Syntactic changes tracking morphological
A language change approach to implicational universals

Alexandra Simonenko (Ghent University)

University of Vienna, Department of Linguistics
January 26, 2023

OUTLINE

GHENT
UNIVERSITY

erc

CLUSTERING OF PROPERTIES

Non-accidental clustering of grammatical properties is key to understanding language mechanisms.

Greenberg (1966), Universal 41 "If in a language the verb follows both the nominal subject and nominal object as the dominant order, the language almost always has a case system."

Dryer (2002): Languages with morphological case marking

SOV	SVO	V-initial
72% (181/253)	**14%** (26/190)	**47%** (28/59)

GHENT
UNIVERSITY

Implicational universals
○○●○○○

Diachronic dimension
○○

Treebank case-studies
○○○○○○○○○○○○○○

Related probabilities?
○○○○○

TYPOLOGICAL PATTERNS

Dryer (2013), Comrie (2013)

SVO	SOV	no dominant
20% (8/41)	**63%** (26/41)	17% (7/41)

Table 1: Nominative-Accusative marking

SVO	SOV	no dominant
4% (1/24)	**66%** (16/24)	30% (7/24)

Table 2: Ergative-Absolutive marking

SVO	SOV	no dominant
48% (39/81)	36% (29/81)	16% (13/81)

Table 3: Neutral marking (no marking)

GHENT
UNIVERSITY

PROBABILISTIC UNIVERSALS

McFadden (2004, 150), "That a correlation of this sort [between case marking and word order type – A.S.] exists is not really a matter of debate. What is controversial is what form it takes and exactly how we are to explain it"

▶ Such correlations are not categorical of the form:
   if A (is present) then B (is present).

▶ Rather, they have a stochastic form: $P(A|B) > P(A|\neg B)$
   (the Probability of the property A (e.g. basic SOV order) if the property B is present (e.g. case marking) is greater than the probability of A when property B is not present)

GHENT
UNIVERSITY

erc

Implicational universals
○○○○●○

Diachronic dimension
○○

Treebank case-studies
○○○○○○○○○○○○○○○

Related probabilities?
○○○○○

CHALLENGE TO GRAMMAR FORMALIZATIONS

▶ Context-free grammars have no room for capturing
stochastic relations, which are inherently
context-dependent.

    ▶ Categorical trade-offs between two grammatical properties
are formalisable. E.g. Kiparsky (1997): MORPHOLOGICAL
CASE and POSITION as alternative strategies of argument
licensing.

GHENT
UNIVERSITY

erc

CHALLENGE TO GRAMMAR FORMALIZATIONS

▶ Typological universals challenge other types of grammars as well: how do we go from the level of a language sample – at which a universal holds – to expressing the dependency at the level of an individual language, if at all?

  ▶ At the synchronic level, in a given language, we don't generally observe stochastic dependencies between the linear order and case marking.

GHENT
UNIVERSITY

CHALLENGE TO GRAMMAR FORMALIZATIONS

Way out of the impasse: project a probabilistic universal onto the **diachronic dimension** of a given language, where both properties undergo a quantitative change and can be quantitatively related.

GHENT
UNIVERSITY

Implicational universals
○○○○○●

Diachronic dimension
○○

Treebank case-studies
○○○○○○○○○○○○○○○

Related probabilities?
○○○○○

PLAN

► Project Universal 41 onto the diachronic dimension of a given language:
  ► Old and Middle French

  ► Old and Middle English

► Suggest a formal meta-grammatical mechanism of expressing the stochastic nature of Universal 41 at the level of an individual speaker.

GHENT
UNIVERSITY

Implicational universals
oooooo

Diachronic dimension
●o

Treebank case-studies
ooooooooooooooooo

Related probabilities?
ooooo

OUTLINE

GHENT
UNIVERSITY

DIACHRONIC PROJECTION OF UNIVERSAL 41

Typological projection

$P(A|B) > P(A|\neg B)$

The probability of a property A (e.g. (S)OV) in a language with property B (e.g. nom-acc marking) is greater than in a language without B (e.g. neutral marking)

GHENT
UNIVERSITY

erc

DIACHRONIC PROJECTION OF UNIVERSAL 41

Diachronic projection (first version)

$P(A|X_{i<k+1}) > P(A|X_{i\geq k+1})$

given language stages X_1 X_2 ... X_k ... X_{k+1} ... X_{k+n},
where acc-nom marking is present for i < k+1

The probability of property A (e.g. (S)OV) at a language stage X
with property B (e.g. nom-acc marking) is greater than at a
language stage X′ without B (e.g. neutral marking)

GHENT
UNIVERSITY

DIACHRONIC PROJECTION OF UNIVERSAL 41

But the diachronic dimension contains more information than that.

- ▶ Unlike in a typological sampling, nom-acc features is not (necessarily) a categorical variable in the diachronic perspective, nor is a basic word order: there exist (many) stages for a given language where there is **some case marking** and **a mixture of word orders**.

- ▶ Typological: "The probability of (S)OV is higher in the presence of case marking in a language."

