desto stærkere markering. Denne prominens afhænger igen af en række træk, som kan resumeres således:

- | |
|---|
| <ol style="list-style-type: none"> 1. Afstanden mellem antecedent og anafor: jo mindre afstand, desto større prominens. 2. Det sproglige materiale mellem antecedent og anafor, herunder mulig interferens med andre potentielle antecedenter. 3. Semantiske og pragmatiske egenskaber ved antecedenten, fx første- vs. højereordensentiter samt referentielle forhold (deixis, (non-)specifik, generisk reference).¹¹ 4. Eventuelle skift i informations- eller kognitive enheder, se afsnit 6.1. 5. Syntaktisk og tekstuel kompleksitet i antecedentens ko-tekst: lav syntaktisk placering medfører lav prominens, se afsnit 6.2. 6. Anafortype: koreferentiel (tro vs. utro) eller resumptiv. |
|---|

Tabel 9. Afgørende faktorer for antecedenters prominens ved anaforisering¹²

Da mit korpus består af nogenlunde lige lange paralleltekster,¹³ kan man antage at der ikke er de store interlingvistiske forskelle mht. de tre første punkter. Derimod viser punkt 4 og 5 sig at variere meget mellem de to sprog.

11 Disse fænomener er indgående behandlet i Korzen (2000: 210-218 og 486-512) og (2001).

12 Jf. bl.a. Givón (1976, 1983a/b), Tomlin (1985, 1987), Fox (1987a/b), Berretta (1990) og Korzen (2000, 2001, 2007).

13 Ved paralleltekster forstås jeg indholdsmæssigt ækvivalente tekster produceret på forskellige sprog uafhængigt af hinanden, men under parallelle betingelser og tilhørende samme teksttype; jf. også Skytte (1999: 12).

6.1. *Periodeskift*

Mht. leksikalsk markering var der en væsentlig forskel i de koreferentielle anaforer med mere end dobbelt så mange tro NP-anaforer i dansk som i italiensk. Dette sammenholdt med en antagelse om et højere antal propositioner i dansk end i italiensk, jf. afsnit 4.0 ovenfor, kunne tyde på en forskel for så vidt angår punkt 4 i tabel 9 med flere skift i informations- eller kognitive enheder i dansk. Sådanne skift tekstualiseres typisk som periode- eller afsnitsskift, og anaforiske relationer hen over sådanne skift kræver stærkere markering.¹⁴ Dette kunne føre til en hypotese om propositionernes tekstualisering i de to korpora: om en højere grad af tekstualisering som selvstændige perioder i dansk end i italiensk.

En sådan hypotese viser sig at holde stik ved en simpel optælling af perioder i de to korpora:

De to undersøgte korpora	Dansk	Italiensk
Antal perioder	678	440
Gennemsnitligt antal ord pr. periode	21,7	33,4

Tabel 10. Periodeantal og antal ord pr. periode i de to undersøgte korpora

Da ordantallet i de to undersøgte korpora stort set var identisk, jf. tabel 3, giver den store forskel i antallet af perioder også en markant forskel i periodelængde målt som antallet af ord pr. periode.¹⁵ Dette fører os automatisk til næste punkt, den syntaktiske kompleksitet, jf. punkt 5 i tabel 9.

14 I det følgende eksempel fra Mr. Bean-korpuset (Skytte et al. (red.) 1999: DSA5) ses hvordan der er brugt NP-anafor ved periodeskift, men pronomenanaforer periodeinternt: “*Mr. Bean* kommer ind på et bibliotek med gamle bøger, *han* får at vide af bibliotekaren, en ældre nobel herre, at *han* ikke må larme. *Mr. Bean* lister hen over gulvet på en mærkelig måde for at undgå at gulvet knirker.” Fænomenet er i øvrigt behandlet og illustreret ved en række andre eksempler i Korzen (2000: 478-486).

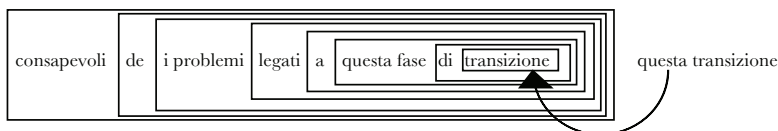
15 Det er klart at man skal være varsom med kvantitative sammenligninger baseret på en så usikker sproglig enhed som “ordet”. Visse typologiske forskelle fører til et lavere ordantal i dansk (fx den ofte enklitiske bestemte artikel og de mange sammensatte substantiver, der svarer til N + præposition + N i italiensk), mens andre medfører et lavere ordantal i italiensk (fx verbum + enklitisk pronomen samt pro-drop-fænomenet: det ikke-eksisterende subjektspronomen i italiensk). Dog er forskellen i de to korpora så markant, at jeg mener at tallene godt kan bruges som udtryk for væsentlige forskelle i periodelængde mellem de to sprog.

