

Demonstrative pronouns in Old English

A textual-structural analysis

Júda Ronén <<https://ac.digitalwords.net/>>

Departmental seminar in Linguistics

2018/5/8

The Hebrew University of Jerusalem

Department of Linguistics

Who am I?

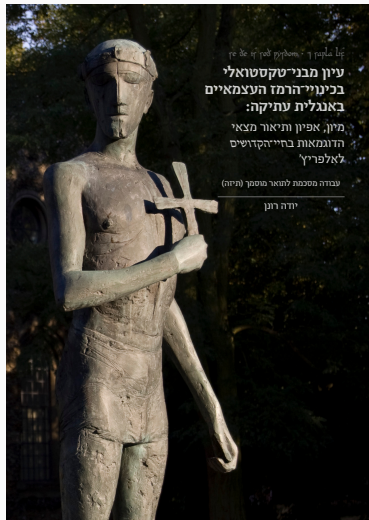
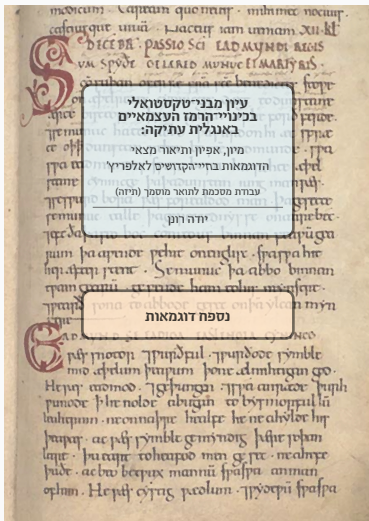
Present (PhD):

Grammar of Textuality: Structural Linguistic Analysis of Text Typology in Modern Welsh

Past (MA):

*se ðe ís soð wysdom · 7 sawla líf — A Textual-Structural Study of the Demonstrative Pronouns in Old English: Classification, characterization and description of the examples in Ælfric's *Lives of Saints**





What is this talk about?

- Two demonstrative:
 - $\overset{M}{se}, \overset{F}{seo}, \overset{N}{þæt}$ ('SE')
 - $\overset{M}{þes}, \overset{F}{þeos}, \overset{N}{þis}$ ('DIS')
- We will survey their functional distribution and structural features.

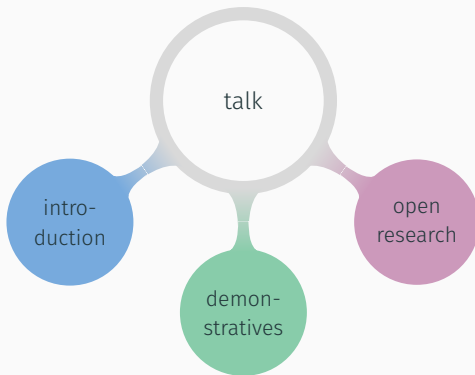
What do I want you to take from this talk?

1. Usefulness for your research
2. Data can be surprising
3. Homonyms should be distinguished
4. Open research

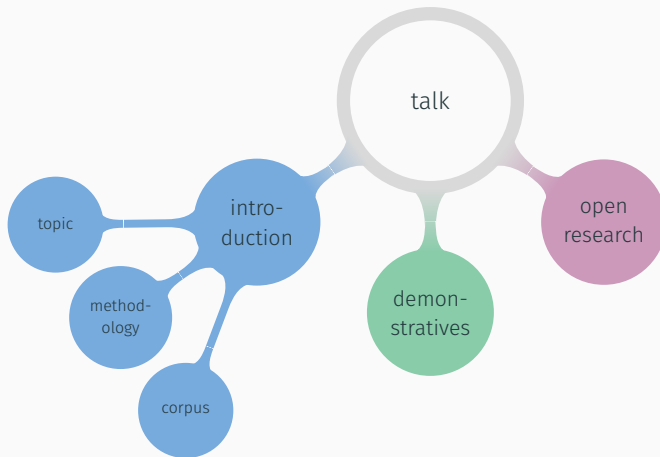
What do I want you to take from this talk?

1. Usefulness for your research
2. Data can be surprising
3. Homonyms should be distinguished
4. Open research
5. Study Old English!

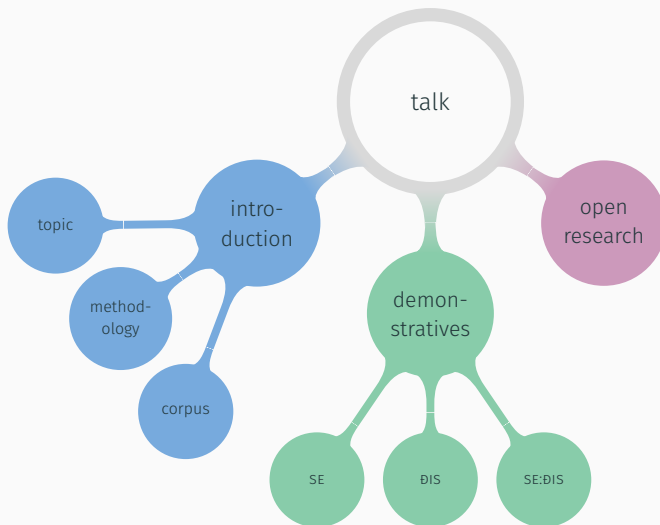
Talk structure



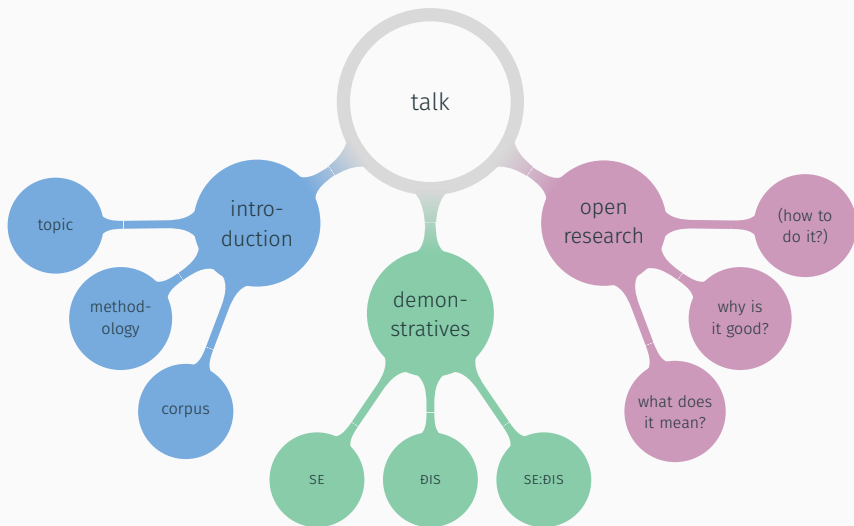
Talk structure



Talk structure



Talk structure



Part I

Introduction

- ⟨þ⟩/⟨p⟩ and ⟨Ð⟩/⟨ð⟩ are in free variation;
phonetic value: *[θ/ð]
- Macron signifies a following nasal (usually -m)
Hence, þam = ðam = þā = ðā
- Scribal abbreviations:
 - þ̅ = þæt 'that'
 - ʒ = and 'and' (Tironian *et*)

Topic

SE and ÐIS: two demonstrative pronouns

	SE	ÐIS
etymology	PIE *so, *seh ₂ , *tod	+NWGmc *-s(i)-
cognates	Skt सः, सा, तत् (sáh, sā́, tát) Grk ὅ, ἡ, τό (ho, hē, tó)	RN sási, súsi, þatsi OFrs <i>this</i> , <i>thius</i> , <i>thit</i>
OldEng	se, seo, þæt	þes, þeos, þis
ModEng	<i>the</i> (< ME <i>þe</i>) <i>that</i> (< ME <i>þat</i>)	<i>this</i> (< ME <i>þis</i>) <i>these</i> (< ME <i>þæs</i>) <i>those</i> (< ME <i>þas</i>)

Two distinct syntactic patterns

- Dependent (determiner): [SE/ÐIS + NOUN]
 - agreement in gender, number and case
 - weak adjectives
 - nucleus and satellite
 - high juncture
- Independent: [SE/ÐIS]

Two distinct syntactic patterns

- Dependent (determiner): [SE/ÐIS + NOUN]
 - agreement in gender, number and case
 - weak adjectives
 - nucleus and satellite
 - high juncture
- Independent: [SE/ÐIS]

Differences between the two

- Internal:
 - different micro- and macro-syntax
 - different functions
 - quantitative difference

⇒ should be considered homonyms (SE_1 , SE_2 , DIS_1 , DIS_2)
- External:
 - scholarly attention

Ex. 1: dependent se (anaphora)

Hwæt þa eadmund clypode ænne bisceop ·
þe him þa gehendost wæs 7 wið hine smeade
hu he þā reþan hinguare and-wyrdan sceolde ·
Ða forhtode ^{NOM.M.SG NOM.M.SG} se bisceop for þā færlican gelimpe ·
7 for þæs cynincges life · 7 cw̃ þ him rād þuhte
þ he to þam gebuge þe hī beað hinguar ·

So then king Edmund called a bishop
who was handiest to him, and consulted with him
how he should answer the savage Hingwar.
Then the bishop feared for this terrible misfortune,
and for the king's life, and said that it seemed best to him
that he should submit to that which Hingwar bade him.

Edmund₃₂, l. 56

Ex. 2: dependent *DIS* (*nynégocentrique* deixis)

MEN ÐA LEOFESTAN HWILON ÁR WE SÆDON EOW HÚ URE hælend
crist on ^{DAT.M.SG}þisum ^{DAT.M.SG}dæge on soðre menniscnysse acenned wæs of þæm
halgan mædene marian.

Men beloved, we told you erewhile how our Saviour Christ was
on **this** **day** born in true human nature of the holy Virgin Mary.

Nativity₁, l. 1

Ex. 3: independent SE (introduction of new characters)

Þa wæs on rome byrig **sum cyne-boren mæden** ·

basilla **gehaten** · on hæðen-scype wunigende ·

NOM.F.SG

seo wolde gehyran þá halgan lare ·

of eugenian muþe · ác heo ne mihte hire genealecean ·

forðan þe cristen-dóm wæs · þær ónscunigend-lic ·

Then was there in Rome **a maiden of royal birth,**

named Basilla, living in heathendom;

she (this one) desired to hear the holy lore

at the mouth of Eugenia, but she could not approach her,

because Christianity was there held in abhorrence.