- ▶ Diachronic?: "The frequency of (S)OV is higher with a higher frequency of case marking at a language stage"

GHENT
UNIVERSITY

Implicational universals
000000

Diachronic dimension
00

Treebank case-studies
●000000000000000

Related probabilities?
00000

OUTLINE

Implicational universals
oooooo

Diachronic dimension
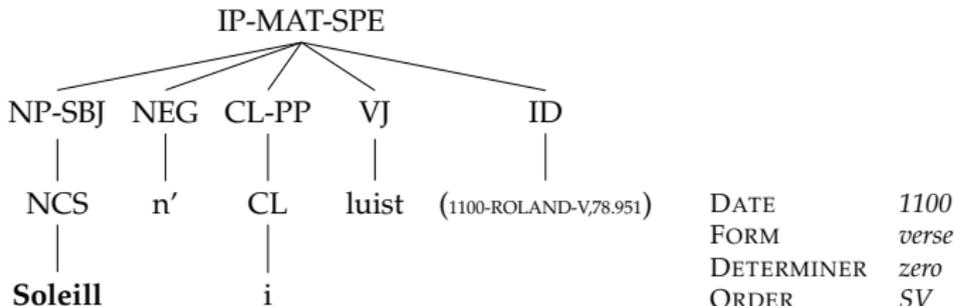oo

Treebank case-studies
o●oooooooooooooo

Related probabilities?
ooooo

EMPIRICAL GOALS

- ▶ Examine the quantitative dimension of case marking in
  - ▶ Old and Middle French
  - ▶ Old and Middle English

- ▶ Examine the quantitative dimension of core argument placement in
  - ▶ Old and Middle French
  - ▶ Old and Middle English

GHENT
UNIVERSITY

PENN TREEBANK FORMAT

Kroch and Santorini (2021): Old and Middle French
50 parsed texts from the X to XVI cc. ($\approx$ 1,5mln words).

IP-MAT-SPE

NP-SBJ   NEG   CL-PP   VJ   ID

NCS   n'   CL   luist   (1100-ROLAND-V,78.951)

Soleill   i

DATE	*1100*
FORM	*verse*
DETERMINER	*zero*
ORDER	*SV*

sun   not   there   shines

"The sun does not shine there."

GHENT
UNIVERSITY

FRENCH CASE SYSTEM

Buridant (2000)

	SG	PL
i		
NOM	reis	rei
ACC	rei	reis
ii		
	fame	fames
iii		
NOM	ber	baron
ACC	baron	barons
invariable		
	païs	païs

Implicational universals
○○○○○○

Diachronic dimension
○○

Treebank case-studies
○○○●○○○○○○○○○○

Related probabilities?
○○○○○

FRENCH CASE SYSTEM

(1) **Rei**s **Chielperic**s tam bien en   fist...
king Chilpéric   so   well of.it made
'King Chilperic dealt with it so well...' (La vie de Saint Léger 980)

(2) ...vus demandez **rei**...
you   ask   king
"...you are asking for a king..."(Li quatre livre des reis 1150)

(3) É   li   nostre **rei** nus   jugerá...
and our king   us will.judge
"And our king will judge us." (Li quatre livre des reis 1150)

FRENCH CASE MARKING

- ▶ all nominal subjects (N = 45314)
- ▶ all nominal subjects with (overt) nominative marking (N = 16955)

FRENCH CASE MARKING

▶ $P(\text{CASE} = nom | \text{DATE} = d) = \frac{e^{\alpha + \beta * Date}}{1 + e^{\alpha + \beta * Date}}$

Implicational universals
000000

Diachronic dimension
00

Treebank case-studies
000000●000000000

Related probabilities?
00000
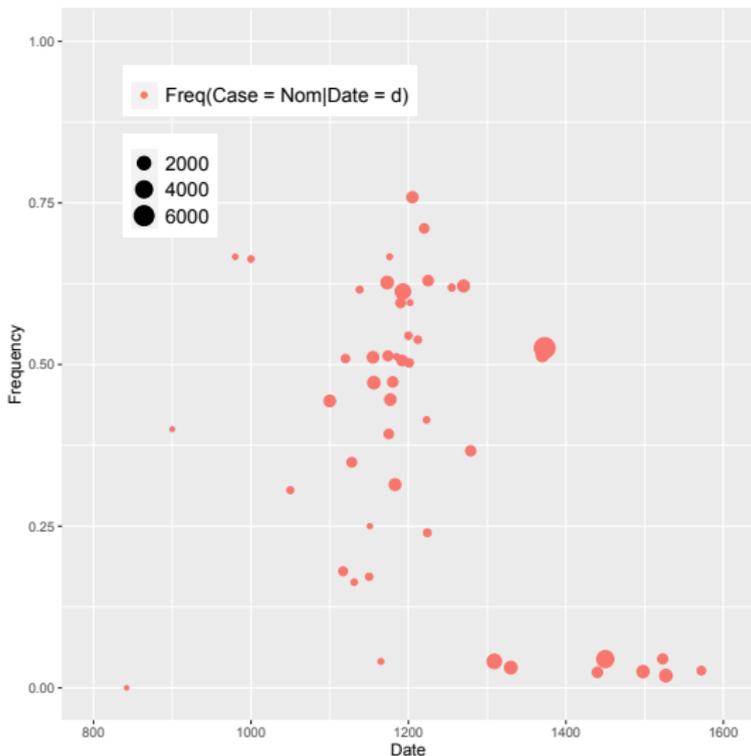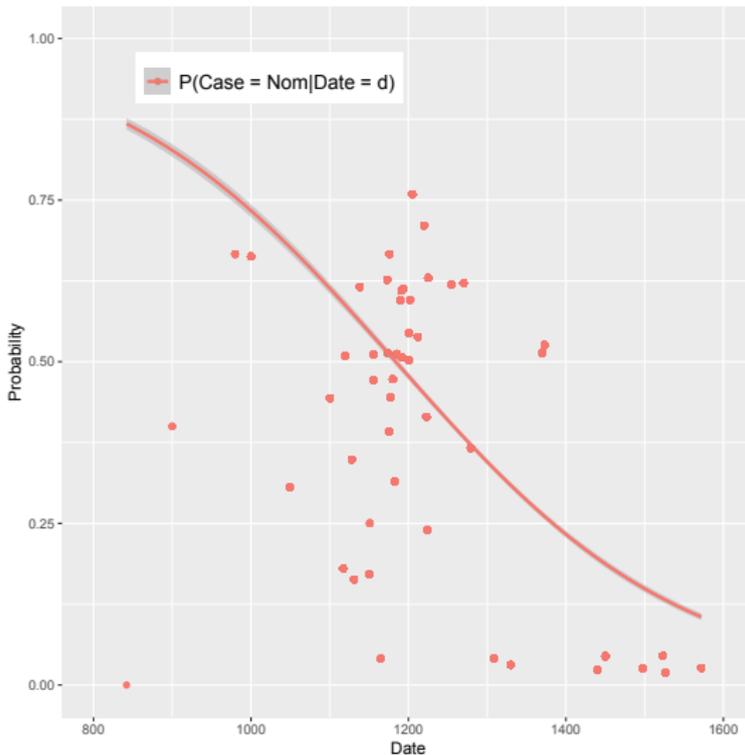
FRENCH ARGUMENT PLACEMENT

