

Information Structure Triggers for Word Order Variation and Change

The OV/VO Alternation in the West Germanic Languages

31-03-2022

Tara Struik

COMPARATIVE LAB

OV/VO

Variation



Dutch OV

....dat hij **het boek** gelezen heeft

German OV

....dass er **das Buch** gelesen hat

English VO

....that he **has read the book**

OV/VO IN EARLY WEST-GERMANIC

OV

English

We nu willaþ **ure saula smerian** mid
mildheortnesse ele
'We now wish to anoint our souls with oil of mercy.'
(HomS_21_[BIHom_6]:73.136.927)

Dutch

die **dat riet ghemaect hadde**
'who made that reed'
Gysseling_1340_1294

High German

daz ir **den ewigin lib befíztín mūzint**
'that you may have eternal life'
(Mitteldeutsche Predigten, 6ra,8)

Low German

Do de greyken **dusse stad vorstort hadden**
'when the Greek had overthrown this city'
(Engelhus_Weltchronik_OF_1435)

VO

Þæt ic **mihte geseon þone scinendan engel**
'That I might see the shining angel'
(ÆLS_[Cecilia]:46.7137)

dat ic **hebbe genomen dat hues terhurst** bi
wille mijns heren
'that I have taken that Huis ter Horst by the will of my lord'
Gysseling_1502A_1296

thaz ih ni **mugi bittan minan fater,**
'that I may not ask my father'
(Tatian 185)

Dat du **scalt hebben eyn kindelin**
'that you will have a child'
(Flos_EE_1401-1450)

COMPARATIVE LAB

Are these languages
the same and
divergent?

OV/VO
Variation

Are these languages
different and further
divergent?



Dutch OV

....dat hij **het boek** gelezen heeft

German OV

....dass er **das Buch** gelesen hat

English VO

....that he **has read the book**

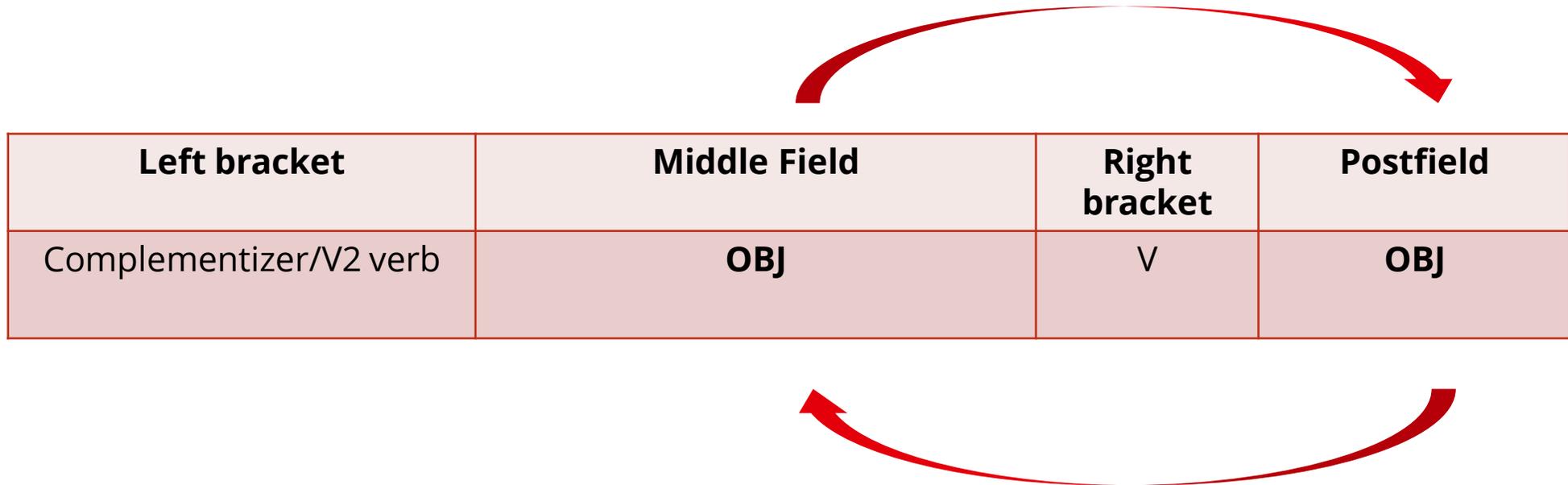
INFORMATION STRUCTURE TRIGGERS FOR WORD ORDER VARIATION AND CHANGE:
THE OV/VO ALTERNATION IN THE WEST GERMANIC LANGUAGES

Roadmap

- **How to analyse OV/VO variation?**
- **One common factor: information structure**
- **One unified methodology**
 - Old English/Middle English**
 - Middle Dutch**
 - Old Saxon/Middle Low German**
 - Old High German/Middle High German**
- **Excursus: scrambling in Dutch**
- **Towards a unified derivation**

HOW TO ANALYSE OV/VO VARIATION?

- The quest for the basic word order...



- ... and how to derive the variation.

PREVIOUS ANALYSES OF VARIATION

- **Old English/Middle English:** long standing debate
 - Basic OV with optional rightward movement (van Kemenade, 1987; Kroch & Pintzuk 1989; Koopman, 1990; Lightfoot, 1991; Stockwell & Minkova, 1991)
 - Basic VO analysis with optional leftward movement (Fischer et. al, 2000; Biberauer & Roberts, 2005; de Bastiani 2019; Struik & van Kemenade 2022)
 - The Double Base Hypothesis; basic OV and basic VO, with rightward and leftward movement whenever relevant (Pintzuk, 1999; Pintzuk, 2005; Taylor & Pintzuk, 2012a,b)
 - Variable position of the object dependent of the existence of/movement of the verb to *vP* (Fuss 2002; Broekhuis 2022)

PREVIOUS ANALYSES OF VARIATION

- **Middle Dutch**
 - VO is “more liberal extraposition” from an OV base (Burridge 1993)
 - Word order is essentially free: VO because overt case marking licenses escape from OV (Weerman 1987, 1989)
- **Old Saxon/Middle Low German**
 - Basic VO. OV derived by leftward movement (Walkden 2014, following Wallenberg 2009).
 - Variable position of the verb might be the result of metre (Dubenion-Smith & Somers 2014)
- **Old High German/Middle High German**
 - OV is basic. VO is extraposition from an OV base (Sapp 2014, Sapp 2016)
 - VO is basic. OV is the result of leftward movement (Petrova 2009, Hinterhölzl 2015)

PREVIOUS ANALYSES OF VARIATION

- Each study with its own selection criteria and assumptions about the structural analysis of the language
 - Distinction between main or subclauses (or not)
 - Clauses with one or two verbs
 - Extraposition: objects DPs or also PPs, subjects or appositions
 - Direct or indirect objects combined
 - VO, but not OV

→ Not readily comparable/compatible

ONE COMMON FACTOR: INFORMATION STRUCTURE

- What motivates the variation?
- **Information Structure**
- “That component of sentence grammar in which propositions as conceptual representations of states of affairs are paired with lexicogrammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts” (Lambrecht 1994)
- In simple terms, the discourse context and the role an object plays in it determines the way an utterance looks.
- Given-before-New Principle (Gundel et al. 1988)
- Speakers and writers have the tendency to place information that is given before information that is new.

