



**University of
Zurich** ^{UZH}

English Department

The effect of register on syntactic alternations

An exploratory study of regional variation in the English dative alternation

Melanie Röthlisberger (University of Zurich)



The English dative alternation

ditransitive dative



Mary gave [**John**] [**the apple**]
recipient theme



prepositional dative



Mary gave [**the apple**] to [**John**]
theme recipient



The English dative alternation

ditransitive dative



prepositional dative



Mary gave **[him]** **[the apple]**
recipient theme

Mary gave **[the apple]** to **[him]**
theme recipient



Why register?

- **regional variation** in the probabilistic constraints that influence grammatical variation, i.e. the choice between two variants (e.g. Bresnan & Hay 2008, Röthlisberger et al. 2017)
 - Text type as important in other syntactic alternations (e.g. Grafmiller 2014: genitive alternation)
- **Inter-register variation:** Different effects of register on the choice of dative alternation
- **Intra-register variation:** Different effects of language-internal constraints depending on register



Research objectives

1. **Inter-register variation:** What is the influence of register/text type on the choice of dative variant in the different varieties?
2. **Intra-register variation:** To what extent do language-internal factors differ in their effect on syntactic variation in different registers?

➔ **regional variation vs. register variation**



Data & Methodology



Corpus data





Corpus data I

- International Corpus of English (ICE) - series
 - 60% spoken (transcriptions), 40% written texts = 1m words per subcorpus
 - 500 texts, 2,000 words per text
 - 12 different registers, same corpus structure

SPOKEN 300	Dialogues	180	Private	100	Face-to-face conversations	90	s1a					
			Public	80	Phonecalls	10						
	Monologues	120	Unscripted	70	Classroom lessons	20	s1b					
					Broadcast Discussions	20						
					Broadcast Interviews	10						
					Parliamentary Debates	10						
Scripted	50	70	70	Legal cross-examinations	10							
				Business Transactions	10							
WRITTEN 200	Non-printed	50	Student Writing	20	Spontaneous commentaries	20	s2a					
					Unscripted Speeches	30						
	Printed	150	Academic writing	40	40	Demonstrations	10					
						Legal Presentations	10					
						Popular Writing	40	40	40	Broadcast News	20	s2b
										Broadcast Talks	20	
						Reportage	20	20	20	Non-broadcast Talks	10	
										Press news reports	20	w2c
	Instructional writing	20	20	20	20	Student Essays	10	w1a				
						Exam Scripts	10					
	Persuasive writing	10	10	10	10	Letters	30	w1b				
						Business Letters	15					
Creative writing	20	20	20	20	Humanities	10	w2a					
					Novels & short stories	20	w2f					



Corpus data II

- Corpus of Global web-based English (GloWbE) (Davies and Fuchs 2015)
 - general websites and blogs
 - data sampled in 2012-13 – <https://corpus.byu.edu/glowbe/>
 - same varieties sampled as represented in ICE
 - 500,000 words per variety sampled from random subsets



