

# Variation in spoken datives and genitives in Britain, North America and the Pacific

Joint work involving many people.

# Objective: wheeling out the electron microscope

What is the extent to which **careful pairwise variety comparisons** can replicate previous findings about probabilistic differences?



# Introduction



# Syntactic variation in spoken English

- (1) a. we sent [the president]<sub>recipient</sub> [a letter]<sub>theme</sub>  
(the ditransitive dative)
- b. we sent [a letter]<sub>theme</sub> to [the president]<sub>recipient</sub>  
(the prepositional dative)
- (2) a. [the president]<sub>possessor</sub>'s [speech]<sub>possessum</sub>  
(the *s*-genitive)
- b. [the speech]<sub>possessum</sub> of [the president]<sub>possessor</sub>  
(the *of*-genitive)



# Countries/varieties

- Canada (CanE)
- New Zealand (NZE)
- Britain (BrE)
- United States (AmE)



# Dative dataset: sources

(version: January 2015)

- **CanE/BrE**: materials were collected in the UK and Canada between 1997 and 2010 according to standard Sociolinguistic procedures by Sali Tagliamonte (Tagliamonte 2014)
- **AmE**: based on data extracted from the Switchboard Corpus of American English (SWBD) (Bresnan et al. 2007; Bresnan and Ford 2010)
- **NZE**: data come from the Origins of New Zealand English corpus (ONZE) (Hay and Bresnan 2006; Bresnan and Hay 2008)



# Genitive dataset: sources

(version: September 2014)

- **BrE**: based on data extracted from the Freiburg English Dialect Corpus (FRED) (Szmrecsanyi 2006)
- **AmE**: based on data extracted from the Switchboard Corpus of American English (SWBD) (Shih et al. 2015)
- **CanE**: data collected between 1997 and 2010 according to standard sociolinguistic procedures by Sali Tagliamonte (see e.g. Jankowski and Tagliamonte 2014)
- **NZE**: data come from the Origins of New Zealand English corpus (ONZE) (Gordon et al. 2007)



# Unitary GLMMs

## (analysis conducted in April 2015)

- the usual predictors are all significant in the expected directions
- cross-variety patterns: reference level 'CanE' (most observations)
  - sig. interactions of country with the internal predictors were fairly minimal in both constructions
  - genitives: possessor animacy appears to play a greater role in genitive choice in CanE than elsewhere
  - datives: end weight effects are weaker in CanE than elsewhere





# Methods & data



# Variable context definitions

- **dative alternation**: *give* only, broadly following guidelines in Bresnan et al. (2007)
- **genitive alternation**: (mostly?/entirely?) compatible with the guidelines in Rosenbach (2014)



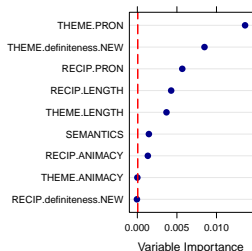
# Annotation: datives

- **SEMANTICS**: transfer of possession (e.g. *give*) vs communication (e.g. *demonstrate*) vs abstract (e.g. *wish*)
- **REC.THEME.DIFF**:  $\log(\text{recipient length}) - \log(\text{theme length})$  (number of words). Bresnan and Ford (2010) report that using a relative measure helps reduce collinearity.
- **RECIP.PRON**: pronoun versus noun
- **THEME.PRON**: pronoun versus noun
- **RECIP.definiteness.NEW**: definite versus indefinite
- **THEME.definiteness.NEW**: definite versus indefinite
- **RECIP.ANIMACY**: animate versus inanimate
- **THEME.ANIMACY**: animate versus inanimate
- **COUNTRY**: US vs UK vs CAN vs NZ
- **SPEAKER.pruned**, **THEME.HEAD.pruned**, **RECIP.HEAD.pruned**

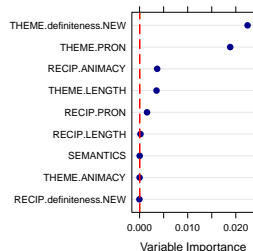


# CRF analysis: dative alternation

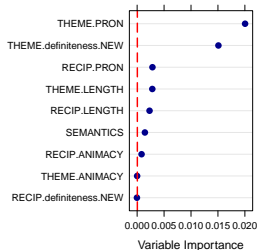
## AmE



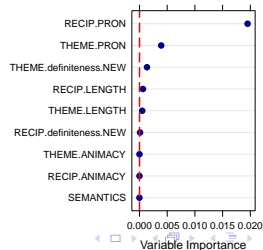
## CanE



## NZE



## BrE



# CRF analysis: dative alternation

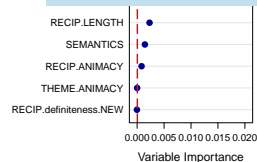
AmE

CanE



observations

- NZE/AmE fairly similar
- animacy comparatively unimportant in AmE, BrE, and NZE
- animacy comparatively important in CanE



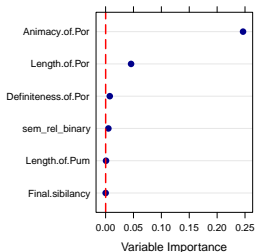
# Annotation: genitives

- **Animacy.of.Por**: animate versus inanimate versus collective versus locative versus temporal
- **Definiteness.of.Por**: definite versus definite (pronoun) versus indefinite
- **sem\_rel\_binary**: prototypical versus non-prototypical
- **Length.of.Por**: number of words, centered
- **Length.of.Pum**: number of words, centered
- **Final.sibilancy**: final sibilant present versus absent
- **COUNTRY**: US vs UK vs CAN vs NZ
- **speaker.pruned**, **Por.head.pruned**, **Pum.head.pruned**

