Detecting Grammatical Properties in Usage Data

Anthony Kroch & Beatrice Santorini
University of Pennsylvania
March 2017

The shift from Infl-final to Infl-medial word order in the history of Yiddish (Santorini 1993)

Unambiguous I-final cases:

(1) ven der vatr   nurt doyts   leyen kan
     if   the father only German read can

(2) ven du mir meyn kop ab shneydst
     if you me   my head off cut

Unambiguous I-medial cases:

(3) ven der vatr kan nurt doyts leyen

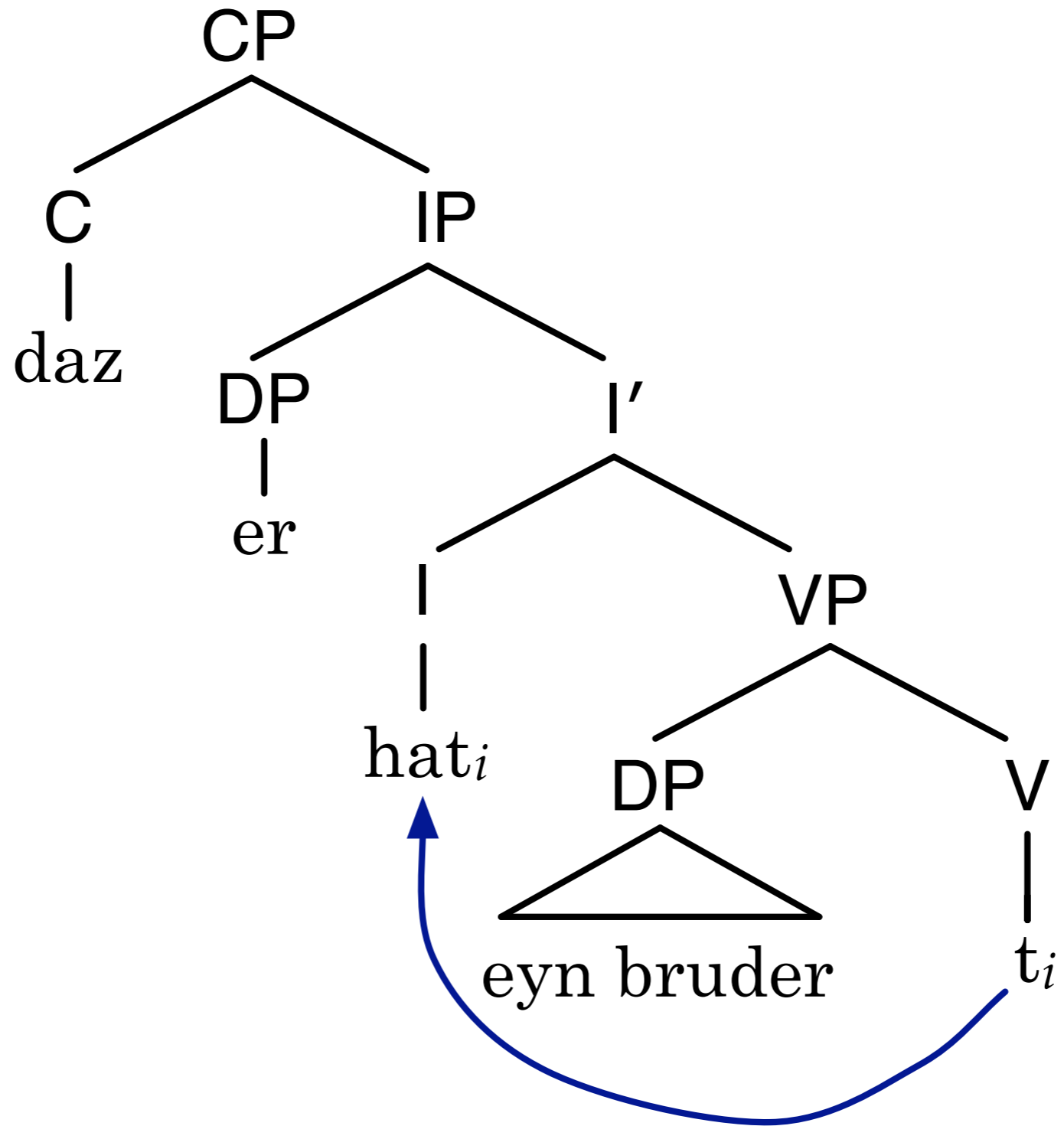(4) ven du shneydst mir meyn kop ab

Extraposition in Yiddish

(5) ven der vatr nurt doyts leyen kan
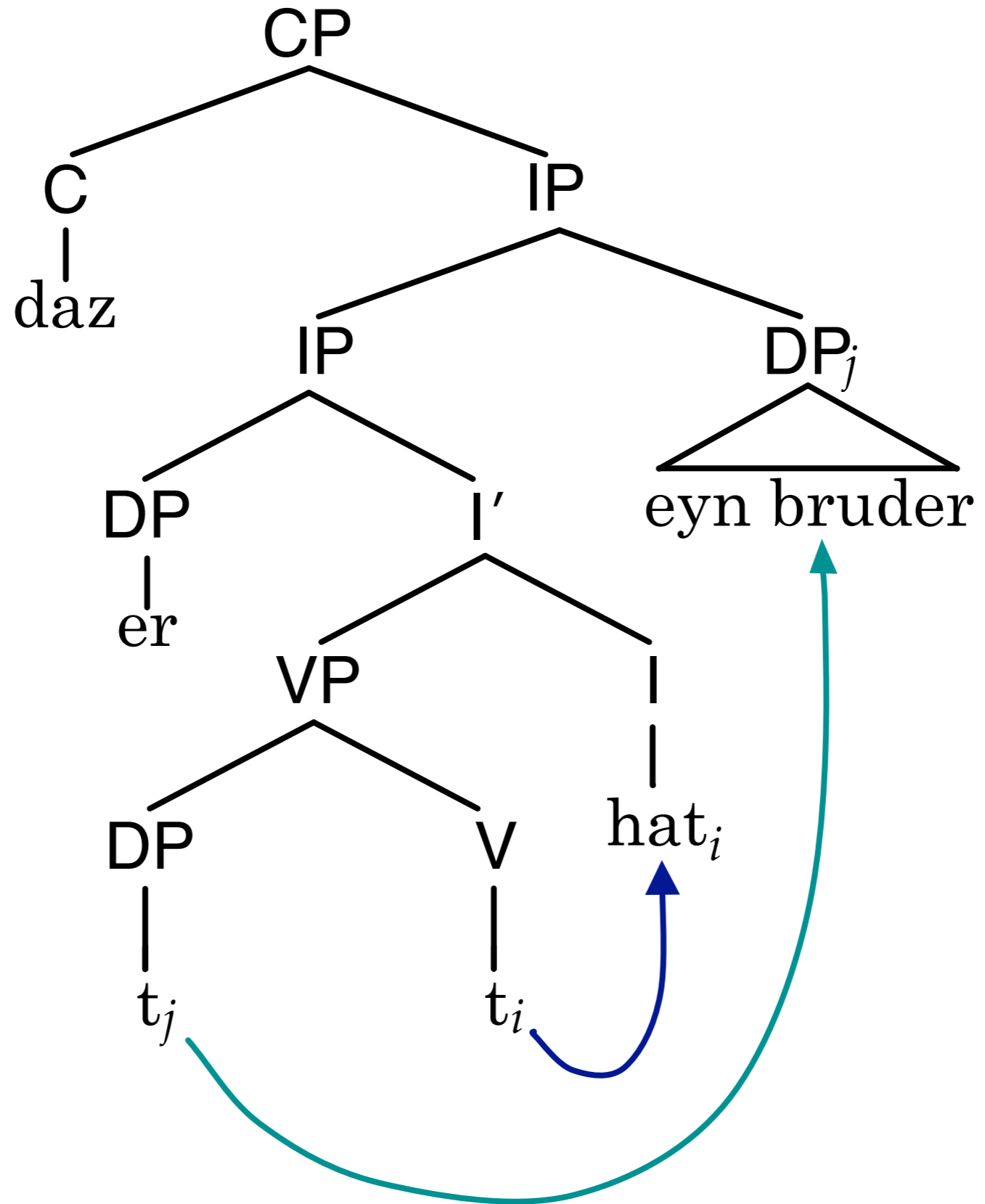
(6) ven der vatr leyen kan nurt doyts

Ambiguity between I-medial and I-final structure

(7) daz   er hat eyn   brudr
     that he had   a   brother

(8) da   ishue kam   in   arts isral
   when Joshua came into   the land of Israel

CP
C — daz
IP
IP
DP — er
I′
VP
DP — t_j
V — t_i
I — hat_i
DP$_j$ — eyn bruder

Frequency of Infl-medial versus Infl-final word order in the history of Yiddish, unambiguous cases

| Date	Simple verb case			Complex verb case		
	INFL-medial	INFL-final	freq.	INFL-medial	INFL-final	freq.
1400-1489	0	27	.00	1	15	.06
1490-1539	5	37	.12	2	35	.05
1540-1589	13	59	.18	4	51	.07
1590-1639	5	81	.06	7	57	.11
1640-1689	13	33	.28	18	44	.29
1690-1739	15	20	.43	25	28	.47
1740-1789	1	1	.50	11	17	.39
1790-1839	54	3	.95	79	0	1.00
1840-1950	90	0	1.00	62	0	1.00

