GHENT UNIVERSITY

European Research Council
Established by the European Commission

A language change approach to probabilistic universals: case and order

Alexandra Simonenko (Ghent University)

DiGS 24, Université Paris Cité
July 4–7, 2023

Outline

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."

erc European Research Council Established by the European Commission

Typological patterns: SOV and case marking

Dryer (2002): Proportion of languages with morphological case marking among language groups with different dominant word orders

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

Restating Universal 41 as a probabilistic rather than implicational generalization:

P(*Case* **= yes** | *Dom. order* **= SOV)** > **P(***Case* **= yes** | *Dom. order* **= ¬SOV)**

Typological patterns: SOV and case marking

The probabilistic universal is bidirectional.

WALS (Dryer 2013, Comrie 2013):

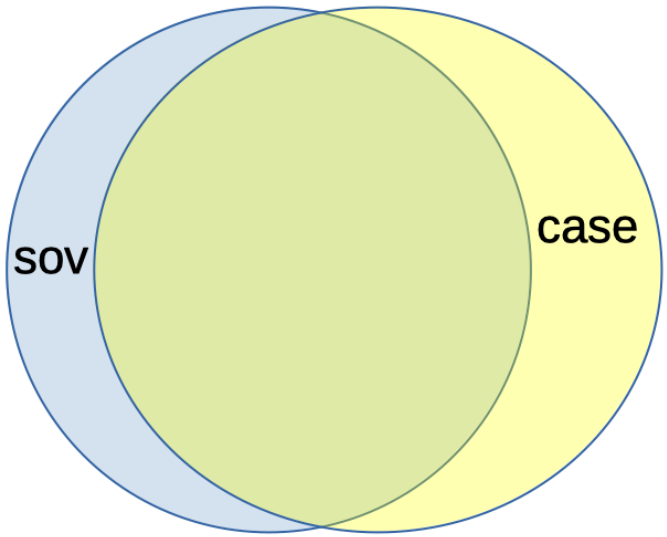
Case	SVO	SOV	no dominant
NOM-ACC	20% (8/41)	**63%** (26/41)	17% (7/41)
ERG-ABS	4% (1/24)	**66%** (16/24)	30% (7/24)
NONE	48% (39/81)	36% (29/81)	16% (13/81)

P(*Dom. order* **= SOV** | *Case* **= yes)** > **P(***Dom. order* **= SOV** | *Case* **= no)**

Typological patterns: SOV and case marking

$$\mathbf{P}(\textit{Case} = \mathbf{yes} \mid \textit{Dom. order} = \mathbf{SOV}) > \mathbf{P}(\textit{Case} = \mathbf{yes} \mid \textit{Dom. order} = \neg\mathbf{SOV})$$

$$\mathbf{P}(\textit{Dom. order} = \mathbf{SOV} \mid \textit{Case} = \mathbf{yes}) > \mathbf{P}(\textit{Dom. order} = \mathbf{SOV} \mid \textit{Case} = \mathbf{no})$$



Typological patterns: SVO and no case marking

A related universal: SVO languages are significantly less likely to feature case marking than not SVO languages (Sinnemäki 2010).[1]

P(*Case* **= yes** | *Dom. order* **= SVO)** $<$ **P(***Case* **= yes** | *Dom. order* **=** $\neg$**SVO)**

[1] A logistic regression model fit to zero/non-zero case marking variable on a sample of 813 languages from 448 genera.

Typological patterns: SVO vs. SOV

SOV correlates with case marking positively, and SVO negatively.

TYPOLOGICAL PATTERNS: RATIO CORRELATION

A Case-Order universal also holds for variables defined as ratios.

Sinnemäki (2008) finds a statistically significant negative correlation between canonical word order and case marking as argument marking strategies.[2]

(Sinnemäki, 2008, 82): "the smaller the functional load [≈ frequency – A.S.] of morphological marking is the more likely word order carries a greater functional load in core argument marking."

$$\text{Freq}(\textit{Case} = \text{no}) \sim \text{Freq}(\textit{Dom. order} = \text{SVO})$$

[2] In a sample of 50 languages, Sinnemäki (2008) assigns Functional Load (from 1 to 4) to word order and case marking in a language, and runs [illegible]all's tau test (tau = −.302, p = 0.017).

Challenge to grammar formalizations

How do we go from a typological sample to expressing the dependency at the level of a given language?

- There are SVO languages with case marking (14% in Dryer's sample).
- There are SOV languages without case marking (28% in Dryer's sample).

At the level of a language

- Case marking correlates with processing speed (Yamashita 1997, Yamashita 2000 for Japanese[3]).
- A canonical word order is preferred in sentences with syncretically marked arguments ("word order freezing" Berdičevskis and Piperski 2020 for German and Russian, Sinnemäki (2008) for Greenlandic, Bouma (2011) for Dutch, Erlewine and Lim (2023) for Pangasinan)
- SVO is used in gestural communication experiments when the roles cannot be distinguished otherwise Gibson et al. (2013).

The typological effect, in contrast, is very strong.

[3] A string with two nominative-marked preverbal arguments is processed slower than a string with distinctly marked preverbal arguments.

At the level of a language

Proposal: to project Case-Order universals onto the **diachronic dimension** of a given language, i.e. a sequence of chronologically ordered stages related by an unbroken succession of child learners (Labov 2007). Advantages:

- observing how minimal (≈generational) frequency changes in one variable are related to minimal frequency changes in the other;
- relating the two variable via a grammatical/pragmatic mechanism assuming all other variables are stable (modulo the generational difference);
- testing (more easily) for a causal relation since at least one of the variables can be assumed to be undergoing a unidirectional change.

Plan

- Project Case-Order universals onto the diachronic dimension of French and English.
- Lay out a Game-Theoretic model of a pragmatic mechanism responsible for the Universal 41 at the level of an individual speaker.
- Compare *Case-Order* relation in the Game-Theoretic simulations and corpus data.

Outline

Diachronic projection of Case-Order universals

Typological projection (paraphrasing Sinnemäki 2008)

P(Freq(SVO | language = i) > Freq(SVO | language = j)) >
P(Freq(SVO | language = i) < Freq(SVO | language = j))

for **all languages** i, j such that
Freq(Case = no | language = i) > Freq(Case = no | language = j)

Diachronic projection

P(Freq(SVO | generation = i) > Freq(SVO | generation = j)) >
P(Freq(SVO | generation = i) < Freq(SVO | generation = j))

for **all generations** i, j such that
Freq(Case = no | generation = i) > Freq(Case = no | generation = j)

erc European Research Council

Outline

Treebanks

- Modéliser le changement. Les voies du français (MCVF, Martineau et al. 2010)
- Penn-BFM Parsed Corpus of Historical French (PPCHF, Kroch and Santorini 2021)
- The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE, Taylor et al. 2003)
- The Penn-Helsinki Parsed Corpus of Middle English, second edition (PPCME2, Kroch et al. 2000)
- The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME, Kroch et al. 2004)

Outline

French: Morphological case

Buridant (2000)