ONE COMMON FACTOR: INFORMATION STRUCTURE

- What motivates the variation?
- **Information Structure** \approx given before new
 \approx given prefers OV, new prefers VO
- Wealth of different information structure annotation schemes (see Féry and Ishihara 2019 for an overview)
 - Given vs. new
 - Focus vs. background
 - Presentational focus vs. contrastive focus
 - Topic vs. focus
 - definiteness marking as a proxy for information status
 - etc.

→ Not readily comparable/compatible

THE OV/VO ALTERNATION IN WEST GERMANIC

Syntax

- Syntactic variation is to a large extent similar
- The different underlying assumptions and selection criteria make it hard to compare across languages

Information structure

- Information structure plays a role in all languages
- The different information structure annotation schemes (in combination with different selection criteria) make it hard to compare across languages
- To really understand the differences and similarities in OV/VO variation in the West Germanic languages – on both a syntactic and functional level - we need a systematic comparison based on uniform selection criteria.

ONE UNIFIED METHODOLOGY

- Subclauses with a finite verb, a non-finite verb and a direct object
Excluded: forms of *to be*, *to*-infinitives, indirect objects, pronouns, quantified and negated objects
- Each object was annotated for:
 - INFORMATION STATUS: Based on anaphoricity and referentiality (following Komen 2013)
 - LENGTH: number of letters
- Statistical validity of the observations was tested using mixed-effects regression modeling.
 - LENGTH significantly increases the likelihood of VO in all languages

INFORMATION STRUCTURE: PENTASET

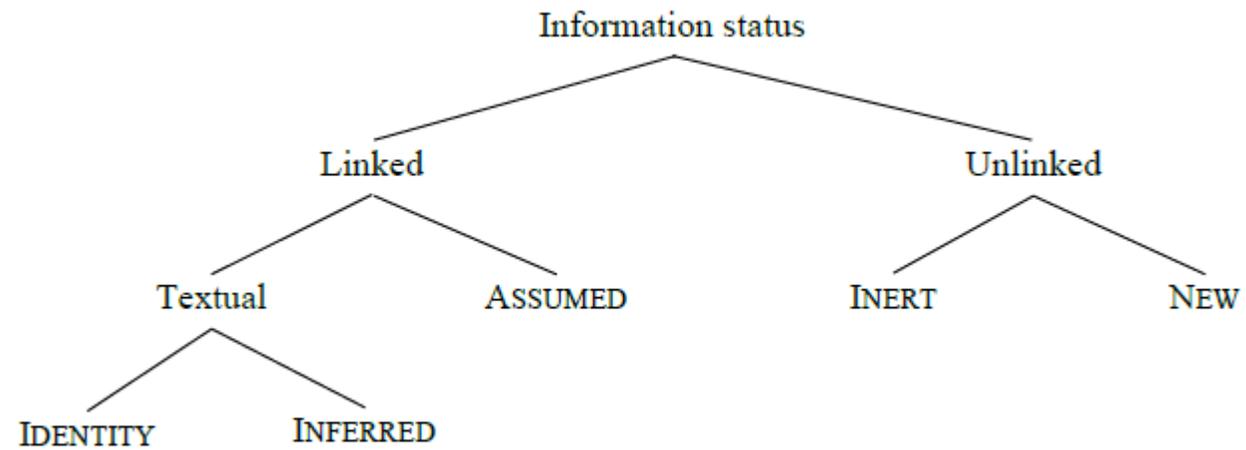


Figure 4. The Pentaset information status labels (Komen 2013: 144)

INFORMATION STRUCTURE: PENTASET

GIVEN

NEW

Identity

Cats deliver more eye narrowing movements when their owners slow blink at them than when the owner did not deliver **this stimulus**

Inferred (elaborating)

until the masters came to reasonable terms with the slaves and confirmed **the agreement** by a solemn oath

Assumed

... and he could read **the stars**.

Discourse-new

I just ate **a delicious pancake**

Inferred (bridging)

