UNIVERSITY OF BIRMINGHAM | COLLEGE OF ARTS AND LAW

Comparative sociolinguistics beyond the vernacular

Applying variationist methods to genre variation in written English

Jason Grafmiller

j.grafmiller@bham.ac.uk

ISLE 5, London

Variationist approach to grammar

- Examines the probabilistic constraints on individuals' linguistic choices: "alternate ways of saying 'the same' thing" (Labov 1972: 188)
  - focus mainly on spoken vernacular data from sociolinguistic interviews
- Internal constraints on variation are generally thought to be independent of stylistic factors (e.g. Labov 2010: 265; Rickford 2014: 601)
- Stylistic effects largely manifest as differences in variant frequencies rather than differences in linguistic (internal) constraints

Different constraints ⟶ Different grammars

"In the community-grammar, variable rule model that I'm endorsing, altering constraint effects beyond minor statistical differences would mean effectively adopting a different grammar […]

What varies from speaker to speaker, and from moment to moment in stylistic practice and bricolage, is…the overall rate of use of a variant […]

[…] using different constraint effects stylistically will be equivalent to diglossic or bilingual behavior, rather than simple stylizing within one language."

Greg Guy (*NWAV*, 2015)

Scope of stylistic variation

Style: any variety of a language that is associated with a particular topic, function, or social/situational context

- encompasses variability across all speaking **and writing** practices within an individual's repertoire
- stylistic variation in writing shares many properties with complex style-shifting in speech, e.g.
  - 'situational' vs. 'metaphorical' shifting (e.g. Rickford 2014)
  - audience design, persona, stance, etc. (Bell 1984; Eckert 2000; Kiesling 2009)

Questions for today

1. To what extent are internal constraints sensitive to stylistic variation?
   a) do we find genre-specific changes in constraint effects?
   b) if so, can we explain these differences?

2. How can we adapt variationist methods to tackle stylistic variation among multiple variables?

Study: English genitive alternation

(1) the best interest of both governments [*of*-genitive]
(2) both governments' best interest [*s*-genitive]

Very well-researched phenomenon (Rosenbach 2014)

- relatively stable across regional varieties (Heller et al. 2017)
- historically quite variable (Wolk et al. 2013)

  ⇨parallel increase in use of *s*-genitives across vernacular speech and newspaper writing in U.S. English (Biber 2003; Hundt & Mair 1999)

- variation in other written styles not so well-studied (cf. Jankowski 2013; Grafmiller 2014)

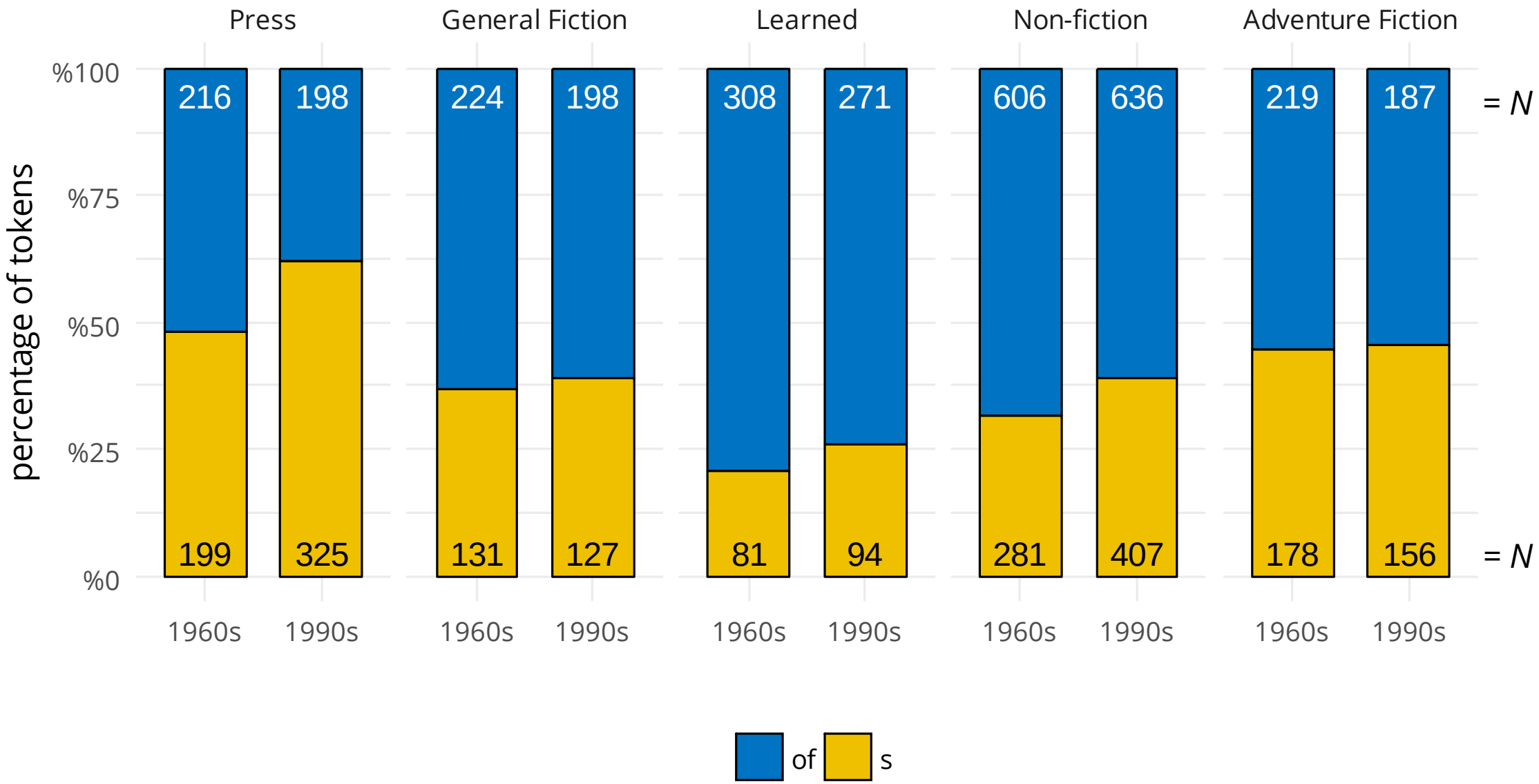
Present dataset (*N* = 5096)