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

erc European Research Council

French: Morphological case

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

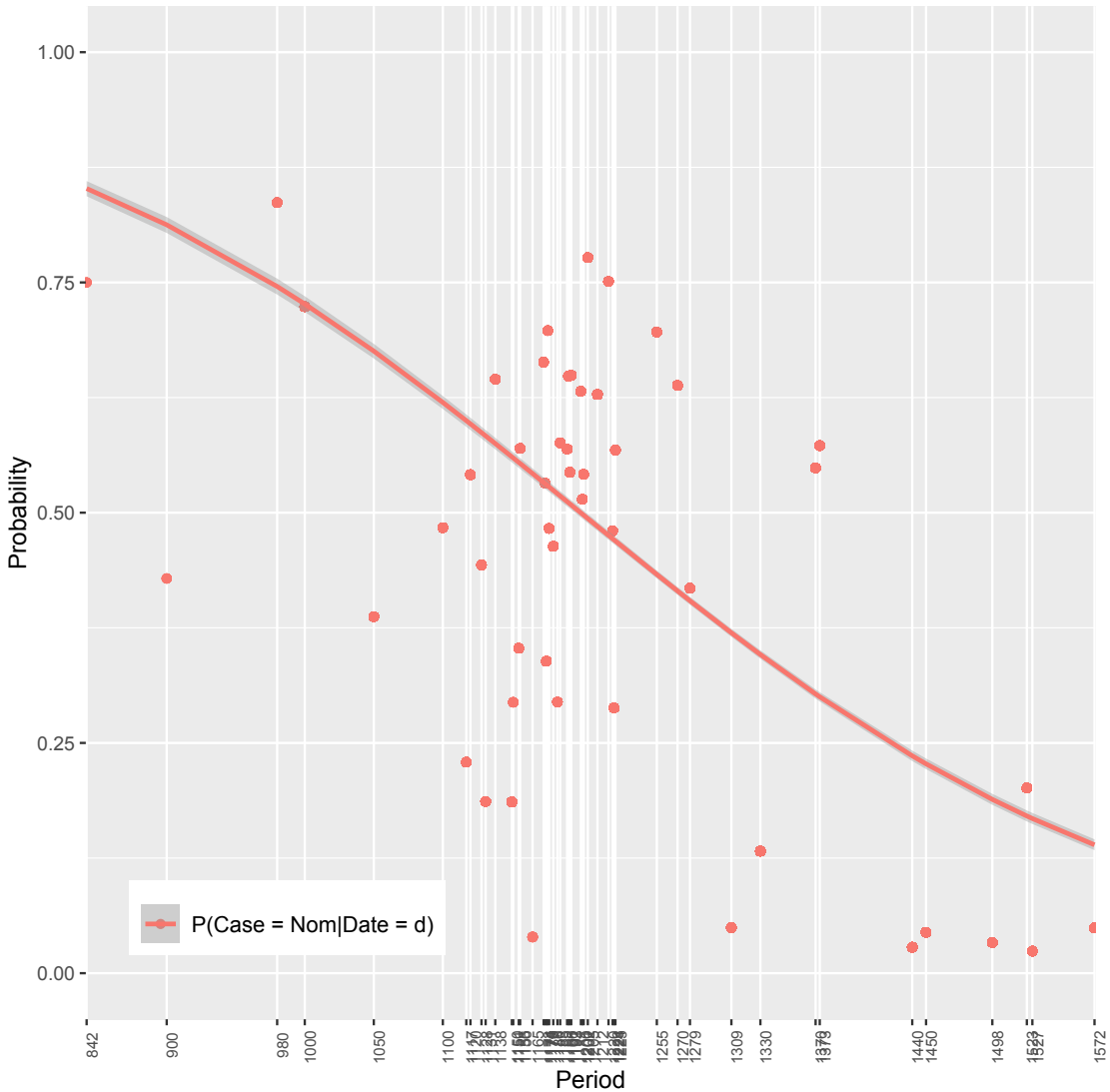
(2) ...vus demandez **rei**...
you ask king
"...you are asking for a king..." (Li quatre livre des reis 1150)

(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: Morphological case

Common & proper nouns with nominative marking (N = 23069, N total = 56068)

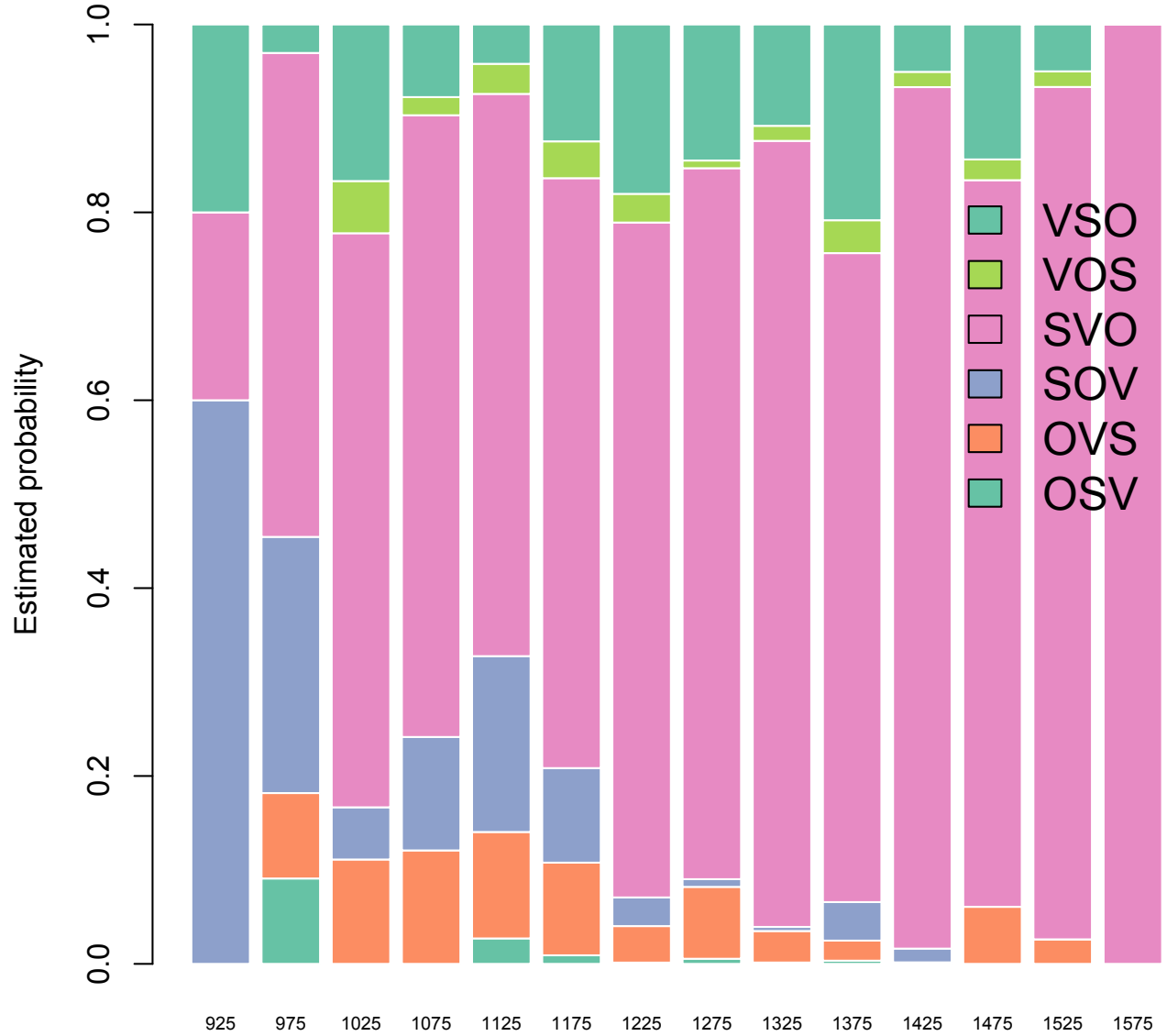
$$P(\text{Case} = Nom | \text{Date} = d) = \frac{e^{\alpha+\beta*Date}}{1+e^{\alpha+\beta*Date}}$$



Outline

French word order

Matrix and subordinate clauses with common/proper noun arguments, N = 13036.