He did not want to sell him **his vinyard**

SOURCE MATERIALS

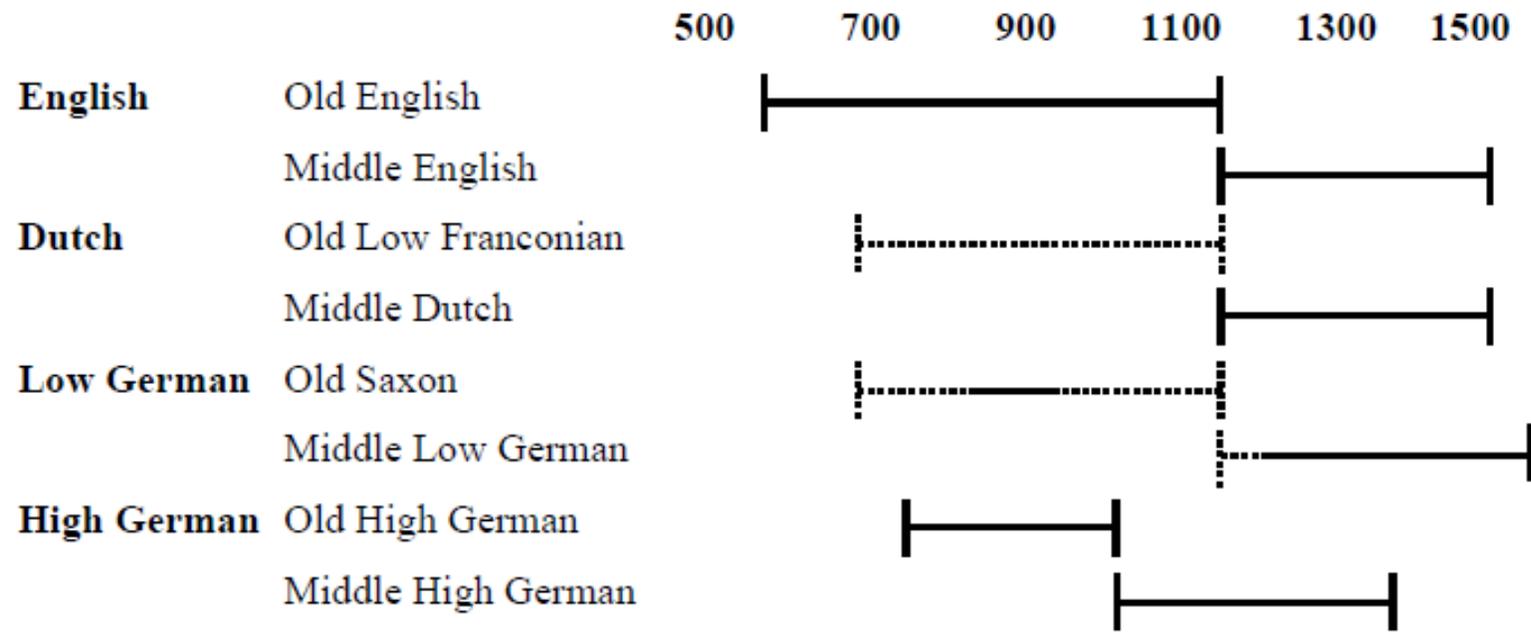


Figure 3. Periodization and attestation of the early Germanic languages.

RESULTS ENGLISH

- Old English (850 -1050, from YCOE (Taylor et al. 2003) – all native OE texts)

	OV	VO	Total
Given	440 (74.8%)	148 (25.2%)	588
New	7 (5.2%)	127 (94.8%)	134

- (early) Middle English (1150 – 1250 from PPCME2, (Kroch et al. 2000) – all native eME texts)

	OV	VO	Total
Given	109 (50.9%)	105 (49.1%)	214
New	0	57 (100%)	57

- Given objects occur freely in OV order; new objects strongly prefer VO.

RESULTS

ENGLISH: TRANSLATIONS FROM LATIN

LATIN ORDER	OV		TOTAL	VO		TOTAL
OE ORDER	OV	VO		OV	VO	
GIVEN	77 81.9%	17 18.1%	94	26 86.7%	4 13.3%	30
NEW	7 33.3%	14 66.7%	21	3 33.3%	6 66.7%	9
	84	31	115	29	10	39

- The number of OV objects is significantly higher (61.9% vs. 73.4%)
- The general information structure pattern is similar, and deviations are mostly in the expected direction, but the number of new OV objects is higher.

RESULTS

LOW GERMAN

- Old Saxon (9th Century, from HeLiPad (Walkden 2015) – one text in metric verse)

	OV	VO	Total
Given	113 (85.4%)	19 (14.6%)	132
New	13 (56.5%)	10 (43.5%)	23

- Middle Low German (1250 -1500; from CHLG (Booth et al. 2020) – native MLG texts, mostly official)

	OV	VO	Total
Given	222 (89.2%)	27 (10.8%)	249
New	21 (41.2%)	30 (58.8%)	51

- New objects occur freely in either OV or VO order, but given objects disprefer VO.

RESULTS

HIGH GERMAN

- Old High German (750-1050; from Referenzkorpus Altdeutsch (Donhauser 2015) – mostly translations)

	OV	VO	Total
Given	33 (57.9%)	24 (42.1%)	57
New	8 (57.1%)	6 (42.9%)	14

- Middle High German (1050-1350; Referenzkorpus Mittelhochdeutsch (Klein et al. 2016), mostly native MHG prose)

	OV	VO	Total
Given	209 (90.9%)	21 (9.1%)	230
New	21 (61.8%)	13 (38.2%)	34

- Old High German: slight preference for OV, no difference between OV and VO
- Middle High German: New objects occur freely in either OV or VO order, but given objects disprefer VO.

RESULTS DUTCH

- Middle Dutch (1200-1500, from Corpus Gysseling (2021); CRM14 (van Reenen & Mulder 1993); CLVN (van der Sijs et al. 2018) – all native MD texts, mostly official)

	OV	VO	Total
Given	508 (86.7%)	78 (13.3%)	586
New	79 (39.3%)	122 (60.7%)	201

- New objects occur freely in either OV or VO order, but given objects disprefer VO.

HOW TO ANALYSE OV/VO VARIATION?

- Information structure influences object position in all languages, but the direction in which this happens is different
 - English: OV reserved for givenness
 - Dutch/High German/Low German: VO is reserved for newness
- Distinction between English as a “VO language” and Dutch and German as “OV languages” already there

HOW TO ANALYSE OV/VO VARIATION: THE 'SIMPLE' SOLUTION

- English

Left bracket	Middle Field	Right bracket	Postfield
Complementizer	OBJ Givenness	V	OBJ

- Dutch/German

Left bracket	Middle Field	Right bracket	Postfield
Complementizer	OBJ	V	OBJ Newness

HOW TO ANALYSE OV/VO VARIATION: THE 'SIMPLE' SOLUTION ...

- ... but maybe too simple?
- In the typological literature OV and VO are generally considered two entirely different language systems, each with their own syntactic correlates (i.e. Dryer 1992; Greenberg 1963; Lehmann 1973; Vennemann 1976, see Haider 2020 for correlates in Germanic OV and VO languages)

A COMMON PAST

- Proto-Indo-European also allowed OV/VO variation
- Presumably basic OV, with optional rightward movement to derive VO (Delbrück 1900; Hock 2015; Viti 2015, but see Pires & Thomason 2008 for a critical discussion regarding the validity of this observation)
- VO is motivated by pragmatic and prosodic factors such as length and emphasis (see for instance West 2011 on Old Avestan)
- Dutch and German as a continuation of Proto-Indo-European; English changes its basic word order
- Why did English shift from an OV language with rightward movement to a VO language with leftward movement without losing the surface variation, and when did this happen?
- The assumption of Grammar Competition (Kroch 1989) or Double Base Hypothesis (Pintzuk 1999, Taylor & Pintzuk 2012a) is conceptually undesirable, and empirically not very well motivated.