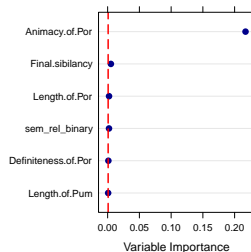


# CRF analysis: genitive alternation

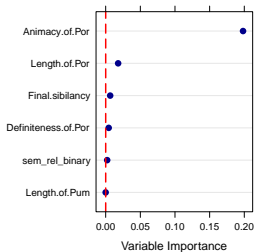
## AmE



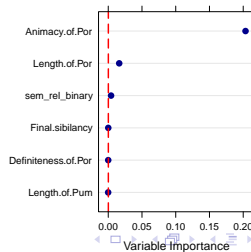
## CanE



## NZE



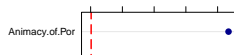
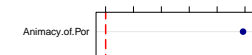
## BrE



# CRF analysis: genitive alternation

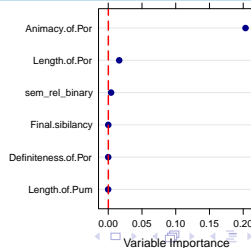
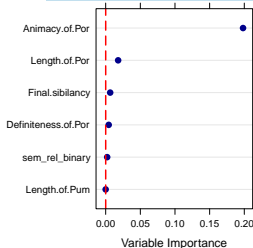
AmE

CanE



observations

- fairly similar in the bird's eye perspective
- differences wrt the extent to which p'or length plays a secondary role





# pairwise GLMMs: model fitting procedure

- Zuur et al. (2009)
- ran the full set of model diagnostics, checking residuals and random effects structure, leverage of certain items, *k*-fold cross-validation etc.
- genitives: include in full model interactions suggested by ctree
- fairly minimal ranef structure: speaker, recipient/theme/possessor/possessum head noun lemmas; intercept adjustments only; re-optimization in each pairwise comparison



# Pairwise comparisons: datives



# Predictions

- [Bresnan and Hay \(2008\)](#): “Non-animate recipients are more likely to be used in the double object construction in the NZ than in US spoken data” (252)  
failure to replicate
- [Wolk et al. \(2013\)](#): end-weight of themes is a stronger effect in AmE than in BrE  
failure to replicate
- [Tagliamonte \(2014\)](#): no substantial differences between BrE and CanE  
we do find differences



# Model quality measures: datives

	$N$	% corr. pred.	$C$	$\kappa$
CanE/BrE	2126	97.0	0.99	8.5
CanE/NZ	2018	96.9	0.98	7.7
CanE/US	2403	96.3	0.98	10.1
NZ/BrE	1808	97.7	0.99	7.7
NZ/US	2077	97.2	0.99	11.0
BrE/US	2190	96.8	0.99	8.0



# Significant cross-variety differences in regression

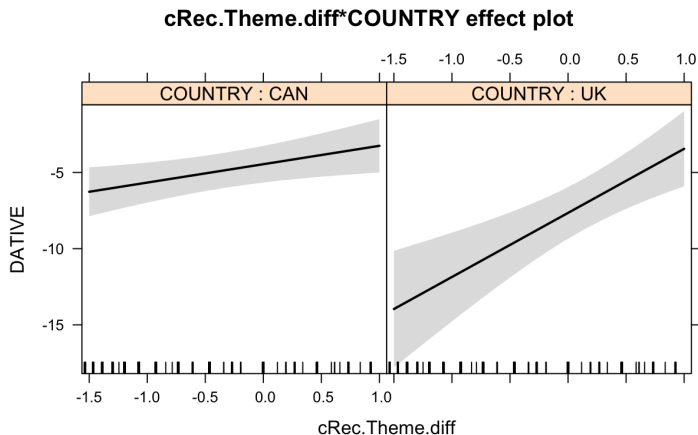
CanE/BrE	REC.THEME.DIFF ( $p = .001$ )
CanE/NZE	REC.THEME.DIFF ( $p = .009$ )
CanE/AmE	<i>nil</i>
NZE/BrE	SEMANTICS ( $p = .02$ ) RECIP.PRON ( $p = .02$ ) THEME.ANIMACY ( $p = .049$ )
NZE/AmE	SEMANTICS ( $p = .004$ )
BrE/AmE	<i>nil</i>



# CanE versus BrE: end weight differences

x-axis: standardized length difference centered around zero

y-axis: log odds (predicted outcome: PD)