SVO/SOV ≈ VO/OV

I *approximate* the relative frequency of SVO with the relative frequency of (finite) VO, as a role-distinguishing order.[34]

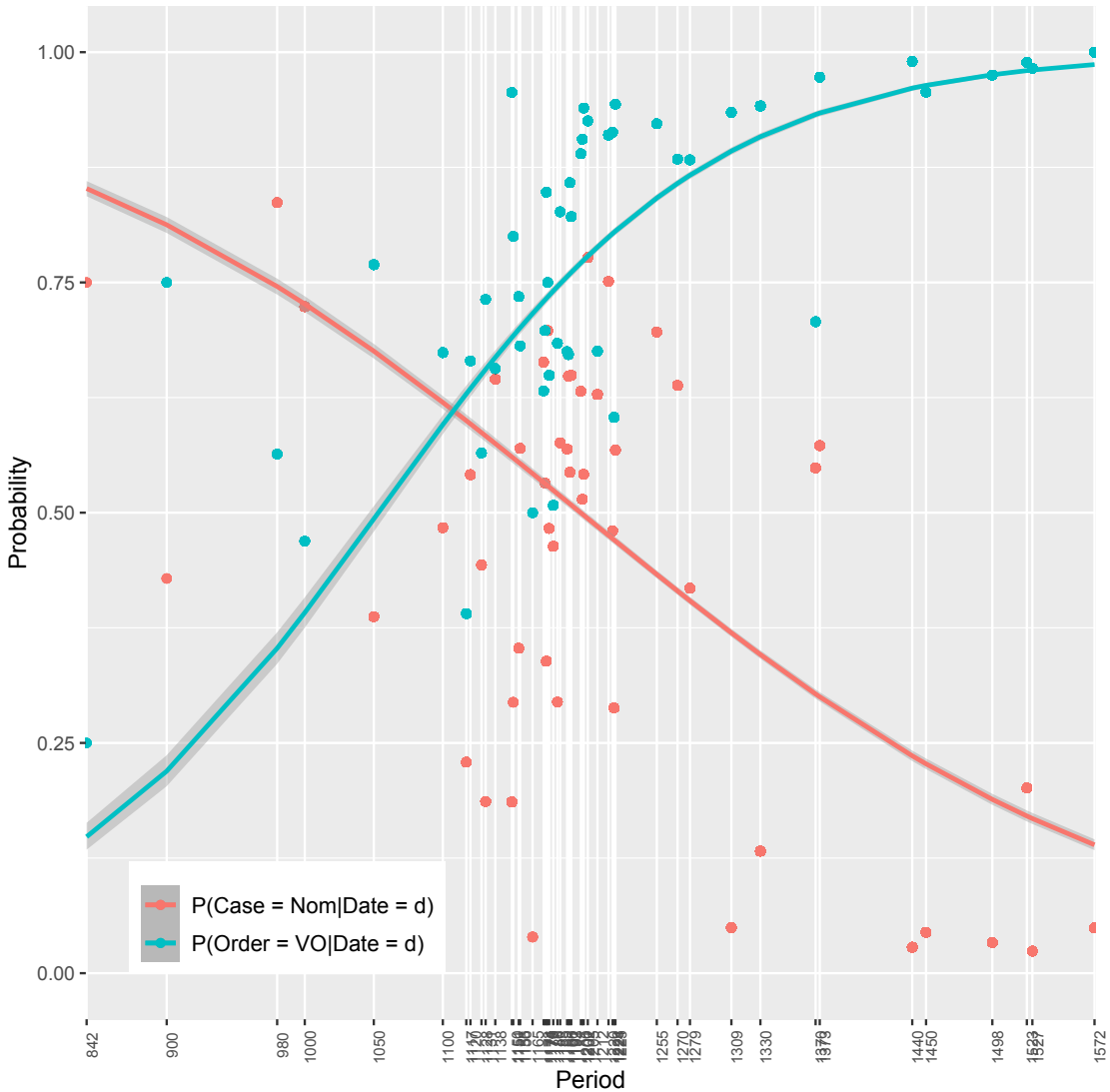
[3] Cf. Gibson et al. (2013), Sinnemäki (2010).

[4] I assume that a) the rates of VO (S**VO**, **VO**S and **VSO**) are stable within the grammar that produces OV ("Scrambling"-grammar) b) the alternative grammar produces *only* SVO.

French word order (and case)

VO (N = 32092, total common and proper noun O N = 39094)

$$P(\text{Order} = VO|\text{Date} = d) = \frac{e^{\alpha+\beta*Date}}{1+e^{\alpha+\beta*Date}}$$



French: (Kendall's) Correlation?

A significant negative correlation between case and VO frequencies.

- At the level of individual text frequencies: $\tau = -0.22$, p = 0.02;
- At the level of 50 year bin frequencies ("generations"): $\tau = -0.47$, p = 0.02.

Outline

English: Morphological case

- 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

English: Morphological case

- 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: Morphological case

- 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

English: Morphological case

- 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: Morphological case

- 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.