Data extraction and annotation

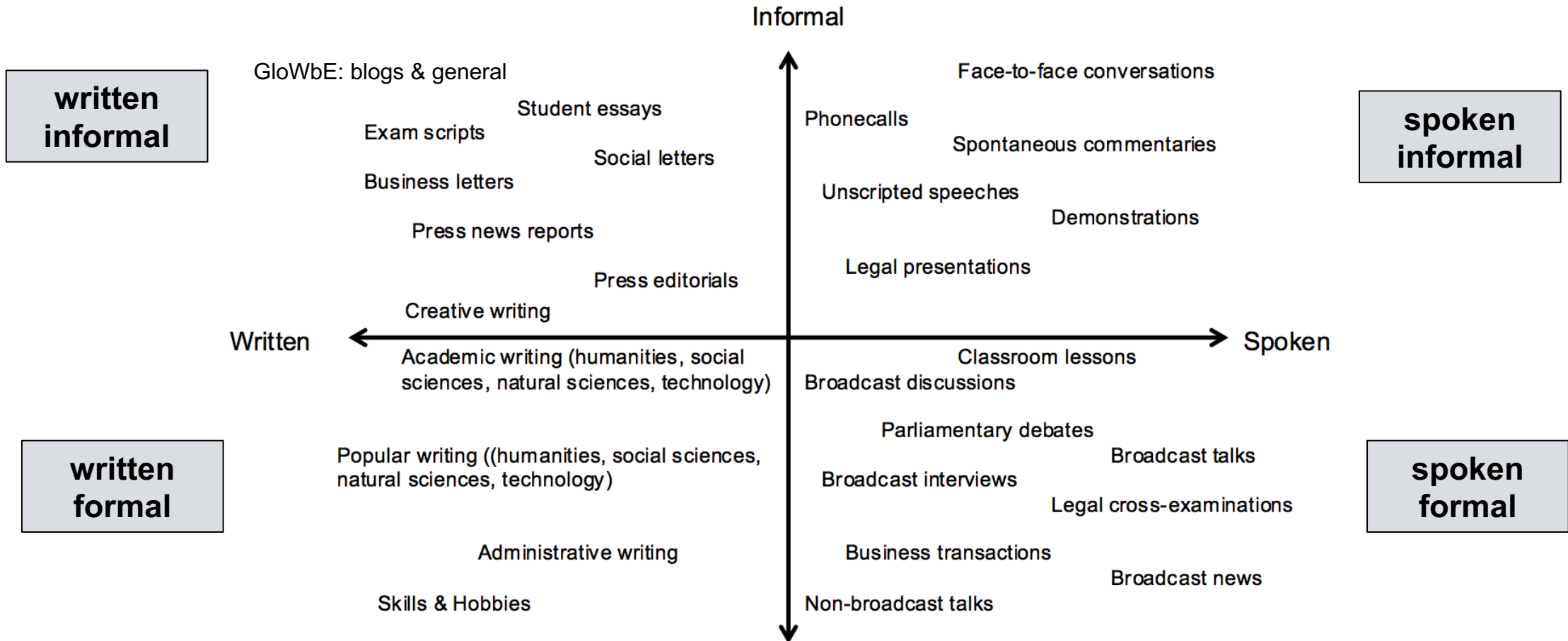
(e.g. Bresnan et al. 2007)

- retrieval of dative variants using verb list and perl script
- restrict to choice context (incl. pronouns)
- code for numerous (language-internal) factors: length (weight ratio), complexity, pronominality, givenness, definiteness, person, animacy, concreteness of theme, verb sense
- code for language-external factors: Register and Mode (spoken vs written)

$N = 13,171$



Register coding



Koch & Oesterreicher (1985)



Mixed-effects logistic regression

1. Inter-register variation (regional variation in register effects):

- fixed effects: 5 most important language-internal predictors (acc. to random forest)
- interaction REGISTER * VARIETY
- random effects: lexical idiosyncrasies, file ID

2. Intra-register variation (internal variation):

- fit one model per variety
- interaction of REGISTER with 5 most important language-internal predictors
- random effects: lexical idiosyncrasies, file ID



Results



RQ 1: Regional variation in register effects

Model formula:

Dative variant ~ (1 | Verb) + (1 | ThemeHeadFilter) + (1 | RecHeadFilter) + (1 | FileID)
+ ThemeHeadFreq + ThemeComplexity + RecPron + ThemePron + WeightRatio
+ Register + Variety + Variety:Register)

- numeric predictors are standardized by two standard deviations (ThemeHeadFreq + WeightRatio)
- WeightRatio is log-scaled (cf. Bresnan et al. 2007)
- Register and Variety are coded with sum coding instead of treatment coding (comparison NOT versus one reference level but versus the grand mean of all other levels)