CanE

summary

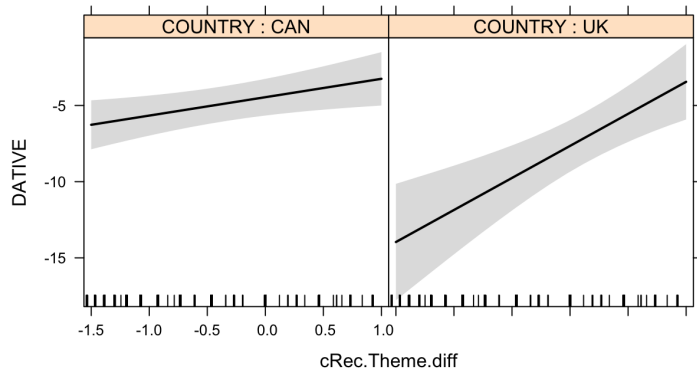
ces

x-axis: stand

y-axis: log o

in BrE, heavy recipients favor PD more strongly than in CanE

⇒ end weight stronger in BrE

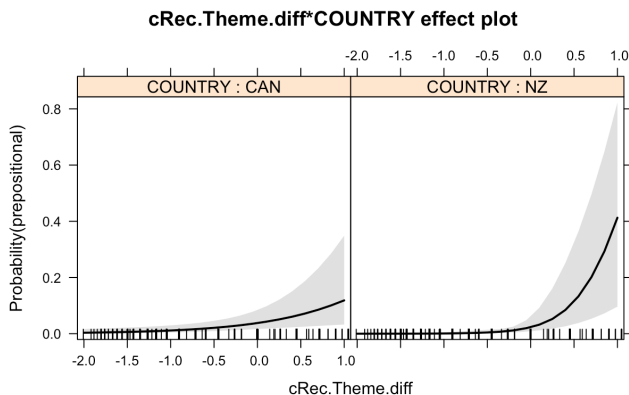


QML

# CanE versus NZE: end weight differences

x-axis: standardized length difference centered around zero

y-axis:  $p$  (predicted outcome: PD)





## summary

CanE

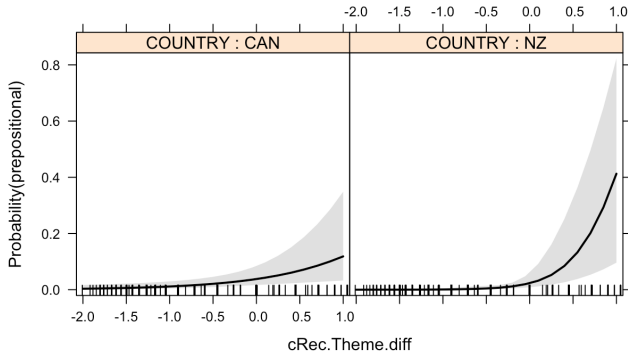
ces

x-axis: stand

y-axis:  $p$  (pro

in NZE, comparatively heavy recipients favor PD more strongly than in CanE  
 ⇨ end weight stronger in NZE

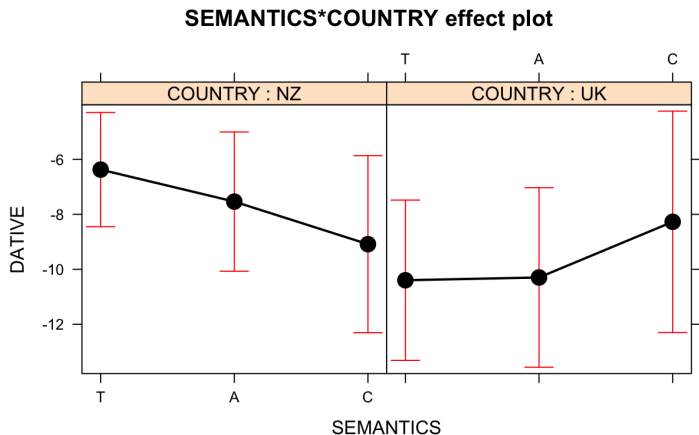
cRec.Theme.diff COUNTRY effect plot



# NZE versus BrE: semantics

x-axis: 'T': transfer, 'A': abstract, 'C': communication

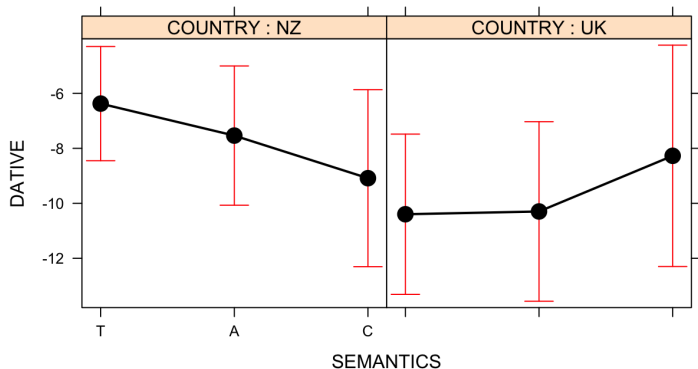
y-axis: log odds (predicted outcome: PD)



## summary

x-axis: 'T': t  
y-axis: log o

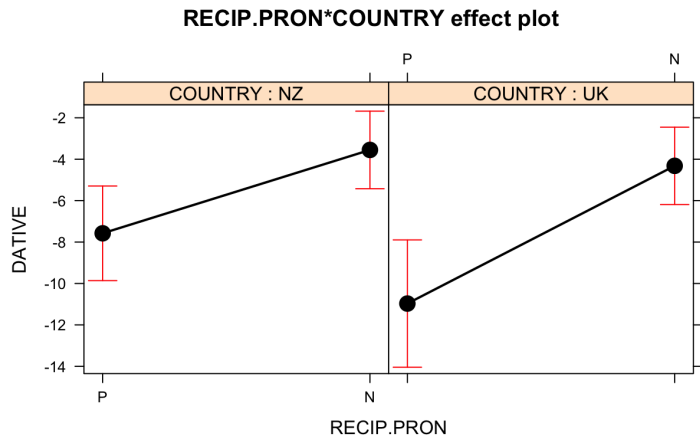
in BrE, comm tokens favor PD vis-a-vis  
transfer; vice versa in NZE  
(similar difference in NZE versus AmE)



# NZE versus BrE: recipient pronominality

x-axis: 'P': pronoun, 'N': noun

y-axis: log odds (predicted outcome: PD)