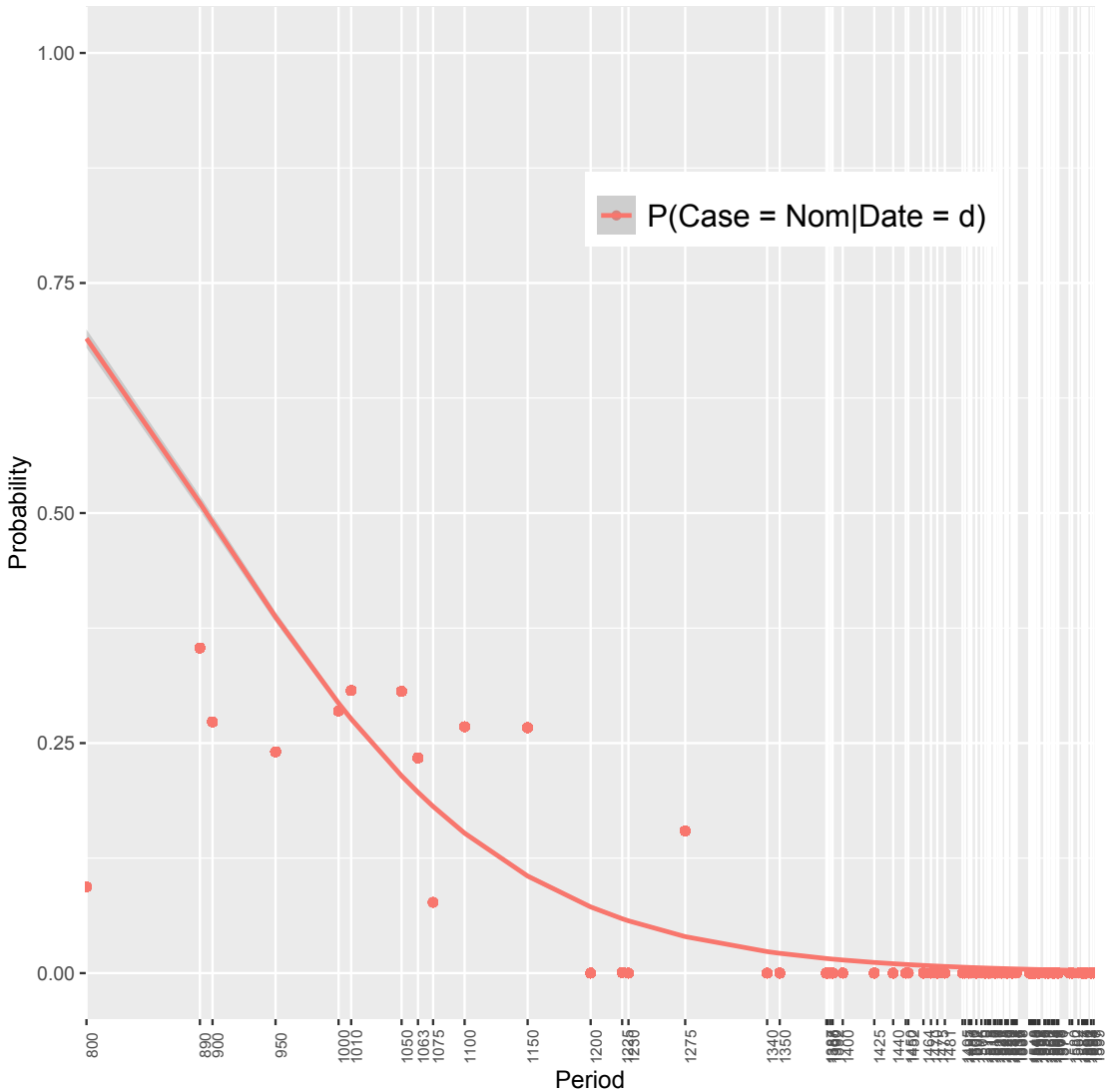
Definite determiner/demonstrative pronoun

	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

English: morphological case

Common and proper noun subjects N = 102806, among which with a def/dem nominative marked determiner 15235

Data from the corpora of Taylor et al. (2003), Kroch et al. (2000)

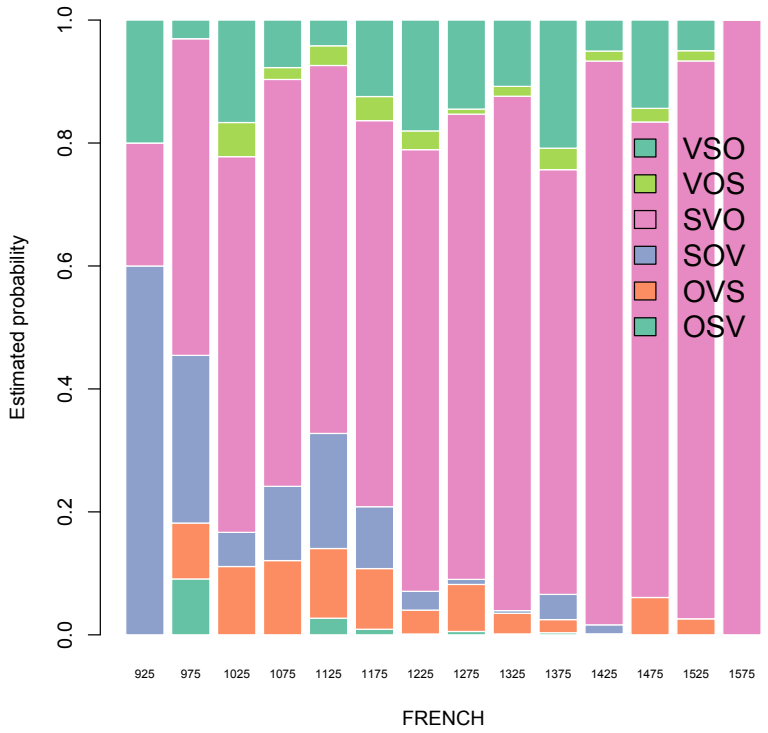


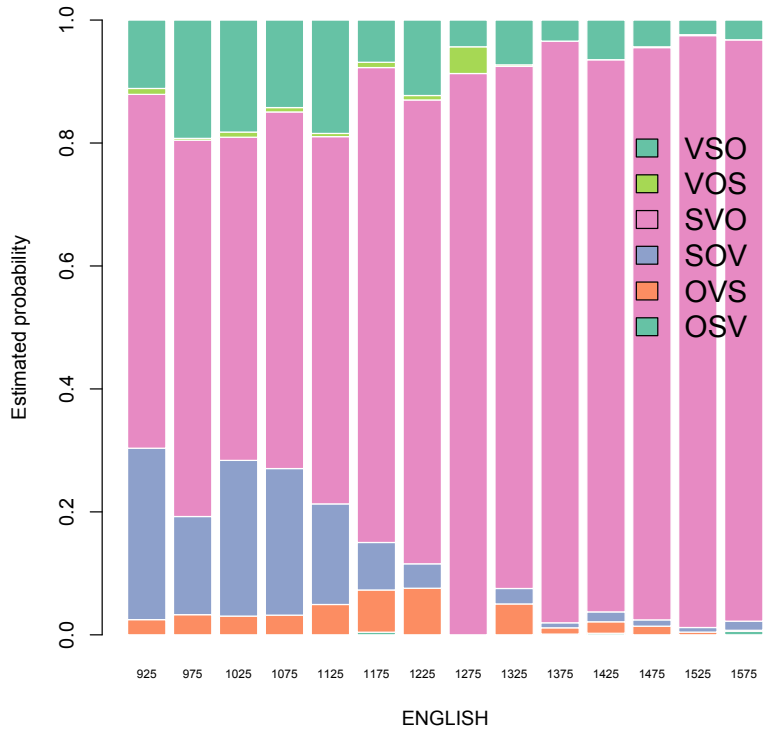
erc European Research Council

Outline

English (and French): word order

Matrix and subordinate clauses with two nominal arguments, N = 13036.