Frequency of Infl-medial versus Infl-final word order
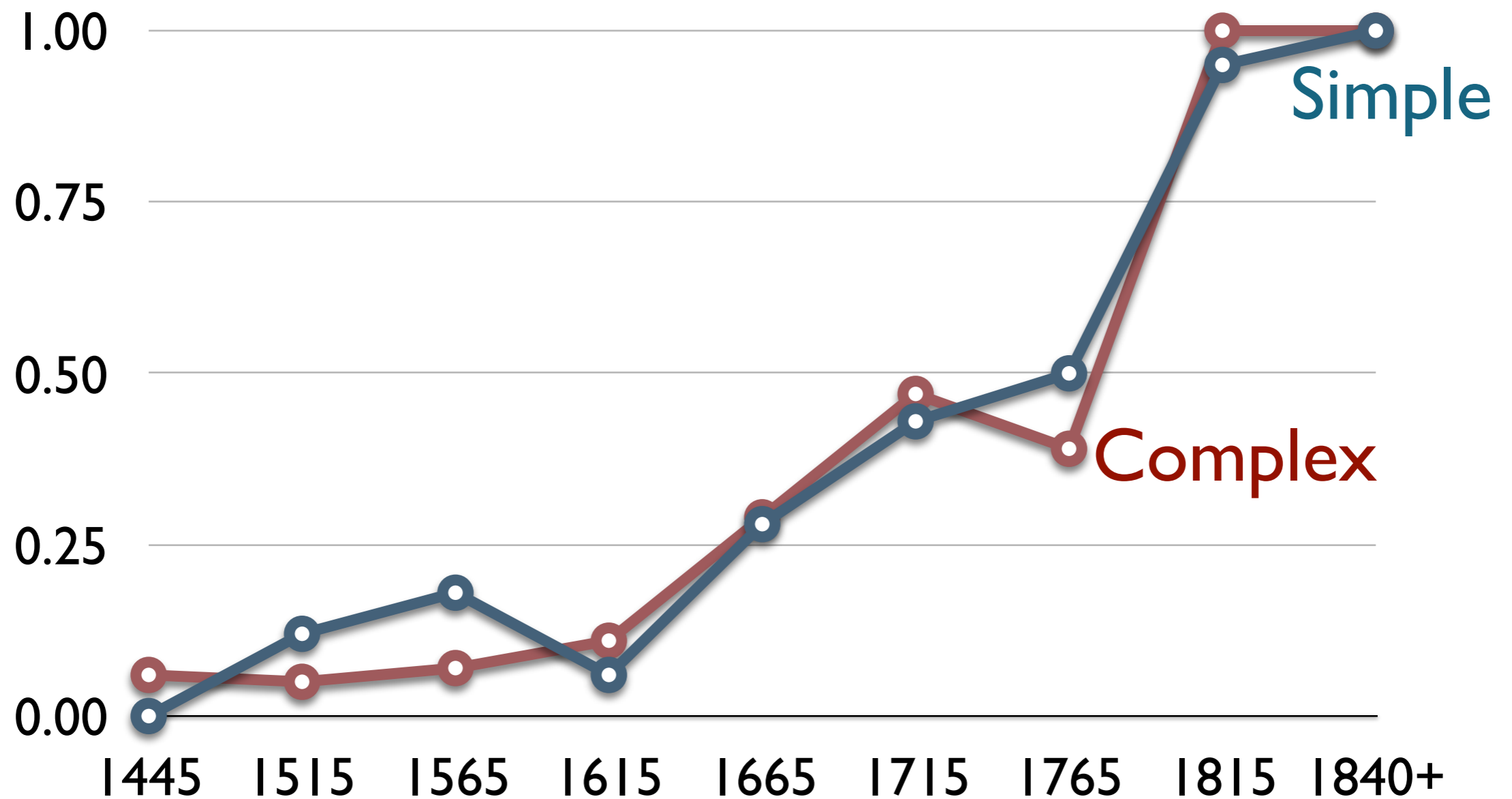
Frequency of Infl-medial versus Infl-final word order in the history of Yiddish, ambiguous case

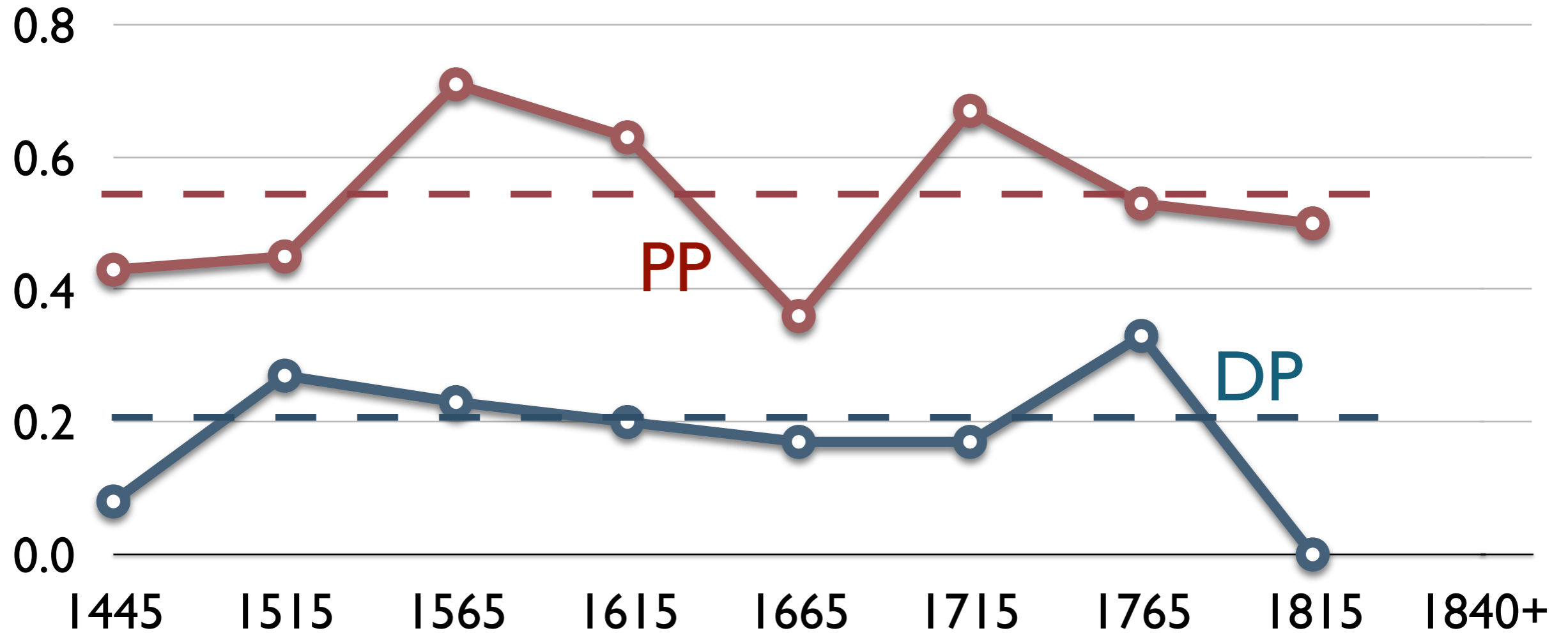| Date	Ambiguous case		
	Possibly INFL-medial	INFL-final	frequency
1400-1489	3	9	.25
1490-1539	13	13	.50
1540-1589	58	39	.60
1590-1639	41	29	.59
1640-1689	32	21	.60
1690-1739	21	11	.66
1740-1789	5	2	.71
1790-1839	58	3	.95
1840-1950	69	0	1.00