Eugenia₂, l. 326

Ex. 4: independent *ðis* (refers to the things said)

Þa astrehte se halga his handa 7 cw̃ ·

[...]

Mid þam ðe he ^{ACC.N.SG} þis cwæð · þa com crist sylf him to ·
[...]

Then the Saint stretched out his hands and said,

‘[...]

While he was saying **this**, there came Christ Himself to him,
[...]

Mark₁₅, l. 68

Inflection

SE				
case	singular			plural
	M	F	N	
NOM	se	seo	þæt	þa
ACC	þone	þa	þæt	þa
GEN	þæs	þære	þæs	þæra
DAT	þam	þære	þam	þam

ÐIS				
case	singular			plural
	M	F	N	
NOM	þes	þeos	þis	þas
ACC	þisne	þas	þis	þa
GEN	þises	þyssere	þises	þyssa
DAT	þysum	þyssere	þysum	þysum

INST forms (synchronically unrelated to SE): þe, þy, þan

Methodology

[...] So we isolated the cases of neutralization and archi-elements first, in order to be able to effectively examine the functions of the forms under scrutiny in conditions of genuine opposition and selectability. [...]

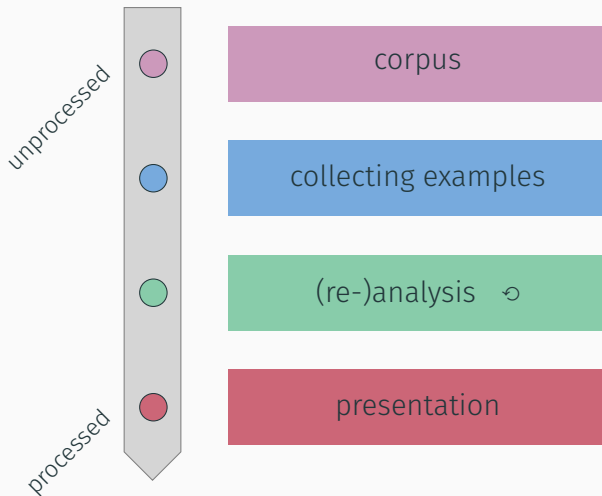
ROSÉN (2005)

Textematic differences

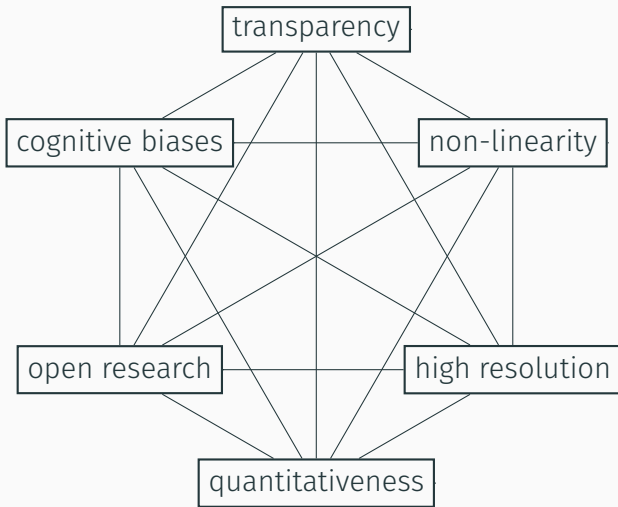
Texteme: a subtextual, componental, signalled and bounded unit manifesting a distinctive idiosyncratic and complete system of grammar; or an emic meta-textual type manifesting a distinctive idiosyncratic and complete system of grammar. The *dialogue* and the *narrative* are among the basic overhead textemes, further refinable and combinable into numerous others; *exposition* is another (in some respects manifesting affinities with the dialogue).

SHISHA-HALEVY (1998)

Process

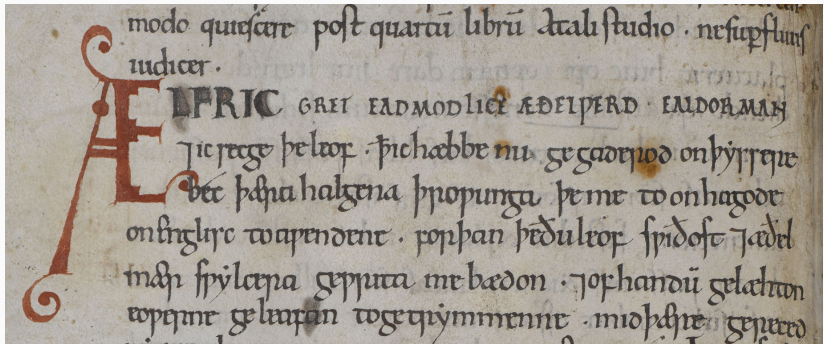


Providing a full, annotated and hyper-linked database



Corpus and author

Ælfric of Eynsham



Cotton MS Julius E VII, f. 3v

- c. 955 – c. 1010
- Benedictine monk, abbot, scholar, writer, translator
- The most prolific writer in Old English
- West-Saxon dialect

Ælfric's Grammar and Glossary

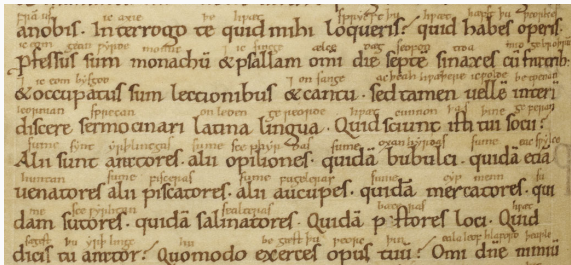
þara naman ꝥpelendora . marculinum .
hiedes . ꝥaminū hēdeor . neutrum . hoc dīr .
cōmune . gemethe . nīar . upst landst man .

odde on lande . Concha . ꝥcýll . nomina sī
Ferpa . pildeor . lupus . pulc . Leo . RARŪ .
Leo . Linx . gemengeo hund . ꝥpuls . om con
nīr . anhyrne deor . þat deor hæfð ane horn .
būpan ðam tꝥam ēgum . sꝥa sꝥrangne . ꝥsꝥa

Oxford, St. John Collage, MS. 154, f. 47v and f. 153r

Main sources: Priscianus (6th century), Aelius Donatus (4th century) and Isidore of Seville (c. 570 — 636).

Ælfric's Colloquy



Teacher: What work do you do?

Student: I am a monk by profession, and I sing every day seven times with the brothers, but nevertheless I wish, in between, to learn to speak Latin.

Teacher: What do you companions know (to do)?

Student: Some are ploughmen, some shepherds, some oxherds, some also hunters, fishers, fowlers, merchants, shoemakers, salters and bakers.

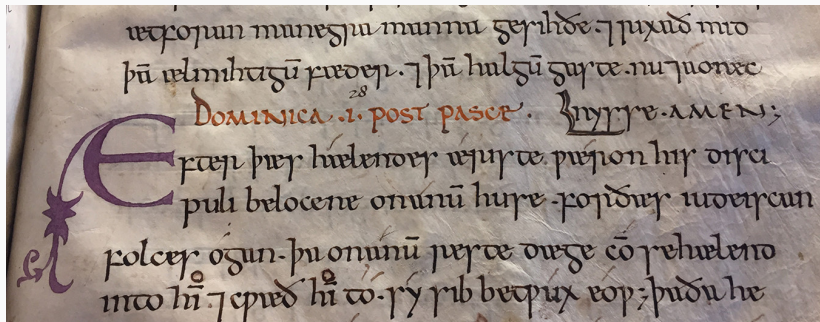
Teacher: What do you say, ploughman? How do you do your work?

Old English Hexateuch



Sume men cpeþað *DE DIVERSIS SYLLIS.*
þæt fceorþian feallað of heofonum ac hit ne sýnt
na fceorþian þ þær feallað ac is fýr of þam fiodene
þe sprungð of þætunglum. swa swa spearcan dōð of fýre
ƿitodlice swa feala fceorþian sýnd gyt on heofonum
swa swa on fýrde ƿær on. þa þa hi god gefceop ƿalle

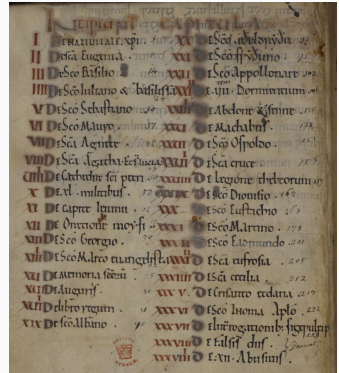
Cotton MS Caligula A XV, f. 151r



Oxford, Bodleian Library, MS Bodley 340

Lives of Saints

- ±40 texts, of two main genres:
 - Saints' Lives and Passions
 - Sermons
- Cotton MS Julius E VII
- ±400 pp. in print;
241 fol. in MS
- Rich and idiomatic Old English
adaption of Latin *vitae
sanctorum*
- Very loosely alliterative
rhythmical prose
- Skeat's edition (1881 and 1900);
trans. by Gunning and Wilkinson



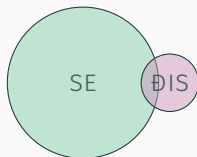
Cotton MS Julius E VII, f. 3v

Part II

Independent demonstrative pronouns

Main differences between SE and DIS in the corpus

	SE	DIS
quantitative	common (821)	less common (119)
syntagmatic	SE + þe	no DIS + þe
referential	nominal and grammatical	mostly grammatical
functional	diverse	quite limited
morphological	almost a full paradigm	partial paradigm



(the illustration isn't precisely proportional)

SE

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+

internal × nominal

- pronominal antecedent

internal × grammatical

- marking time

external × nominal

- introduction of new characters
- focality

external × grammatical

- meta-linguistic reference
- translation and interpretation
- marking time
- grammatical anchor in apposition

SE: types of reference (examples)

nominal grammatical

internal

+

+

external

+

+

internal × nominal

NOM.M.SG

se **ðe** wile drincan

he **who** wish to drink

external × nominal

[...] · ȝ þá wæron þa synna

calle adilegode butan **anre synna** ·

NOM.F.SG

Seo wæs seo mæste · [...]

