

Spatiotemporal variation in the dative alternation: a study of four corpora of British and American English

Scott Grimm and Joan Bresnan

Stanford University

September 24, 2009

Background: Hinrichs and Szmrecsányi (2007)

The availability of matching corpora and increasingly sophisticated computational techniques have increased our ability to detect change over shorter time periods

Hinrichs and Szmrecsányi (2007) showed a change in the English genitive alternation over a 30-year period in both the UK and US:

- ▶ the US leads the UK in moving toward the preposed ('s) genitive

We base our study on Hinrichs and Szmrecsányi (2007) and examine if similar changes can be detected in an entirely different construction, the dative alternation

Background: Hinrichs and Szmrecsányi (2007)

Hinrichs and Szmrecsányi (2007) used the Brown 'family' of corpora

- ▶ 4 corpora differing in the variety of English spoken and the time period of the sampling:

	1960's	1990's
US	Brown	Frown
UK	LOB	F-LOB

Background: Bresnan et al. (2007)

Previous work on the dative alternation has emphasized that multiple variables, such as argument length and pronominality, contribute to the speakers choice between forms (Bresnan et al. (2007) inter alia)

Background: Harmonic alignment

The study of Bresnan et al. (2007) demonstrated that the choice between different forms in the dative alternation manifest *harmonic alignment*

Prominence scales align harmonically with syntactic position:

shorter	>	longer
pronoun	>	non-pronoun
more thematic	>	less thematic
more persistent (primed)	>	less persistent (primed)

V NP NP

V recipient theme

V NP PP

V theme recipient

Background: Harmonic alignment

A similar pattern to harmonic alignment in the dative occurs in the genitive (Hinrichs and Szmrecsányi 2007)

- ▶ shorter, animate, topical (in terms of text frequency of head) before longer, inanimate, non-given

Hinrichs and Szmrecsányi (2007) found that the 30-year historical changes are increasing alignment of length, e.g. longer possessums favor 's genitives

Are these changes occurring across constructions?

- ▶ and if so, why?

Data Collection: NLP tools

Methodological question: How can this case study on the dative be achieved in an *efficient* manner?

We employ state-of-the-art tools from NLP parsing technology to aid in extracting the relevant dative constructions from the corpora

Data

The corpora in the Brown family contain 500 text samples (2000 words each) across 15 genres, tallying to 1 million words per corpus

We extracted datives from the entire corpora in a departure from Hinrichs and Szmrecsányi (2007) who limited their study to journalistic text (sections A and B)

Parsing the corpora

Used the Klein-Manning parser provided by the Stanford NLP group (Klein and Manning 2003)

Input: manually corrected POS-tagged by Freiburg group; cooperation with Benedikt Szmrecsányi and Lars Hinrichs

Output: The parser provides Stanford Dependencies output as well as phrase structure trees. Typed dependencies are otherwise known grammatical relations. These are produced using hand-written tregex patterns as described in de Marneffe et al. (2006, 2008)

The methodology is not dependent on parser choice—other parsers (Charniak, Collins) are capable of being enriched by the typed dependencies

Parser output example

Sentence string:

This would give the hydrogen atom a slight charge-excess .

Parse tree:

(ROOT (S (NP (DT This)) (VP (MD would) (VP (VB give) (NP (DT the) (NN hydrogen) (NN atom)) (NP (DT a) (JJ slight) (NN charge-excess)))) (. .)))

Typed dependencies (grammatical relations):

subj(give-3, This-1), aux(give-3, would-2), det(atom-6, the-4),
nn(atom-6, hydrogen-5), iobj(give-3, atom-6), det(charge-excess-9,
a-7), amod(charge-excess-9, slight-8), dobj(give-3, charge-excess-9

Extraction and Filtering

Developed Python script which ran over the parsed output and extracted the structures which had both the relevant grammatical relations (*obj*; *prep-to*) and a verb known to alternate

Principles of exclusion

Were interested in dative constructions where the alternation was in theory possible

Followed methodological decisions in Bresnan et al. (2007) and excluded:

- ▶ transformed datives (with questioned, passivized, or topicalized recipients or themes)
- ▶ dative verbs with sentential complements
- ▶ sentences which lacked two overt complements.
- ▶ Occurred in non-alternating fixed expressions, such as the parentheticals “(I’ll/I) tell you what” and “To tell you the truth.”