Rise of I-medial word order, II
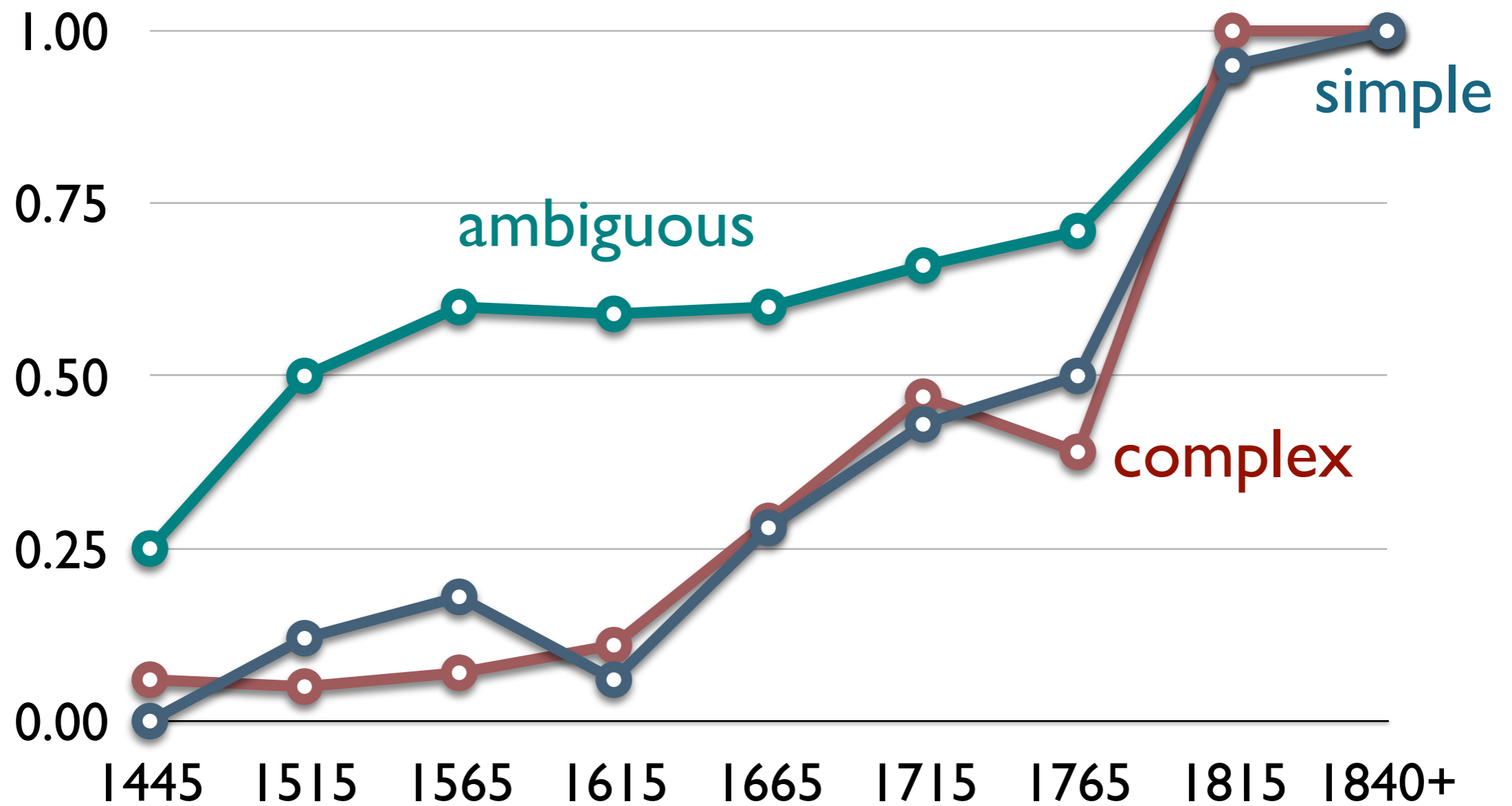
Frequency of DP and PP postposing
in the history of Yiddish (Santorini 1993)

Date	DP postposing			PP postposing		
	Postposed	Not postposed	freq.	Postposed	Not postposed	freq.
1400-1489	1	12	**.08**	9	12	**.43**
1490-1539	7	19	**.27**	13	16	**.45**
1540-1589	7	24	**.23**	52	21	**.71**
1590-1639	10	40	**.20**	39	23	**.63**
1640-1689	4	19	**.17**	17	30	**.36**
1690-1739	1	5	**.17**	6	3	**.67**
1740-1789	1	2	**.33**	8	7	**.53**
1790-1839	0	1	**.00**	1	1	**.50**
1840-1950	no INFL-final data		–	no INFL-final data		–