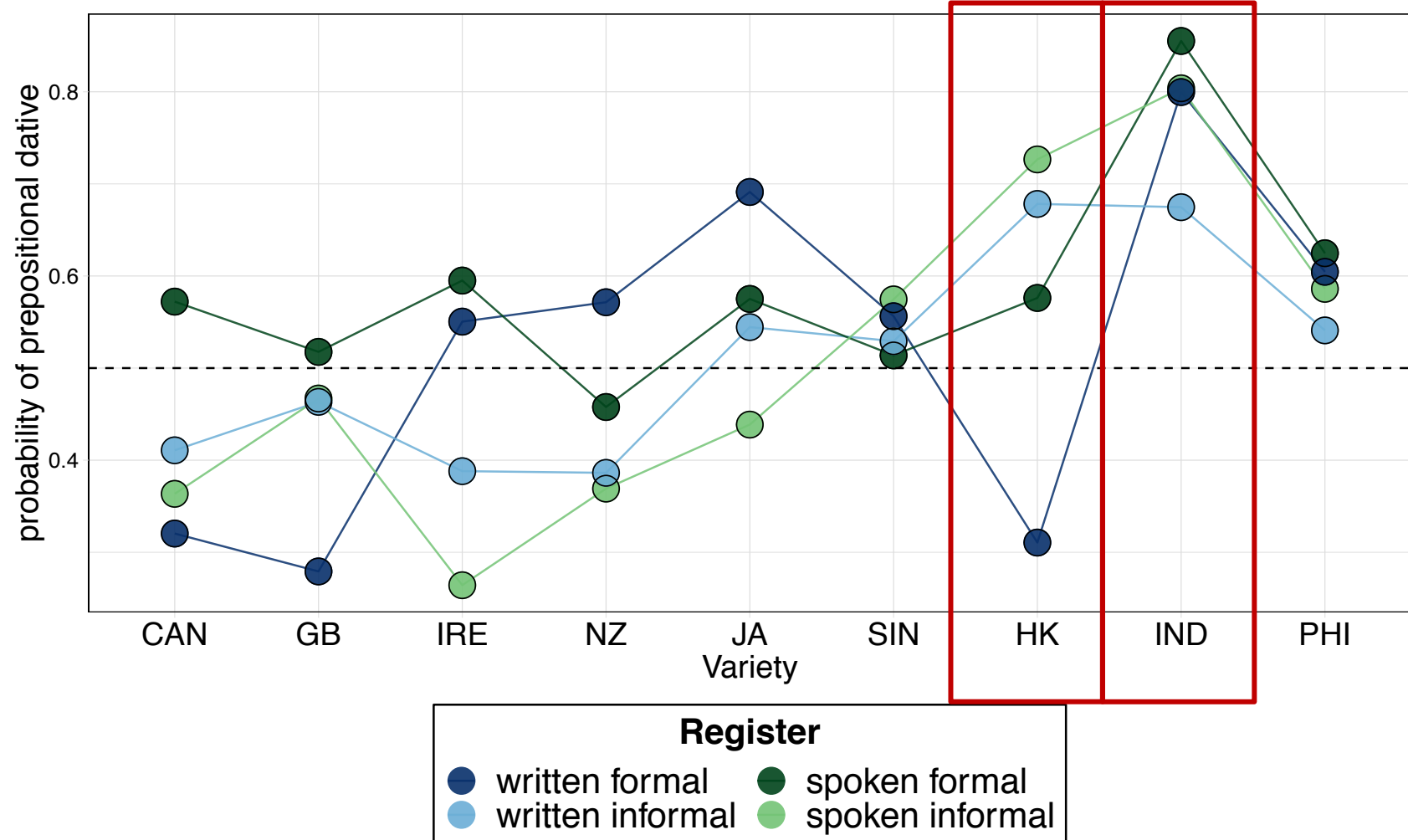


Regional variation in register effects: Register * Variety interaction

Evaluation

C-statistic: 0.98
Accuracy: 93.3%