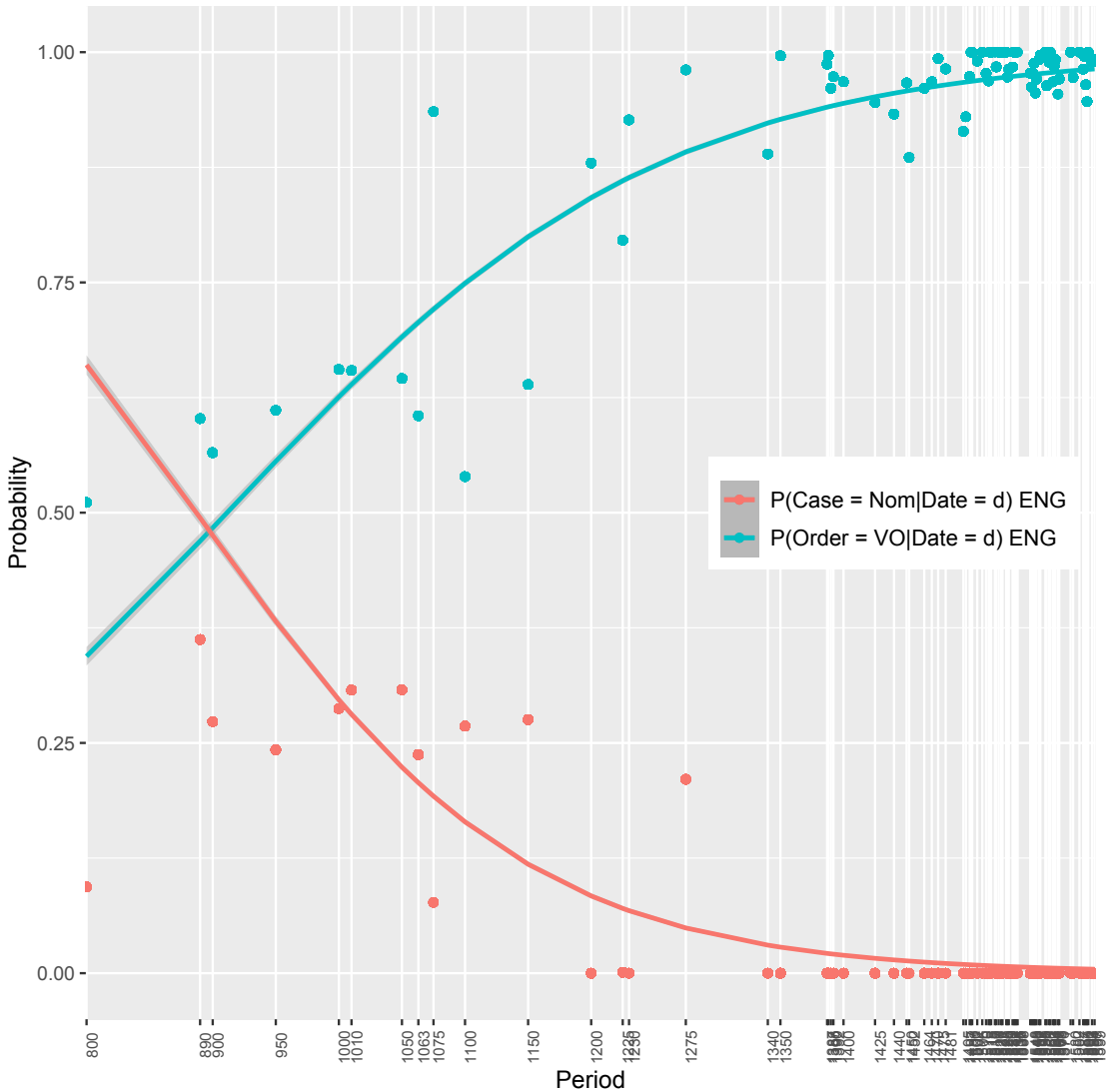


erc European Research Council

English: word order (and case)

Common and proper noun objects N = 64706, of which postverbal 52338)

$$P(\text{Order} = VO|\text{Date} = d) = \frac{e^{\alpha+\beta*Date}}{1+e^{\alpha+\beta*Date}}$$

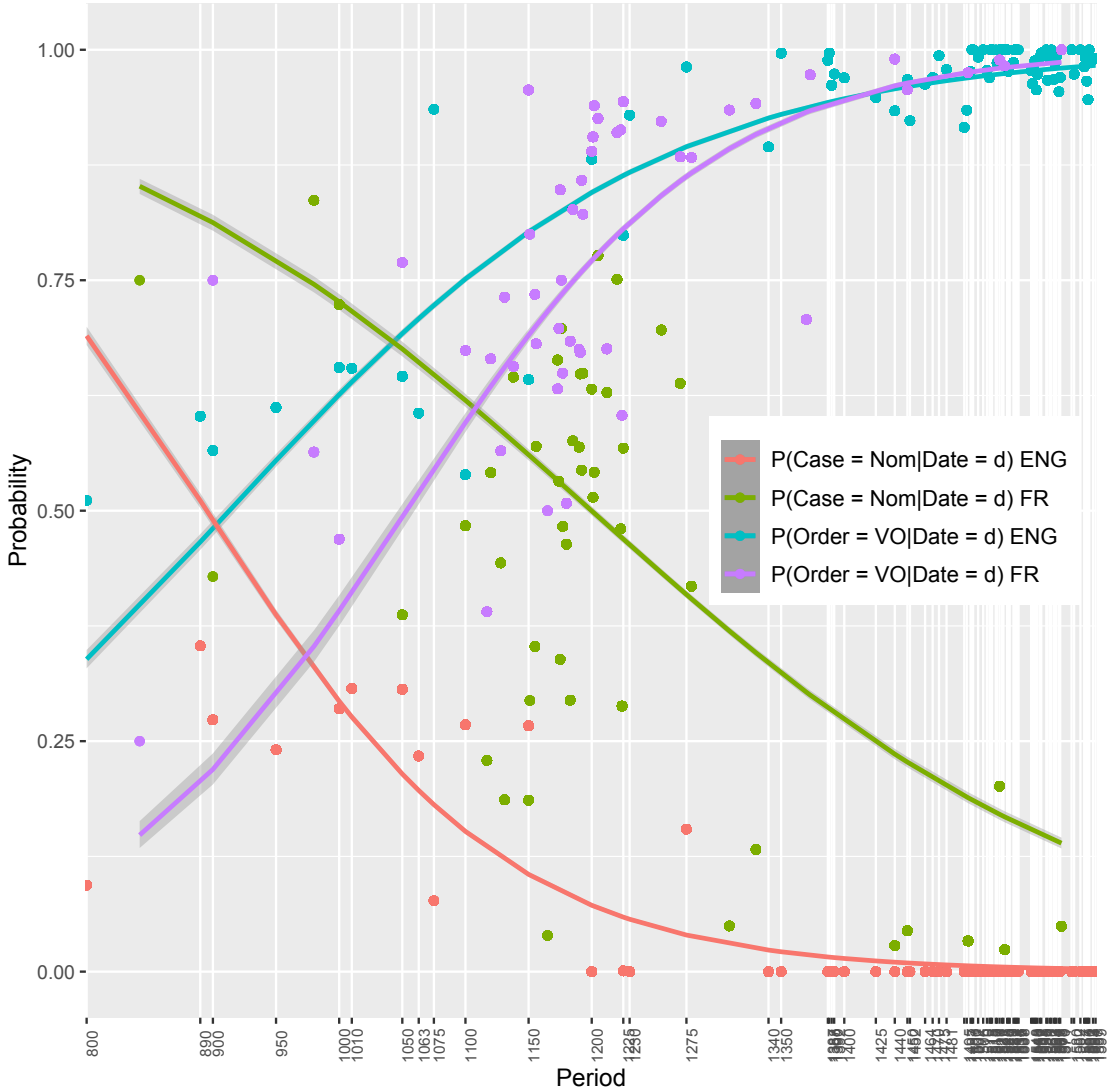


English: (Kendall's) correlation?