EARLY SIMILARITIES: ONE SYNTACTIC SYSTEM?

	AuxOV	OAuxV
Old English/Middle English	Frequent	Minority
Old Saxon/Old High German	Fairly frequent	Minority
Middle Low German/High German	Infrequent	Frequent
Middle Dutch	Non-existent	Frequent

þurh þa heo **sceal** **hyre** **scippend** **understandan**
through which it must its creator understand
'through which it must understand its creator'
(ÆLS_[Christmas]:157.125)

EXCURSUS

SCRAMBLING IN PRESENT-DAY DUTCH

Left bracket	Middle Field			Right bracket	Postfield
Complementizer	OBJ Given	ADV	OBJ New	V	

Verhagen 1986; Neeleman & Reinhart 1998; Broekhuis 2008; Neeleman & van de Koot 2008; but see Schoenmakers et al. 2021

dat Jan (**het boek**) waarschijnlijk (**het boek**) las.

that Jan the book probably the book read

'...that Jan probably read the book.'

EXCURSUS

SCRAMBLING IN EARLY DUTCH

Left bracket	Middle Field			Right bracket	Postfield
Complementizer	OBJ3 ?	ADV	OBJ2 ?	V	OBJ1 New

*naedat sij op ten xviii. julij **haer legher te Heyloe**_{OBJ2} opghebroken hadden*

after they on the 18 july their army at Heiloo dissolve had

'after they had broken up their army stationed at Heiloo on 18 July'

CLVN_Nanning van Foreest_1573-83

*dat diegene die dat bijer buten vueren sellen **dat teyken daeraf***_{OBJ3}

that the one who that beer outside carry will that proof thereof

in den poerten toenen sellen,

in the gates show will

'that the one who transports the beer out will show proof thereof at the gates'

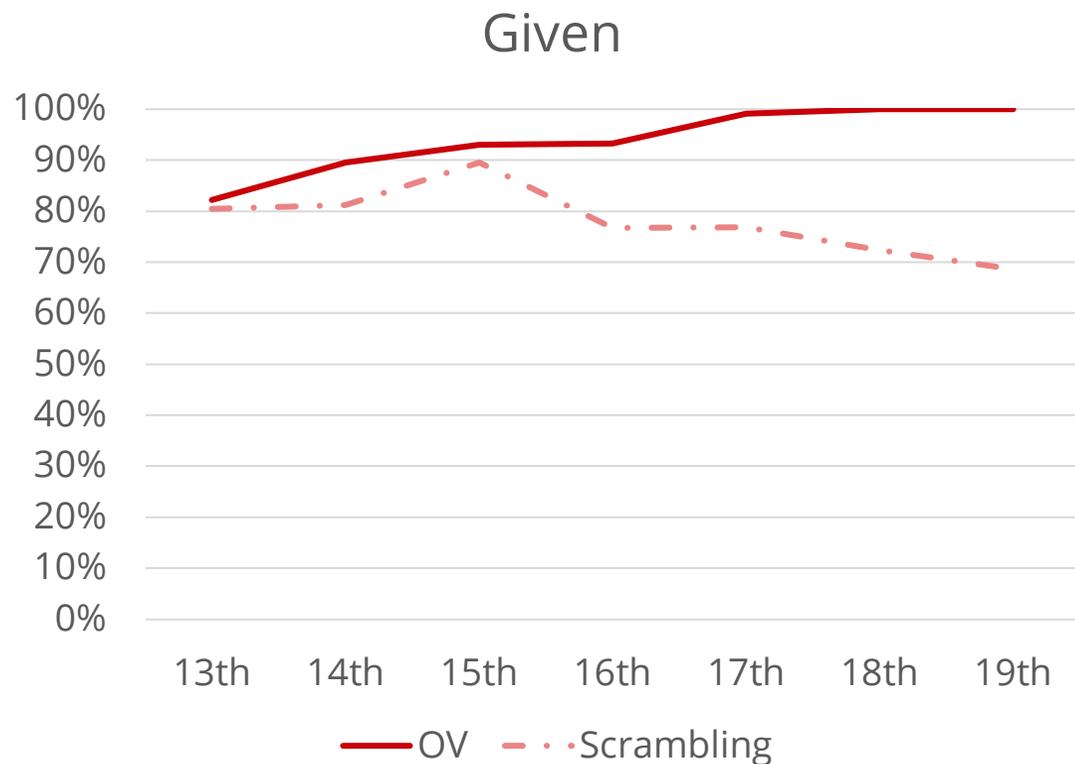
EXCURSUS

SCRAMBLING IN EARLY DUTCH

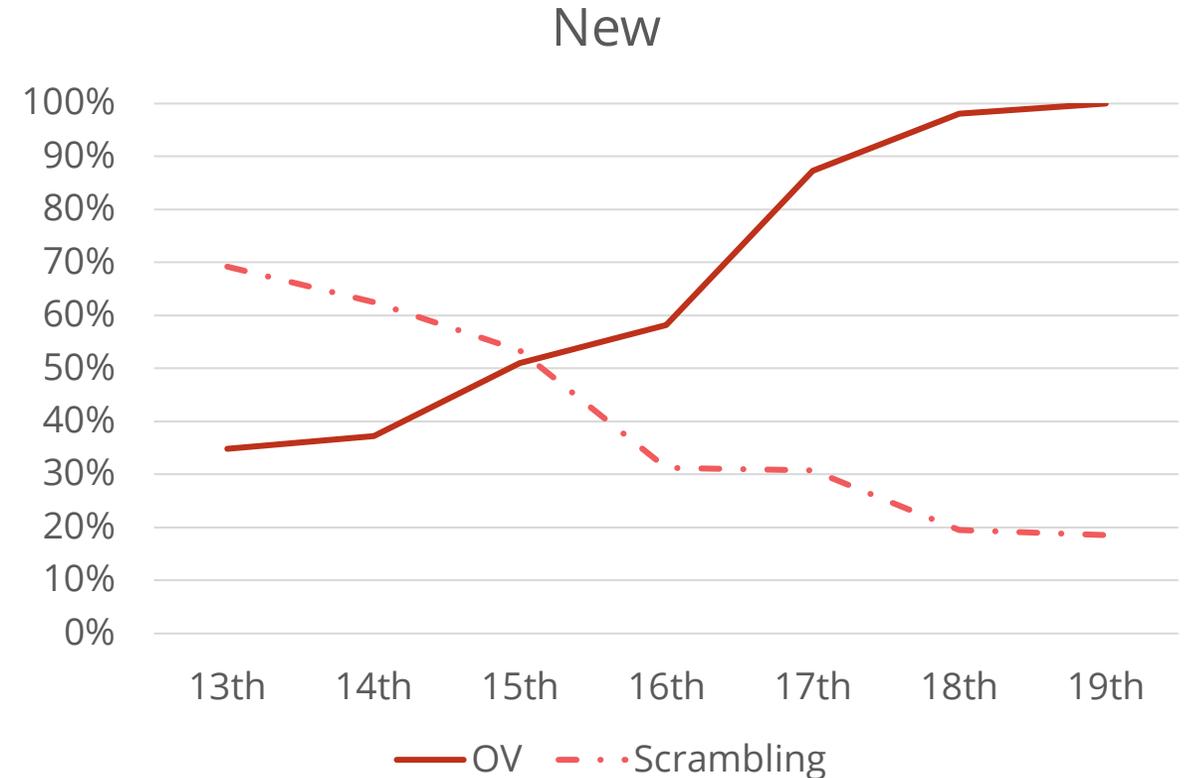
- Is there a diachronic relation between OV/VO variation, the loss of VO and scrambling as a means to mark information structure?