Stability of DP and PP postposing

mean rate of PP postposing = .56
mean rate of DP postposing = .20

Correcting for postposing in ambiguous cases

For each time period, let
 A = # surface ambiguous I-medial cases
 F = # surface I-final cases that would be
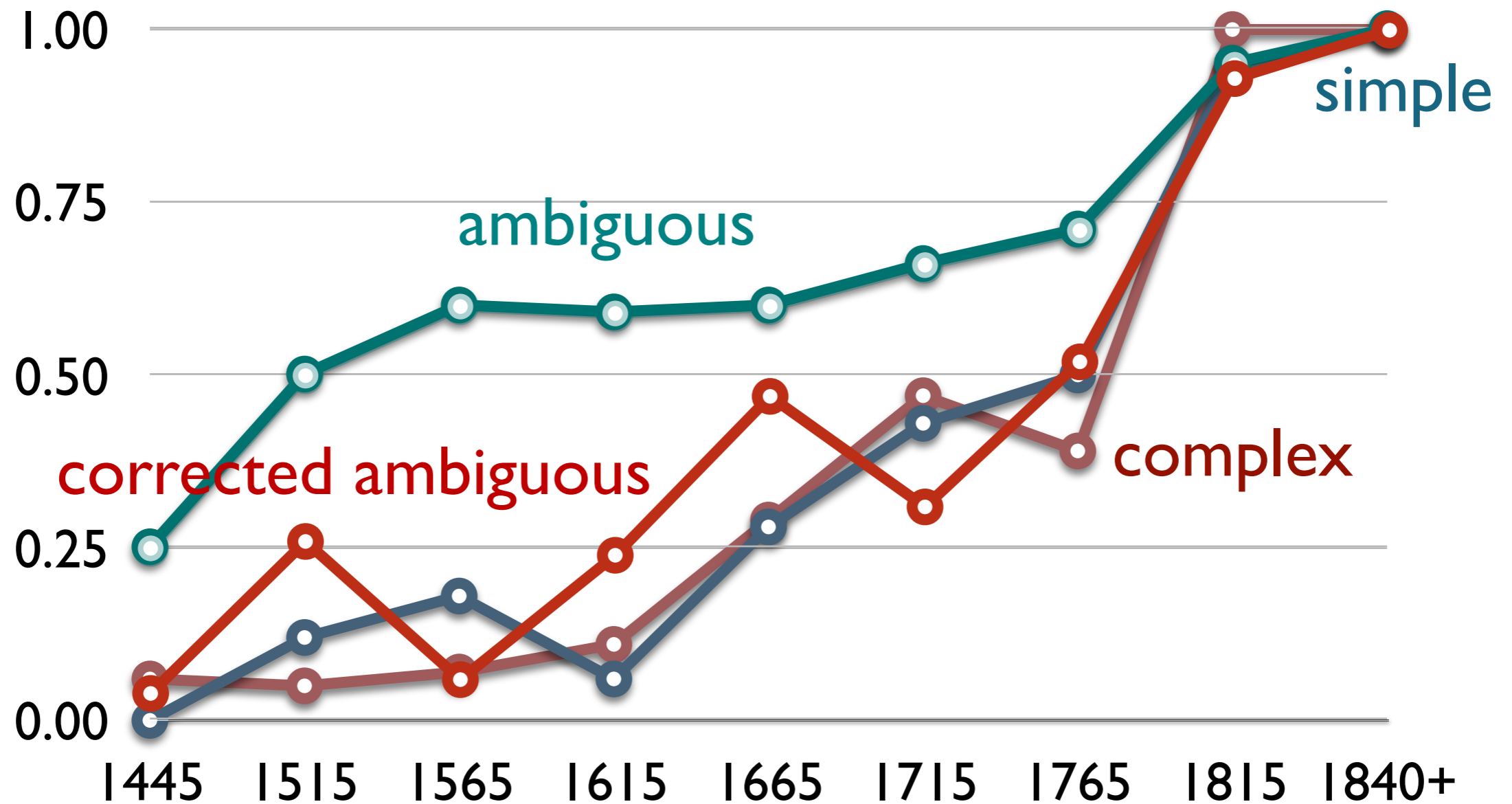   ambiguous if medial
 p = rate of postposing

Then the estimated true # of I-medial cases in A is

 Estimated I-medial = A−F•p

This calculation is done separately for sentences with postverbal DPs and PPs, with the results then summed.

Rise of I-medial word order, III

Detecting stages in the transition from OV to VO in four languages: English, French, Yiddish and Icelandic

Data sources: English

- Anthony Kroch and Ann Taylor. 2000. Penn-Helsinki Parsed Corpus of Middle English, second edition (PPCME2).

Data sources: French

- France Martineau et al. 2005. Corpus du projet Modéliser le changement: les voies du français (MCVF).

- Anthony Kroch and Beatrice Santorini. 2016. Penn supplement to the MCVF corpus.

- Alexei Lavrentiev, Christiane Marchello-Nizia, Céline Guillot and Serge Heiden. 2014. BFM – Base de Français Médiéval [En ligne].

Data sources: Yiddish

- Beatrice Santorini. 2008. Penn Yiddish Corpus.