A significant negative correlation between case and VO frequencies.

- At the level of individual text frequencies: $\tau = -0.47$, $p < 0.001$;
- At the level of 50 year bin frequencies ("generations"): $\tau = -0.5$, $p = 0.02$.

French and English: comparison



Outline

Outline

"Correlative drift"

- There is a negative correlation between the estimated probabilities of case marking and VO order.

(Sapir, 1921, 178-181) "as the inflected forms of English became scantier, as the syntactic relations were more and more inadequately expressed by the forms of the words themselves, position in the sentence gradually took over functions originally foreign to it. ... The drift toward the abolition of most case distinctions and the correlative drift toward position as an all-important grammatical method."

How can this relationship be captured?

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

erc European Research Council

"Correlative drift"

Proposal:

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

"META-GRAMMAR": A GAME-THEORETIC MODEL

Core insight: the probability of nominative-marked subjects is related to the probability of (the grammar allowing for) object scrambling via a *surprisal-penalty*.[5]

Plan:

- implement this relation is a game-theoretical setup;
- compare the results with the empirical rates of VO.

[5] The more rare a marker is, the more its use is penalized ($\log_2 \frac{1}{p}$). Thanks to Joel Wallenberg for the suggestion.

"META-GRAMMAR": A GAME-THEORETIC MODEL

- To express a given meaning, the Speaker picks a sentence stochastically, deriving probabilities from their model of the Listener (a Bayesian inference over a probability distribution over sentences corresponding to the meaning in question).
- A Pragmatic Speaker, when assigning probs to utterances u_1 ... u_n to convey an information state s_i, takes into account how **likely** the Listener is to get s_i upon hearing a given u_j, and how costly u_j is.

(4) $U_S(u_j; s_i) = \mathbf{log}\ \mathbf{L}(\mathbf{s}_i \mid \mathbf{u}_j) - C(u_j)$ Scontras et al. (2017)

- The Speaker is rational and wants to maximize the probability of correct interpretation (minimizes ambiguity).
- The nominative marking comes with a surprisal cost, inversely proportional to its frequency (estimated from the corpus data).

"META-GRAMMAR": A GAME-THEORETIC MODEL

Meanings: {"X is the object", "X is the subject"}

Sentences that can mean "X is the object":[5]

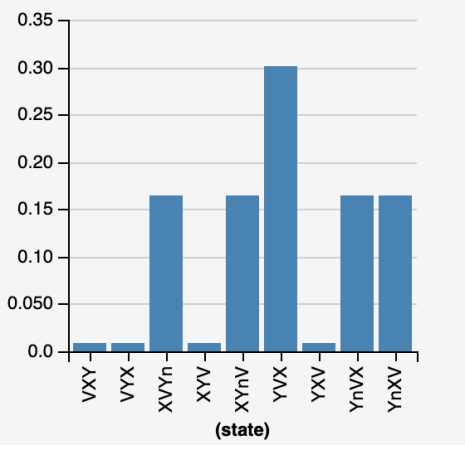
	SENTENCE	LITERAL MEANING	GRAMMAR
OV	$XY_{nom}V$	X is the object	"Scrambl."
	$Y_{nom}XV$	X is the object	"Scrambl."
	XVY_{nom}	X is the object	"Scrambl."
	YXV	X is an object or X is the subject	"Scrambl."
	XYV	X is an object or X is the subject	"Scrambl."
VO	VXY	X is an object or X is the subject	"Scrambl."
	VYX	X is an object or X is the subject	"Scrambl."
	$Y_{nom}VX$	X is the object	"Scrambl."/"SVO"
	YVX	X is the object	"Scrambl."/"SVO"

[5] "X is the subject" is expressed symmetrically.

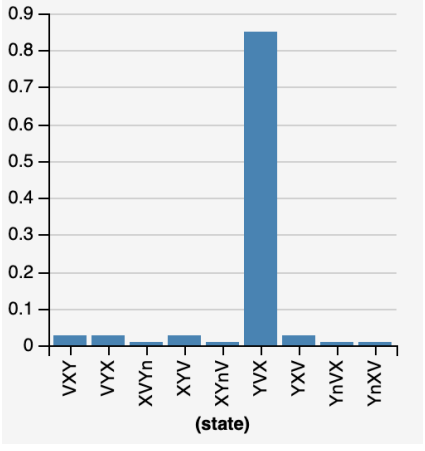
"META-GRAMMAR": A GAME-THEORETIC MODEL

WebPPL by Goodman and Stuhlmüller (2014)

Expressing "X is the object" when nom marking Freq = 31%



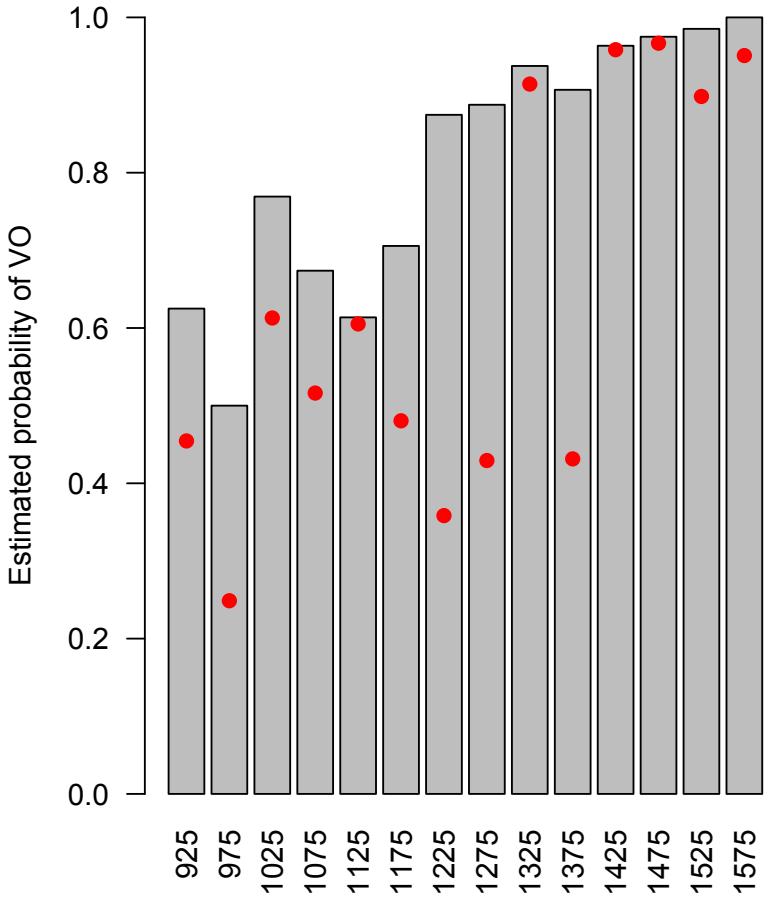
Expressing "X is the object" when nom marking Freq = 4%