Focus on 5 genres of US English from the 1960s (Brown) & 1990s (Frown)

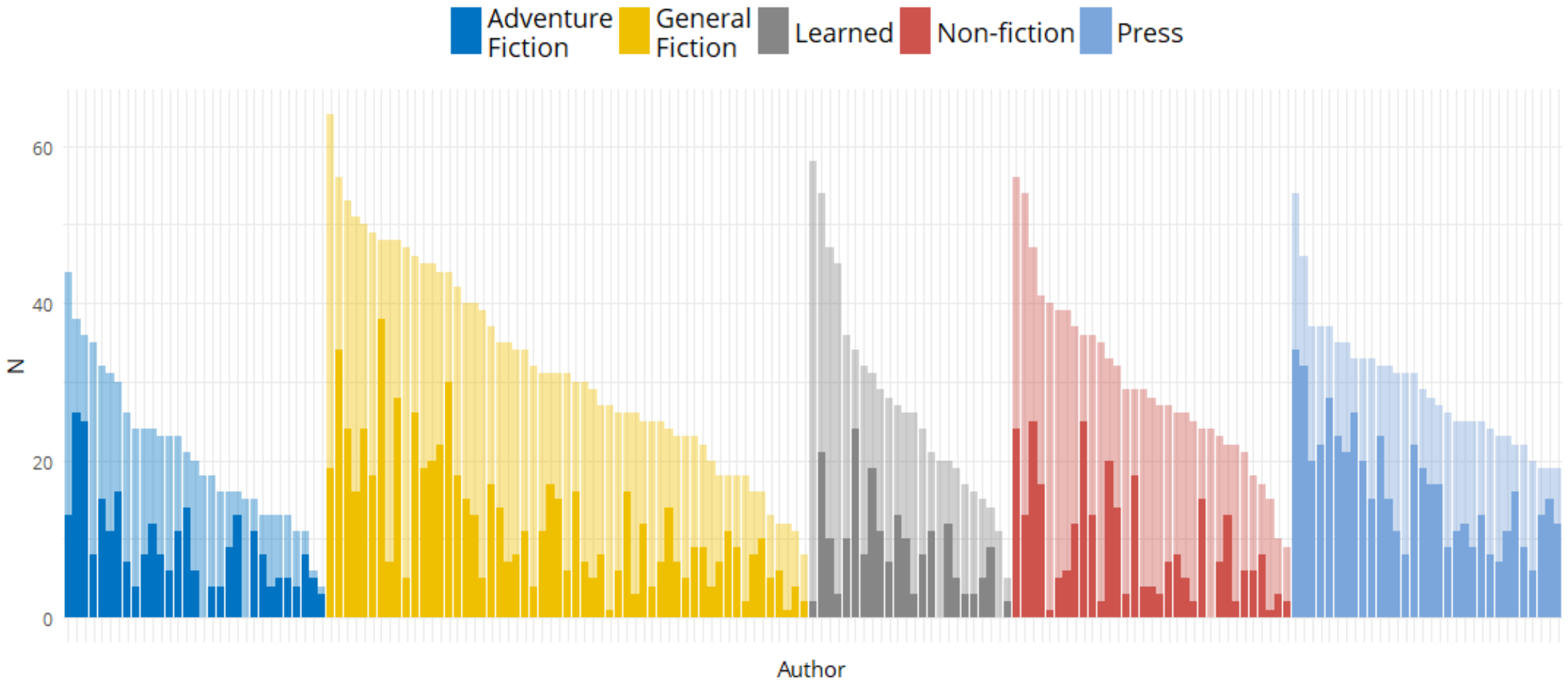
- Press: newspaper reportage (A)
- Learned: academic books & papers (J)
- Non-fiction: memoirs, biographies, letters (G)
- General fiction: 'literary' works (K)
- Adventure fiction: e.g. westerns (N)

Extract all instances of interchangeable genitives (Rosenbach 2002; Heller 2018)

Distribution of genitives by time and genre



By-author (*N* = 177) variability in Brown/Frown



Light bars = total number of tokens; **Dark bars** = number of *s*-genitives

Factors coded for

- Possessor animacy (animate vs. inanimate)
- Possessor/Possessum length (number of words)
- Possessor NP type (proper N vs. common N)
- Possessor ends in a sibilant? (yes / no)
- Possessor givenness (given vs. new)
- Lexical density of local context (type-token ratio)
- Semantics (prototypical vs. non-prototypical)
- Prior genitive (*s*-gen vs. other)

Comparative Sociolinguistic Method

Adapt 3 ‘lines of evidence’ (Poplack & Tagliamonte 2001; Taglimonte 2013)

Looking across the genres individually…

1. How are constraints ranked in terms of overall explanatory power?
2. What is the strength and ordering of the levels within the constraints (the size and direction of the effects)?
3. Which constraints are significant?

Analysis: Assessing explanatory importance

Do certain constraints vary across genres in their relative importance?