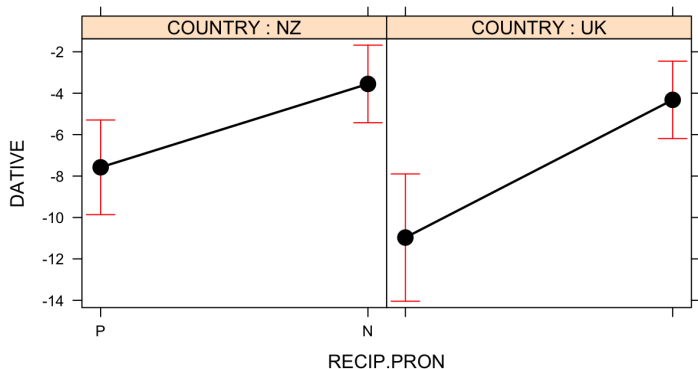


NZE

summary

x-axis: 'P': p  
y-axis: log o

in BrE, nominal recipients have a more strongly favoring effect, compared to pronominal recipients, than in NZE

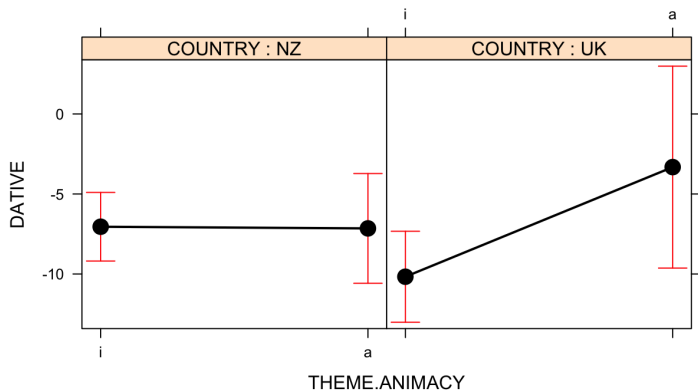


# NZE versus BrE: theme animacy

x-axis: 'i': inanimate, 'a': animate

y-axis: log odds (predicted outcome: PD)

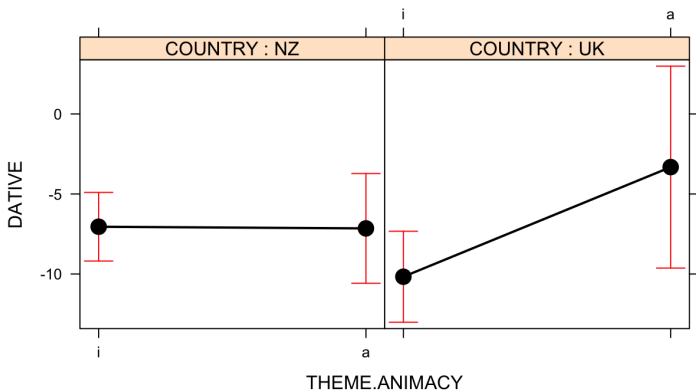
**THEME.ANIMACY\*COUNTRY effect plot**



## summary

x-axis: 'i': inanimate  
y-axis: log odds

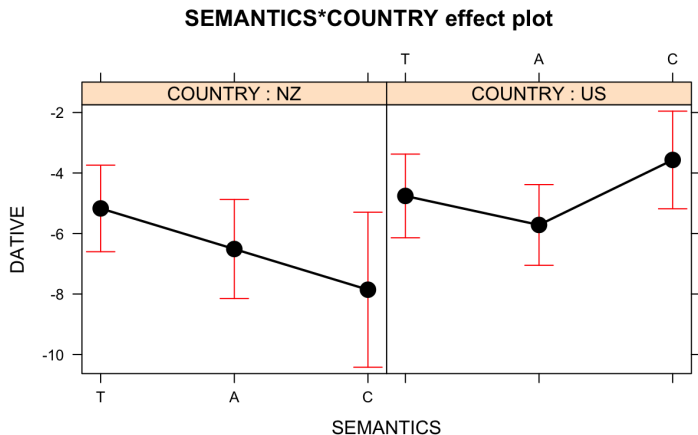
in BrE, animate themes favor the PD (as they should); in NZE, no effect.



# NZE versus AmE: semantics

x-axis: 'T': transfer, 'A': abstract, 'C': communication

y-axis: log odds (predicted outcome: PD)



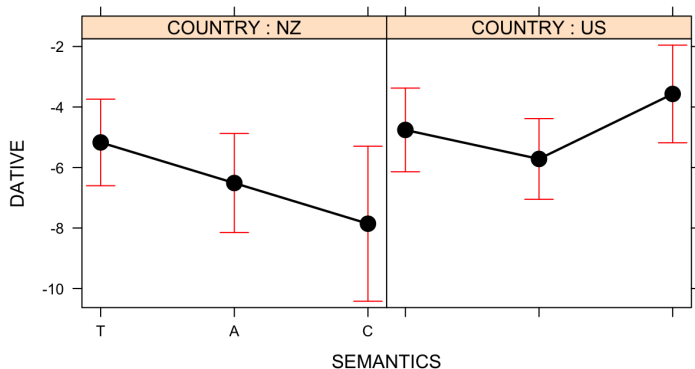


## summary

x-axis: 'T': t

y-axis: log o

in AmE, comm tokens favor PD vis-a-vis  
transfer; vice versa in NZE  
(similar difference in NZE versus BrE)



## Interim discussion: datives

- compared to BrE & NZE, end weight weak in CanE
- recipient animacy: stable across the board
- specifically, no recipient animacy difference NZE/AmE (contra Bresnan and Hay 2008)  
note: in the NZE part of the dataset, 42/61 (68%) of inanimate recipients appear in the ditransitive pattern; in AmE the figure is 135/187 (72%)
- semantics difference NZE/BrE and NZE/AmE: in NZE, communicative uses of *give* disfavor the PD