Clauses with nominal subjects and objects[1]

Period	OSV	OVS	SOV	SVO	VOS	VSO
X–XI	0	20	21	119	5	16
XII	5	48	130	467	29	80
XIII	2	50	23	547	21	136
XIV	4	36	45	903	37	250
XV	3	10	9	505	11	49
XVI	0	10	0	281	6	15
XVII–XVIII	0	0	0	290	2	0

---

[1]Counts are based on the 1 mln word corpus version

GHENT
UNIVERSITY

FRENCH ARGUMENT PLACEMENT

(4) [lei]$_{obj}$ consentit$_v$ et observat$_v$
     law agreed   and observed
     'he respected and observed the law' (0980-LEGER-V,XII.82)

(5) [Li quens Rollant]$_{sbj}$ [Gualter de l' Hum]$_{obj}$ apelet$_v$...
     the king Roland   Walter of the Hum   called
     'The king Roland called Walter of Hum...' (1100-ROLAND-V,65.779)

(6) ...vus demandez **rei**...
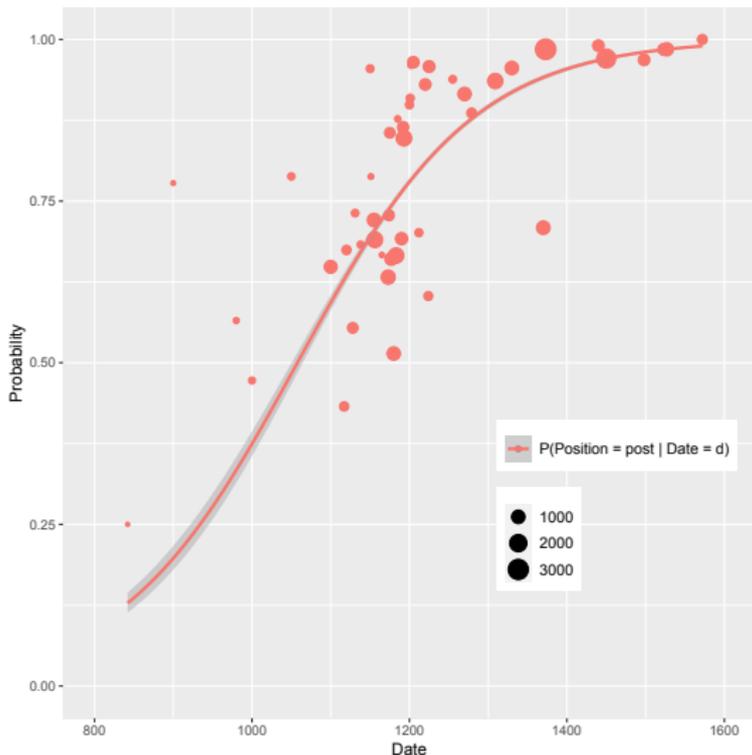   you ask   king
     "...you are asking for a king..." (Li quatre livre des reis 1150)