6.2. Syntaktisk kompleksitet og anafortype

Som sagt kan flere sproglige markeringer optræde i samme anafor. Et godt eksempel på dette er:

- (12) Siamo consapevoli dei problemi legati a questa fase di *transizione* ed è **questa transizione** che ha creato più preoccupazione nel nostro personale [...] <ep-00-03-01.txt:16>
 ‘Vi er klar over de problemer, som er forbundet med denne *overgangs-*fase, og det er **denne overgang**, der har vakt størst bekymring blandt vores personale.’

I *questa transizione* optræder tre forskellige markeringer: den morfofonologiske (tabel 8, A2) i den demonstrative artikel, den leksikalske (tabel 8, B1) i NP-formen og den positionelle (tabel 8, C1) i sætningskløvningen.¹⁶ Samtidig er der meget kort afstand mellem antecedent og anafor (jf. tabel 9, punkt 1), så forklaringen skal findes i den syntaktisk komplekse ko-tekst, antecedenten befinder sig i, en kompleksitet som kan beskrives som i Figur 1.



Figur 1. Den syntaktiske struktur i antecedentens ko-tekst i eks. (12)

Antecedenten *tra nsizione* er styrelse for et præpositionssyndagma, som står attributivt til (*questa*) fase, som er styrelse i et præpositionssyndagma med *a*, som er sekundært valensled til *legati*, som er attributivled til (*i*) *problemi*, som er styrelse for et *di*-præpositionssyndagma, som er sekundært valensled til *consapevoli*; altså en meget lav (underordnet) syntaktisk position, hvilket medfører en meget lav prominens.¹⁷

16 Det skal tilføjes at sætningskløvningen kan optræde af andre årsager end (alene) antecedentens komplekse ko-tekst, fx ønsket om en særlig fokusering af *transizione*.

17 Det er muligvis ikke udelukket at se *questa fase di transizione* som antecedent, hvorved der vil være tale om en utro anafor og stadig en lav syntaktisk position, om end ikke helt så lav som *transizione* alene.

En samlet optælling af både morfofonologisk og leksikalsk markerede anaforer i de to korpora giver resultatet i tabel 11.

Anafortype	Markering	Danske	Italienske
Koreferentielle tro anaforer	Definit NP	216	88
	Demonstrativ NP	16 = 6,9 %	22 = 20 %
Koreferentielle utro anaforer	Definit NP	27	34
	Demonstrativ NP	17 = 38,6 %	34 = 50 %
Resumptive anaforer, NP'er	Definit NP	6	6
	Demonstrativ NP	24 = 80 %	33 = 84,6 %
Resumptive anaforer, pro-former	Personligt/klitisk pron.	26	7
	Nulsubjekt (pro-drop)	–	3
	Demonstrativt pron.	77 = 74,8 %	48 = 82,8 %

Tabel 11. Morfofonologisk og leksikalsk markering af koreferentielle og resumptive anaforer

Tabellen viser hvordan visse anaforiske relationer kan kaldes “ukomplicerede” og andre “komplicerede”. Markering med demonstrativ, i funktion som pronomen eller artikel, ekspliciterer at den angivne størrelse (antecedenten) skal findes inden for samme tekst, og behovet for denne markering er generelt, dvs. sproguafhængigt, større ved “komplicerede” anafortyper, dvs. især resumptive og, i lidt mindre grad, utro koreferentielle anaforer.¹⁸

Men herudover ses en generel forskel de to sprog imellem: Den morfofonologiske markering med demonstrativ er i samtlige tilfælde forholdsmæssigt hyppigere i italiensk end i dansk. Dette skyldes især forskelle i den syntaktiske kompleksitet. Syntaktisk kompleksitet kan måles kvantitativt på forskellige måder, hvoraf én er periodelængde, som viste betydeligt længere perioder i italiensk end i dansk, jf. tabel 10. En anden, som kan ses i lyset af den netop nævnte, er antallet af propositioner pr. periode, jf. også Korzen & Gylling (2012) og Korzen (2015).

18 Også associative anaforrelationer er relativt komplicerede og kræver normalt leksikalsk markering. Om undtagelser fra denne regel, se Korzen (2009).