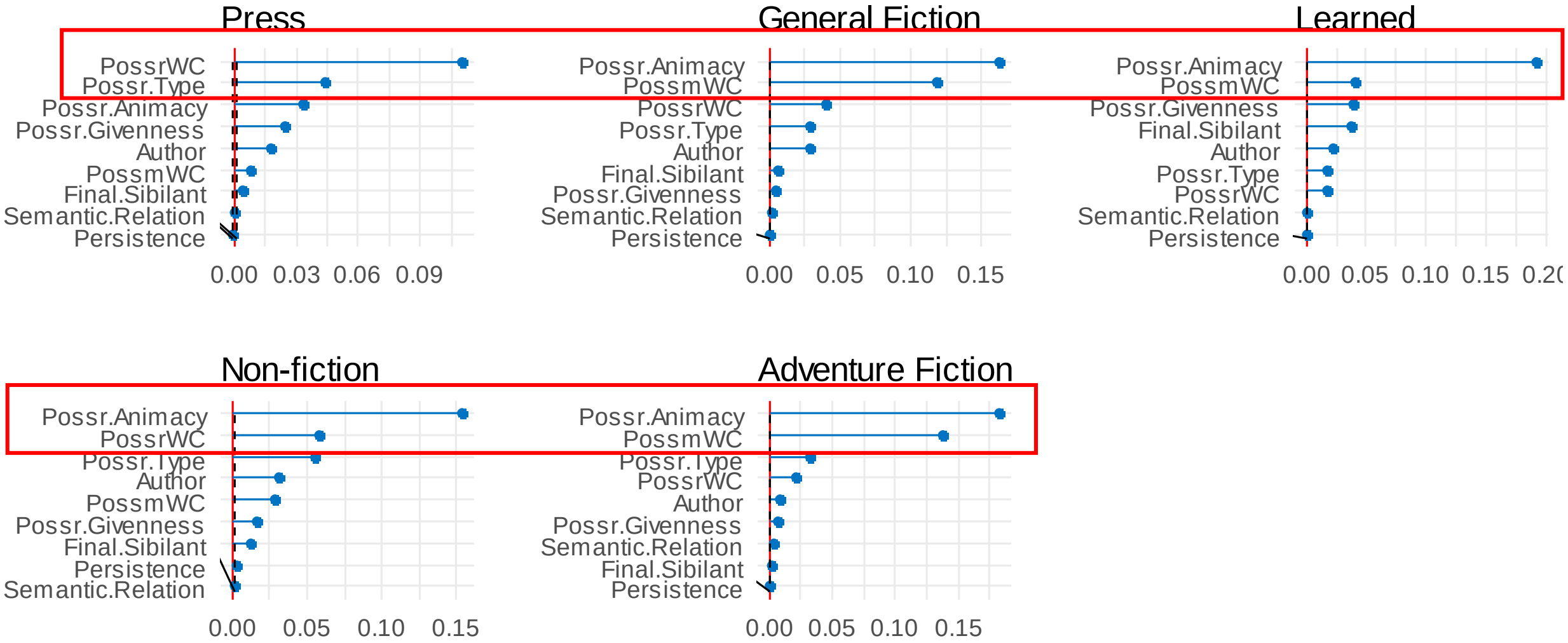
- compute variable importance rankings for individual genre models
  - compare accuracy of original model to model with predictor randomly permuted
  - different rankings reflect different degrees of constraint importance

Model specs

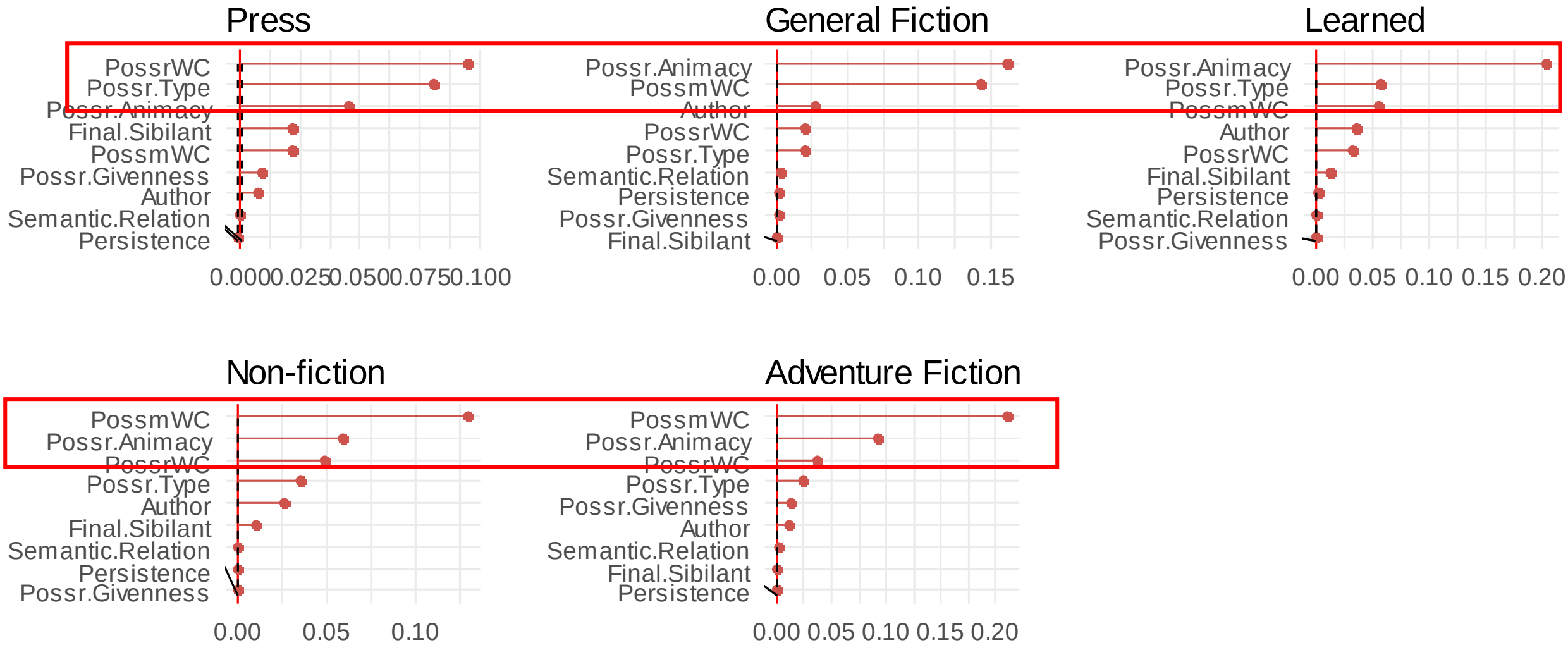
- Random forest variable importance
  - 10 models: one for each genre in each time period
  - compute permutation variable importance measures for predictors in each model

```
Type ~ Author + PossrAnimacy + PossrLength +
   PossmLength + PossrGiven + PossrNP + FinalSibilant
   + SemanticRelation + TypeTokenRatio
```