no indication of overdispersion / multicollinearity

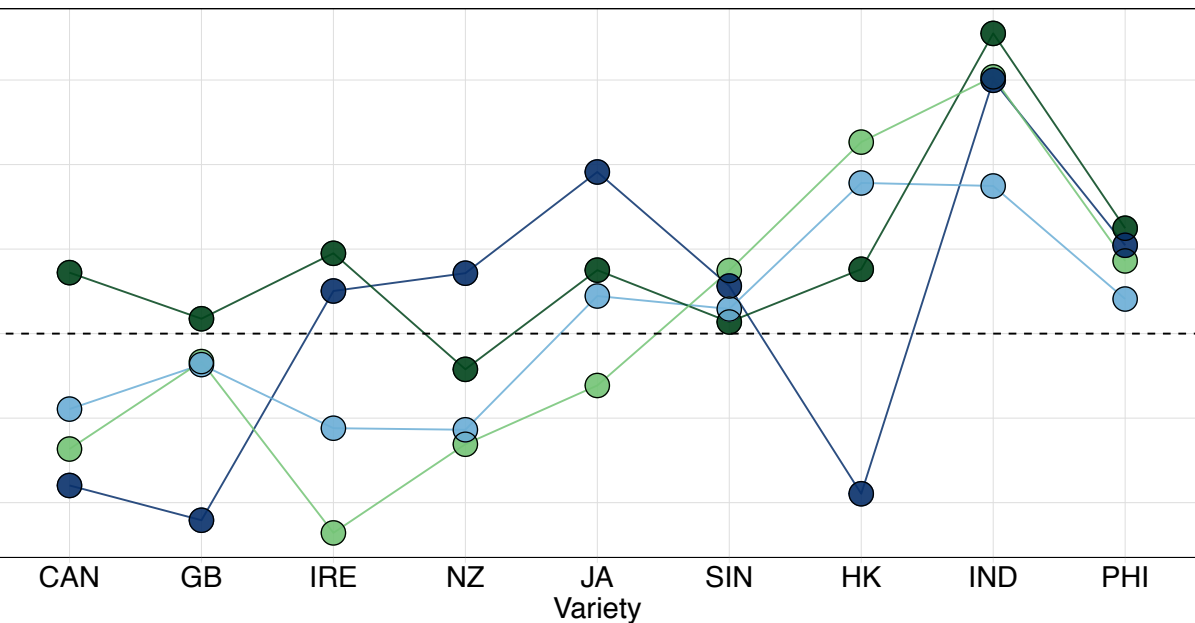




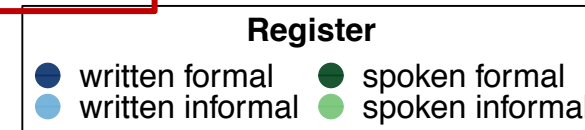
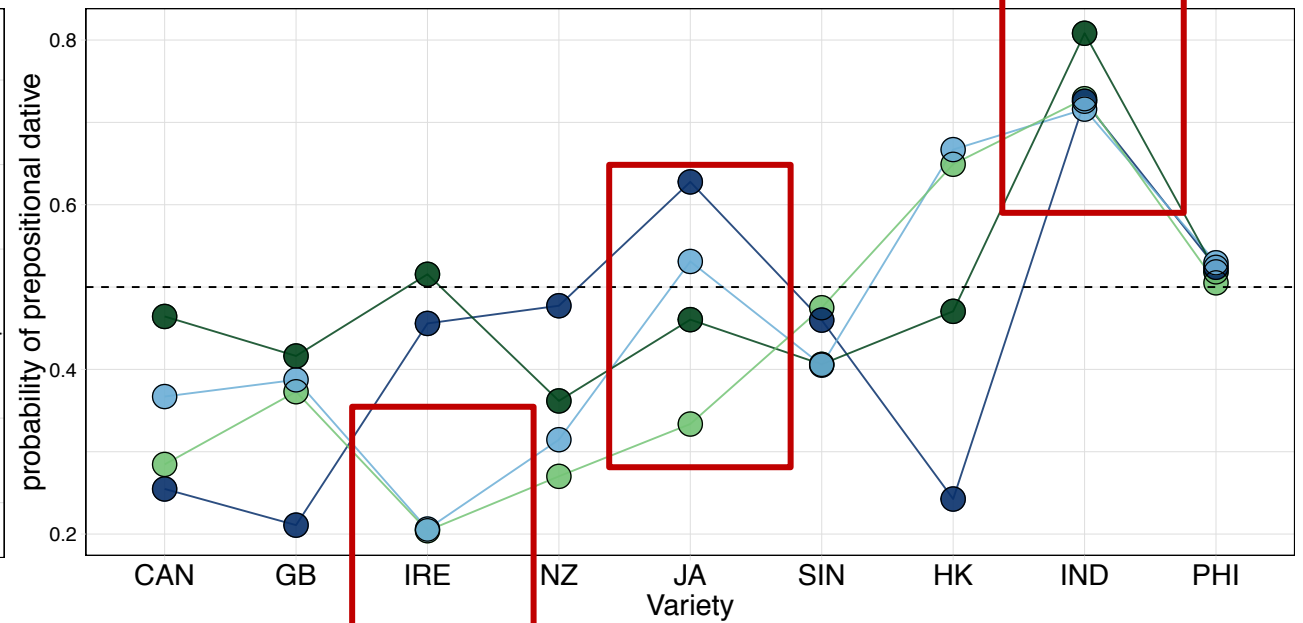
Two corpora - same register?