Principles of exclusion

- ▶ Occurred in other contexts in which no alternation between the NP PP and NP NP construction was possible. These contexts include:
 - ▶ Instances of ditransitive *make* where the theme is not an NP with *offer* or *promise*. These generally undergo the benefactive, rather than the dative, alternation.
 - ▶ Instances of "concealed questions" with *tell*, as in "I'll tell you another plant that is purple".
 - ▶ Instances with (unambiguous) spatial goals
 - ▶ Non-alternating idioms

Evaluation of the Automated Extraction: False Positives

	NP NP	NP PP	Total	After Filtering	% retained
Brown	854	666	1520	819	53%
Frown	992	773	1765	759	43%
LOB	805	759	1564	765	49%
FLOB	834	1076	1910	771	40 %

Evaluation of the Automated Extraction: False Negatives

Took a random sample of 100 instances of sentences with *give* in Brown using a simple regular expression search

dative sentences found in database	45
false negatives (datives missed)	3
false positives (non-datives)	52

From this estimate, there are around 6% of the datives that are missed by the above procedure

Annotation: automatic approximation

The model in Bresnan et al. (available in the languageR package) coded 14 different factors, including:

Speaker

Modality

Verb

Semantic Class of Verb

Length (Theme / Recipient)

Animacy (Theme/ Recipient)

Definiteness (Theme / Recipient)

Pronominality (Theme / Recipient)

Accessibility (or Information Structure) (Theme / Recipient)

Annotation: automatic approximation

Annotating some of the variables are costly in human hours, viz. animacy, accessibility

Part of the methodological experiment was to determine what one could do efficiently and automatically

The data points in the resulting database were then automatically labeled for thematicity, length of arguments, persistence and pronominality

Annotation: evaluation of automatic approximation

Although we work with an approximation of the Bresnan et al. factors, the resultant models are still highly accurate.

Evaluate with C statistic

C (concordance) statistic is a measure of the discriminative power of the logistic equation

A value above .8 shows discriminative power (Harrell 2001)

Modeling the Switchboard and Wall Street Journal data

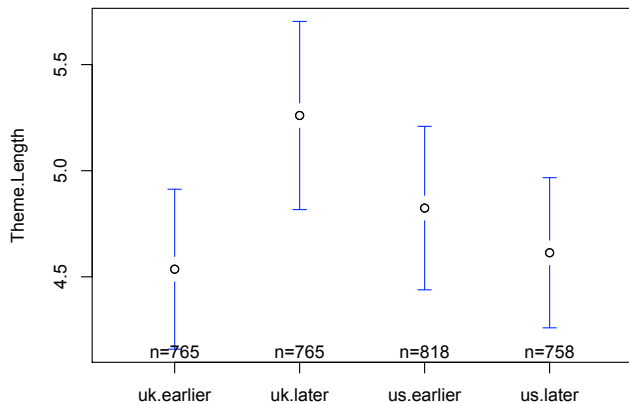
- ▶ with all the factors in Bresnan et al. results in a C score of .96
- ▶ with only the factors *length* and *pronominality* results in a C score of .95