Data sources: Icelandic

- Joel C. Wallenberg, Anton Karl Ingason, Einar Freyr Sigurðsson and Eiríkur Rögnvaldsson. 2011. Icelandic Parsed Historical Corpus (IcePaHC).

Preliminaries

- Only non-finite VPs are considered to avoid interference from V-to-C and V-to-T.

- Modals are treated as auxiliary verbs in all four languages.

- Sentences in which the direct object moves further left than T are also excluded since the "in situ" position is not recoverable.

English

Evidence for VO word order in Early Middle English

(1) oðet he habe iȝetted ou al þet ȝe wulleð
    'until he has granted you all that you want'
                    (CMANCRIW,I.68.229)

(2) þt he schulde in huden him ȝef he walde libben
    'that he should hide himself if he would live'
                    (CMANCRIW,II.132.1744)

More evidence for VO word order in Early Middle English

(1) hwaso mei gan in
    'whoever may go in'
                    (CMANCRIW,II.60.5)

(2) ha wes sone ibroht forð
    'she was soon brought forth'
                    (CMKATHE, B.827)

More evidence for VO word order in Early Middle English

(1)   worþy mennes sones þat sche myȝte han be maried to

'worthy men's sons that she might have been married to'

(CMAELR3-M23,33.189)

(2)   þe terme, þe which hij ne shul nouȝt passe over

'the limit which he should not pass over'

(CMEARLPS-M2,125.5471)

Possible evidence for OV word order in Early Middle English

(1) þeos ne schulen neaver song singen song

'these should never sing songs'
(CMHALI,142.222)

(2) þat ne have noht here sinnes forleten here sinnes

'who have not their sinnes forsaken.'
(CMTRINIT,67.934)

More evidence for OV word order in Early Middle English

(1) al þe blodi sunnen þet ha is wið iwundet
     'all the bloody sins that she is wounded with'
                              (CMANCRIW,I.62.202)

(2) sumþing þet god maȝe of arisen
     'something that good may arise from'
                              (CMANCRIW,I.74.296)

Two-argument VPs: OOV word order

(1) Ne durste nauere gume nan oðerne ufele igreten
    'Nor did a man ever dare to afflict evil on another'
                    (1200-BRUT,564.1322)

(2) þatt icc have ʒuw summ del nu spelledd offe
    'which I have now told you something of '
                    (CMORM-M1,1,221.1820)

Two-argument VPs: OVO word order

(1) For all þeo the habbeð any good idon me

'For all those who have done me any good'

(CMANCRIW,I.64.212)

(2) I sal yu lere þe dute of god

'I shall teach you the fear of God'

(CMBENRUL-M3,2.20)

Two-argument VPs: VOO word order

(1) oðet he habe iȝetted ou al þet ȝe wulleð
     'until he has granted you all that you want'
                              (CMANCRIW,I.68.229)

(2) and wile ȝelden eche men his mede efter his werke
    'and will pay each man his reward by his work'
                              (CMLAMB1-M1,143.310)

Distribution of Full DP Objects in Double Object Clauses in Early Middle English (<1420)

	IO>V	V>IO	rate of IO scrambling
DO>V	1	1	
V>DO	2	31	0.06
rate of DO scrambling		0.03	N=35

Chi-square: .731 (ns)

Expected rate of OOV based on rates of IO and DO scrambling	.03 ✕ .06 = **0.002**
Actual rate of OOV	1/35 = **0.03**

Distribution of All Objects in Double Object Clauses in Early Middle English (<1420)

	IO>V	V>IO	rate of IO scrambling
DO>V	20	11	
V>DO	30	260	0.1034
rate of DO scrambling		0.037	N=321

Chi-square: 62.498

Expected rate of OOV based on rates of IO and DO scrambling	.037 **x** .1034 = **0.0038**
Actual rate of OOV	20/321 = **0.06**

French

VO & OV word order: modal + infinitive

(1) Je veul avoir mon loier
     'I want to have my pay.'
              (127X-CASSIDORUS-P,164.1546)

(2) Kar ne poeit le jur choisir le jur
                              ?
     'For he cannot choose the day.'
              (116X-MARIE-DE-FRANCE-R,111.2262)

VO & OV word order: *avoir* + participle

(1) Rollant ad mis l' olifan a sa buche
'Roland raised the ivory horn to his mouth.'
(1100-ROLAND-V,133.1772)

(2) Li reis Marsilie out sun cunseill finet sun cunseill

?

'King Marsilla had adjourned his council.'
(1100-ROLAND-V,5.53)

Two-argument VPs: OOV word order

(1) Or ad Deus saint Thomas cel' ampole donee
'Now God gave Saint Thomas this phial'
(1173-becket-p-bfm,182.14984)

(2) ainsi pourroit Grace a Dieu querre
'In this way, he could ask God for grace'
(1190-BORON-R-PENN,7.88)

Two-argument VPs: OVO word order

(1) Tu auoiz dous choses amises al creator
'You had presented two things to the creator'
(1190-SBERNAN-P-BFM,10.325)

(2) Ancor uolt plus grant honor faire a nostre lum
'He wished to do our man an even great honor'
(1190-SBERNAN-P-BFM,37.1192)

Two-argument VPs: VOO word order

(1) Et Pilates a douné le cors Joseph
    'and Pilate gave the body to Joseph'
                    (1210-BORON-P-PENN,24.230)

(2) É Deu ad dune le regne a Absalon tun fils
    'and God has given the kingdom
        to your son Absalom'
            (1150-QUATRELIVRE-P-PENN,88.3317)