Constraint ranking: 1960s



Constraint ranking: 1990s



Analysis: Assessing strength and direction

Do certain constraints vary across genres in the strength and/or direction of their effects?

- measure constraint effect sizes and direction in each genre using regression coefficients
  - use coefficients to generate probabilistic distance measures between genres
  - visualize distances with multidimensional scaling maps, neighbor nets, etc.

Model specs

- Mixed-effects logistic regression[1]
  - 10 models: one for each genre in each time period
  - by-author intercepts and slopes for Possessor Animacy

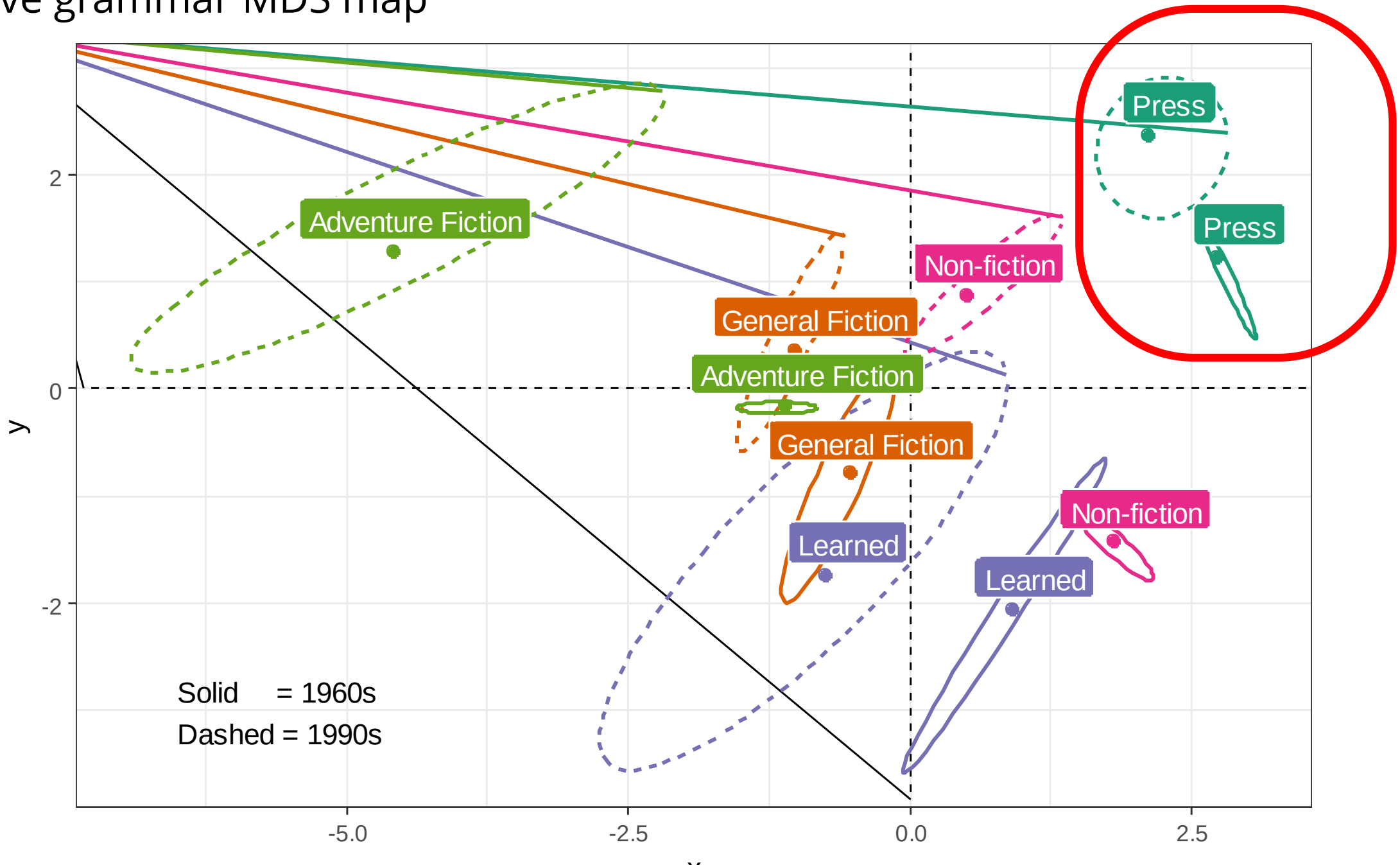
```
Type ~ (1|Author) + (0 + PossrAnimacy|Author) +
   PossrAnimacy + PossrLength + PossmLength +
   PossrGiven + PossrNP + FinalSibilant +
   SemanticRelation + TypeTokenRatio
```

[1]Bayesian models using `brms` package with standardized predictors and weakly informative priors