What is the effect of adding GloWbE data to the written informal register?

with GloWbE:



without GloWbE:





RQ 2: Register-internal variation (intra-register variation)

Model formula:

Dative variant \sim (1|Verb) + (1|ThemeHead) + (1|RecHead) + (1|FileID)
+ Register * (WeightRatio + RecPron + ThemeComplexity + ThemePron
+ ThemeHeadFreq)

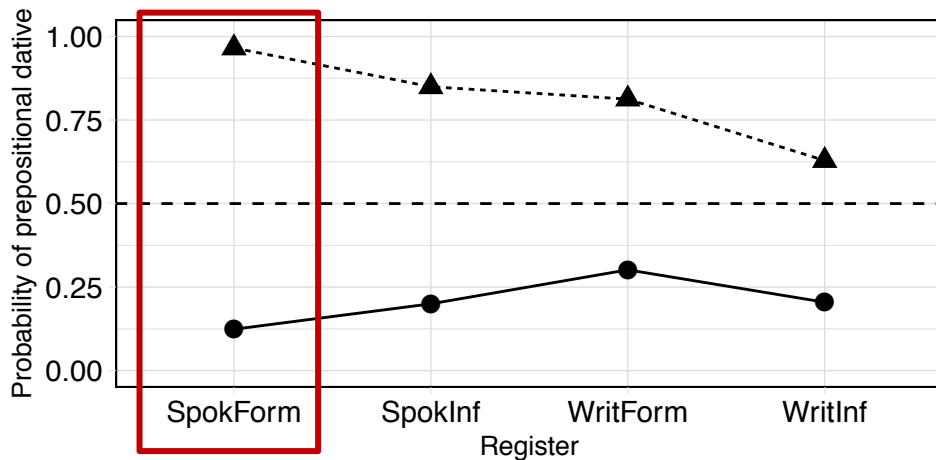
- One model per variety (9 models in total)
- numeric predictors are standardized by two standard deviations (ThemeHeadFreq + WeightRatio)
- binary predictors are transformed into numbers and centered around the mean
- WeightRatio is log-scaled (cf. Bresnan et al. 2007)
- Register is again coded with sum coding instead of treatment coding