Distribution of Objects in Double Object Clauses
in Early Old French (<1260)

	IO>V	V>IO	rate of IO scrambling
DO>V	11	6	
V>DO	9	55	0.14
rate of DO scrambling		0.10	N=81

Chi-square: 18.52

Expected rate of OOV based on rates of IO and DO scrambling	.14 **x** .10 = **0.014**
Actual rate of OOV	11/81 = **0.14**

Distribution of Objects in Double Object Clauses
in late Old French (<1460)

	IO>V	V>IO	rate of IO scrambling
DO>V	2	17	
V>DO	31	176	0.15
rate of DO scrambling		0.09	N=226

Chi-square: .276

Expected rate of OOV based on rates of IO and DO scrambling	.15 x .09 = **0.013**
Actual rate of OOV	2/226 = **0.01**

Yiddish

VO & OV word order: modal+infinitive

(1) da velin mir vermisin di khasene
     'Then we will ruin the wedding.'
                    (1615E-COURT,108.80)

(2) …ver nur kan zayn gezind farshiken zayn gezind
                                              ?
     'who ever can send away his servants'
                    (1619W-LETTERS,.16)

VO & OV word order: *avoir*+participle

(1) …vau min hat fergebin unzi zind
'where they have forgiven our sins'
(1704E-ELLUSH,.16)

(2) di hbn eyn yudn drmurt eyn yudn
?
'They murdered a Jew.'
(1465W-COURT,16.67)

Two argument VPs: OOV word order

(1) ikh hab den isral eyn tubh gtan
     'I have done the Israelites a good turn'
                              (1579E-SHIR,10.60)

(2) un mustn imrdarn dem mtsraim ir fikh hitn
      'and always had to guard the animals
         for the Egyptians'
                              (1589E-ESTER,7.123)

Two argument VPs: OVO word order

(1) sukhr habn unzri bridr gigebn fil gelt
   'Merchants gave our brothers much money'
   (1692E-VILNA,217.134)

(2) drum hat er dem menshn gebn di turh …
   'therefore has he the people given the Torah'
   (1620E-LEVTOV1,4l.47)

Two argument VPs: VOO word order

(1) hat gibrakht meyn oybrstn alirley shpetsirey
    '[who] brought my boss all kinds of spices'
                                (1665W-COURT,221.246)

(2) mer haben unzer formuner gegeben meinem
    stieffater tsvay hundert gulden
    'but our guardians gave my stepfather 200 guilders'
                                (1518W-GOETZ,.137)

Distribution of Objects in Double Object Clauses in early East Yiddish (<1800)

	IO>V	V>IO	rate of IO scrambling
DO>V	24	4	
V>DO	5	3	0.62
rate of DO scrambling		0.57	N=36

Chi-square: 2.14

Expected rate of OOV based on rates of IO and DO scrambling	.57 x .62 = **0.357**
Actual rate of OOV	24/36 = **0.667**

Distribution of Objects in Double Object Clauses
in pre-contemporary East Yiddish (<1900)

	IO>V	V>IO	rate of IO scrambling
DO>V	10	1	
V>DO	3	8	0.27
rate of DO scrambling		0.11	N=22

Chi-square: 9.1

Expected rate of OOV based on rates of IO and DO scrambling	.11 x .27 = **0.030**
Actual rate of OOV	10/22 = **0.454**

Distribution of Objects in Double Object Clauses in contemporary East Yiddish (>1900)

	IO>V	V>IO	rate of IO scrambling
DO>V	2	6	
V>DO	6	25	0.19
rate of DO scrambling		0.19	N=39

Chi-square: .124

Expected rate of OOV based on rates of IO and DO scrambling	.19 x .19 = **0.037**
Actual rate of OOV	2/39 = **0.051**

Icelandic

Distribution of Objects in Double Object Clauses in pre-contemporary Icelandic (<1900)

	IO>V	V>IO	rate of IO scrambling
DO>V	41	27	
V>DO	8	192	0.04
rate of DO scrambling		0.12	N=268

Chi-square: 107.6

Expected rate of OOV based on rates of IO and DO scrambling	.12 **x** .04 = **0.005**
Actual rate of OOV	41/268 = **0.152**

Distribution of Objects in Double Object Clauses
in contemporary Icelandic (>1900)

	IO>V	V>IO	rate of IO scrambling
DO>V	0	3	
V>DO	2	47	0.04
rate of DO scrambling		0.06	N=52

Chi-square: .127

Expected rate of OOV based on rates of IO and DO scrambling	.06 **x** .04 = **0.002**
Actual rate of OOV	0/52 = **0.000**

Fronted VPs after the subject in OF

(1) car nous riens feire ne devons
     'for we must do nothing'
                    (1190-BORON-R-PENN,100.1575)

(2) li rois le cors mener an puet
     'the king can take the body from there'
                    (1170-YVAIN-R,81.2816)