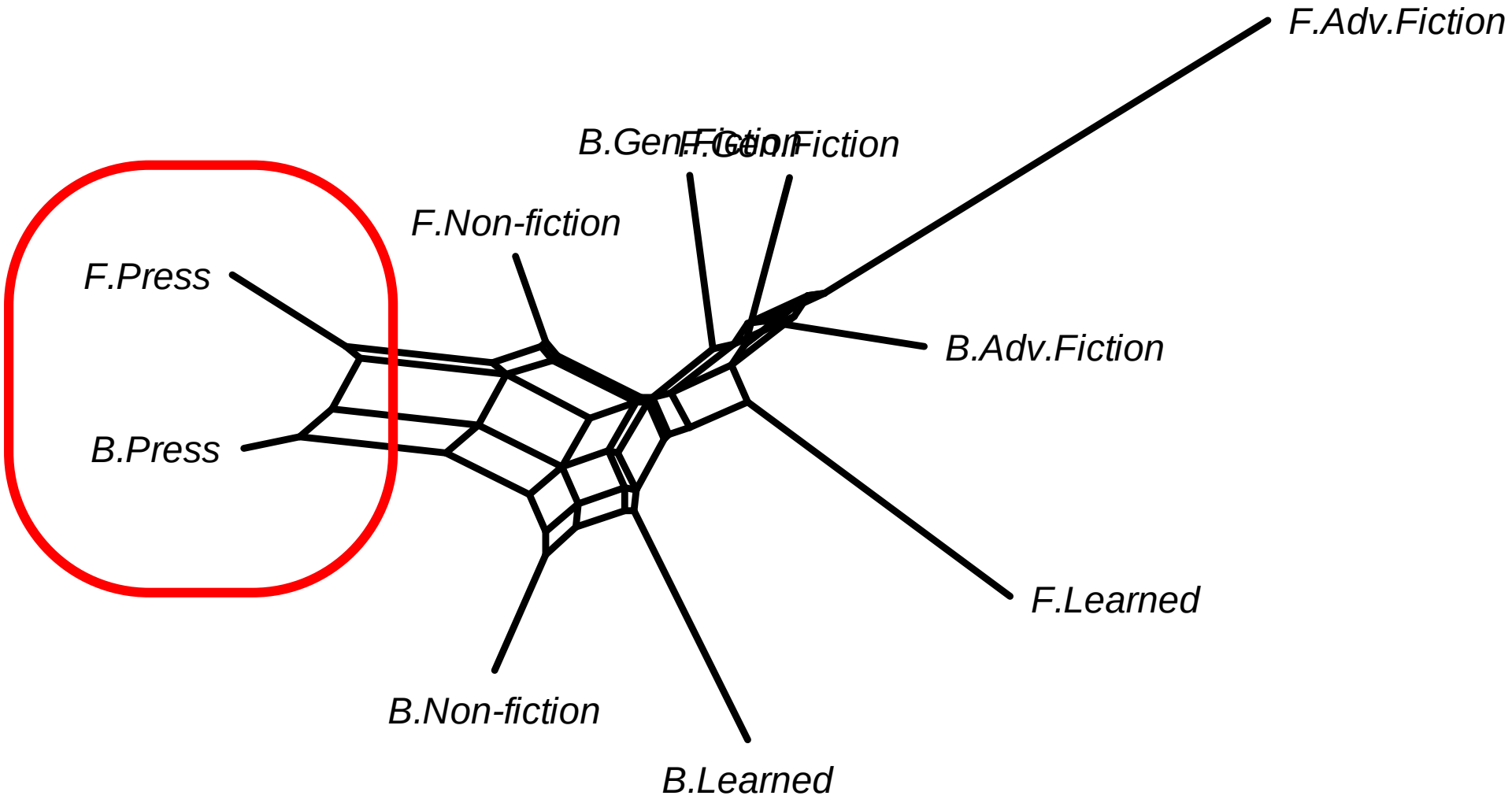
Model coefficients

	Brown (1960s)					Frown (1990s)				
	Press	Gen. Fiction	Learned	Non-fiction	Adv. Fiction	Press	Gen. Fiction	Learned	Non-fiction	Adv. Fiction
Intercept	-0.27	-1.88	-3.15	-2.07	-0.66	0.86	-1.36	-2.73	-0.94	-0.61
Possr = Animate	1.40	3.96	3.31	3.26	4.99	1.31	3.52	4.34	2.15	3.68
Possr length	-3.26	-4.42	-2.69	-3.78	-3.43	-2.31	-1.69	-2.71	-2.55	-4.88
Possm length	0.21	-2.04	-0.75	0.40	-2.50	-0.77	-3.37	-2.35	-2.08	-6.35
Final Sibilant = Y	-0.74	-1.02	-2.55	-1.88	-1.27	-1.27	-1.26	-2.77	-1.45	-2.44
Possr = Given	-0.96	-1.24	-2.33	-0.85	-1.85	-0.84	-0.48	0.89	0.04	-2.05
Possr = Proper N	1.07	0.49	0.01	1.07	0.83	1.28	1.28	2.84	1.09	1.32
Prior = s-genitive	0.04	-0.35	-0.43	0.70	0.12	0.06	0.69	0.47	0.37	0.66
Semantics = Proto	0.16	1.35	0.25	0.72	-0.01	-0.26	0.67	-0.38	0.28	0.78

Genitive grammar MDS map



Genitive grammar neighborNet



Colloqualization or economization?