[...] and then were the sins

all blotted out, save **one of the sins**.

This one was the greatest, [...]

internal × grammatical

DAT.N.SG

[...] **mid** **ðam** **þe** basilus · to-bræc þ
husel · [...]

[...] while (lit. **with** **that** **that**) Basil
divided the housel [...]

external × grammatical

NOM.N.SG

Hluttur offrung **þæt** is · ȝ licwurðe gode ·

þæt mann wydewan geneosige · [...]

A pure offering is **this**, and
acceptable to God,

that one should visit widows, [...]

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+
internal × nominal		internal × grammatical
<ul style="list-style-type: none">• pronominal antecedent		<ul style="list-style-type: none">• marking time
external × nominal		external × grammatical
<ul style="list-style-type: none">• introduction of new characters• focality		<ul style="list-style-type: none">• meta-linguistic reference• translation and interpretation• marking time• grammatical anchor in apposition

SE

Pronominal antecedent

261 exx.

Pronominal antecedent

- nominal relative clause, fused relative construction, independent relative clause, free relative clause; CHALKER and WEINER (1998)
- **pronominal antecedent**; CAMPBELL (1913) and POLOTSKY (1991)
- translating *independent relative clause*: מְשֻׁפֵּט זֶקֶה מְעַצֵּם; ACADEMY OF THE HEBREW LANGUAGE (1984)
- antecedent-less relative / substantival relative; SHISHA-HALEVY (1976)
- pseudo/unecht/uneigentlich relative clause; TILL (1955, II. §475)

Pronominal antecedent types

Type	exx.
Free	154
Topic	27
Apposition	80
	261

Pronominal antecedent types

Type	exx.
Free	154
Topic	27
Apposition	80
	261

When are free pronominal antecedents used?

Two primary motivations:

- lack of a simple expression
 - text-specific entities
 - lexicon
- focussing on a specific aspect

Ex. 5: generic reference

Hi eodon þa secende · ȝ symle clypigende ·
swa swa hit gewunelic is ^{DAT.PL.} **þā** **ðe** on wuda gað oft ·
hwær eart þu nu gefera · ȝ him ȝwyrde þ heafod ·
Hér · hér · hér · ȝ swa gelome clypode [...]

They went on seeking and always crying out,
as is often the wont of **those** **who** go through woods;
'Where art thou now, comrade?' And the head answered them,
'Here, here, here.' [...]

Edmund₃₂, l. 148

Ex. 6: text-specific reference

Ic wille wyrcean me naman ⁊ ofer-winnan iudan ·

⁊ ^{ACC.PL} þa ðe him mid synd þe forsawon ðone cyning ·

I will get me a name and overcome Judas,
and **them** **that** are with him, who despised the king.

Machabees₂₅, l. 300

Pronominal antecedent types

Type	exx.
Free	154
Topic	27
Apposition	80
	261

Ex. 7: bipartite construction

[...]

NOM.M.SG

7 **se** **ðe** on geleafan þurhwunað · **se** leofað á on ecnysse ·

[...]

and **he (this one)** **who** abideth in the faith, **he (this one)** shall live for ever.

Apollinaris₂₂, l. 244

Ex. 8: bipartite construction

Eall swa gelice ^{NOM.M.SG} **se** **ðe** gelyfð wiglungum
oððe be fugelum · oððe be fnorum ·
oððe be horsum · oððe be hundum ·
ne bið **hé** ná cristen · ac bið for-cuð wiðer-saca ·

So likewise **he** **who** trusteth in auguries,
either from birds, or from sneezings,
either from horses or from dogs,
he is no Christian, but is an infamous apostate.

Auguries₁₇, l. 88

Features of the bipartite construction

- | | | | | | |
|-----|------|-----|--|-----------|-----------------|
| se/ | þa | þe | | se/þa / | he/hi |
| he/ | they | who | | this one/ | these / he/they |
- Only NOM: M.SG and PL
- Only non-past; no past
- No *ge-* prefix in verbs
- Collocation with lexical items of ‘eternality’: *á, on ecnyse, æfre, ...*
- *bið* is chosen over *is*
- Semantics

Pronominal antecedent types

Type	exx.
Free	154
Topic	27
Apposition	80
	261

Agreement

how case is chosen

not in apposition externally

in apposition internally

how case is chosen

not in apposition externally

Se apostol **behét** ^{DAT.PL} **þam** **ðe** healdað clænnysse ·
þ hi synd godes templ · 7 þæs halgan gastes wunung ·
The apostle **promised** **those** **who** preserve chastity,
that they are God's temple, and the Holy Ghost's habitation.

in apposition internally

He wearð þa beheafðod for ^{GEN.M.SG} **ðæs hælendes** naman ·
^{ACC.M.SG} **þone** **þe** he ær **gewundode** wælhreowlice on rode ·
Then was he beheaded for the name **of the Saviour**,
him **who** he had before cruelly **wounded** on the cross,

Ex. 9: pronominal antecedent in apposition

Þa beseah dioclitian² se deofollica cwellere
to ðam halgan were · þe ðær swa heage stód ·
ꝛ cwæð orgællice · ne eart þu lá **sebastianus** ·
ACC.M.SG
þone **ðe** ic gefyrn hét mid flanū acwellan ·

Then looked Diocletian, the fiendish murderer,
towards the holy man, who stood there so loftily,
and said haughtily, 'Art not thou **Sebastian**,
he **whom** I before commanded to be slain with arrows?'

Sebastian₅, l. 447

Ex. 10: pronominal antecedent in apposition

Soðlice ic ge-seah þone sylfan apostol

NOM.M.SG

se

þe

halwende rād æghwylcū men forgið ·

Verily I saw the apostle himself,

he

who

giveth saving counsel unto every man.

Thomas₃₆, l. 355

Hwæt is gód butan **gode** anum ^{NOM.M.SG} **se** **þe** is healic godnisse · [...]

What is good but **God** only, **he** **who** is supreme goodness, [...]

Nativity₁, l. 91

Ex. 12: referring to God

Se halga sang geswutelað þa halgan þrynnysse ·
on anre godcundnysse æfre wunigende ·

NOM.F.SG

seo ðe æfre wæs · ⁊ eac nu wunað ·
⁊ æfre is towerd butan ateorunge ·

This holy song signifieth the Holy Trinity
in One Godhead, ever abiding,
(‘she’) who ever was, and also now continueth,
and ever is to come, without ceasing.

Mark₁₅, l. 215

He wearð þa gehæled · 7 hæfde his gesihðe ·

7 ham eode blyðe butan latteowe ana

NOM.M.SG

Se ðe lytle ær þanon wæs gelæd þurh ðone oþerne ·

7 his magas ðancodon mycclū ðæs gode ·

Whereupon **he** was made whole, and had his sight,

and without a guide he blithely went home alone,

he who a little time before was led by another;

and his kinsmen greatly thanked God for this.

Swythun₂₁, l. 219

Ex. 14: reversal of a situation

[...]

7 þone bisceop cyste · ablicged þurh god ·

ACC.M.SG

þone þe he ær geteohhode mid teonan to for-seonne ·

[...]

and, being divinely terrified, kissed the bishop,
him who he had before determined to scorn with insult.

Martin₃₁, l. 671

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+
internal × nominal		internal × grammatical
<ul style="list-style-type: none">• pronominal antecedent		<ul style="list-style-type: none">• marking time
external × nominal		external × grammatical
<ul style="list-style-type: none">• introduction of new characters• focality		<ul style="list-style-type: none">• meta-linguistic reference• translation and interpretation• marking time• grammatical anchor in apposition

SE

Introduction of new characters

102 exx.

BASILIVS **WES GE-HATEN** **SVM HALIG BISCOP** ·

NOM.M.SG

se wæs fram cyld-hade · swiðe ge-healdsum ·
 þeah þe he to langū fyrste · ún-ge-fullod wære ·

A certain holy bishop was named Basil;
 he (this one) from childhood was very continent,
 although for a long period he was unbaptized.

Basil₃, l. 1

Who is introduced like this?

Basilus, Iosep, **Iulianus**, Celsus, **Sebastianus**, Fabianus, **Maurus**, Florus, Galicanus, Terrentianus, Marcellus, Saul, Nabođ, Mennases, Dioclitianus, Bonifacius, Abgarus, Antiochus, Mathathias, Seron, Iohannes, **Oswold**, Aidan, Cynegyls, Maximianus, Anatolius, Auitianus, Leofstan, Ualerianus, Almachius, Polemius, Basilla, (Antiochia), **Agnes**, Constantia, Afrodosia, Thabitas, Gezabel, Oswyn, Sintice, Philippus, Gaius, Matheus, **Basilissa**, **Georius**, **Apollonaris**, Claudia, (Ioppe), (Bosór).

þa be-tæhte chromatius · se heah-gerefa þa cnihtas ·
 sumū þegne to ge-healdenne · ^{NOM.M.SG}se hatte nicostratus ·

Then the prefect Chromatius delivered the youths
 to a certain officer to keep; he (this one) was hight
 Nicostratus.

Sebastian₅, l. 32

Who is introduced like this?

Effrem, Nicostratus, Hilarianus, Pascaius, Flaccus,
Náámán, Taurus, Tatheus, Iudas, Eleazarus, Uictor,
Gad, Euthicia, (Ethna), Protus and Iacintus.

Ex. 17: after a new character is introduced by name (15 exx.)

On þære ylcan scole wæs **seo wælreowa iulianus** ·
cristen frā cyld-hade · ^{NOM.M.SG}**se** wearð casere siððan ·
ȝ awarep his ge-leafan · ȝ ge-wende to deofle ·

In the same school was **the cruel Julian**,
a Christian from childhood; **he (this one)** afterwards became
Emperor,
and renounced his faith, and turned to the devil.

Basil₃, l. 16

Who is introduced like this?