EXCURSUS

DIACHRONIC DEVELOPMENT OF OV/VO AND SCRAMBLING



OV/VO N = 1015
Scrambling N = 469



OV/VO N = 404
Scrambling N = 138

EXCURSUS

SCRAMBLING IN DUTCH

- The boundary between the given and new domain shifts from the verb to the adverbial in the middle field

	Middle Field			Right bracket	Postfield
Historical Dutch	OBJ₃ Given	ADV	OBJ₂	V	OBJ₁ New
Present-day Dutch	OBJ₃ Given	ADV	OBJ₂ New	V	



- If VO is rightward movement in historical Dutch, it is unclear how scrambling is motivated in historical Dutch.

TOWARDS A UNIFIED ANALYSIS

- The structural variation is **similar** in the earliest stages of English, Dutch, Low German and High German
 - OV/VO variation
 - Aux-O-V “Verb Projection Raising” allowed in all languages
 - Scrambling
- The frequency and motivation for the variation is **different**
 - New = VO in English, but given may be OV – majority of objects is VO
 - Given = OV in Dutch/German, but new may be VO – majority of objects is OV
- **Aim:** a syntactic analysis in which the variation is derived uniformly, but which is flexible enough to allow pragmatic interface effects to take place and which allows for different diachronic trajectories.

TOWARDS A UNIFIED ANALYSIS

Left bracket	Middle Field			Right bracket	Postfield
Complementizer	OBJ3	ADV	OBJ2	V	OBJ1
	Spec,vP		Spec,VP		

- Antisymmetric analysis of early West-Germanic clause structure (cf. Kayne 1994)
- All phrase structures are head-initial. English, Dutch and German are **structurally** VO.
- **Surface order** is derived by leftward movement.
- OV word order and scrambling is derived by movement to Spec,VP (OBJ₂) and subsequently Spec,vP (OBJ₁), leaving behind a copy in each intermediate position
 - Obligatory in Dutch/German
 - (Becomes) restricted in English

TOWARDS A UNIFIED ANALYSIS: INTEGRATION OF INFORMATION STRUCTURE

- In early English, only given objects move to OBJ₃
- Result of morphosyntactic properties/phi-features on the object that facilitate reference:

CASE	MASCULINE	NEUTER	FEMININE	PLURAL
NOMINATIVE	se	þæt	sēo	
ACCUSATIVE	þone	þæt	þā	þā (þe)
GENITIVE	þæs	þāra	þære	
DATIVE	þæm þām	þæm, þām	þære	

Table 6. Old English demonstratives

- Grammaticalization of the definite determiner means loss of case and gender marking and hence of syntactically visible referential marking
- Result: loss of VO

TOWARDS A UNIFIED ANALYSIS: INTEGRATION OF INFORMATION STRUCTURE

- In early Dutch and German, all objects obligatorily move to the highest object position
- Phonetics and semantics interfaces subtract possible spell-out positions (cf. Struckmeier 2017)
 - a. Cécile heeft waarschijnlijk een roos geplant.
Cécile has probably a rose planted
'Cécile probably planted a(ny) rose.'
 - b. Cécile heeft een roos waarschijnlijk geplant.
Cécile has a rose probably planted
'Cécile probably planted a (certain) rose.'

TOWARDS A UNIFIED ANALYSIS: INTEGRATION OF INFORMATION STRUCTURE

- In early Dutch and German, all objects obligatorily move to the highest object position
- Phonetics and semantics interfaces subtract possible spell-out positions (cf. Struckmeier 2017)
- Remaining optionality is governed by pragmatic preferences, such as given-before-new and the end-weight principle
- Information structure not a cue for *differential movement*, but for *differential pronunciation*
- Spell Out of the object in OBJ1 (VO) becomes impossible (due to a complex of changes)

TOWARDS A UNIFIED ANALYSIS

Left bracket	Middle Field			Right bracket	Postfield
Complementizer	OBJ3	ADV	OBJ2	V	OBJ1
	Spec,vP		Spec,VP		

- Uniform syntactic derivation of OV/VO variation in languages which are closely related
- Allows for a coherent account of diachronic changes in individual languages, which may lead to different outcomes

COMPARATIVE LAB

Are these languages
the same and
divergent?

OV/VO
Variation

Are these languages
different and further
divergent?



Dutch OV

....dat hij **het boek** gelezen heeft

German OV

....dass er **das Buch** gelesen hat

English VO

....that he **has read** the book

THANK YOU!

DERIVING OLD/MIDDLE ENGLISH WORD ORDER

- Variation in the history of English can be analysed from a VO base (see also Fischer et. al, 2000 Biberauer & Roberts 2005 et seq.; Wallenberg 2009; de Bastiani 2019)

Struik & van Kemenade (in press):

- Optional pied-piping/DP movement of the object to spec, νP_{emb}
- attracted by a “picky” EPP/Edge-feature. This feature only attracts “big-DPs”, i.e. objects which have an additional (φ -related) feature layer to signal referentiality (cf. Biberauer & van Kemenade 2012)
 - Only given objects undergo movement, which results in OV
- Optional pied-piping of the object to Spec, νP_{mat}
- Strategy to defocus entire clause (work in progress with Chiara de Bastiani, cf. also Milicev 2016), which results in V-Aux
- Phase-based: object must move to spec, νP_{emb} in order to be raised to νP_{mat}
 - Postverbal spellout of unraised objects (due to the PIC) (Biberauer & Roberts 2005 et seq.)