# Pairwise comparisons: genitives



# Predictions

- **Rosenbach (2002), Hinrichs and Szmrecsanyi (2007):** animate possessors discourage *s*-genitive usage less strongly in AmE than in BrE  
(but: mostly written data)  
failure to replicate
- **Hinrichs and Szmrecsanyi (2007):** long possessums favor *s*-genitive usage in AmE but not in BrE  
(but: written data)  
failure to replicate
- **Hundt and Szmrecsanyi (2012):** animacy is overall a more important factor for predicting genitive variation in (earlier, written) NZE than in (earlier, written) BrE  
replication successful



# Model quality measures: genitives

	$N$	% corr. pred.	$C$	$\kappa$
CanE/BrE	3652	93.5	0.98	12.4
CanE/NZE	3946	94.9	0.99	13.8
CanE/AmE	3104	93.3	0.98	12.9
NZE/BrE	3616	94.4	0.99	11.1
NZE/AmE	3068	94.6	0.98	8.0
BrE/AmE	2774	92.0	0.98	7.7



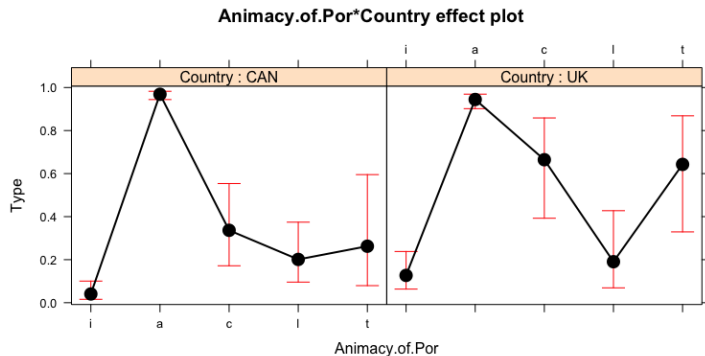
# Significant cross-variety differences in regression

CanE/BrE	Animacy.of.Por ( $p < .001$ ), Final.sibilancy ( $p = 0.09$ )
CanE/NZE	Animacy.of.Por ( $p = .09$ ), sem_rel_binary ( $p < .07$ )
CanE/AmE	Animacy.of.Por ( $p = .04$ ), Length.of.Pum ( $p = .03$ )
NZE/BrE	Animacy.of.Por ( $p = .04$ ), Final.sibilancy ( $p = .02$ )
NZE/AmE	<i>nil</i>
BrE/AmE	<i>nil</i>



# CanE versus BrE: possessor animacy

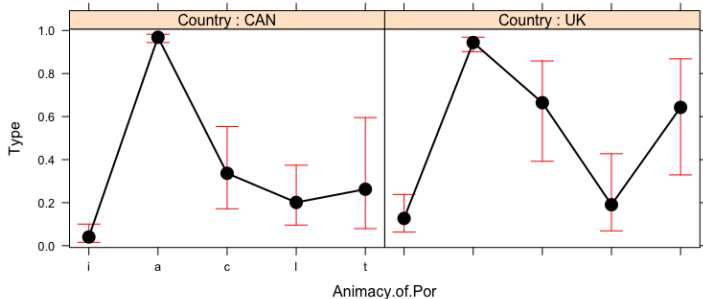
x-axis: 'i' inanimate, 'a' animate, 'c' collective, 'l' locative, 't' temporal  
 y-axis:  $p$  (predicted outcome: s-genitive)



# CanE vs BrE summary

x-axis: 'i' inanimacy  
 y-axis:  $p$  (probability)

in BrE, the *s*-genitive-favoring effect of locative possessors, relative to inanimate possessors, is weaker than in CanE

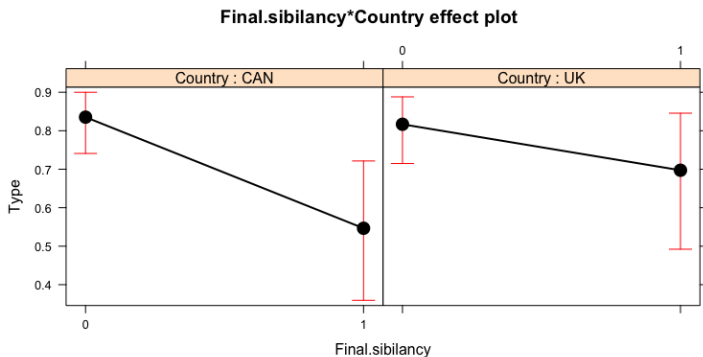




# CanE versus BrE: final sibilancy

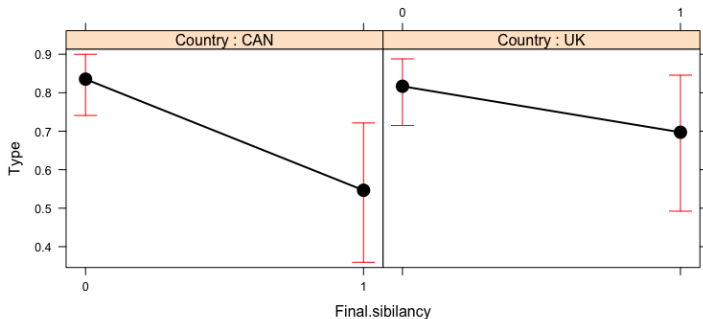
x-axis: '0' final sibilant absent, '1' final sibilant present

y-axis:  $p$  (predicted outcome: s-genitive)