Annotation: length of arguments

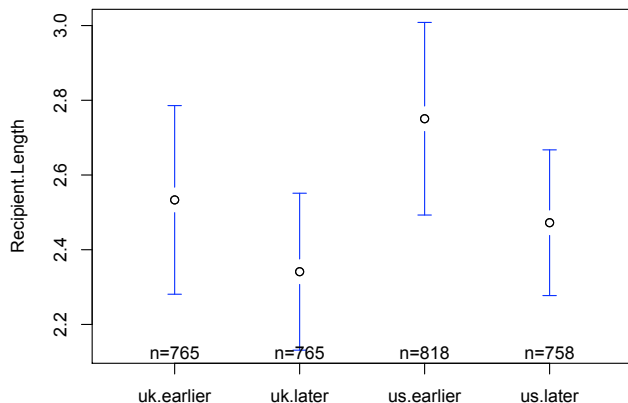
Length in words is a convenient proxy for syntactic complexity (Szremścányi 2004, Wasow 1997), which has in turn been argued to be the driving force in the choice of alternative word orders (Hawkins 1994;2004, Gibson 2000)

The number of space-delimited words encodes the length.

Distribution: length of theme arguments



Distribution: length of recipient arguments



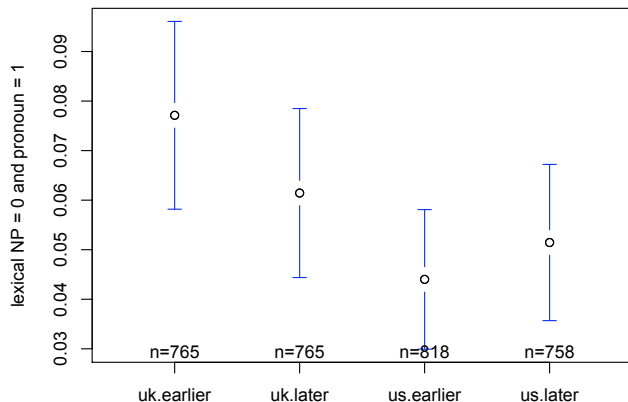
Annotation: pronominality

Pronominality is a simplification of nominal expression type (Ariel 1990, Silverstein, et al.), which also influences word order via pragmatic or possibly prosodic effects (Behagel 1909, Anttila 2008, Shih et al. 2009).

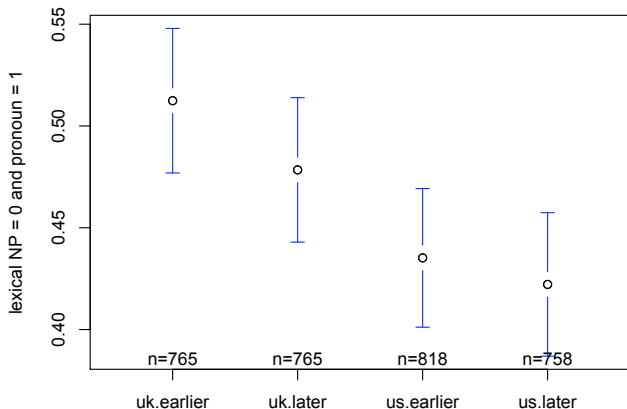
The following were marked pronominal:

- ▶ definite pronoun *it*
- ▶ demonstrative pronoun *that*
- ▶ personal pronoun *me*
- ▶ reflexive pronoun *myself*
- ▶ personal pronoun followed by a lexical NP *she gave them all her children a spanking*

Distribution: theme pronominality



Distribution: recipient pronominality



Annotation: thematicity

“Thematicity”, or the overall topicality of a noun, has been taken as a driving factor in the genitive alternation

The term thematicity is used by Hinrichs and Szmrecsányi (2007) in reference to Osselton (1988): “according to Osselton, while *sound*, *soil*, and *fund* will not normally take the s-genitive, in a book on phonetics, *sound* will get its genitive, in one on farming, *soil* will do so, and in a book on economics you can expect to find a *fund*’s success (Osselton 1988: 143).”