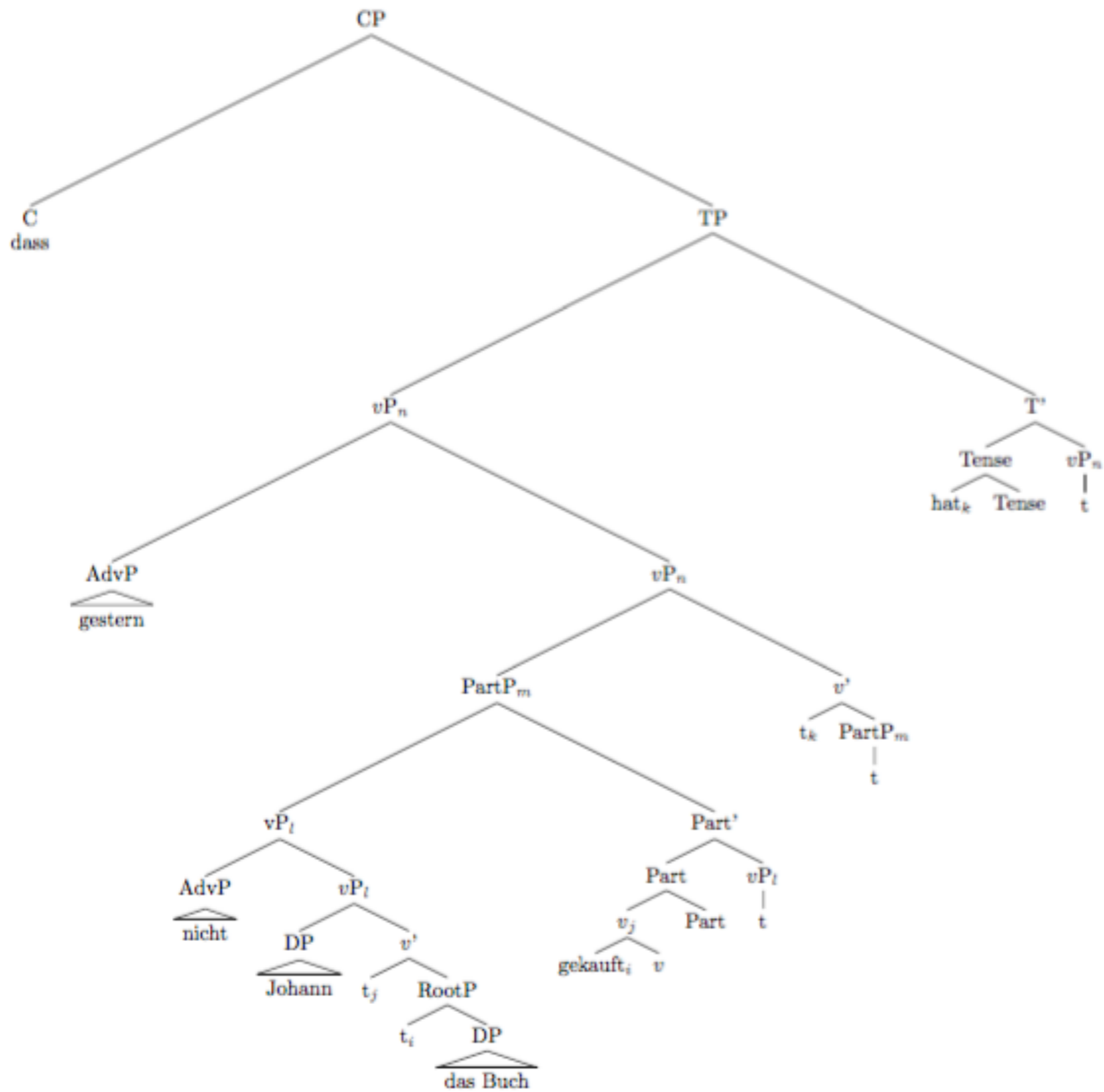# Fronted VPs before the subject in OF

(1) grant demi pied mesurer i pout hom
'one could measure a full half foot there'
(1100-ROLAND-V,94.1189)

(2) seignur servir bien deit l'um tel
'one should serve such a lord well'
(1120-BRENDAN-R,55.665)

A roll-up derivation of OV word order

Biberauer, Theresa. 2003. Reconsidering the EPP and Spec-TP in Germanic. In Astruc, Luisa & Marc Richards, eds., Cambridge Occasional Papers in Linguistics, 1:100-120.

Biberauer, Theresa. 2005. Changing EPP-parameters in the history of English: accounting for variation and change. *English Linguistics* 9, 1:5-46.

Joel Wallenberg. 2009. Antisymmetry and the Conservation of C-Command: Scrambling and Phrase Structure in Synchronic and Diachronic Perspective. Penn Dissertation.

CP
├─ C
│ dass
└─ TP
   ├─ vP$_n$
   │ ├─ AdvP
   │ │ gestern
   │ └─ vP$_n$
   │   ├─ PartP$_m$
   │   │ ├─ vP$_l$
   │   │ │ ├─ AdvP
   │   │ │ │ nicht
   │   │ │ └─ vP$_l$
   │   │ │   ├─ DP
   │   │ │   │ Johann
   │   │ │   └─ v'
   │   │ │   ├─ t$_j$
   │   │ │   └─ RootP
   │   │ │   ├─ t$_i$
   │   │ │   └─ DP
   │   │ │   das Buch
   │   │ └─ Part'
   │   │   ├─ Part
   │   │   │ ├─ v$_j$
   │   │   │ │ ├─ gekauft$_i$
   │   │   │ │ └─ v
   │   │   │ └─ Part
   │   │   └─ vP$_l$
   │   │   └─ t
   │   └─ v'
   │   ├─ t$_k$
   │   └─ PartP$_m$
   │   └─ t
   └─ T'
      ├─ Tense
      │ ├─ hat$_k$
      │ └─ Tense
      └─ vP$_n$
         └─ t

The end