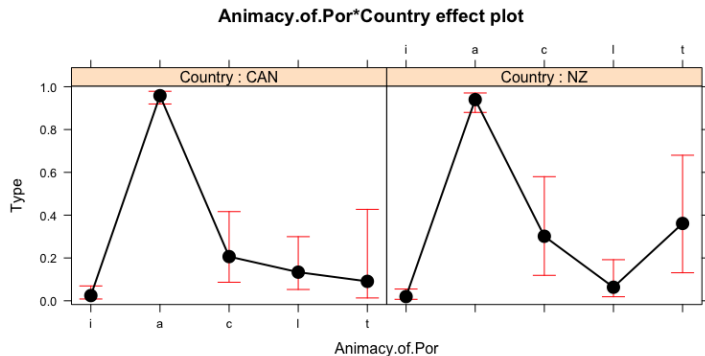
## summary

x-axis: '0' final sibilant in BrE, the s-genitive-disfavoring effect of  
 y-axis:  $p$  (probability) final sibilancy is weaker than in CanE



# CanE versus NZE: possessor animacy

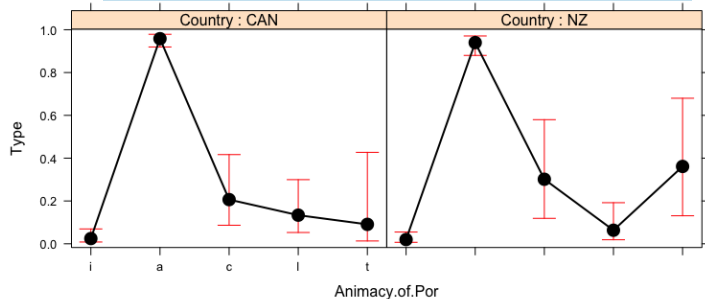
x-axis: 'i' inanimate, 'a' animate, 'c' collective, 'l' locative, 't' temporal  
 y-axis:  $p$  (predicted outcome: s-genitive)



# CanE vs NZE summary

x-axis: 'i' animacy  
 y-axis:  $p$  (probability)

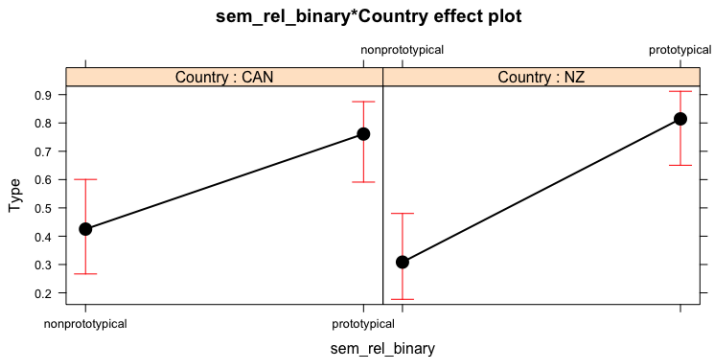
in NZE, temporal possessors favor the s-genitive more strongly (relative to inanimate possessors) than in CanE



# CanE versus NZE: semantic relation

x-axis: semantic relation

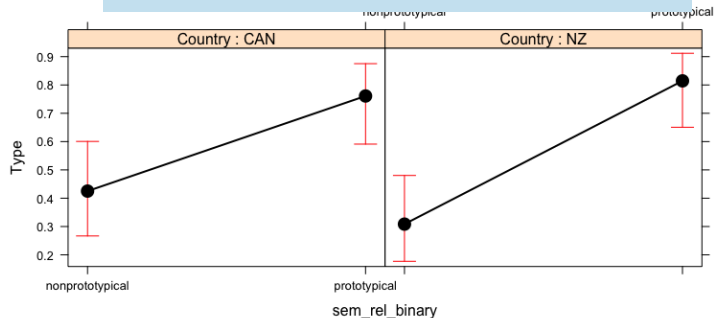
y-axis:  $p$  (predicted outcome: s-genitive)



# CanE vs NZE summary

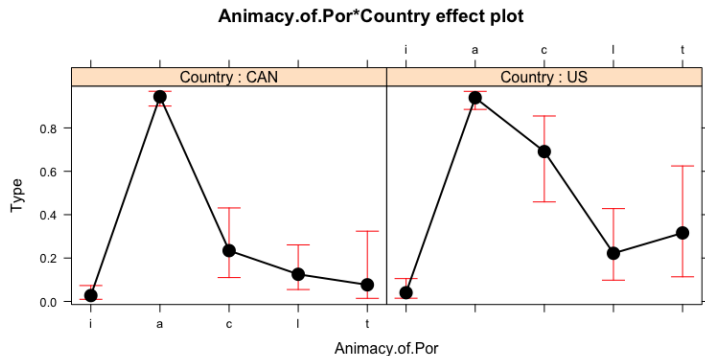
x-axis: semantic relation  
 y-axis:  $p$  (probability)

in NZE, prototypical semantic relations favor the *s*-genitive more strongly than in CanE



# CanE versus AmE: possessor animacy

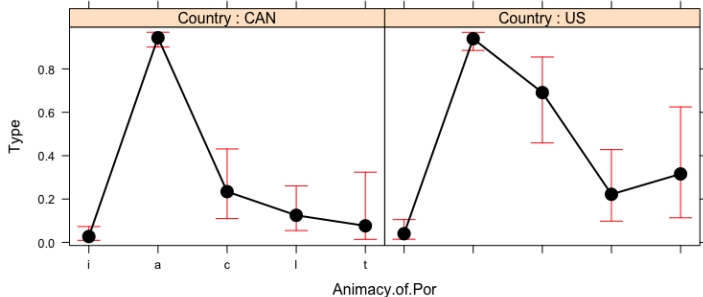
x-axis: 'i' inanimate, 'a' animate, 'c' collective, 'l' locative, 't' temporal  
 y-axis:  $p$  (predicted outcome: s-genitive)