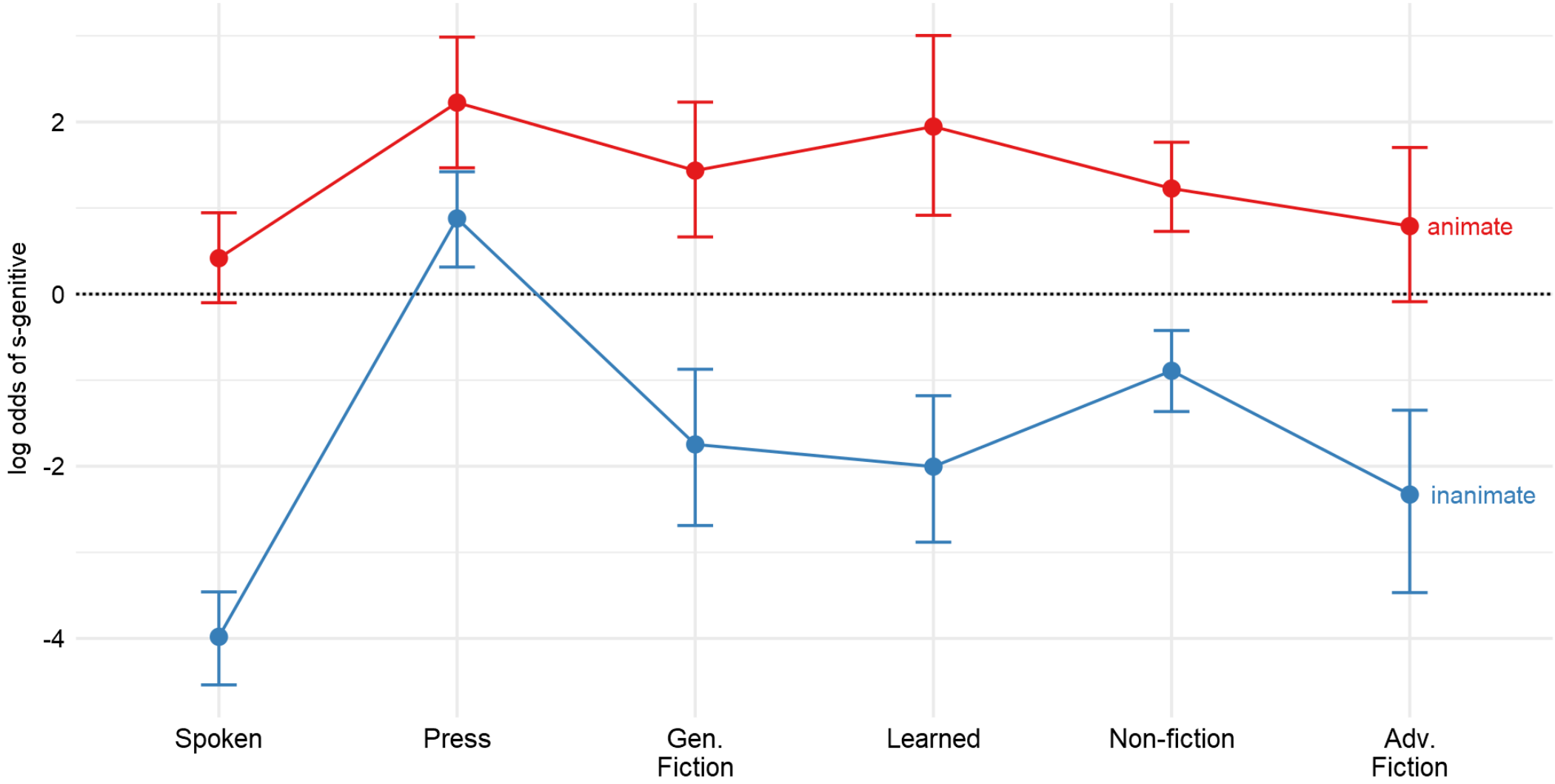
Colloquialization

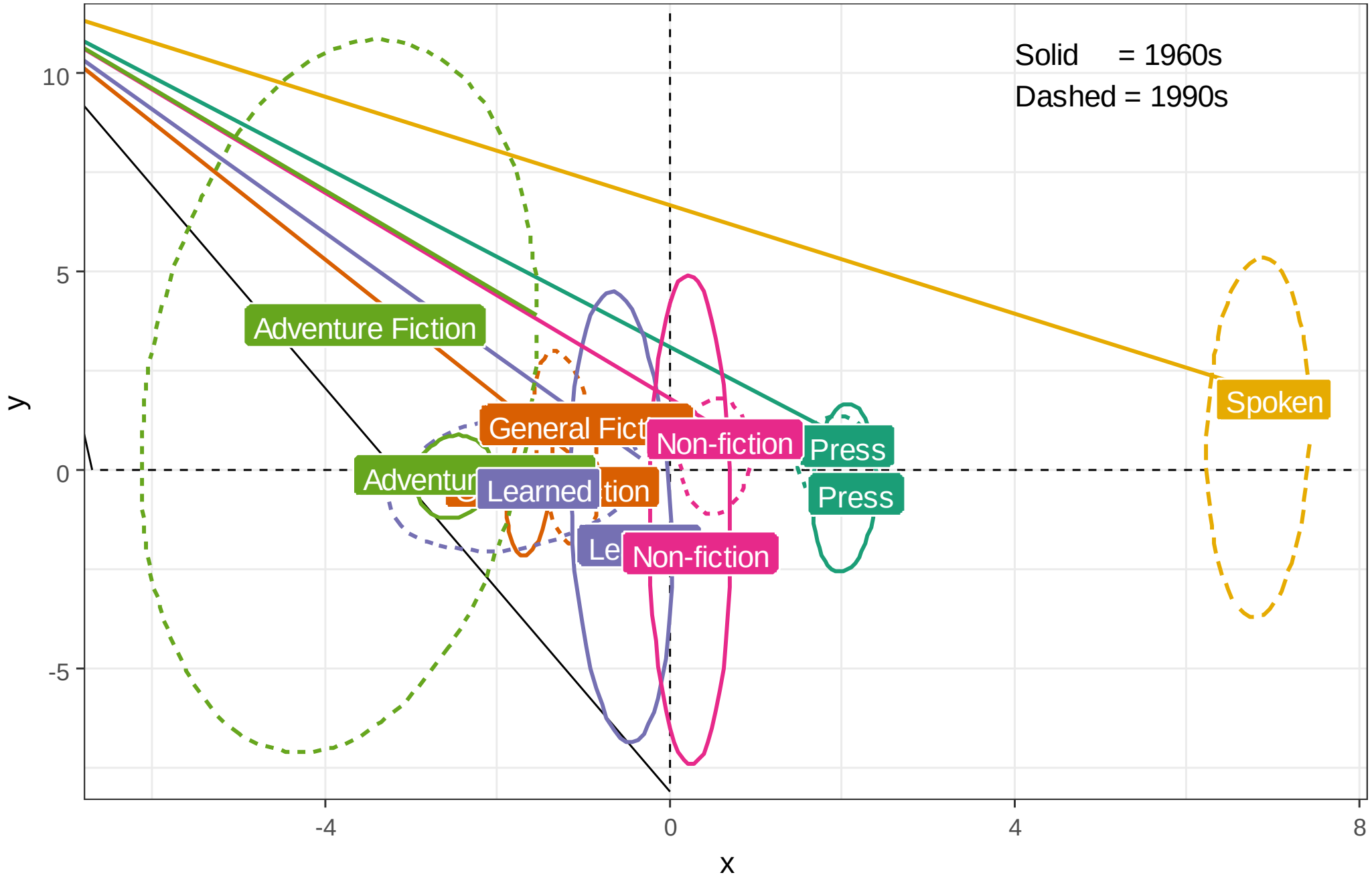
- s-genitives have been increasing spoken U.S. English over the late 20th century (Biber & Finegan 1989; Rosenbach 2002)
- News texts are becoming more overall more colloquial and conversational (Hundt & Mair 1999; Rühlemann & Hilpert 2017)