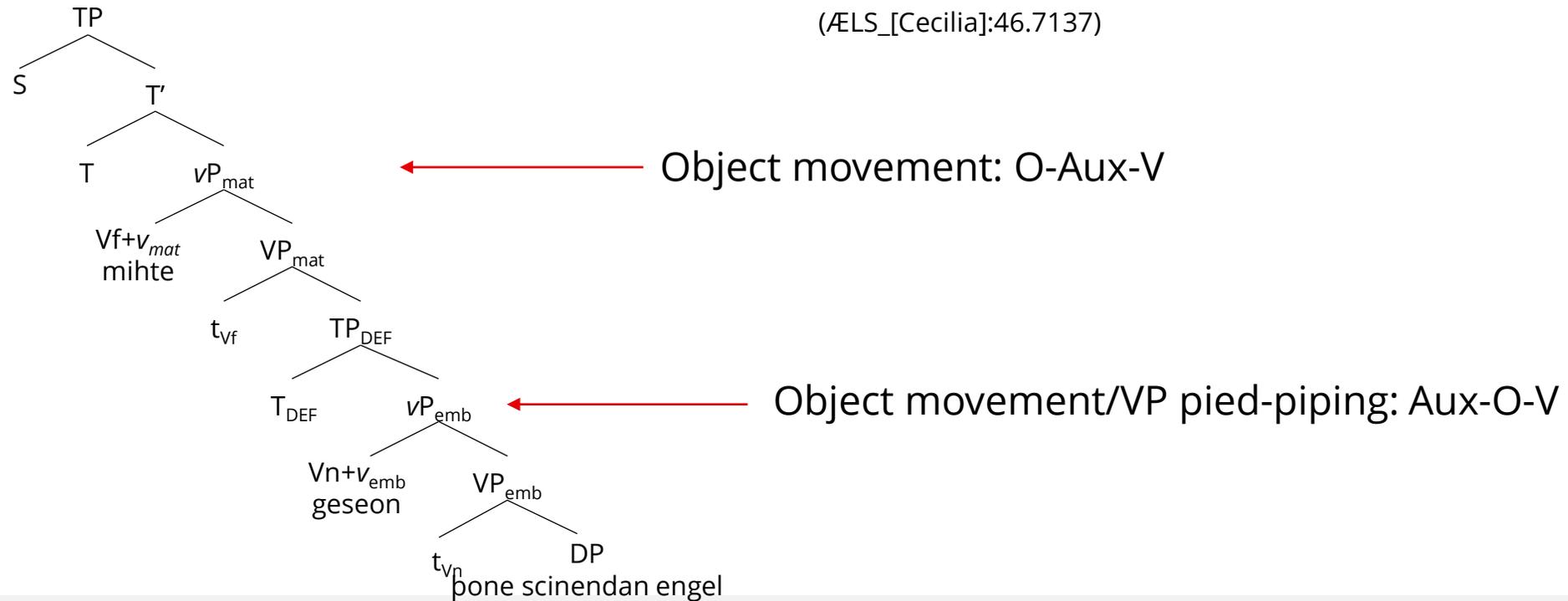
DERIVATION OF OLD/MIDDLE ENGLISH WORD ORDER

Aux-V-O

þæt ic mihte geseon þone scinendan engel
that I might see that shining angel

'that I might see the shining angel'

(ÆLS_[Cecilia]:46.7137)



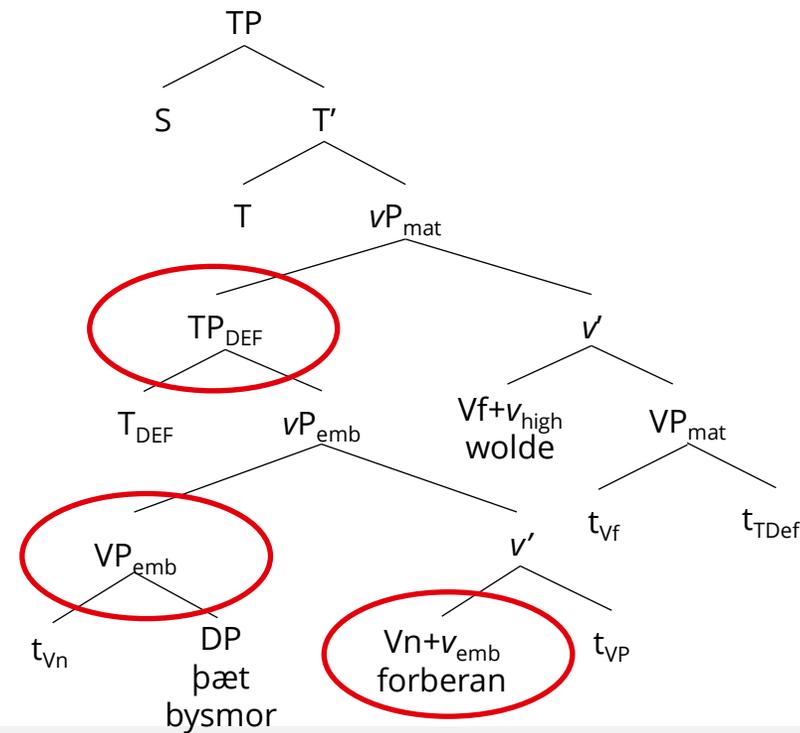
DERIVATION OF OLD ENGLISH WORD ORDER

O-V-Aux

gif heo **þæt** **bysmor** forberan wolde
if she that disgrace tolerate would

'if she would tolerate that disgrace'

(ÆLS_[Eugenia]:185.305)