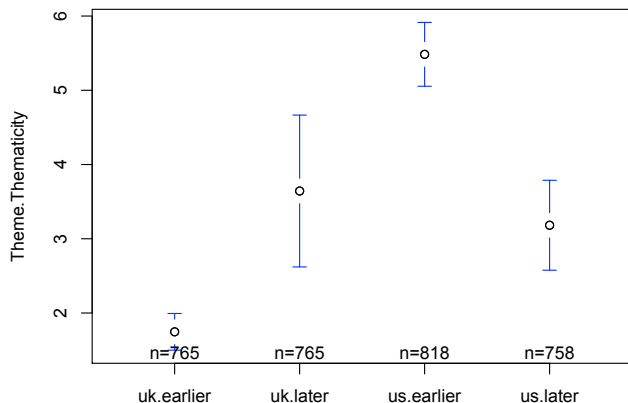
Annotation: thematicity

Hinrichs and Szmrecsányi (2007) measure thematicity via text frequency of the possessor head in the genitive construction.

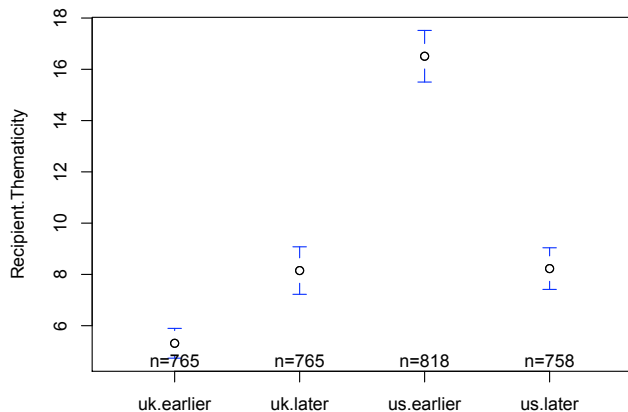
We follow and measure the text frequency of the head noun of both Recipient and Theme arguments

- ▶ estimated by the total occurrences of head noun in the entire document, e.g. in an entire article, capturing the overall salience of a lexical item in the text

Distribution: theme thematicity



Distribution: recipient thematicity



Annotation: persistence

Persistence is a measure of production priming:
speakers reuse what they have just heard or just used.

Szmrecsányi (2005) found persistence to play a highly significant
role in linguistic choice for different English alternations.

Annotation: persistence

We coded for α persistence (exact match), whereby we located the first previous dative construction within a range of 10 sentences:

NP = previous NP NP in a dative construction

PP = previous NP PP in a dative construction

0 = no previous dative construction

Logistic regression modeling

Logistic regression models controls simultaneously for multiple factors giving a binary response.

$$P(\text{Response} = \text{NP PP} | X) = \frac{1}{1 + \exp(-(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots))}$$

where X is the model matrix of independent variables $[x_1, x_2, \dots]$ and β s are their coefficients.

Mixed Effect Modeling

Mixed effect modeling includes both *fixed* and *random* effects
(Baayen et al. 2008)

Treating verbs as a random effect allows us to adjust for variance due to individual verbs and generalize beyond them

We use a generalized linear mixed-effect model (implemented with *lmer* function in R)

Model Results: US corpora

Across the US corpora, main effects of thematicity of recipient and length and pronominality of recipient and theme are all harmonically aligned with construction choice,

i.e. values indicative of higher prominence aligned with the more prominent syntactic position, and conversely

Additionally, there is significantly greater preference for the double object construction in the 1990s than in the 1960s

Model Results: US corpora main effects

Length and Pronominality have larger effect sizes

Predicted odds are for the NP PP construction

Factor	Odds	P-Value
Log of Recipient Length	6.67	0.000
Log of Theme Length	.15	0.000
Recipient Pronoun	0.09	0.000
90's Corpus	0.63	0.032
Theme Pronoun	3.48	0.059
Recipient Thematicity	0.98	0.052

Model Results: US corpora interaction effects

Factor	Odds	P-Value
Theme Pronoun * 90's Corpus	10.42	0.014

Stronger dispreference for *V NP Pron*, i.e. **V NP Pron*, in the Frown corpus

- ▶ Consistent with increasing harmonic alignment

Model Results: UK corpora

Across the UK corpora, main effects of thematicity, length, and pronominality of recipient and theme also show harmonic alignment with construction choice.