Economization

- weak effect of animacy is the result of pressures on journalists to write more economically (Biber 2003; Hinrichs & Szmrecsanyi 2007)
- *s*-genitives are more compact, thus preferred when writers need to minimize text length and maximize information content

Animacy effects across genres and speech




Solid = 1960s
Dashed = 1990s
Adventure Fiction
General Fict
Non-fiction
Press
Adventur
Learned
tion
Press
Le
Non-fiction
Spoken
10
5
0
-5
-4
0
4
8
y
x

Stylistic covariation

- Style is not just about 1 variable
- Want to examine covariation among several variables, à la Biber & colleagues
- Take same approach to measuring cross-varietal grammars as Szmrecsanyi et al. (earlier today)

Intend to examine 4 variables in Brown corpora

Four variables exhibiting decreasing sociolinguistic awareness…

1. restrictive relativizer alternation (*that/which/ZERO*)
2. particle placement (*wrap this talk up/wrap up this talk*)
3. dative alternation (*give me the book/give the book to me*)
4. genitive alternation

Systematic covariation

- Does combined analysis of several variables identify reasonable groupings?
  - e.g. different variants associated with varying degrees of formality;
  - *of*-genitives, *which* relativizers, *to* datives(?) all index formality

- But are these variations conditioned in the same way across genres?
  - probabilistic distances measure **covariation in the underlying variable grammars, not just the relative frequencies of variants**

Challenges

- Collecting data
  - carving up datasets into small subsets (severely) limits analysis
  - what other kinds of data or phenomena can we use?
- Choosing variables
  - focus on variables above/below level of consciousness
- Incorporating phonetic/phonological variables
  - in principle, we can combine patterns of any kind of variables to examine stylistic patterns

Thank You!

Contact: j.grafmiller@bham.ac.uk

Data, code & slides: https://osf.io/tkfnc/

References

- Biber, Douglas. 2003. Compressed noun-phrase structures in newspaper discourse. In Jean Aitchison & Diana M. Lewis (eds.), *New Media Language*, 169–181. London: Routledge.
- Burnham, K. P. & D. R. Anderson. 2014. P values are only an index to evidence: 20th- vs. 21st-century statistical science. *Ecology* 95(3). 627–630.
- Forrest, Jon. 2015. Community rules and speaker behavior: Individual adherence to group constraints on (ING). *Language Variation and Change* 27(3). 377–406. doi:10/gdjzpk.
- Grafmiller, Jason. 2014. Variation in English genitives across modality and genres. *English Language and Linguistics* 18(3). 471–496. doi:10.1017/S1360674314000136.
- Guy, Gregory R. 2015. Coherence, constraints and quantities. *Paper presented at NWAV 44*. Toronto. October 23.
- Heller, Benedikt, Benedikt Szmrecsanyi & Jason Grafmiller. 2017. Stability and Fluidity in Syntactic Variation World-Wide: The Genitive Alternation Across Varieties of English. *Journal of English Linguistics* 45(1). 3–27. doi:10.1177/0075424216685405.
- Heller, Benedikt. 2018. *Stability and Fluidity in Syntactic Variation World-wide*. Leuven, Belgium: KU Leuven Ph.D. Thesis.