Iulianus, Ualens, Dioclitianus, Quintianus, Lucas,
Iohannes, Helias, Abdias, Eupator, Bachides, Clemens,
Ualérius, (Atheniscre byrig), **Eugenia**, (Hierusalē).

Ex. 18: after a new unnamed character is introduced (24 exx.)

[...]

ȝ maurus ȝa gemette ær he to mynstre come ·

æinne dumbne cnapan · ȝ ^{NOM.M.SG}**se** wæs creopere eac ·
ac se fæder · ȝ seo modor · hine feredon þider ·

[...]

and Maurus met there, before he came to the monastery,
a dumb boy; **he (this one)** was also a cripple,
but the father and mother carried him thither.

Maurus₆, l. 15

Who is introduced like this?

A dumb and cripple boy; a blind man; a bedridden man; a man with cancer; a cripple; a buffoon (truð); a man who would drink during the fast; a priest; a thane who went blind; a man who was blind for seven years; a man whose head was bound; a thane; a scholar from Ireland; a blind man; a mass-priest; a servant; a heretic; a dumb child; (a temple); a thane's daughter; God's angel; a poor churl.

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+

internal × nominal

- pronominal antecedent

external × nominal

- introduction of new characters
- **focality**

internal × grammatical

- marking time

external × grammatical

- meta-linguistic reference
- translation and interpretation
- marking time
- grammatical anchor in apposition

SE

Focality

~79 exx.

Ex. 19: focality

He cwæð þ̅ lucas sæde swa swa hit ful soð is ·
þ̅ manega **menn** ongunnon godspel to writenne
butan þam halgan gaste 7 þæs hælendes wissunge ·
7 be heora gewille sædon swa swa him geþuhte ·
7 ^{DAT.PL} **þam** gelamp seo awyrigung þe se witega cwæð ·
[...]

He quoth that Luke said, even as it is very true,
that many **men** began to write the Gospel
without the direction of the Holy Ghost, and of the Saviour,
and according to their own will said even as it seemed to them,
and on **them** fell the curse which the prophet spake,
‘[...]

Ex. 20: focality

[...]

se god þe me forgeaf þis gode geþanc ·

NOM.M.SG

sé wyle þe ge-hyran me þingiende to him ·

the god who gave me this good thought,

he will hear thee, interceding to him for me.

Basil₃, l. 539

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+
internal × nominal	internal × grammatical	
• pronominal antecedent	• marking time	
external × nominal	external × grammatical	
• introduction of new characters	• meta-linguistic reference	
• focality	• translation and interpretation	
	• marking time	
	• grammatical anchor in apposition	

SE

Meta-linguistic reference

84 exx.

Ex. 21: referring to an earlier syntagm

[...] ác **se mán ana gæð úprihte** · ^{NOM.N.SG}**þæt** getacnað · þ hé sceall má
þæncan úpp · þonne nyðer · þelæs þe þ mód sy neoðer · þonne se
lichoma [...]

[...] but **man alone goeth upright**; **that** signifieth that his
thoughts should be more upward than downward, lest the
mind be lower than the body [...]

Nativity₁, l. 57

Ex. 22: referring to the what the allocutor said

Se wyþersaca cwæð · ic wylle georne leof ·
ác ic ne mæg · þeah ic wylle · forþan þe **ic wið-sóc criste ·**
on ge-wryte afæstnode ic wære þæs deofles ·
Se halga wer him cwæð tó · Ne hoga þu embe ^{ACC.N.SG} **þæt** · [...]

The apostate said, ‘I earnestly desire it, my lord;
but I cannot though I wish, because **I denied Christ,**
and confirmed in writing that I was the devil’s.’

The holy man said to him, ‘Be not anxious about **that**; [...]

Basil₃, l. 413

Ex. 23: referring to an event

Þa wæs seo wund gehæled · þe se læce worhte ár ·
eac swilce þa ge-wæda · þe heo bewunden wæs mid ·
wæron swa ansunde · swylce hi eall niwe wæron ·

Sexburh þa hyre swuster swiðe ^{GEN.N.SG} þæs fægnode ·

The wound which the leech had once made was healed;
likewise the linen clothes in which she had been wound
were as fresh as if they had been all new.

Thereat her sister Sexburh was very glad,
[...]

Æðeldryðaða₂₀, l. 93

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+
internal × nominal		internal × grammatical
<ul style="list-style-type: none">• pronominal antecedent		<ul style="list-style-type: none">• marking time
external × nominal		external × grammatical
<ul style="list-style-type: none">• introduction of new characters• focality		<ul style="list-style-type: none">• meta-linguistic reference• translation and interpretation• marking time• grammatical anchor in apposition

SE

Translation and interpretation

84 exx.

Structural features

A þæt is B

A, that is B

	Translation	Interpretation
Occurrences	46 exx.	38 exx.
þæt is in PL ref.	+	+
^{NOM.N.SG} þæt ^{3.PL} synd/wæron	+	+
Tense	only is (no PST or bið)	both PST and NPST
Case of B	only NOM	other cases as well
Text-types	sermons narratives	sermons (narratives, 6 exx.) (dialogue, 1 ex.)

Ærest abel adames sunu wæs gode swa gecweme
þurh unsceðþignysse 7 rihtwisnysse · þ̅ crist sylf hine
het **abél iustus** · ^{NOM.N.SG}þæt is **se rihtwisa abél** ·

First Abel, Adam's son, was so pleasing to God,
through innocence and righteousness, that Christ Himself
called him **Abel Justus**, **that** is, **the Righteous Abel**.

Memory of the Saints₁₆, l. 13

Ex. 25: translation of terms and names

Heo is on bocū manegū naman gecyged · be hyre weorces þenungum · Hyre nama is ^{NOM.N.SG}anima þæt is sawul ⁊ seo naman
gelympð to hire life · ⁊ ^{NOM.N.SG}spiritus gast · belimpð to hire ymb-wlátunge · Heo is ^{NOM.N.SG}sensus · þæt is andgit · oððe fel-nyss ·
þonne heo gefret · Heo is ^{NOM.N.SG}animus · þæt is mód · þonne heo wát · Heo is ^{NOM.N.SG}mens · þæt is mód þonne heo under-stent ·
Heo is ^{NOM.N.SG}memoria · þæt is gemynd · þonne heo gemanð · Heo is ^{NOM.N.SG}ratio · þæt is gescead · þonne heo tó-scaet · Heo is
^{NOM.N.SG}uoluntas · þæt is wylla · þonne heo hwæt wyle · Ac swa þeah ealle þas naman syndon sawul · ælc sawul is gást · ác swá
þeah nis na ælc gast sawul ·

It is called by various names in books, according to its offices. Its name is **anima**, **that** is **soul**, and the name
befitteth its life; and **spiritus**, **spirit**, which appertaineth to its contemplation. It is **sensus**, **that** is **perception** or
sensation, when it perceiveth. It is **animus**, **that** is **intellect**, when it knoweth. It is **mens**, **that** is **mind**, when it
understandeth. It is **memoria**, **that** is **memory**, when it remembereth. It is **ratio**, **that** is **reason**, when it
reasoneth. It is **voluntas**, **that** is **will**, when it willeth anything. Nevertheless all these names are one soul.

Nativity₁, l. 180

Ex. 26: translation of terms and names

[...]

ȝ god him gesette þa oðerne naman **israhel** .

NOM.N.SG

þæt is **uir uidens deum**; ^{NOM.N.SG} **Ðæt** is ongliscre spræce; **se wer þe god**
gesihð .

[...]

[...]

and God gave him that second name, of **Israel**,

that is, '**Vir uidens deum**', (**that** is in English speech) '**The man**
who seeth God',

[...]

Memory of the Saints₁₆, l. 32

ƿa awrát se hælend him sylf ƿis gewrit ·
 ƿ asende ðam cynincge ðus cwæðende him to ·

Beatus es qui credidisti in me · [...]

NOM.N.SG

ƿæt is on engliscum gereorde · **Eadig eart ðu abgar ·**
ƿu ƿe gelyfdest on me · [...]

Then the Saviour Himself wrote this letter,
 and sent it to the king, thus saying to him;

'Beatus es qui credidisti in me, [...]

That is, in the English language, **'Blessed art thou, Abgar,**
thou who believedst on Me [...]

Abdon and Sennes₂₄, l. 102

We andettað mid muðe · ȝ on mode gelyfað
 on þa halgan ðrynnysse · þe is heofonlic god ·
NOM.N.SG
 þæt is fæder *and* sunu · *and* se frefrigende gast ·
 [...]

We confess with mouth and believe in mind
 in the Holy Trinity, Who is heavenly God,
 that is Father and Son, and the Comforting Spirit,
 [...]

Dionysius₂₉, l. 219

Ex. 29: interpreting a nominal syntagm (þæt synd)

[...]

⁊ we sceolon winnan wið þa wælhreowan fynd ·

NOM.N.SG

þæt synd ða ungesewenlican · þa swicolan deofla

þe willað ofslean ure sawla mid leahtrum ·

[...]

[...]

and we ought to strive against the cruel enemies,

that is (lit. are), the invisible ones, and the deceitful devils,

that wish to slay our souls with vices.

[...]

Machabees₂₅, l. 690

Ðæs halgan maures líf wæs þus gelogod ·
 Ða þa hé twelf wintra wæs · he wæs betæht benedicte ·
 he wunode mid him twentig wintra siððan ·
 on his agenū mynstre em feowertig geara ·
NOM.N.SG
 þæt synd eall to-gedere twá hund-seofontig geare ·

This holy Maurus' life was thus divided;
 when he was twelve winters old, he was committed to Saint
 Benedict,
 and he remained with him twenty winters thereafter,
 and in his own monastery just forty years;
 that is (lit. are) are altogether seventy-two years.

Maurus₆, l. 357

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+
internal × nominal		internal × grammatical
<ul style="list-style-type: none">• pronominal antecedent		<ul style="list-style-type: none">• marking time
external × nominal		external × grammatical
<ul style="list-style-type: none">• introduction of new characters• focality		<ul style="list-style-type: none">• meta-linguistic reference• translation and interpretation• marking time• grammatical anchor in apposition

SE

Marking time

81 exx.