De to undersøgte korpora	Dansk	Italiensk
Antal perioder	678	440
Antal propositioner	1.455	1.473
Gennemsnitligt antal propositioner pr. periode	2,15	3,35
Perioder med fem eller flere propositioner	5,5 %	21,1 %
Underordnede propositioner med infinit el. nominaliseret verbal	13,4 %	25,5 %

Tabel 12. Antal propositioner pr. periode og syntaktisk kompleksitet

Tallene i tabel 12 giver et klart billede af større syntaktisk kompleksitet i italiensk end i dansk. Forskellen i det gennemsnitlige antal propositioner pr. periode (2,15 vs. 3,35) udgør hele 55,8 %, og heraf er en langt højere procentdel tekstualiseret med infinit eller nominaliseret verbal, altså med en højere grad af deverbalerisering, eller desententialisering, i italiensk end i dansk, med deraf følgende lavere syntaktisk placering end ved finite verber (de infinitte eller nominaliserede verber er i sig selv underordnet en anden struktur), et karakteristisk træk i den meget udbredte romanske nominalstil. Især af de citerede eksempler på resumptive anaforer, eksemplerne (5) – (8), fremgår det også klart at de italienske antecedenter er markant længere og mere komplekse end de danske.

7. Konklusioner

Vi har set at der i de typiske anaforforekomster i det undersøgte korpus for det første er en interessant parallelitet med de typiske leksikalske strukturer i endo- og exocentriske sprog, hvorved de leksikalske forskelle forstærkes i tekststrukturen: De hyppigste danske nominalanaforer er tekstsemantisk “tomme”: De tilføjer ikke yderligere leksikalsk-semantisk indhold til teksten dér hvor de forekommer; en ny beskrivelse ligger i den til anaforen knyttede prædikation. Parallelt hermed er substantiverne leksikalsk generiske og abstrakte i de endocentriske sprog, hvorimod verbet har den største leksikalske informationsvægt. Det omvendte ses i de exocentriske sprog, jf. Tabel 7.

For det andet var der klare relationer mellem de hyppigste anaforforekomster og den typiske tekststruktur. Det højere antal tro koreferentielle NP-anaforer i dansk kan – sammen med det højere samlede antal anaforer – relateres til det højere antal perioder i det danske korpus, og det generelt højere antal NP-anaforer med demonstrativ artikel i italiensk kan relateres til en generelt større syntaktisk og tekstuel kompleksitet end i dansk.

Fælles for de to sprog er dog at visse anafortyper er mere “komplicerede” end andre og kræver større sproglig markering. De mest “ukomplicerede” anaforer er de tro koreferentielle, lidt mere “komplicerede” er de utro koreferentielle, og de mest “komplicerede” er de resumptive anaforer.

Henvisninger

- Baron, I. & M. Herslund (2005). Langues endocentriques et langues exocentriques, i M. Herslund & I. Baron (red.) *Le génie de la langue française, Langue française* 145, 35-53.
- Berretta, M. (1990). Catene anaforiche in prospettiva funzionale: antecedenti difficili, *Rivista di Linguistica* 2/1, 91-120.
- Børsen (2013). Verdensledere samlet bag irsk jernring, *Børsen*, fundet 6/10-2014 (http://borsen.dk/nyheder/politik/artikel/1/260052/verdensledere_samlet_bag_irsk_jernring.html)
- Cresti, E. & I. Korzen (red.). (2010). *Language, Cognition and Identity. Extensions of the endocentric/exocentric language typology*. Firenze: Firenze University Press.
- Diritto di critica (2013). Siria, Cameron il mediatore prepara la guerra, *Diritto di critica*, fundet 6/10-2014 (<http://www.dirittodicritica.com/2013/05/14/siria-cameron-obama/>)
- Fox, B.A. (1987a). Morpho-syntactic markedness and discourse structure, *Journal of Pragmatics* 11, 359-375.
- Fox, B.A. (1987b). *Discourse Structure and Anaphora. Written and conversational English*. Cambridge University Press.
- Givón, T. (1976). Topic, Pronoun and Grammatical Agreement, i C.N. Li (red.) *Subject and Topic*, New York/San Francisco/London: Academic Press, 149-188.
- Givón, T. (1983a). Topic continuity in discourse: an introduction, i Givón (red.) (1983), 1-41.
- Givón, T. (1983b). Topic continuity in spoken English, i Givón (red.) (1983), 343-363.
- Givón, T. (red.). (1983). *Topic Continuity in Discourse: A Quantitative Cross-language Study*. Amsterdam/Philadelphia, John Benjamins.
- Herslund, M. (red.). (1997). *Det franske sprog. Kapitel I. Grundlag*. København: Handelshøjskolen i København.
- Herslund, M. (2009). Dal linguaggio alla cultura. Presentazione del progetto danese, i I. Korzen & C. Lavinio (red.) *Lingue, culture e testi istituzionali*, Firenze: Franco Cesati, 17-29.