Register-internal variation

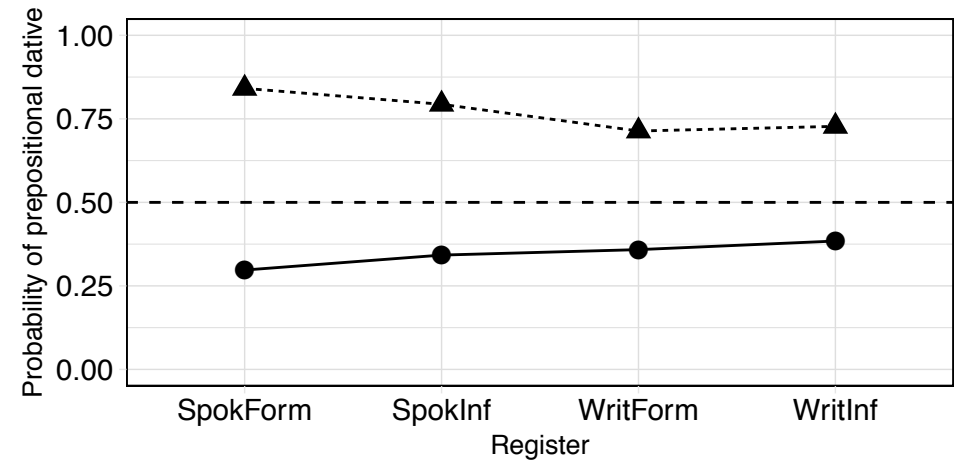
- No statistically significant interactions in: BrE, HKE, IrE, JamE, NZE, SinE
- Theme complexity & theme head frequency significantly different by Register in CanE
- Recipient pronominality significantly different by Register in IndE

Interaction REGISTER : RECPRON in Indian E.



recipient pronominality ● pron ▲ non-pron

Interaction REGISTER : RECPRON all varieties



recipient pronominality ● pron ▲ non-pron



In sum

1. **Inter-register variation:** What is the influence of register/text type on the choice of dative variant in the different varieties?

Register differs in its effect across different varieties of English. Overall,

- formal registers are more likely to use the prepositional dative (with the exception of HKE)
- In **PhiE**, **SinE** the differences between the registers are negligible → homogeneous registers
- In **HKE**, **BrE**, **IrE** and **JamE**, register effects are diverging
- other varieties lay somewhere in between

The addition of GloWbE data results in slightly different results, regarding the effect of spoken informal vs written informal data:

GloWbE data might be more different from the written informal spectrum than assumed.



In sum

2. **Intra-register variation:** To what extent do language-internal factors differ in their effect on syntactic variation in different registers?

- Language-internal predictors differ in their effect size, notably recipient pronominality and theme complexity
- Apart from CanE and IndE, all varieties are very homogeneous regarding intra-register variation



to conclude...



Take-home messages

- registers display variety-specific patterns (at least wrt the dative alternation) (Neumann, this conference)
- comparison between registers of ICE corpora requires reconsiderations
- probabilistic grammars are fairly stable between varieties (we already know that) but are even more homogeneous among registers (sic!)



Open questions

- Change the coding of register to a more fine-grained coding (e.g. text type)
- Investigate competition between formality and mode (are they as similarly distinct as the terminology implies?)
- ...



Thank you!

melanie.roethlisberger@es.uzh.ch

[@m_roethlisberger](#)

slides available at: www.melanie-roethlisberger.ch



References

Bresnan, J., A. Cueni, T. Nikitina, and R. H. Baayen (2007). Predicting the dative alternation. In G. Boume, I. Kraemer, and J. Zwarts (Eds.), *Cognitive foundations of interpretation*, pp. 69–94. Amsterdam: Royal Netherlands Academy of Science.

Bresnan, J. and J. Hay. (2008). Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118, 245-259.

Davies, M. and R. Fuchs (2015). Expanding horizons in the study of World Englishes with the 1.9 billion word Global Web-Based English Corpus (GloWbE). *English World-Wide* 36(1), 1-28.

Grafmiller, J. (2014). Variation in English genitives across modality and genres. *English Language and Linguistics* 18(3), 471-496.

Koch, P. and W. Oesterreicher (1985). Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36, 15-43.

Röthlisberger, M., Grafmiller J. and B. Szmrecsanyi. (2017). Cognitive indigenization effects in the English dative alternation. *Cognitive Linguistics* 28(4), 673-710.