Implicational universals
○○○○○○

Diachronic dimension
○○

Treebank case-studies
○○○○○○●○○○○○○○○

Related probabilities?
○○○○○

FRENCH DIRECT OBJECT PLACEMENT

- ▶ all nominal direct objects (N = 30365)
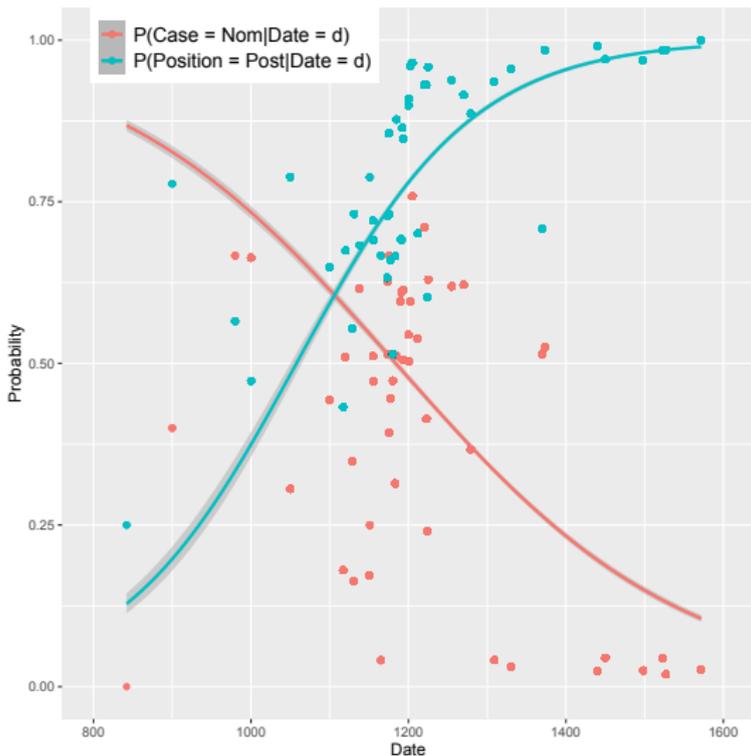- ▶ all nominal direct objects following the finite verb (N = 25101)
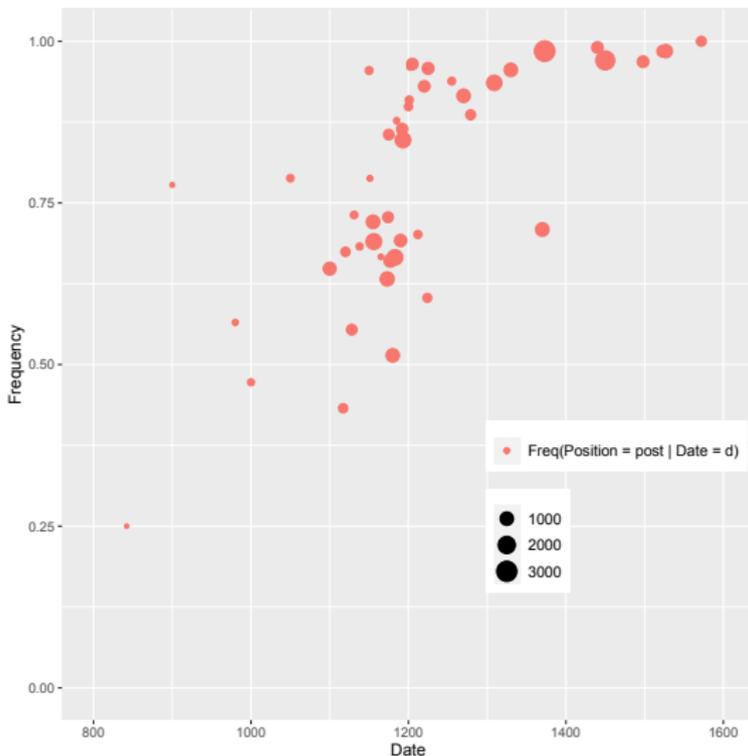
Implicational universals
oooooo

Diachronic dimension
oo

Treebank case-studies
oooooo●oooooooo

Related probabilities?
ooooo

FRENCH DIRECT OBJECT PLACEMENT

▶ $P(\text{POSITION} = post | \text{DATE} = d) = \frac{e^{\alpha + \beta * Date}}{1 + e^{\alpha + \beta * Date}}$

Implicational universals
○○○○○○

Diachronic dimension
○○

Treebank case-studies
○○○○○○○○●○○○○○○

Related probabilities?
○○○○○

SUBJECT MARKING AND DIRECT OBJECT PLACEMENT: FRENCH

Implicational universals
oooooo

Diachronic dimension
oo

Treebank case-studies
oooooo●oooooo

Related probabilities?
ooooo

CAPTURING THE RELATION?

► The two sets of *frequencies* are not significantly (inversely) correlated ($r(48)$ = -0.71, p = 0.47). It is not the case that in a given text, the frequency of nominative marking significantly (inversely) correlates with the frequency of object scrambling.

► In contrast, the two predicted *probability* trends have virtually the same rate, modulo the sign ($\beta = -0.005$ for nominative marking, $p < 10^{-16}$; $\beta = 0.008$ for object preposing, $p < 10^{-16}$).

GHENT
UNIVERSITY

erc
European Research Council

CAPTURING THE RELATION?

Revising the initial frequency-based hypothesis as a probability-based hypothesis:

▶ Diachronic: "The predicted probability of OV is higher with a higher predicted probability of case marking at a language stage" (which is not the same as a higher frequency of OV with a higher frequency of case marking in a given text).

Next: "testing" the hypothesis on the English data.

GHENT
UNIVERSITY

ENGLISH CASE SYSTEM

▶ Already in Old English, the nom-acc morphological contrast is absent in 2 out of 4 major declension classes. Pintzuk (2002)

Masculine *a*-stems. Example: *stān* 'stone'

	Singular	Plural
NOM	stān	stānas
ACC	stān	stānas
GEN	stānes	stāna
DAT	stāne	stānum

GHENT
UNIVERSITY

ENGLISH CASE SYSTEM

▶ Already in Old English, the nom-acc morphological contrast is absent in 2 out of 4 major declension classes. Pintzuk (2002)

Neuter *a*-stems. Example: *scip* 'ship'

	Singular	Plural
NOM	scip	scipu
ACC	scip	scipu
GEN	scipes	scipa
DAT	scipe	scipum

ENGLISH CASE SYSTEM

▶ Already in Old English, the nom-acc morphological contrast is absent in 2 out of 4 major declension classes. Pintzuk (2002)

Feminine *o*-stems. Example: *giefu* 'gift'

	Singular	Plural
NOM	giefu	giefa,-e
ACC	giefe	giefa,-e
GEN	giefe	giefa,-ena
DAT	giefe	giefum

GHENT
UNIVERSITY

ENGLISH CASE SYSTEM

▶ Already in Old English, the nom-acc morphological contrast is absent in 2 out of 4 major declension classes. Pintzuk (2002)

Weak masculine nouns. Example: *nama* 'name'

	Singular	Plural
NOM	nama	naman
ACC	naman	naman
GEN	naman	namena
DAT	naman	namum

ENGLISH CASE SYSTEM

- ► Already in Old English, the nom-acc morphological contrast is absent in 2 out of 4 major declension classes. Pintzuk (2002)

- ► It is still robust in the definite/demonstrative determiner system.