Seo eadiga (sic) lucia on þære ylcan stowe wunode ·
 þe heo ofslagen wæs · oðþ̅ sacerdas coman ·
 ȝ hī gehusloden · mid haligre (sic) gerynu ·
 ȝ heo gewát to criste · **mid** ^{DAT.N.SG} **þam** **þe** hī cwædon amen.

The blessed Lucy remained in the same place
 where she was struck down, until priests came
 and houselled her with the holy mysteries,
 and she departed to Christ as (lit. **with that that**) they were
 saying ‘Amen’.

Lucy, l. 146

Wunodon ða on sibbe sume hwile **æfter** ^{DAT.N.SG}ðam ·
 7 iudas þa sende mid sibbe to rome
 gecorene ærendracan [...]

They dwelt then in peace some time **after** **that**,
 and Judas then sent, with peace, to Rome
 chosen messengers; [...]

Machabees₂₅, l. 642

Þa wurdon þa mynster-men micclum afyrhte ·

GEN.N.SG
 7 bædon **þæs** **on mergen** þ̅ hí moston þone sanct
 mid arwurðnysse underfón · þone þe hi ær forsocon ·
 [...]

Then the monks were much affrighted,
 and prayed **then** **in the morning** that they might reverently
 receive
 the Saint, him whom they had before refused.
 [...]

Oswold₂₆, l. 550

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+

internal × nominal

- pronominal antecedent

external × nominal

- introduction of new characters
- focality

internal × grammatical

- marking time

external × grammatical

- meta-linguistic reference
- translation and interpretation
- marking time
- grammatical anchor in apposition

SE

Grammatical anchor in apposition

35 exx.

Ex. 34: representing a content clause

Witodlice ^{NOM.N.SG} þæt ís soþ wysdóm · þæt man gewylnige þ soðe lif
[...]

Verily this is true wisdom, that a man desire the true life [...]

Nativity₁, l. 235

Ex. 35: representing a content clause: rhetoric usage

Hat hí swa yfele · ȝ swa unclæne ·

þ gif ðu hwylcne wyrige · þ ðu gewisce him ^{GEN.N.SG} þæs ·
þæt his líf beo gelíc þinum laðum godum ·

Call them [i.e. the gods] so evil and so unclean,
that if thou wouldest curse any one, thou shouldest wish him
thus,
that his life be like to thy loathly gods.

Agatha₈, l. 79

SE

Doxologies

22 exx.

What is a doxology?

- Greek δοξολογία (from δόξα ‘glory’ and -λογία ‘saying’).
- Hebrew שִׁירָה.
- Characteristic vocabulary:
 - æfre ‘ever’, on ecnysse ‘forever’, á ‘always’, to worulde ‘forever, לעולם’, on ealra worulda woruld ‘for ever and ever, לעלמי כל עלמיא’
 - wuldor ‘glory’, lof ‘praise’, wurðmynt ‘honour’
 - rixian ‘to rule’, (æ)lmihtig ‘(al)mighty’
 - amen ‘amen, אמן’

Quantitative information

Vast majority of texts close with a doxology:



Of which, $\sim 3/4$ connect the text and doxology:



Of which, $3/4$ use SE as their connective element:



Ex. 36: not connected

[...]

Swylce synd þa dydrunga þære dry-manna;

Sy wuldor 7 lof þam wel-willendan scyppende
þe his halgan mærsað mihtiglice mid wundrum ·
seðe á rixað on ecnysse · AMEN ·

[...]

Such are the delusions of the sorcerers.

Be glory and praise to the benevolent Creator,
who magnifieth His saints mightily by miracles,
who reigneth ever to all eternity. Amen.

Swythun₂₁, l. 496

The non-connected doxologies' opening

Sy wuldor 7 lof þam welwillendan { drihtne
gode } [...] AMEN.
scyppende }

Be glory and praise to the benevolent { Lord
God } [...] AMEN.
Creator }

Ex. 37: connected with SE (se + þe)

Þa arærde þa leoda · þær heo læg cyrcan ·

⁊ on hire naman gehalgodon · þam hælende to wurðmynte

NOM.M.SG

se ðe æfre rixað on ecnysse god. AMEN.

Then the people reared a church on the spot where she lay,
and hallowed it in her name, for honour for the Saviour,
He who ever reigneth as God throughout eternity. Amen.

Lucy₉, l. 150

Ex. 38: connected with **SE** (double; **se** + **þe** and **se**)

[...] ac we cweðað to soðum se þe synnum gehyrsumað ·
ȝ godes beboda forsyhð nu on þæs godspelles timan ·
þ he bið þam cynincgū gelic ðe gecuron deofolgild ·
ȝ heora **scyppend** forsawon ·

NOM.M.SG

Se **ðe** soþlice is

ana god ælmihtig · æfre rixigende ·

DAT.M.SG

þam sy wulder on ealra worulda woruld. AMEN.

[...] but we say in truth, that he who obeyeth sins
and despiseth God's commands, now in the gospel's age,
is like the kings who chose idolatry,
and despised their **Creator**; **He** **who** is verily
One God Almighty, ever reigning;
to **Him** be glory to all ages of ages. AMEN.

Ex. 39: connected with SE (þæs)

[...]

ac crist geswutelað mannū hwær se soða geleafa is ·
þonne he swylce wundra wyrçð þurh his halgan
wide geond þas eorðan ·

GEN.N.SG

þæs him sy wuldor

á mid his heofonlican fæder · and þam halgan gaste (a buton
ende)· AMEN ·

[...] but Christ manifesteth to men where the true faith is,
since He worketh such miracles by His saints
widely throughout the earth;

wherefore to Him be Glory

ever with His Heavenly Father, and with the Holy Ghost, for
ever and ever. AMEN.

Ex. 40: connected without SE

His lic wearð þa gefeod to geleaffulre cyrcan
mid mycelre arwyrðnesse · 7 þar on innan bebyriged ·
7 his halgan ban wúrdon wide to-dælede ·
7 mid mycelre lufe hi man wyrðað ge-hwær
swa swa us secgað béc for his soðan geleafan
þā hælende to lofe þe leofað á on ęcnesse · *amen*.

His body was then borne to a holy church
with much veneration, and therein buried;
and his holy bones were widely distributed,
and with much love men revere them everywhere,
as the books tell us, for his true faith,
to the praise of the Saviour, who liveth aye in eternity. Amen.

Vincent₃₇, l. 276

SE: types of reference

	nominal	grammatical
internal	+	+
external	+	+

internal × nominal

- pronominal antecedent

external × nominal

- introduction of new characters
- focality

internal × grammatical

- marking time

external × grammatical

- meta-linguistic reference
- translation and interpretation
- marking time
- grammatical anchor in apposition

DIS

DIS: types of reference

	nominal	grammatical
internal	-	-
external	(+)	+

- **external × nominal**

- concrete reference

- **external × grammatical**

- representation of content
- marking time
- meta-textual reference
- topic under discussion

DIS: types of reference

	nominal	grammatical
internal	-	-
external	(+)	+

- **external × nominal**

- concrete reference

- **external × grammatical**

- representation of content
- marking time
- meta-textual reference
- topic under discussion

DIS

Representation of content

43 exx.

[...]

NOM.N.SG

þis wearð þa gesæd sona þā fæder ·
⁊ he beweop þone sunu swilce he dead wære ·

[...]

Then **this** was soon told to the father,
and he bewept his son as if he were dead.

Crisantus and Daria₃₅, l. 73

An gecyndelic · æ · is eallum mancynne gesett ·
 þ̅ nan man ne gedó dare oðrum menn ·
 swa swa se hælend **cwæð** on his halgan godspelle ·
Þæt þæt ðu þe sylfū nelt on þinū life becuman ·
ACC.N.SG
ne do ðu þæt oðrum menn · þis **cwæð** drihten sylf ·

One natural law is appointed to all mankind,
 that no man may do harm to another man,
 even as the Saviour **said** in His holy gospel;
'That which thou desirest not to befall thyself in thy life,
that do not to another man.' **This** **said** the Lord Himself.

Forty Soldiers₁₁, l. 348

Þa astrehte se halga his handa 7 **cwæð** ·

[...]

Mid þam ðe he ^{ACC.N.SG} **þis** **cwæð** · þa com crist sylf him to ·
on þære ylcan gelicnysse þe he leofode on worulde ·
7 hine mid sibbe gegrette · 7 sæde him þas word ·

Then the Saint stretched out his hands and **said**,

[...]

While he was **saying** **this**, there came Christ Himself to him,
in the same likeness in which He had lived in the world,
and gave him the greeting of peace, [...]

Læt þine godas geyrsian · gif hi aht magon ·

Læt hi sylfe beodan þæt we us to him gebiddan ·

gif þu ^{ACC.N.SG} þis dón ne miht · drece us loca hu þu wylle ·

Let thy gods be angry if they can do aught.

Let themselves command us to worship them;

if thou canst not accomplish **this**, afflict us, lo! how thou wilt.

Agnes₇, l. 113

DIS: types of reference

	nominal	grammatical
internal	-	-
external	(+)	+

- **external × nominal**

- concrete reference

- **external × grammatical**

- representation of content
- marking time
- meta-textual reference
- topic under discussion

DIS

Marking time

22 exx.

^{DAT.N.SG}
 Æfter þysū wearð ge-leaht seo geleaffulla eugenia ·
 7 to þā hæðenan temple getogen mid ge-ðreate ·
 þ heo þære gydenan diane · godes wurðmynt gebude ·

After this the faithful Eugenia was caught,
 and dragged, with threatening, to the heathen temple,
 that she might offer the worship, due to God, to the goddess
 Diana.

Eugenia₂, l. 383

æfter þysum 'after this': characteristics

- Initial; capital letter in MS
- Evolution mode
- Not compatible with the narrative chaining marker þa

⇒ a boundary signal

He cw̃ þa to þā arleasan · ^{ACC.N.SG} **oð þis** ic cāpode þe ·
ge-þafa nu þ̃ ic gode campige heonon-forð ·
ȝ under-fó þine gife · se ðe feohte mid ðe
ic eō godes cempa ne mot ic na feohtan ·

He said to the impious one: ‘**till now** I have fought for thee;
suffer me henceforth to fight for God,
and let him who fights for thee receive thy gift.
I am God’s champion; I must not fight.’