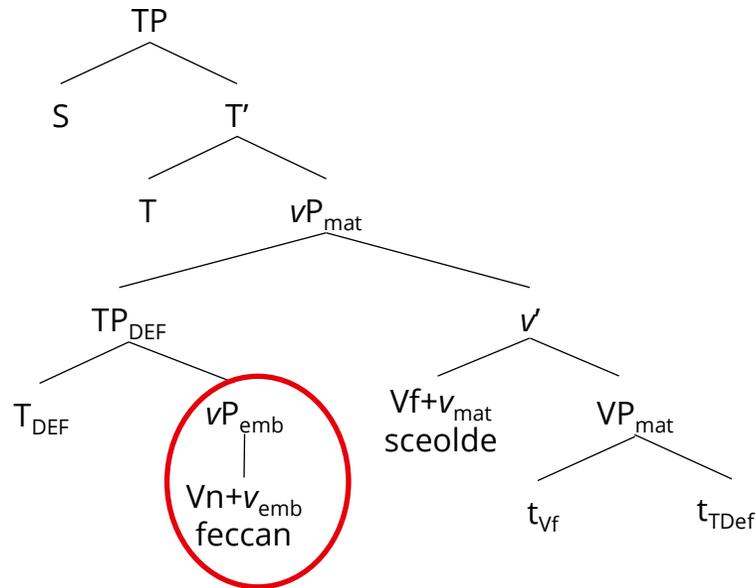


DERIVATION OF OLD ENGLISH WORD ORDER

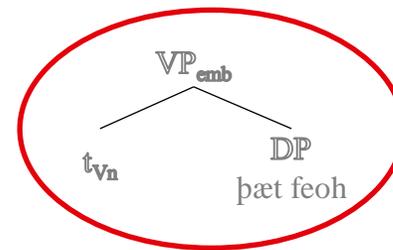
V-Aux-O

þæt he feccan sceolde þæt feoh mid reaflice
 that he *fetch* *should* *the* *goods* *with* *robbery*
 'that he should steal the goods'

(ÆLS_[Maccabees]:760.5327)



Spell-out



SOME FURTHER DIFFERENCES IN CLAUSES WITH 2 VERBS

	AuxOV	OAuxV
Old English/Middle English	Frequent	Minority
Old Saxon/Old High German	Fairly frequent	Minority
Middle Low German/High German	Infrequent	Frequent
Middle Dutch	Non-existent	Frequent

→ Two-verb clauses demonstrate (incipient) transparency effects in Dutch and German, which are not observed in English

DERIVING CONTINENTAL WEST-GERMANIC

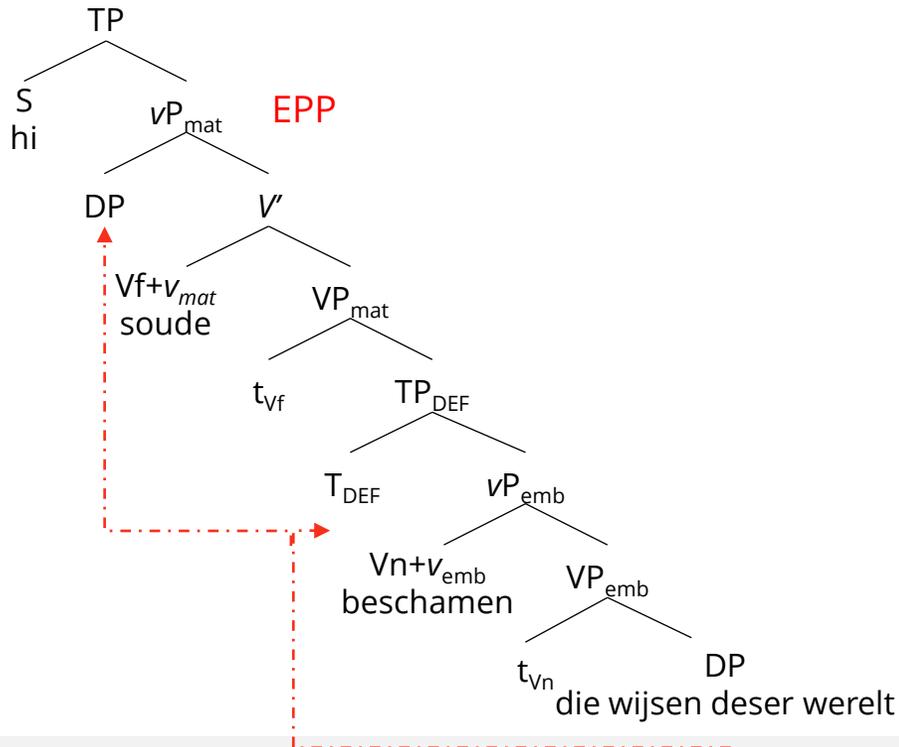
- EPP feature is not as selective as in Old/Middle English; it attracts all objects
- EPP feature can be satisfied by optional pied-piping
- Variation in object position (the information structure effect) is the result of differential Spell-Out of copies of the object (Struik & Schoenmakers subm.)
- vP_{emb} loses possibility to license the object (cf. Ter Beek 2008)
- Licensing duties shift to vP_{mat} : transparency effect

DERIVING CONTINENTAL WEST-GERMANIC

Vf-Vn-O

<i>op</i>	<i>dat</i>	<i>hi</i>	<i>soude</i>	<i>beschamen</i>	<i>die</i>	<i>wijzen</i>	<i>deser</i>	<i>werelt</i>	
so	that	he	would	shame	that	wise	this.GEN	world	
'so	that	he	would	shame	the	wise	of	this	world'

Peerle_1537-38



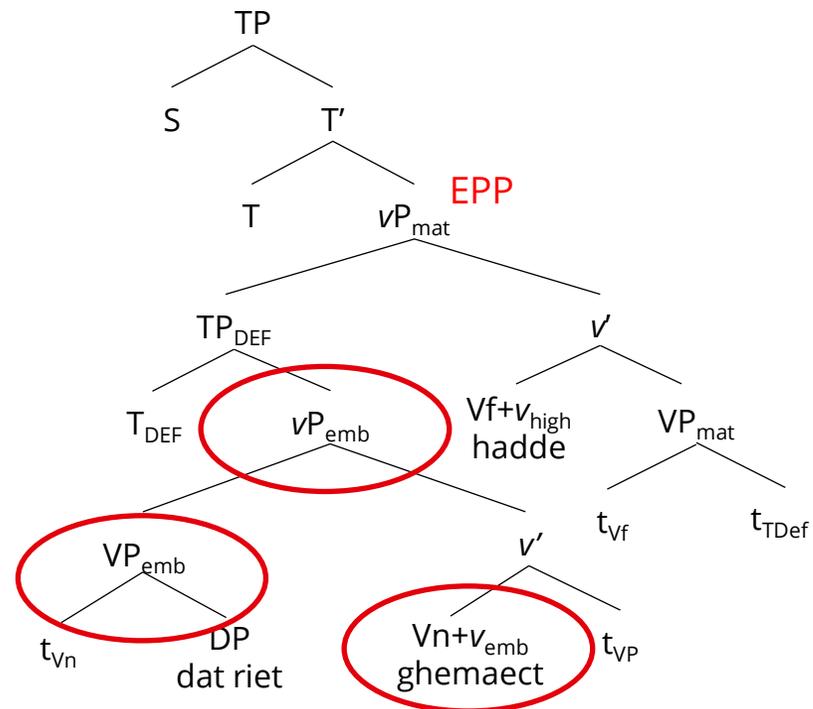
Object movement to spec, vP_{mat},
but low Spell-out