Definite determiner/demonstrative pronoun

	Singular			Plural
	MASC	NEUT	FEM	All genders
NOM	sē	þæt	sēo	þā
ACC	þone	þæt	þā	þā
GEN	þæs	þæs	þǣre	þāra
DAT	þǣm	þǣm	þǣre	þǣm

GHENT
UNIVERSITY

erc

ENGLISH CASE MARKING

▶ all subjects with a definite/demonstrative determiner (N = 40436)
▶ all subjects with a definite/demonstrative determiner with a nom marking (N = 22142), from Taylor et al. (2003), Kroch and Taylor (2000)

GHENT
UNIVERSITY
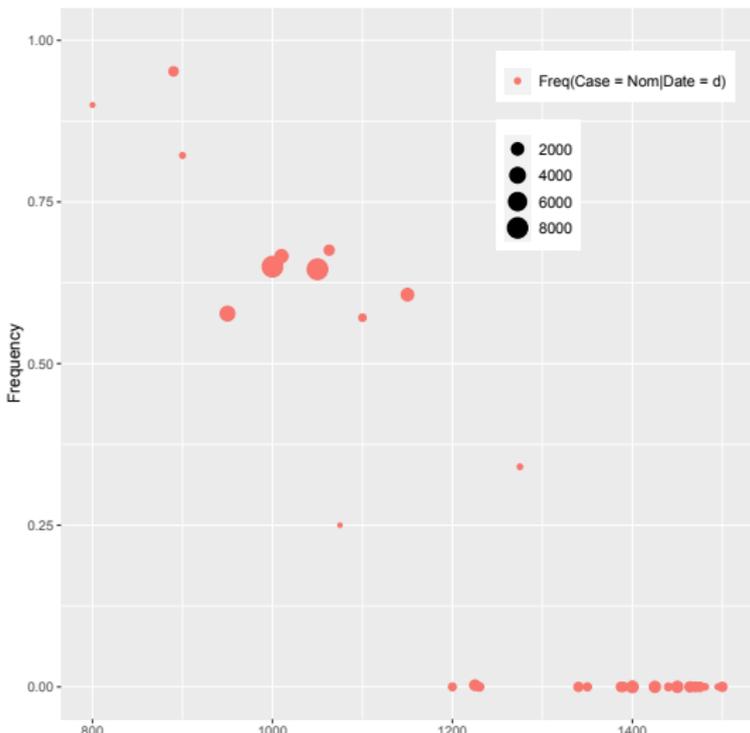
ENGLISH CASE MARKING

▶ $P(\text{CASE} = nom | \text{DATE} = d) = \frac{e^{\alpha + \beta * Date}}{1 + e^{\alpha + \beta * Date}}$

GHENT
UNIVERSITY
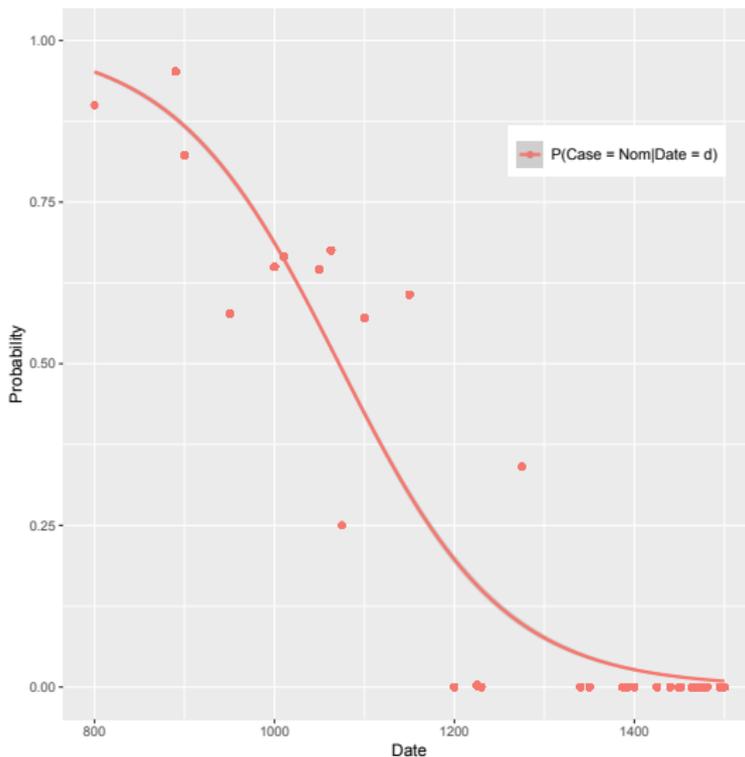
ENGLISH DIRECT OBJECT PLACEMENT

► Old English features both OV_{fin} and $V_{fin}O$. Pintzuk (2002)

(1) O Aux V

    a. Ne God þonne ane hwile his mihta ne his wundra sylf nele cyðan
       Nor God then a while his powers nor his wonders self NEG-would reveal
       'Nor would God himself then reveal his powers or his wonders for a while...'
       (WHom, 138.64–65)

    b. swa hwider swa se cining Oswi his rice mihte þennan
       as far as the king Oswy his kingdom could stretch
       '... as far as the king Oswy could stretch his kingdom.'
       (Chad, 44)

GHENT
UNIVERSITY

erc
European Research Council

ENGLISH DIRECT OBJECT PLACEMENT

▶ Old English features both OV$_{fin}$ and V$_{fin}$O. Pintzuk (2002)

(2) Aux V O

   a. þu hafast gecoren þone wer
      you have chosen the man
      'You have chosen the man.' (ApT 23.1)

   b. þæt he mot ehtan godra manna
      that he might persecute good men
   '. . . that he might persecute good men . . .'
      (WHom, 130.37–38)