Martin₃₁, l. 103

DIS: types of reference

	nominal	grammatical
internal	-	-
external	(+)	+

- **external × nominal**

- concrete reference

- **external × grammatical**

- representation of content
- marking time
- meta-textual reference
- topic under discussion

DIS

Meta-textual reference

15 exx.

Ex. 47: referring to the whole narrative

NOM.N.SG

Þis wæs geworden ær ðæt gewinn come
ðurh hengest · 7 horsan þe hyndon ða bryttas ·
7 se cristen-dom wearð ge-unwurðod syððan ·
oðþ agustinus hine eft astealde ·
be gregories lare þæs geleaffullan papan ·

This was done before that strife came
through Hengest and Horsa who defeated the Britons,
and Christianity was again dishonoured,
until Augustine re-established it,
according to the instruction of Gregory, the faithful pope.

Alban₁₉, l. 147

Meta-textual references

- The whole narrative (2 exx.)
- A sub-narrative (3 exx.)
- A narrative event (9 exx.; cf. ex. 23)
- A Narrator's Channel segment (1 ex.)

DIS: types of reference

	nominal	grammatical
internal	-	-
external	(+)	+

- **external × nominal**

- concrete reference

- **external × grammatical**

- representation of content
- marking time
- meta-textual reference
- **topic under discussion**

DIS

Topic under discussion

6 exx.

Be ^{DAT.N.SG} **þysum** cwæð se ælmihtiga god · to moyse on þā wæstene ·
Gif ge on minum bebodum farað · 7 mine behoda healdað ·
Ðonne sende ic eow rén-scuras on rihtne timan symble ·

Concerning **this** spake the Almighty God to Moses in the wilderness,
'If ye walk in my statutes and keep my commandments,
then will I always send you rain-showers in due time, [...]

Prayer of Moses₁₃, l. 156

Ðær wæron ge-hælde þurh ða halgan femnan
 fela adlige menn · swa swa we gefyrn gehyrdon ·
 7 eac ða þe hrepodon þæs reafes ænigne dæl ·
 þe heo mid bewunden wæs · wurdon sona hale ·
 7 manegum eac fremode seo cyst micclū ·
 þe heo ærest on læg · swa swa se lareow beda
 on ðære bec sæde · þe he ge-sette be ^{DAT.N.SG} **ðysum** ·

By means of this holy woman were healed
 many sick men, as we have heard of old;
 those also who touched any part of the shroud
 in which she had been wound, were instantly cured;
 and likewise the coffin wherein she had first lain
 greatly benefited many persons, as the teacher Beda
 said in the book which he wrote concerning **this**.

DIS: types of reference

	nominal	grammatical
internal	-	-
external	(+)	+

- **external × nominal**

- concrete reference

- **external × grammatical**

- representation of content
- marking time
- meta-textual reference
- topic under discussion

DIS

Concrete reference

21 exx.

Ða be-láf **sebastianus** on þære byrig mid **þam papan** · 7 se geonga **tiburtius** · 7 þa twægen
 gebroðra · **marcus** · 7 **marcellianus** · mid heora fæder **tranquilline** · **Nicostratus** mid **his**
breðer · 7 his gebeddan **Zoe** · **Uictorinus** mid **his breðer** · 7 **his broðor suna** ·

NOM.PL

Ðas belifon on rome · on þære reðan eht-nysse ·
 7 þa oþre ealle endemes ferdon awæg ·

mid chromatia · swa swa him crist gewissode ·

There remained in the city, with **the pope**, **Sebastian**, and the young **Tiburtius**, and the
 two brothers **Marcus** and **Marcellianus**, with their father **Tranquillinus**, **Nicostratus**,
 with **his brother** and his wife **Zoe**, **Victorinus**, with **his brother**, and **his brother's son**;
these remained in Rome in the fierce persecution,
 and all the others at last went away

with Chromatius, even as Christ instructed them.

Sebastian₅, l. 339

Decius þa het **þa halgan cyningas**

his godū geoffrian · ac hi ȝwyrdon þus ·

We offriað ure lac þam lyfigendan gode ·

hælendū criste · ȝ we hopiað to him ·

geoffra ðu sylf þinū sceandlicum godū ·

þa cwæð decius se deofles biggenga ·

DAT.PL

Þysum is to gearcigenne þa reþestan wita ·

Then Decius bade **the holy kings**

to sacrifice to his gods, but they answered thus:

‘We offer our sacrifices to the living God,

Jesus Christ, and we hope in Him;

do thou thyself sacrifice to thy shameful gods.’

Then said Decius, the devil’s worshipper;

‘For **these (men)** must be prepared the sharpest punishments.’ 110

DIS: types of reference

	nominal	grammatical
internal	-	-
external	(+)	+

- **external × nominal**

- concrete reference

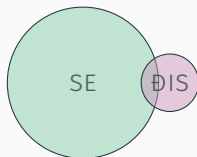
- **external × grammatical**

- representation of content
- marking time
- meta-textual reference
- topic under discussion

SE:DIS

Main differences between SE and DIS in the corpus

	SE	DIS
quantitative	common (821)	less common (119)
syntagmatic	SE + þe	no DIS + þe
referential	nominal and grammatical	mostly grammatical
functional	diverse	quite limited
morphological	almost a full paradigm	partial paradigm



(the illustration isn't precisely proportional)

Quantitative imbalance

	SE	DIS
with verbs of report	3 exx.	18 exx.

Not enough examples to draw conclusions

	SE	DIS
betwux þam:þysum	2 exx.	2 exx.
Affective use of DEM → þ	2 exx.	2 exx.

Complementary distribution

	SE	DIS
oð þæt:þis	10 exx. — indir. sp. (time of speech) — narrative	9 exx. — dir. sp. (<i>nynégocentrique</i>)
æfter þam:þysum	1 ex. modifying <i>sume hwile</i>	11 exx. boundary signal
concrete phys. ref.	indir. sp.	dir. sp

Conclusion

- Relatively few core functions, fitting in a 2×2 matrix.
 - Some of these functions are unexpected.
- SE:DIS opposition is quite limited:
 - overlapping functions
 - number of occurrences

So, what do I want you to take from this talk?

1. Usefulness for your research
2. Data can be surprising
3. Homonyms should be distinguished
4. Open research
5. Study Old English!

So, what do I want you to take from this talk?

1. Usefulness for your research
2. Data can be surprising
3. Homonyms should be distinguished
4. Open research
5. Study Old English!

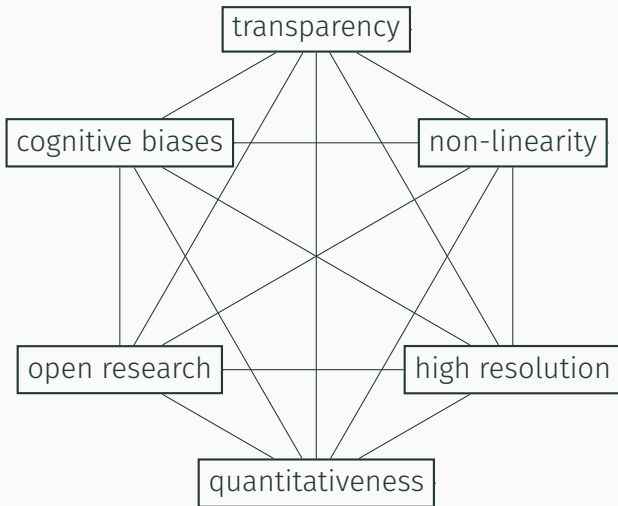
Part III

Open Research in linguistics

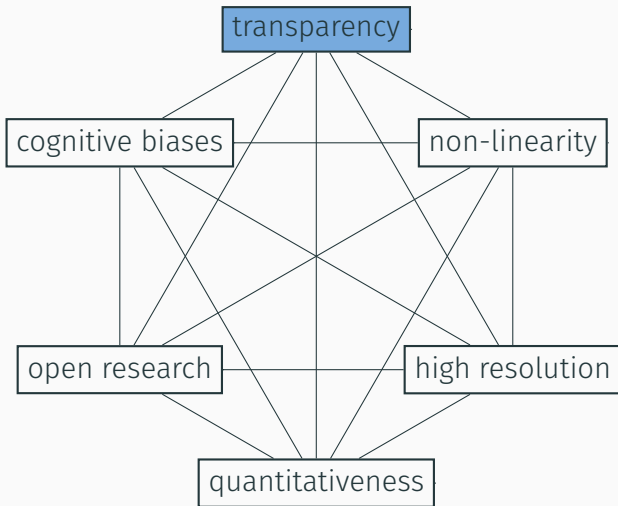
General principles

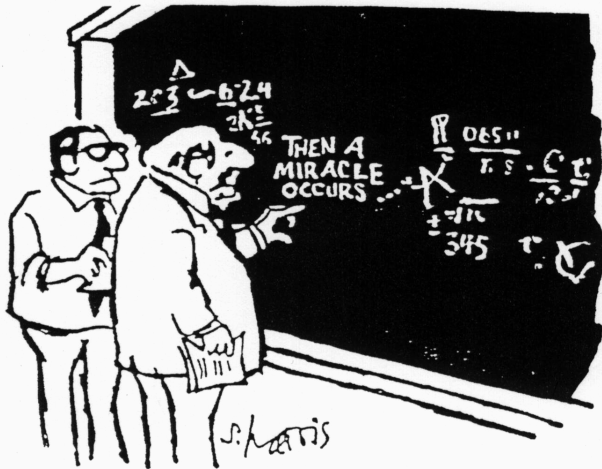
- Principle 0:
 - The corpus should be available to all.
- Principle I:
 - Make all collected examples available to the readers (not just a representative sample).
 - Don't overwhelm them.
- Principle II:
 - Structured data are more valuable than unstructured data.
 - Annotate, tag and categorize.