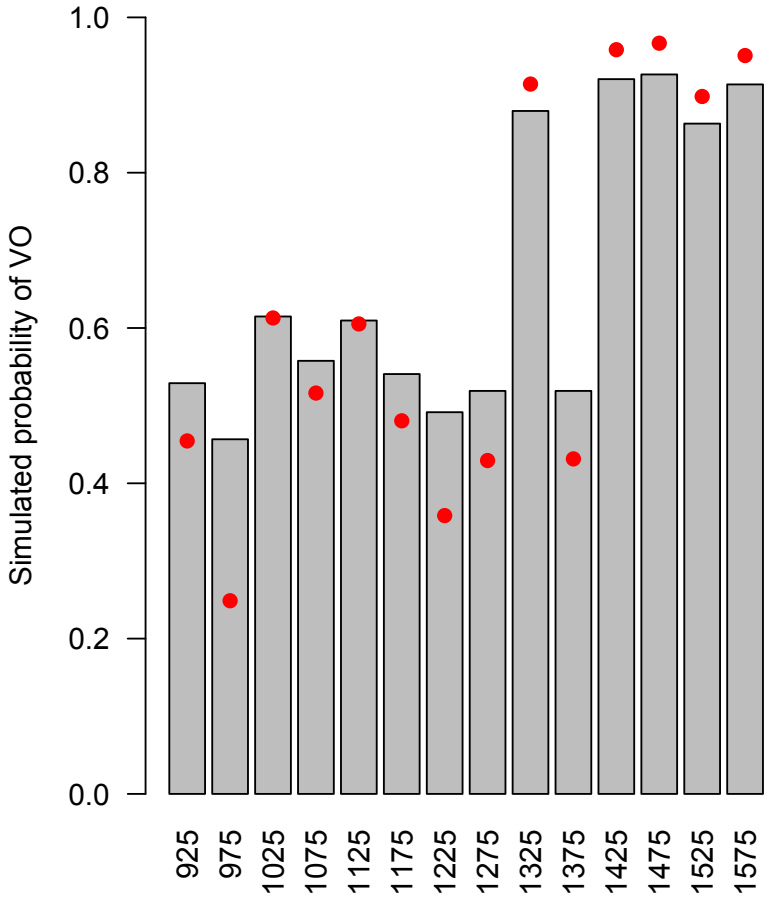


Outline

GT SIMULATIONS: FRENCH

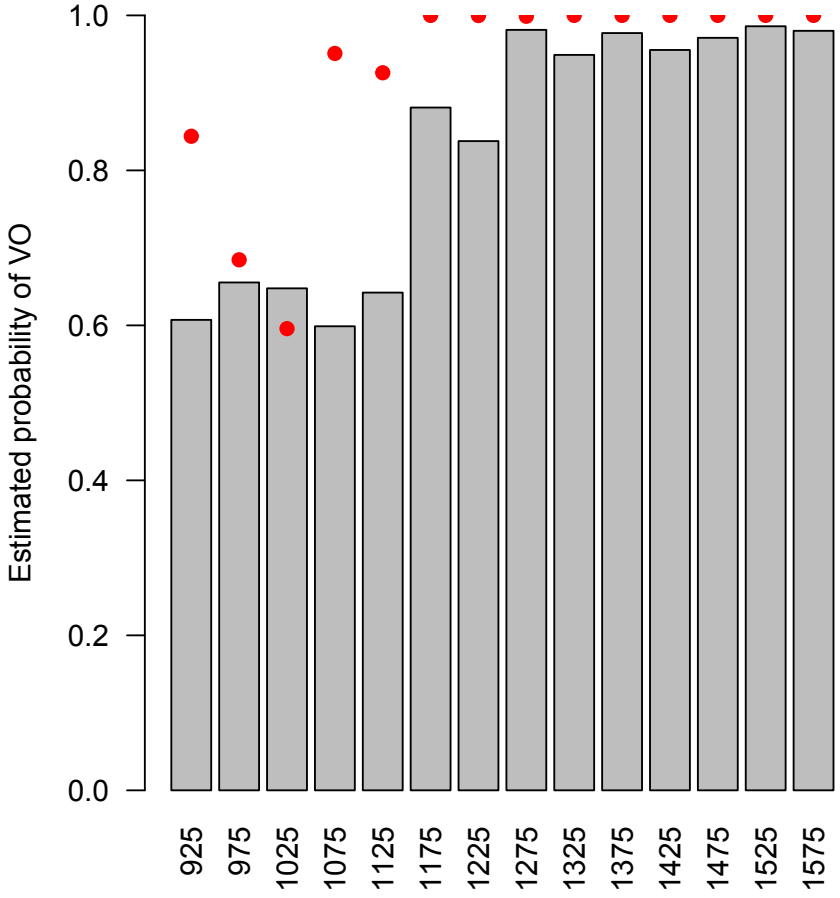
Kendall's correlation between empirical vs. simulated VO frequencies: tau = 0.46, p = 0.02