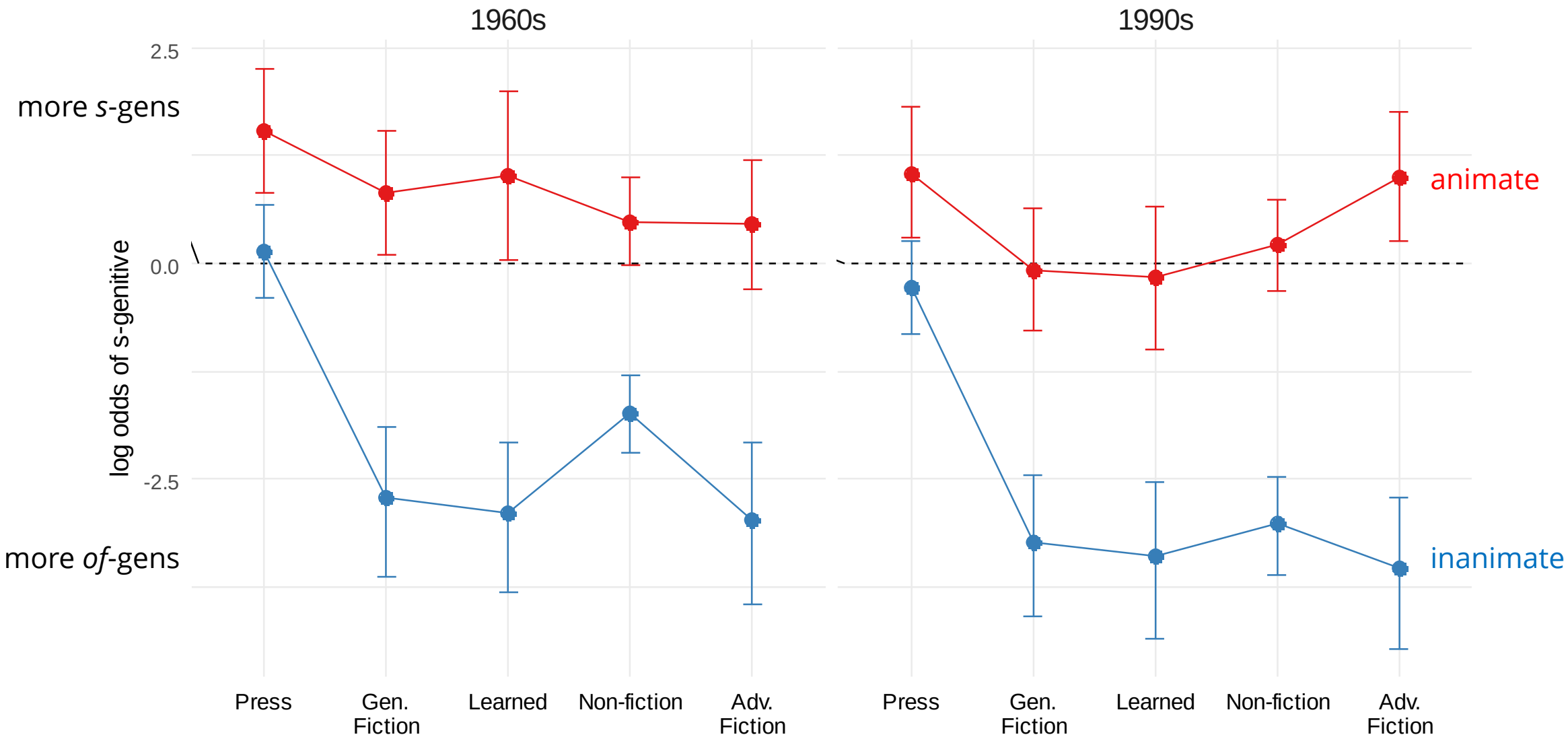
References cont.

- Hinrichs, Lars & Benedikt Szmrecsanyi. 2007. Recent changes in the function and frequency of Standard English genitive constructions: A multivariate analysis of tagged corpora. *English Language and Linguistics* 11(3). 437–474. doi::10.1017/S1360674307002341.
- Hundt, Marianne & Christian Mair. 1999. 'Agile' and 'uptight' genres: The corpus-based approach to language change in progress. *International Journal of Corpus Linguistics* 4. 221–242.
- Jankowski, Bridget L. & Sali A. Tagliamonte. 2014. On the genitive's trail: data and method from a sociolinguistic perspective. *English Language and Linguistics* 18(2). 305–329. doi:10.1017/S1360674314000045.
- Labov, William. 2010. *Principles of linguistic change. Vol. 3: Cognitive and cultural factors*. Malden, MA Oxford Chichester, West Sussex: Wiley-Blackwell.
- Labov, William. 2014. What is to be learned: The community as the focus of social cognition. In Martin Pütz, Justyna A. Robinson & Monika Reif (eds.), Current Topics, vol. 59, 23–51. Amsterdam: John Benjamins.
- Mair, Christian. 2006. *Twentieth-Century English: History, variation, and standardization*. Cambridge, UK ; New York: Cambridge University Press.
- Poplack, Shana & Sali Tagliamonte. 2001. *African American English in the diaspora*. Malden, MA: Blackwell.

References cont.

- Rickford, John R. 2014. Situation: Stylistic variation in sociolinguistic corpora and theory. *Language and Linguistics Compass* 8(11). 590–603. doi:10.1111/lnc3.12110.
- Rosenbach, Anette. 2002. *Genitive variation in English: Conceptual factors in synchronic and diachronic studies*. Berlin: Mouton de Gruyter.
- Rosenbach, Anette. 2007. Emerging variation: Determiner genitives and noun modifiers in English. *English Language and Linguistics* 11(1). 143–189. doi:10.1017/S1360674306002140.
- Rosenbach, Anette. 2014. English genitive variation – the state of the art. *English Language and Linguistics* 18(2). 215–262. doi:10.1017/S1360674314000021.
- Röthlisberger, Melanie. 2018. *Regional variation in probabilistic grammars: A multifactorial study of the English dative alternation*. Leuven, Belgium: KU Leuven Ph.D. Thesis.
- Tagliamonte, Sali. 2013. Comparative Sociolinguistics. In J. K. Chambers & Natalie Schilling (eds.), *Handbook of Language Variation and Change*, 130–156. 2nd ed. Chichester, UK: John Wiley & Sons Inc.

Effect of possessor animacy by Genre and Time



Inanimate possessors in Frown press