Six advantages



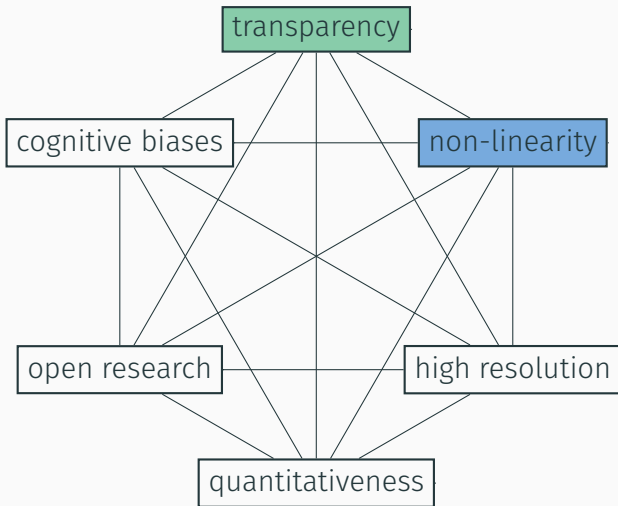
Six advantages: transparency





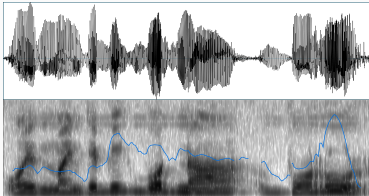
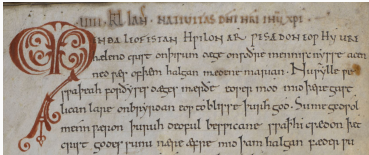
'I think you should be more explicit here in step two.'

Six advantages: non-linearity



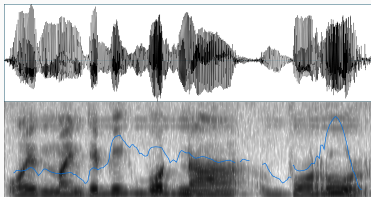
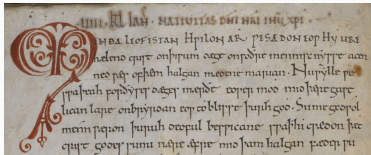
Non-linearity of the *langue*

Parole (linguistic output) is linear.

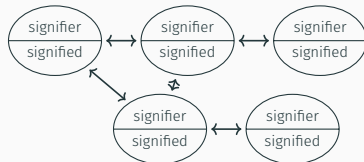


Non-linearity of the *langue*

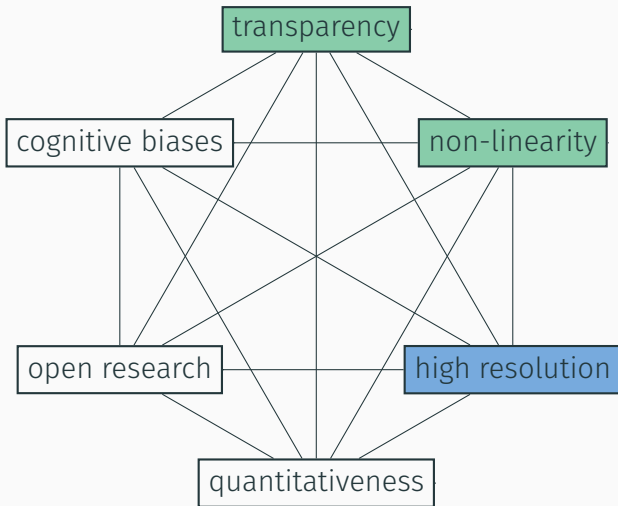
Parole (linguistic output) is linear.



Langue (linguistic system) is not.



Six advantages: high resolution



Zoom, context and the linguistic system



Abstract shape

Istvan Banyai, *Zoom*, 1995

Zoom, context and the linguistic system



Abstract shape

Istvan Banyai, *Zoom*, 1995

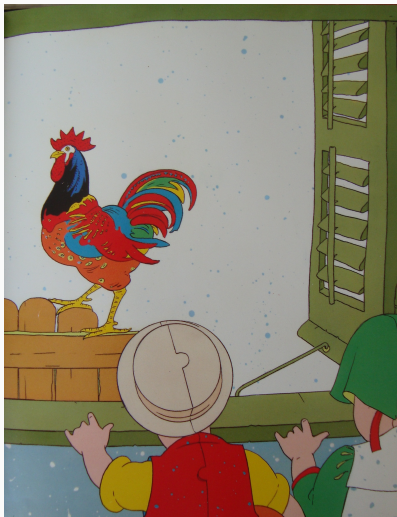
Zoom, context and the linguistic system



Gallus gallus domesticus

Istvan Banyai, *Zoom*, 1995

Zoom, context and the linguistic system



With children watching

Istvan Banyai, *Zoom*, 1995

Zoom, context and the linguistic system



In a house

Istvan Banyai, *Zoom*, 1995

Zoom, context and the linguistic system



In an idealized farm

Istvan Banyai, *Zoom*, 1995

Zoom, context and the linguistic system



Within a village

Istvan Banyai, *Zoom*, 1995

Zoom, context and the linguistic system



A toy set

Istvan Banyai, *Zoom*, 1995

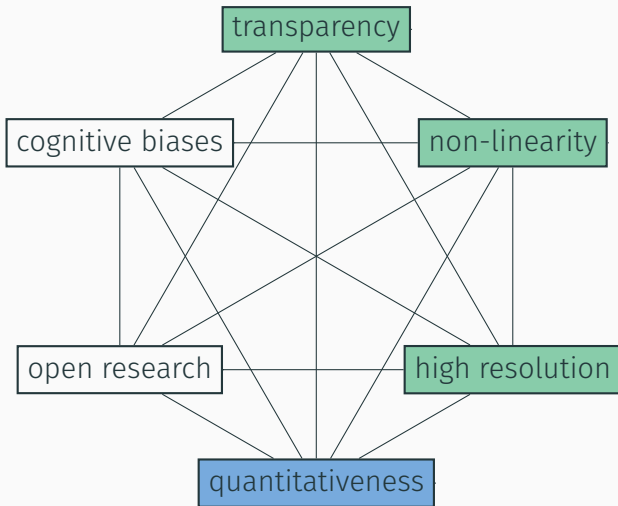
Zoom, context and the linguistic system



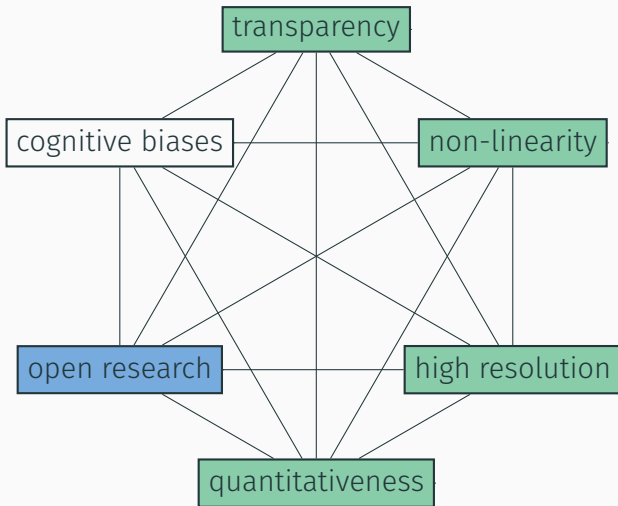
An advertisement

Istvan Banyai, *Zoom*, 1995

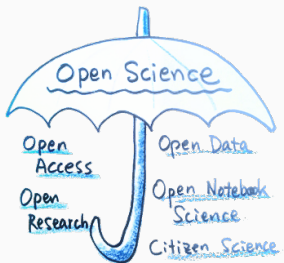
Six advantages: quantitiveness

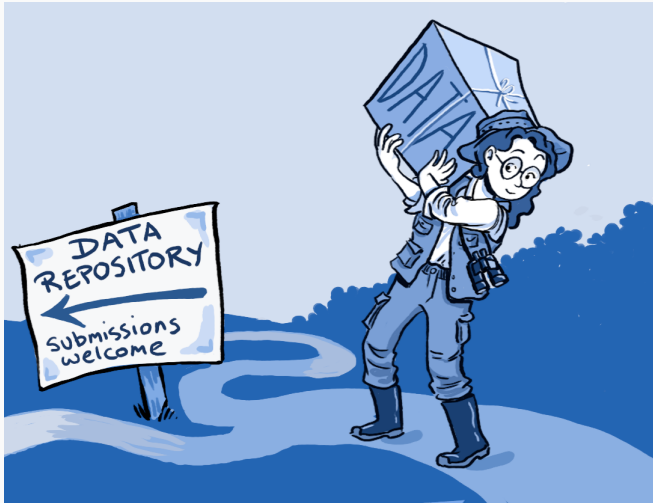


Six advantages: open research

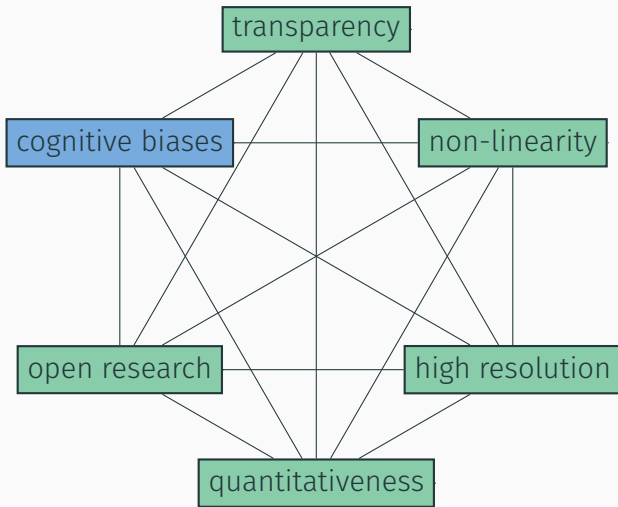


Open Research is more than Open Access



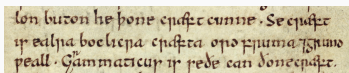
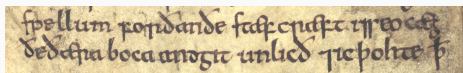


Six advantages: cognitive biases



[...] ordonner les faits linguistiques au point de vue de la langue même.

MEILLET (1948)



[...] grammar is the key that unlocks
the understanding of these books [...]

This craft (grammar) is the
source and foundation of
all literary crafts.