# Canada vs. USA: animacy summary

x-axis: 'i' inanimacy  
y-axis:  $p$  (probability)

in AmE, collective possessors favor the s-genitive more strongly (compared to inanimate possessors) than in CanE

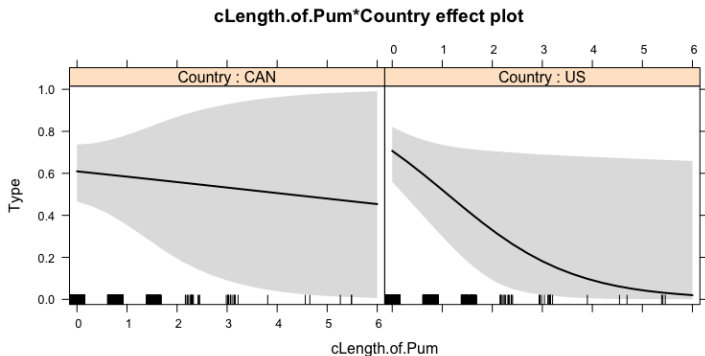




# CanE versus AmE: possessum length

x-axis: centered possessum length (in words)

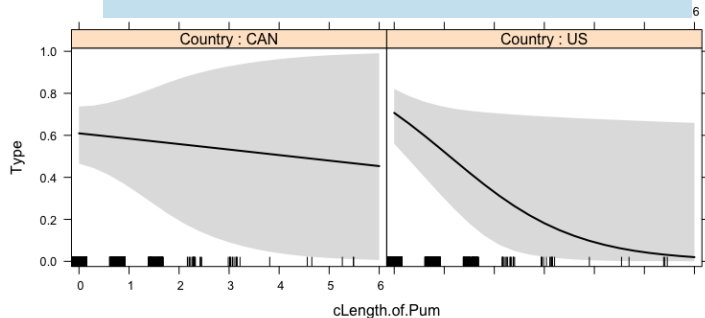
y-axis:  $p$  (predicted outcome: s-genitive)



# CanE vs. AmE summary

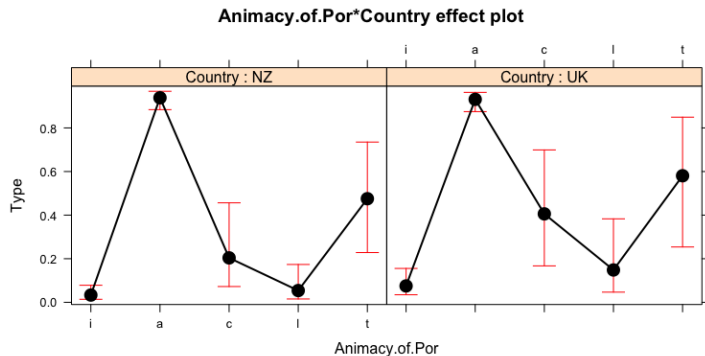
x-axis: center  
y-axis:  $p$  (prob)

in AmE, the *s*-genitive-disfavoring (no, that's not a typo!) effect of long possessums is even stronger than in CanE



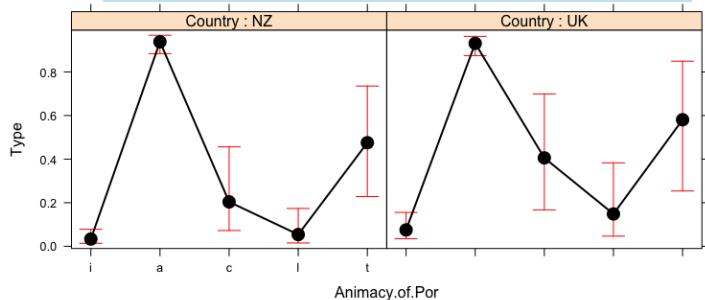
# NZE versus BrE: possessor animacy

x-axis: 'i' inanimate, 'a' animate, 'c' collective, 'l' locative, 't' temporal  
 y-axis:  $p$  (predicted outcome: s-genitive)



# NZE vs BrE summary

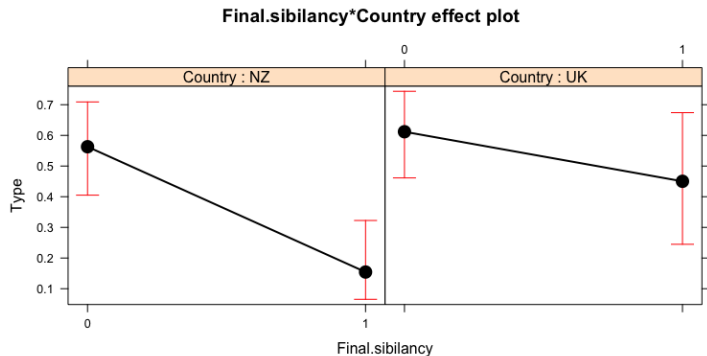
x-axis: 'i' inanimacy in BrE, the s-genitive-favoring effect of temporal  
 y-axis:  $p$  (probability) of animate possessors (compared to inanimate possessors) is weaker than in NZE ( $p < .1$ )