GHENT
UNIVERSITY

Implicational universals
○○○○○○

Diachronic dimension
○○

Treebank case-studies
○○○○○○○○○○○○○●○○

Related probabilities?
○○○○○

ENGLISH DIRECT OBJECT PLACEMENT

- ▶ all nominal direct objects (N = 62089)
- ▶ all nominal direct objects following the finite verb (N = 49571)
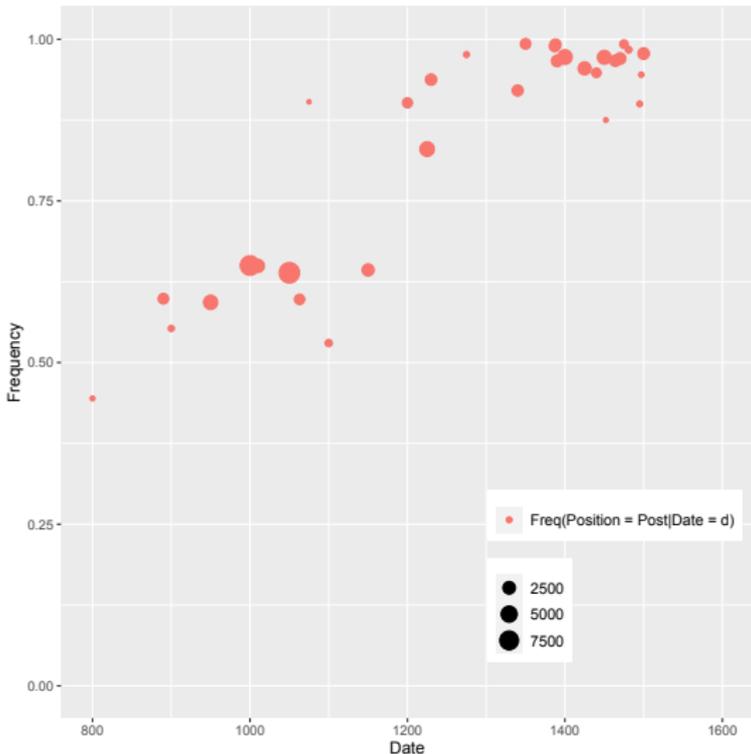
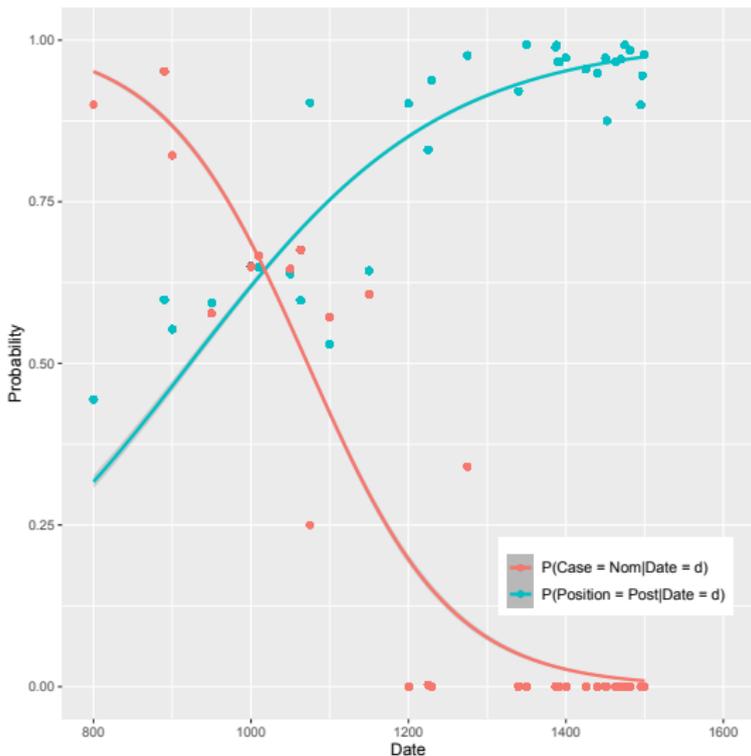# ENGLISH DIRECT OBJECT PLACEMENT

▶ $P(\text{POSITION} = post | \text{DATE} = d) = \frac{e^{\alpha + \beta * Date}}{1 + e^{\alpha + \beta * Date}}$


GHENT
UNIVERSITY
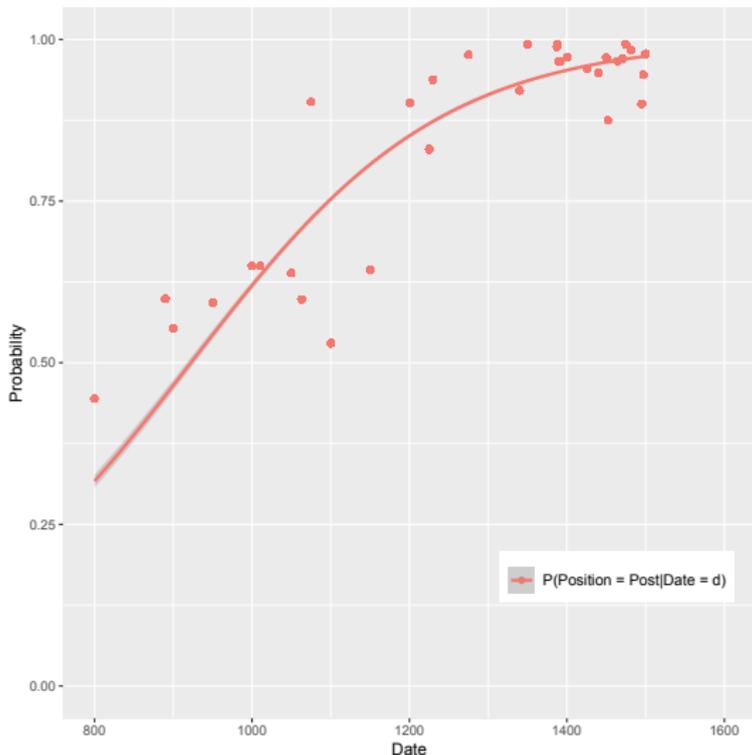
SUBJECT MARKING AND DIRECT OBJECT PLACEMENT: ENGLISH

SUBJECT MARKING AND DIRECT OBJECT PLACEMENT: ENGLISH

As expected, the predicted probability trends have very similar rates (non-similarity to be tested formally), modulo the sign (β = −0.01 for nominative marking, p < 10^{-16}; β = 0.006 for object preposing, p < 10^{-16}).