Ælfric's Grammar, Oxford, St. John Collage, MS. 154, ff. 1v and 143r

Júda Ronén

<https://ac.digitalwords.net>



Creative Commons Attribution

Part VI

Technology, linguistic research and collaboration

Plain text vs. binary blobs

	text files	binary files
human-readable	yes	no
machine-readable (programmatically)	easily!	not easily / not possible

Text file:

```
\appendix
\part{Technology, linguistic research\and collaboration}
\frame{\partpage}

\begin{frame}{General guidelines}
  \begin{itemize}
    \item Learning good tools is cost-effective.
    \item Always use plain-text files.
    \item Automatize using generic modular tools
      (scripted text processing).
    \item Share source files.
  \end{itemize}
\end{frame}

\begin{frame}{Why?}
  \begin{itemize}
    \item Time-saving
      \begin{itemize}
        \item Write once, use in diverse ways.
        \item Automatize repetitive mindless work.
      \end{itemize}
    \item Future-proof
    \item Scientific collaboration
  \end{itemize}
\end{frame}
```

Binary file:

```
0503 0c33 0100 0211 ...3....
0304 2112 3105 4151 ..!.1.AQ
6113 2271 8132 0614 a."q.2..
91a1 b142 2324 1552 ...B#$..R
c162 3334 7282 d143 .b34r..C
0725 9253 f0e1 f163 .%.S...c
7335 16a2 b283 2644 s5....8D
9354 6445 c2a3 7436 .TdE..t6
17d2 55e2 65f2 b384 ..U.e...
c3d3 75e3 f346 2794 ..u..F'.
a485 b495 c4d4 e4f4 .....
a5b5 c5d5 e5f5 5666 .....VF
7686 96a6 b6c6 d6e6 v.....
f637 4757 6777 8797 .7Gwgw..
a7b7 c7d7 e7f7 1100 .....

```

General guidelines

- Learning good tools is cost-effective.
- Always use plain-text files.
- Automatize using generic modular tools (scripted text processing).
- Share source files.

Why?

- Time-saving
 - Write once, (re-)use in diverse ways.
 - Automatize repetitive mindless work.
- Future-proof
- Scientific collaboration

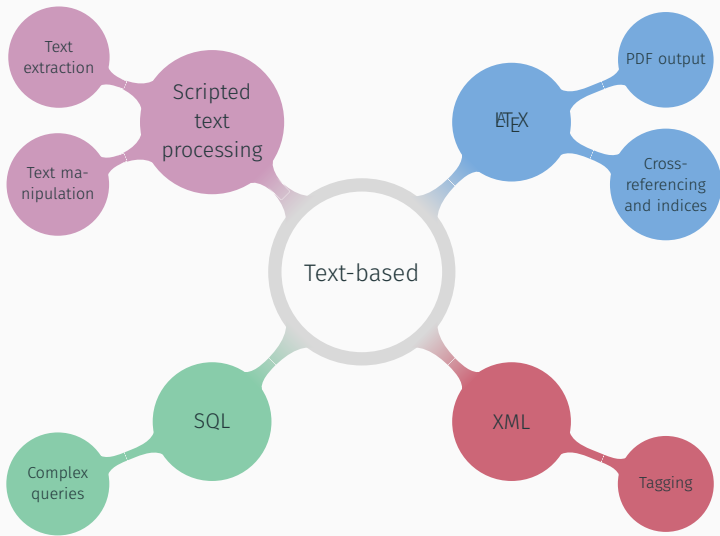
Sharing written corpora

- Unicode (it's 2018!)
 - MUFI (Medieval Unicode Font Initiative, folk.uib.no/hnooh/mufi)
 - MdC (Manuel de Codage)
- TEI (Text Encoding Initiative, tei-c.org)
 - XML
- Wikisource (Wikimedia Foundation, wikisource.org)
 - MediaWiki syntax
 - en.wikisource.org/wiki/Ælfric's_Lives_of_Saints
- Internet Archive (archive.org)
- Git repositories (hosting: GitHub, GitLab, ...)

Sharing audio corpora

- Lossless files
 - WAV (Waveform Audio File Format)
 - FLAC (Free Lossless Audio Codec, xiph.org/flac)
- ELAN (EUDICO Linguistic Annotator, tla.mpi.nl/tools/tla-tools/elan)
- Internet Archive (archive.org)

Text-based processing



Tagging examples

```
5126 1 Are <you id="784" type="ti" speaker="Firenze" addressee="Harry" page="203">you</you> all right?' said the centaur, pulling Harry to his feet.
1
2 'Yes - thank <you id="785" type="-" speaker="Harry" addressee="Firenze" page="203" tags="idiomatics">you</you> - what was that?'
3
4 The centaur didn't answer. He had astonishingly blue eyes, like pale sapphires. He looked carefully at Harry, his eyes lingering on the scar which stood out, livid, on Harry's forehead.
5
6 '<you id="786" type="chi-sg" speaker="Firenze" addressee="Harry" page="203">You</you> are the Potter boy,' he said. '<you id="787" type="chi-sg" speaker="Firenze" addressee="Harry" page="203">You</you> had better get back to Hagrid. The Forest is not safe at this time - especially for <you id="788" type="chi-sg" speaker="Firenze" addressee="Harry" page="203">you</you>. Can <you id="789" type="chi-sg" speaker="Firenze" addressee="Harry" page="203">you</you> ride? It will be quicker this way.
7
8 'My name is Firenze,' he added, as he lowered himself on to his front legs so that Harry could clamber on to his back.
```

Scripted text processing

```
1 ack "<[a-z][^>]*>" english.xml -o |\
2 sed 's/<[a-z]* id="[0-9]*" type="\([^"]*\)" speaker="\([^"]*\)" addressee="\([^"]*\)".*/\1\t\2\t\3/g' |\
3 sort | uniq |\
4 sed '/^(-|<|am|\t|=|[^ct]\|chi-pl\).*/d' |\
5 sort -t'\t' -k2 -nr
```

- Python, Perl, *sh, ...
- Unix utilities: sed, grep, awk, ...
- Pipeline (|)

Data extraction I

⋮		
chi	Doris Crockford	Harry
ti	Draco	Harry
chi	Draco	McGonagall
ti	Draco	Neville
ti	Draco	Ron
ti	Dudley	Harry
ti	Dumbledore	Hagrid
ti	Dumbledore	Harry
ti	Dumbledore	McGonagall
ti	Filch	Harry
ti	Filch	Peeves
chi	Filch	Snape
ti	Firenze	Bane
chi	Firenze	Harry
ti	Firenze	Harry
⋮		

Data extraction II

:					
yourself	0	chi-sg	ni	a man wearing a violet cloak	Vernon
yourself	1	ti	nin	Dumbledore	McGonagall
yourself	2	-	ni	McGonagall	Hagrid
your	0	ti	ni	Vernon	Petunia
you	0	chi-ambiguous	-	FIS (The Dursleys)	audience
you	1	neb	-	author	audience / generic
you	2	chi-ambiguous	-	FIS (Vernon)	audience / generic
you	3	chi-sg	nin	a man wearing a violet cloak	Vernon
you	4	ti	-	Vernon	Petunia
you	5	-	-	Vernon	Petunia
you	6	ti	ni	Vernon	Petunia
you	7	-	-	Petunia	Vernon
you	8	impersonal	-	author	audience
:					

ACADEMY OF THE HEBREW LANGUAGE (1984). דקדוק לבית ספר, תשמ"ד.

האקדמיה ללשון העברית. URL: <http://hebrew-terms.huji.ac.il/>.

ÆLFRIC (1881–1900). *Ælfric's Lives of Saints: being a set of sermons on saints' days formerly observed by the English Church*. Ed. by Walter W. SKEAT. Trans. by GUNNING, WILKINSON, and Walter W. SKEAT. Early English Text Society 76 and 82. London: N. Trübner and co. for the Early English Text Society. URL: <https://archive.org/search.php?query=Aelfric%27s%20Lives%20of%20Saints>.

- ÆLFRIC (2013). *Ælfric's Lives of Saints: being a set of sermons on saints' days formerly observed by the English Church*. Wikisource digital edition. Ed. by Walter W. SKEAT. Digitized by Júda RONÉN. URL: https://en.wikisource.org/wiki/%C3%86lfric's_Lives_of_Saints.
- CAMPBELL, Charles Boyle (1913). "Concerning the Pronominal Antecedent and the Form of the Accompanying Relative Pronoun in Modern German Prose". PhD thesis. Berlin: University of Chicago. URL: <https://archive.org/details/concerningpronom00camprich>.
- CHALKER, Sylvia and Edmund WEINER (1998). *The Oxford Dictionary of English Grammar*. Oxford University Press. ISBN: 9780192800879.

- MEILLET, Antoine (1948). *Linguistique Historique et Linguistique Générale*. Vol. 1. 1st ed. 2 vols. Paris: Edouard Champion.
- POLOTSKY, Hans Jakob (1991). ““Article” and “Determinative Pronoun” in Coptic”. In: *Lingua Aegyptia* 1, pp. 241–242.
- ROSÉN, Haiim B. (2005). “The Jerusalem School of Linguistics and the Prague School”. In: SWIGGERS, Pierre. *Haiim B. Rosén: bio-bibliographical sketch*. Ed. by Pierre SWIGGERS. Vol. 8 (N.S.) Biobibliographies et exposés. Leuven: Peeters. ISBN: 90-429-1695-8.
- SHISHA-HALEVY, Ariel (1976). “The Circumstantial Present as an Antecedent-less (i.e. Substantival) Relative in Coptic”. In: *Journal of Egyptian Archaeology* 62, pp. 134–137. JSTOR: 3856353.

- SHISHA-HALEVY, Ariel (1998). *Structural Studies in Modern Welsh Syntax: aspects of the grammar of Kate Roberts*. Studien und Texte zur Keltologie. Münster: Nodus Publikationen. ISBN: 3-89323-612-0.
- TILL, Walter Curt (1955). *Koptische Grammatik (Saidischer Dialekt)*. Mit Bibliographie, Lesestücken und Worterverzeichnissen. Leipzig: O. Harrassowitz. 360 pp.