- Herslund, M. (2010). Predicati e sostantivi complessi: complementarità e isomorfia, i E. Cresti & I. Korzen (red.) *Language, Cognition and Identity. Extensions of the endocentric/exocentric language typology*, Firenze: Firenze University Press, 1-8.
- Herslund, M. & I. Baron. (2003). Language as World View. Endocentric and exocentric representations of reality, i I. Baron (red.) *Language and Culture, Copenhagen Studies in Language* 29, 29-42.
- Koehn, P. (2005). Europarl: A parallel corpus for statistical machine translation. *Conference Proceedings: the tenth Machine Translation Summit*, Phuket, 79-86.
- Korzen, I. (1996). *L'articolo italiano fra concetto ed entità*, *Etudes Romanes* 36. København: Museum Tusulanum Press.
- Korzen, I. (2000). Reference og andre sproglige relationer, i G. Skytte & I. Korzen. *Italiensk-dansk sprogbrug i komparativt perspektiv. Reference, konnexion og diskursmarkering*, København: Samfundslitteratur, 161-619.
- Korzen, I. (2001). Anafore e relazioni anaforiche. Un approccio pragmatico-cognitivo, *Lingua nostra* LXII (3-4), 107-126.
- Korzen, I. (2005a). Lingue endocentriche e lingue esocentriche: lessico, testo e pensiero, i I. Korzen & P. D'Achille (red.), *Tipologia linguistica e società*, Firenze: Franco Cesati, 31-54.
- Korzen, I. (2005b). Struttura linguistica e schema cognitivo: tipologie a confronto, i I. Korzen (red.) *Lingua, cultura e intercultura: l'italiano e le altre lingue*, *Copenhagen Studies in Language* 31, 123-134.
- Korzen, I. (2006). Tipologia anaforica: il caso della cosiddetta "anafora evolutiva", *Studi di grammatica italiana* XXV, 323-357.
- Korzen, I. (2007). Linguistic typology, text structure and anaphors, i I. Korzen & L. Lundquist (red.) *Comparing Anaphors. Between Sentences, Texts and Languages*, *Copenhagen Studies in Language* 34, 93-109.
- Korzen, I. (2009). Anafora associativa: ulteriori associazioni, i F. Venier (red.) *Tra pragmatica e linguistica testuale. Ricordando Maria-Elisabeth Conte*, Alessandria: Edizioni dell'Orso, 307-326.
- Korzen, I. (2014). Implicit association in political discourse. On associative anaphors in Italian and Danish EU proceedings, i I. Korzen, A. Ferrari & A.-M. De Cesare (red.) *Tra romanistica e germanistica: lingua, testo, cognizione e cultura / Between Romance and Germanic: Language, text, cognition and culture*. Bern et al.: Peter Lang, 217-236.
- Korzen, I. (2015). Frasi complesse e complessità frasale: il discorso

- politico in un'ottica tipologico-comparativa, i C. Bruno et al. (red.) *Plurilinguismo/Sintassi. Atti del XLVI Congresso Internazionale SLI*, Roma: Bulzoni, 631-648.
- Korzen, I., A. Ferrari & A.-M. De Cesare (red.). (2014). *Tra romanistica e germanistica: lingua, testo, cognizione e cultura / Between Romance and Germanic: Language, text, cognition and culture*. Bern et al.: Peter Lang.
- Korzen, I. & M. Gylling (2012). Text Structure in a Contrastive and Translational Perspective. On Information Density and Clause Linkage in Italian and Danish, *Translation: Computation, Corpora, Cognition* 2/1, 23-46.
- Lundquist, L. (2005). Noms, verbes et anaphores (in)fidèles. Pourquoi les Danois sont plus fidèles que les Français, i M. Herslund & I. Baron (red.) *Le génie de la langue française, Langue française* 145, 73-91.
- Skytte, G. (1999). Projektet "Mr. Bean på dansk og italiensk", i G. Skytte et al. (red.) (1999: 10-33).
- Skytte, G. et al. (red.). (1999). *Tekststrukturering på italiensk og dansk. Resultater af en komparativ undersøgelse*, København: Museum Tusulanums Forlag.
- Tomlin, R.S. (1985). Foreground-background information and the syntax of subordination, *Text* 5/1-2, 85-122.
- Tomlin, R.S. (1987). Linguistic reflections of cognitive events, i R.S. Tomlin (red.), *Coherence and grounding in discourse*, Amsterdam/Philadelphia: John Benjamins, 455-479.