# NZE versus BrE: final sibilancy

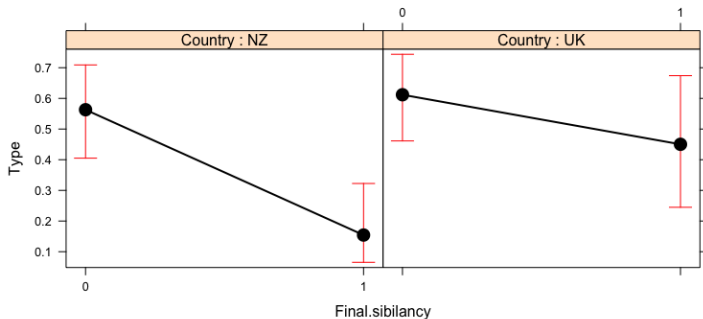
x-axis: '0' final sibilant absent, '1' final sibilant present

y-axis:  $p$  (predicted outcome: s-genitive)



## summary

x-axis: '0' final sibilancy in BrE, the *s*-genitive-disfavoring effect of  
 y-axis:  $p$  (probability) final sibilancy is weaker than in NZE



## Interim discussion: genitives

- possessor animacy: most variable constraint  
(but no difference in the AmE versus BrE pairing!)
- final sibilancy also surprisingly variable  
(2 significant interaction terms)
- more often than not, increasing possessum length predicts  
*of*-genitive usage  
(anti-end weight/Easy First)



# About possessor animacy

What happens if we run models with a binary animacy distinction?





## About possessor animacy

What happens if we run models with a binary animacy distinction?

- still not the slightest trace of an animacy difference between BrE and AmE ( $p = .82$ )



## About possessor animacy

What happens if we run models with a binary animacy distinction?

- still not the slightest trace of an animacy difference between BrE and AmE ( $p = .82$ )
- in BrE and AmE, the *s*-genitive-disfavoring effect of inanimate p'ors is weaker than in CanE
- in BrE and AmE, the *s*-disfavoring effect of inanimate p'ors is weaker than in NZE ( $p < .1$ )



## About possessor animacy

What happens if we run models with a binary animacy distinction?

- still not the slightest trace of an animacy difference between BrE and AmE ( $p = .82$ )
- in BrE and AmE, the *s*-genitive-disfavoring effect of inanimate p'ors is weaker than in CanE
- in BrE and AmE, the *s*-disfavoring effect of inanimate p'ors is weaker than in NZE ( $p < .1$ )

⇒ CanE and NZE have strong animacy constraints



# Synopsis



# Summary

- no obvious cross-constructional parallelisms according to regression analysis  
But CRF: animacy comparatively important in CanE in both alternations
- surprisingly hard to replicate previous findings
- CRF: genitive grammars more homogeneous than dative grammars



# To do / issues

- Annotation problem concerning theme pronominality (datives)  
*give them [so much ice cream] = pronominal?*
- dative extraction criteria for CanE/BrE differ from those for AmE/BrE, according to the documentation ⇨ [link to pairwise differences we are observing?](#)



# References I

- Bresnan, J., A. Cueni, T. Nikitina, and H. Baayen (2007). Predicting the Dative Alternation. In G. Boume, I. Krämer, and J. Zwarts (Eds.), *Cognitive Foundations of Interpretation*, pp. 69–94. Amsterdam: Royal Netherlands Academy of Science.
- Bresnan, J. and M. Ford (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1), 168–213.
- Bresnan, J. and J. Hay (2008, February). Gradient grammar: An effect of animacy on the syntax of give in New Zealand and American English. *Lingua* 118(2), 245–259.
- Gordon, E., M. Maclagan, and J. Hay (2007). The ONZE corpus. In J. C. Beal, K. P. Corrigan, and H. L. Moisl (Eds.), *Creating and Digitizing Language Corpora. Volume 2: Diachronic Databases*, pp. 82–104. Basingstroke: Palgrave Macmillan.
- Hay, J. and J. Bresnan (2006, January). Spoken syntax: The phonetics of giving a hand in New Zealand English. *The Linguistic Review* 23(3).

## References II

- Hinrichs, L. and B. 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, 437–474.
- Hundt, M. and B. Szmrecsanyi (2012). Animacy in early New Zealand English. *English World-Wide* 33, 241–263.
- Jankowski, B. L. and S. A. Tagliamonte (2014, July). On the genitive's trail: data and method from a sociolinguistic perspective. *English Language and Linguistics* 18(02), 305–329.
- Rosenbach, A. (2002). *Genitive variation in English: conceptual factors in synchronic and diachronic studies*. Berlin, New York: Mouton de Gruyter.
- Rosenbach, A. (2014, July). English genitive variation – the state of the art. *English Language and Linguistics* 18(02), 215–262.





## References III

- Shih, S., J. Grafmiller, R. Futrell, and J. Bresnan (2015, January). Rhythm's role in genitive construction choice in spoken English. In R. Vogel and R. Vijver (Eds.), *Rhythm in Cognition and Grammar*. Berlin, München, Boston: DE GRUYTER.
- Szmrecsanyi, B. (2006). *Morphosyntactic Persistence in Spoken English*. Berlin / New York: Mouton de Gruyter.
- Tagliamonte, S. A. (2014). A comparative sociolinguistic analysis of the dative alternation. In R. Torres-Cacoullos, N. Dion, and A. Lapierre (Eds.), *Linguistic variation: Confronting fact and theory*, pp. 297–318. London, New York: Routledge.
- Wolk, C., J. Bresnan, A. Rosenbach, and B. Szmrecsanyi (2013). Dative and genitive variability in Late Modern English: Exploring cross-constructural variation and change. *Diachronica* 30(3), 382–419.
- Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith (2009). *Mixed Effects Models and Extensions in Ecology with R*. New York: Springer.