DERIVING CONTINENTAL WEST-GERMANIC

O-Vn-Vf

<i>Die</i>	<i>dat</i>	<i>riet</i>	<i>ghemaect</i>	<i>hadde</i>
who	that	reed	made	had
'who	made	that	reed'	

Gysseling_1340_1294



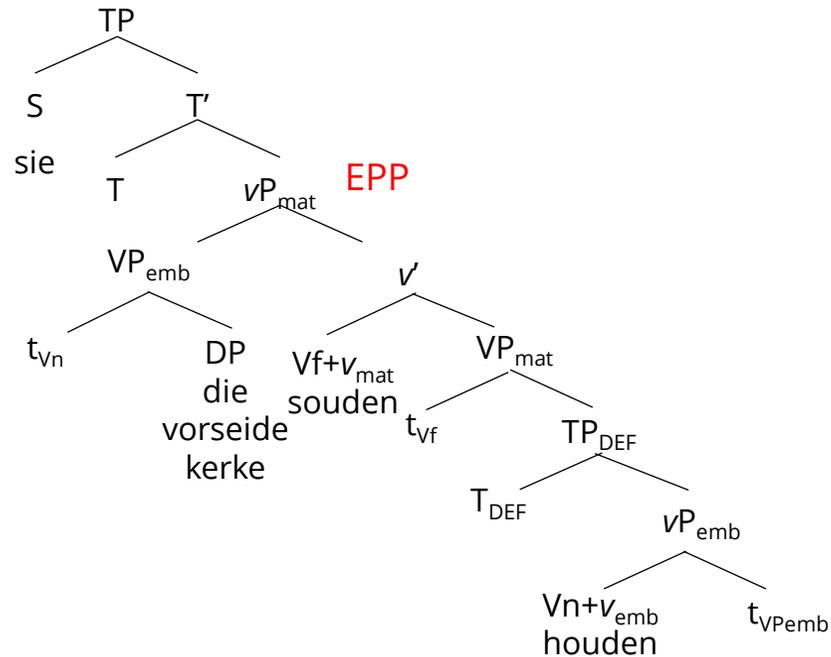
DERIVING CONTINENTAL WEST-GERMANIC

O-Vf-Vn

<i>dat</i>	<i>sie</i>	<i>die</i>	<i>vorseide</i>	<i>kerke</i>	<i>daer</i>	<i>scadeloes</i>		<i>ende</i>	<i>vri</i>	<i>souden</i>	<i>houden</i>
that	they	that	aforementioned	church	there	without	damage	and	free	would	keep

'that they would indemnify the aforementioned church'

Gys_0681_1286



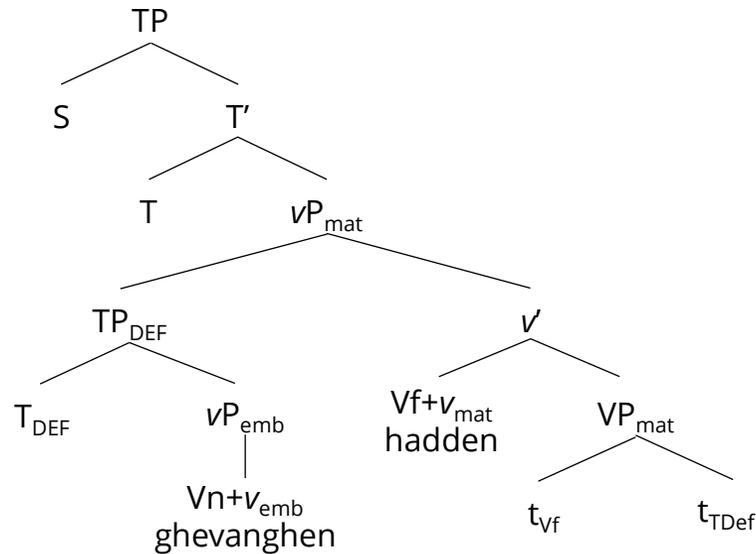
DERIVING CONTINENTAL WEST-GERMANIC

V-Aux-O

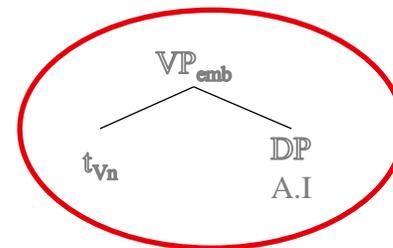
<i>Dat</i>	<i>Ferry</i>	<i>Bertram</i>	<i>...</i>	<i>ghevanghen</i>	<i>hadden</i>	<i>Anthuenis</i>	<i>Inffroot,</i>	<i>poortere</i>	<i>der</i>	<i>voorsejde</i>	<i>stede</i>
that	F.	B.	...	captured	had	A.	I.	citizen	the.GEN	aforementioned	city

'that F. B. ...captured A. I., citizen of the aforementioned city.'

CLVN_Brugge_1510-1520



Spell-out



DIACHRONIC DEVELOPMENT

- English: towards no movement
 - Information structural trigger for V-Aux is lost → strict Aux-V order
 - DP objects lose additional feature layer (possibly related to the grammaticalisation of the determiner) → strict VO order

- Dutch/German: towards obligatory movement
 - EPP feature attracts all objects
 - Licensing of objects becomes strictly local and spell-out of VO objects is lost
 - Objects must be licensed in a position above the finite verb → transparency effects and strict OV
 - Pied-piping remains possible → VAux – AuxV alternation (without apparent semantic/pragmatic effect)