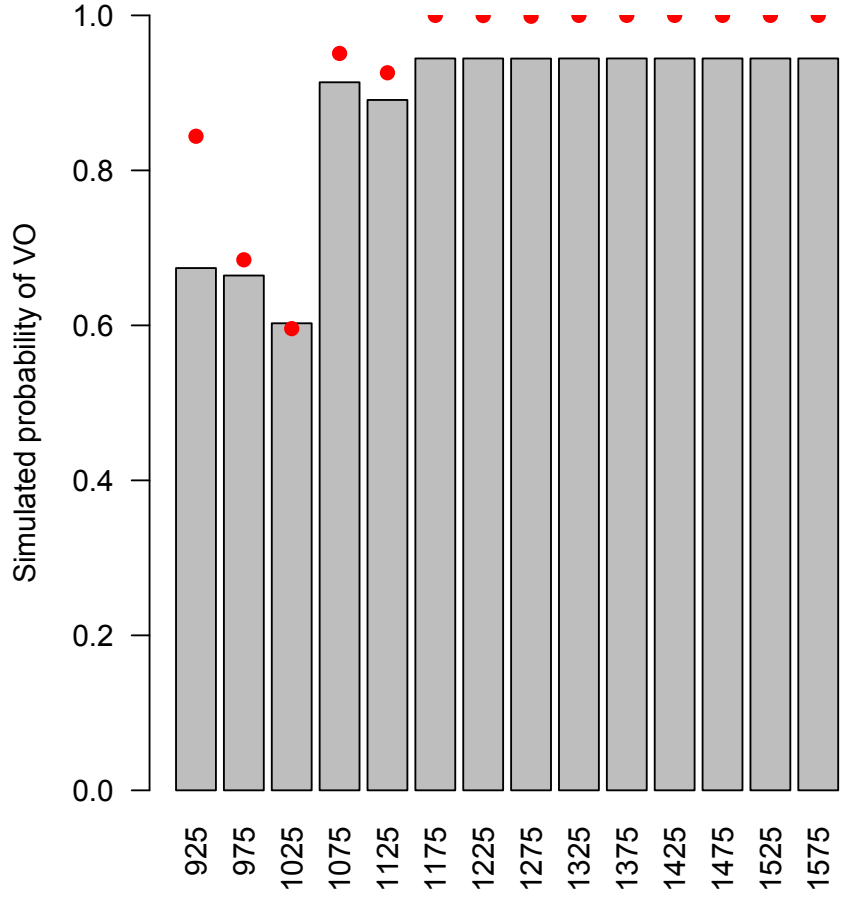




GT SIMULATIONS: ENGLISH

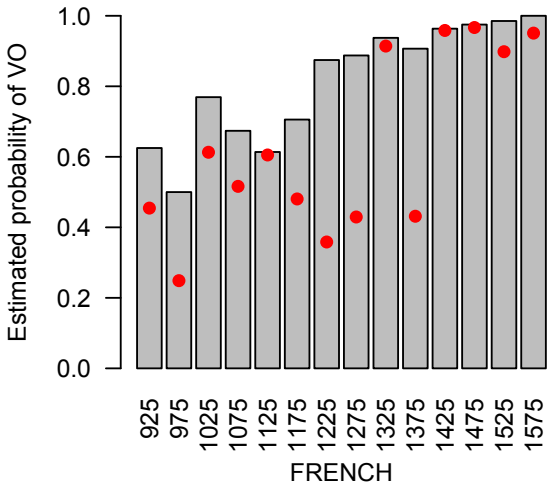
Kendall's correlation between empirical vs. simulated VO frequencies: tau = 0.5, p = 0.02

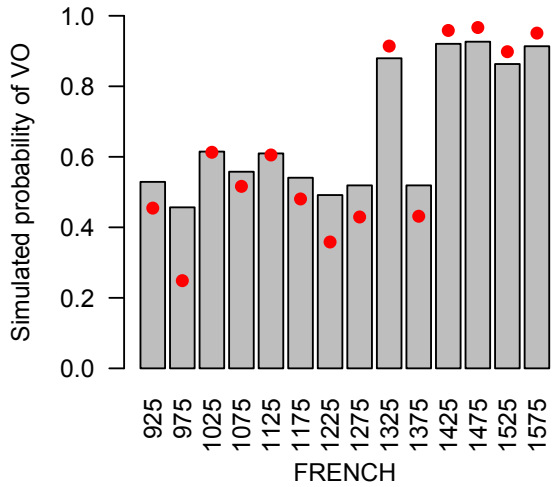


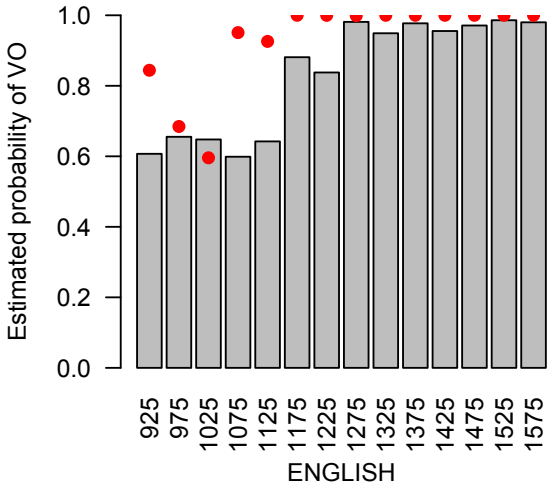


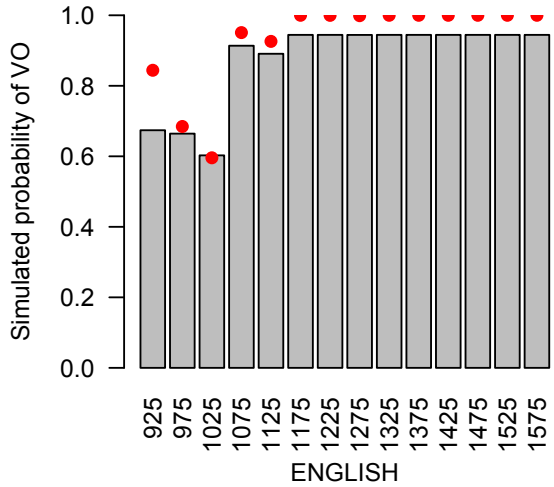
Outline

GT simulations: French and English









Conclusions

Capturing Case-Order universals via language change perspective:

The disappearance of morphological case brings about "a change in the fitness landscape for another strategy" (Jäger, 2007, 93), i.e. for the "Scrambling"-grammar.

- Alternatively, we can conclude that significant correlations in empirical and simulated results for two languages are coincidences, accidentally related to a robust typological pattern.

Berdičevskis, Aleksandrs, and Alexander Piperski. 2020. Corpus evidence for word order freezing in Russian and German. In *Proceedings of the Fourth Workshop on Universal Dependencies (UDW 2020)*, 26–33.

Bouma, Gerlof. 2011. Production and comprehension in context. *Bidirectional optimality theory* 180:169.

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`.

Erlewine, Michael Yoshitaka, and Joey Lim. 2023. Deriving scrambling with word order freezing. *Glossa: a journal of general linguistics* 8.

Gibson, Edward, Steven Piantadosi, Kimberly Brink, Leon Bergen, Eunice Lim, and Rebecca Saxe. 2013. A noisy-channel account of crosslinguistic word-order variation. *Psychological science* 24:1079–1088.

Goodman, Noah D, and Andreas Stuhlmüller. 2014. The Design and Implementation of Probabilistic Programming Languages. `http://dippl.org`. Accessed: 2023-6-30.

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, Beatrice Santorini, and Lauren Delfs. 2004. URL https://www.ling.upenn.edu/ppche/ppche-release-2016/PPCEME-RELEASE-3, The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME). Department of Linguistics, University of Pennsylvania.

Kroch, Anthony, Ann Taylor, and Beatrice Santorini. 2000. URL https://www.ling.upenn.edu/ppche/ppche-release-2016/PPCME2-RELEASE-4, The Penn-Helsinki Parsed Corpus of Middle English (PPCME2). Department of Linguistics, University of Pennsylvania. CD-ROM, second edition.

Labov, William. 2007. Transmission and diffusion. *Language* 83:344–387.

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*.

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.

Sapir, Edward. 1921. *Language, an introduction to the study of speech*. New York: Harcourt, Brace and Co.

Scontras, Gregory, Michael Henry Tessler, and Michael Franke. 2017. Probabilistic language understanding: An introduction to the Rational Speech Act framework. Technical report, https://michael-franke.github.io/probLang.

Sinnemäki, Kaius. 2008. Complexity trade-offs in core argument marking. *Language complexity: Typology, contact, change* .

Sinnemäki, Kaius. 2010. Word order in zero-marking languages. *Studies in Language* 34:869–912.

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.

Yamashita, Hiroko. 1997. The effects of word-order and case marking information on the processing of japanese. *Journal of Psycholinguistic Research* 26:163–188.

Yamashita, Hiroko. 2000. Structural computation and the role of morphological markings in the processing of Japanese. *Language and speech* 43:429–455.