GHENT
UNIVERSITY

erc
European Research Council

SUBJECT MARKING AND DIRECT OBJECT PLACEMENT: COMPARISON
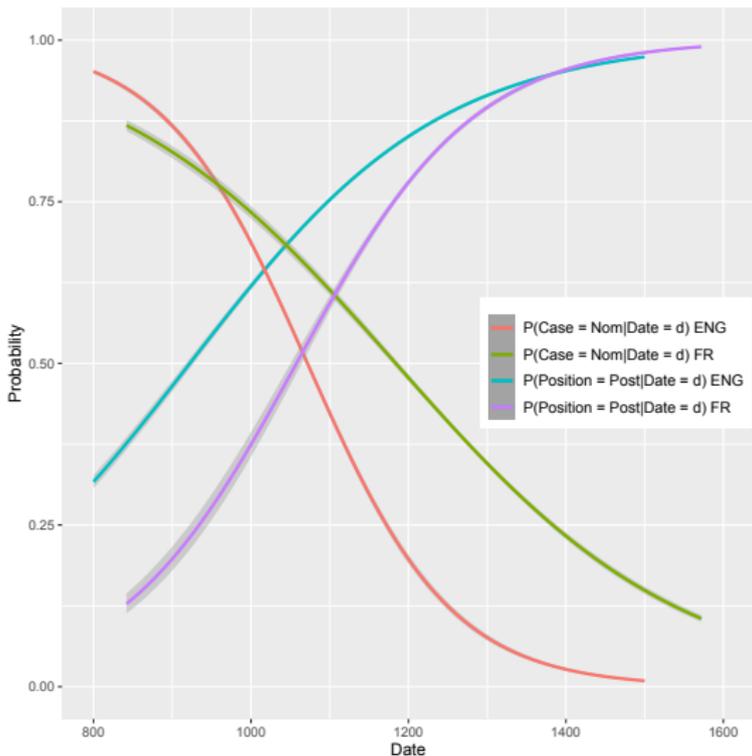
OUTLINE

Related probabilities?

CAPTURING THE RELATION, AGAIN

The probability-based diachronic hypothesis

▶ "The predicted probability of OV is higher with a higher predicted probability of case marking at a language stage"

Where can the predicted probabilities be related?

▶ Not in a given grammar: no formal room
▶ Not in a given usage: Pintzuk (2002) shows that in Old English there is no dependency between case marking of a direct object and *that argument's* linear position (contra Roberts (1997) and Weerman (1997)).

GHENT
UNIVERSITY

CAPTURING THE RELATION, AGAIN

▶ A tentative answer: in the Speaker's "meta-grammar",
defined as the space of stochastically used competing
grammatical representations (Tony Kroch's competing
grammars)

GHENT
UNIVERSITY

CAPTURING THE RELATION, AGAIN

A possible concrete implementation: acquisition-based approach to grammar competition (Yang 2010).

▶ Core insight: the probability of accusative-marked subjects adversely affects the probability of (the grammar allowing for) object scrambling during language acquisition.[1]

▶ Modeling using a reinforcement learning algorithm.

------

[1]We cannot of course assume that the speakers literally have access to logistic models... Predicted probabilities need to be generated in a psychologically plausible way.

GHENT
UNIVERSITY

Implicational universals
oooooo

Diachronic dimension
oo

Treebank case-studies
ooooooooooooooo

Related probabilities?
o●ooo

CAPTURING THE RELATION, AGAIN

Penalty c_i of a grammar G_i:

$$c_i = P(G_i \not\to x | x \in E)$$

c_i is the probability that G_i fails to analyse an example in a dataset X, estimated by the relative frequency of failures (Narendra and Thathachar 1989).

$$\lim_{t \to \infty} P(\mathcal{G} = G_1 | E = e, T = t) = \frac{c_2}{c_1 + c_2}$$

$$\lim_{t \to \infty} P(\mathcal{G} = G_2 | E = e, T = t) = \frac{c_1}{c_1 + c_2}$$

CAPTURING THE RELATION, AGAIN

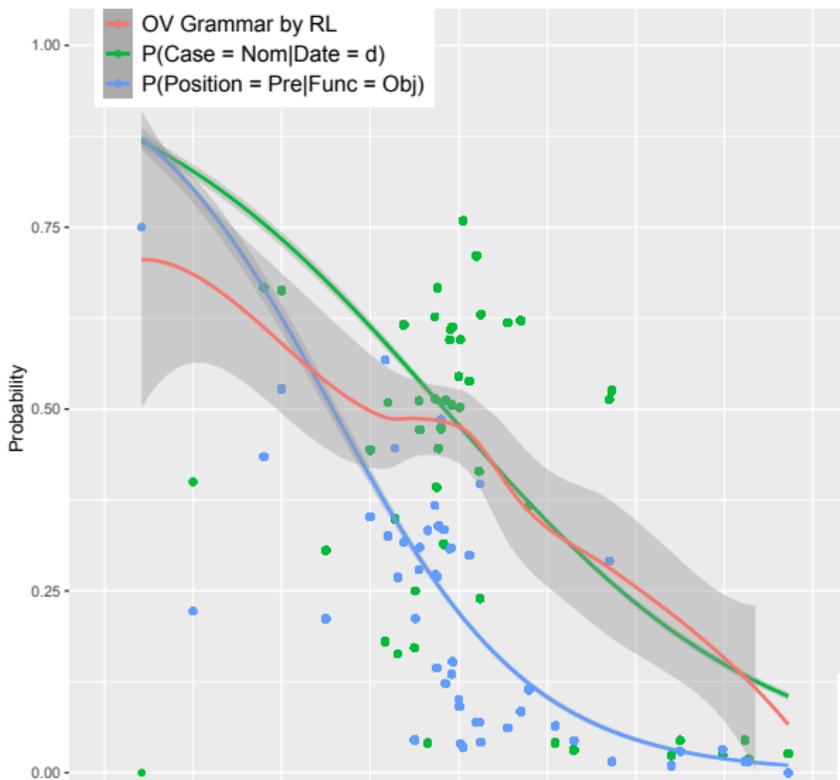Crucial: probabilities of grammars depend on their mutual probabilities of failure.

$$\lim_{t \to \infty} P(\mathcal{G} = G_1 | E = e, T = t) = \frac{c_2}{c_1 + c_2}$$

$$\lim_{t \to \infty} P(\mathcal{G} = G_2 | E = e, T = t) = \frac{c_1}{c_1 + c_2}$$

▶ For Gr$_{OV}$, c$_{OV}$ is the probability of accusative-marked subject.

▶ For Gr$_{VO}$, c$_{VO}$ is the probability of OV plus the probability of VS.

GHENT
UNIVERSITY

erc
European Research Council

CAPTURING THE RELATION, AGAIN

Comparing the probability of the Gr_{OV} predicted by the RL algorithm with with the probability of OV as predicted by a logistic regression.

GHENT
UNIVERSITY

Implicational universals
○○○○○○

Diachronic dimension
○○

Treebank case-studies
○○○○○○○○○○○○○○

Related probabilities?
○●○○○

CAPTURING THE RELATION, AGAIN

▶ "Gr$_{OV}$" needs a lot of refinement. It has to be defined as a grammar *that allows for object scrambling*. Tentatively, it is a grammar allowing for Spec TP to be occupied by any argument, not just subjects. Once properly defined, the contexts of its failure will be also refined and penalty probability recalculated.

GHENT
UNIVERSITY

erc
European Research Council

CAPTURING THE RELATION, AGAIN

Another avenue to explore to related the two trends in the mental representations: Game Theoretic framework (along the lines of Jäger (2007)).

CONCLUSIONS

- ► Typological Universal 41, projected onto the diachronic dimension, corresponds to the probabilities of 1) nominative marking and 2) object scrambling growing/descending at virtually the same rates in both French and English.

- ► It seems important to have a language model that would capture this. Alternatively, we can conclude that it's a coincidence, which is accidentally observed in the history of different languages and accidentally relates to a statistically robust typological pattern...

- ► Predicted probabilities of competing grammars in the reinforcement learning acquisition model depend on their failure probabilities, which correspond to an abstraction based on the behaviour of the whole population (of tokens) at a given time period.

A suitable framework would need to assume that abstractions, approximating the behaviour of the

Buridant, Claude. 2000. *Nouvelle grammaire de l'ancien français*. Paris: Sedes.

Comrie, Bernard. 2013. Alignment of case marking of full noun phrases. In *The world atlas of language structures online*, ed. Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL `https://wals.info/chapter/98`.

Dryer, Matthew S. 2002. Case distinctions, rich verb agreement, and word order type. *Theoretical Linguistics* 28:151–157.

Dryer, Matthew S. 2013. Order of subject, object and verb. In *The world atlas of language structures online*, ed. Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL `https://wals.info/chapter/81`.

Greenberg, Joseph H. 1966. *Universals of language: Report of a Conference held at Dobbs Ferry, New York, April 13-15, 1961 (2nd Ed.)*. Cambridge, MA: MIT press.

Jäger, Gerhard. 2007. Evolutionary game theory and typology: A case study. *Language* 74–109.

Kiparsky, Paul. 1997. The rise of positional licensing. In *Parameters of morphosyntactic change*, ed. Ans van Kemenade and Nigel Vincent, 460–494. Citeseer.

Kroch, Anthony, and Beatrice Santorini. 2021. *Penn-BFM Parsed Corpus of Historical French (PPCHF)*. URL `https://github.com/beatrice57/mcvf-plus-ppchf/`.

Kroch, Anthony, and Ann Taylor. 2000. URL `http://www.ling.upenn.edu/hist-corpora/`, The Penn-Helsinki Parsed Corpus of Middle English (PPCME2). Department of Linguistics, University of Pennsylvania. CD-ROM, second edition.

Martineau, France, Paul Hirschbühler, Anthony Kroch, and Yves Charles Morin. 2010. *Corpus MCVF annoté syntaxiquement, (2005–2010), dirigé par France Martineau, avec Paul Hirschbühler, Anthony Kroch et Yves Charles Morin*.

McFadden, Thomas. 2004. The positon of morphological case in the derivation: A study on the syntax-morphology interface. Doctoral Dissertation, University of Pennsylvania.

Narendra, Kumpati S., and Mandayam A.L. Thathachar. 1989. *Learning Automata: An Introduction*. Englewood Cliffs, NJ: Prentice-Hall.

Pintzuk, Susan. 2002. Morphological case and word order in Old English. *Language Sciences* 24:381–395.

Roberts, Ian. 1997. Directionality and word order change in the history of English. In *Parameters of morphosyntactic change*, ed. Ans Van Kemenade and Nigel Vincent, 397–426. Cambridge University Press.

Taylor, Ann, Anthony Warner, Susan Pintzuk, and Frank Beths. 2003. URL `http://www-users.york.ac.uk/~lang22/YcoeHome1.htm`, The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE). Department of Linguistics, University of York. Oxford Text Archive, first edition.

Weerman, Frederikus Pieter. 1997. On the relation between morphological and syntactic case. In *Parameters of morphosyntactic change*, ed. Ans van Kemenade and Nigel Vincent, 427–459. Cambridge: CUP.

Yang, Charles. 2010. Three factors in language variation. *Lingua* 120:1160–1177.