There is also a significant change toward greater preference of the double object construction, in parallel to the US corpora.

Model Results: UK corpora main effects

Factor	Odds	P-Value
Log of Recipient Length	7.00	0.000
Log of Theme Length	0.21	0.000
Recipient Pronoun	0.03	0.000
Theme Pronoun	8.18	0.000
90's Corpus	0.44	0.001
Theme Thematicity	1.15	0.002
Recipient Thematicity	0.90	0.030

Model Results: UK corpora interaction effects

Factor	Odds	P-Value
Recipient Pronoun * 90's Corpus	6.12	0.000
Theme Thematicity * 90's Corpus	0.86	0.001
Recipient Thematicity * 90's Corpus	10.42	0.014

Stronger preference for *V Pron NP* in the FLOB corpus

- ▶ Consistent with increasing harmonic alignment

Diachronic Change Constant across Text Varieties

The observed changes are general across the different varieties of text in the Brown corpora:

- ▶ Added as a random effect general text type
 - ▶ 4 super-types: Fiction, Learned, Press and Prose
- ▶ The fixed effects hold after adjusting for the different sections as random effects

Changes do not appear to be just a reflection of journalism or of one particular text type

Model Results: Spatial differences across 1960's corpora

Across the 1960's corpora, main effects of thematicity, length, and pronominality of recipient and theme also show harmonic alignment with construction choice.

There is also a greater preference in the US 1960's corpus towards thematic themes

Model Results: Spatial differences across 1990's corpora

Across the 1990's corpora, there are main effects of length, and pronominality of recipient and theme.

Here too the effects display harmonic alignment with construction choice.

The difference between 90's UK and US English parallels the difference between 60's and 90's US English: the *V NP Pron effect is stronger in the Frown than in the FLOB dataset

Model Results: Global trends

Increased preference for double object construction

Consistent with increased preference observed by Hinrichs and Szmrecsányi (2007) for the 's genitive

- ▶ More economic form preferred
- ▶ Possibly additional influence from increased informal language (Kroch and Small 1978)

Model Results: Global trends

Harmonic alignment effects across time and space

Where comparable (length/thematicity), the direction of the effects parallel those of Hinrichs and Szmrecsányi (2007)

- ▶ Harmonic alignment effects appear *across constructions*

Harmonic alignment effects are commonly related to processing/comprehension ease

- ▶ Increase in harmonic alignment may indicate historical change patterns towards what is easier to comprehend/process

Accuracy of model

The C statistic (discriminatory accuracy) for the models was quite high .96

Gold Standard

In order to assess the accuracy of the method, we created gold standard versions of the four corpora:

- ▶ 200 sentence randomized samples of original corpora (800 in total)
- ▶ Hand-corrected
- ▶ Aids in error analysis
- ▶ Serves to test predictive power of the previous model
 - ▶ Do the model trends remain consistent with corrected data?
 - ▶ Can we can predict the clean data from the noiser data?

Gold Standard: Model Results

For US corpora:

- ▶ The later period favors the V NP NP construction;
- ▶ shorter themes contribute independently to V NP PP constructions, while shorter recipients and pronominal recipients favor V NP NP.

For UK corpora:

- ▶ Length, Expression Type significantly go in direction of harmonic alignment, same as US
- ▶ period shows a movement toward V NP NP

Accuracy of the Noisy Version

Check the predictions of the model:

Given fixed effects of the first (noisy) model, calculate predicted response values given the input of gold standard, and compare with the actual response values

84.5% accuracy (against a baseline of 59%) (based on the Brown)

Outlook

This paper demonstrates how regularities of syntactic change can be observed across a set of large corpora in an efficient manner via exploiting NLP tools

This leads to a model of probabilistic changes in linguistic choices across space and historical time

Further research may elucidate how change proceeds across constructions

Thank you

Thanks to:

Marie-Catherine de Marneffe, Chris Manning, Rachel Cristy, Nick Romero, Benedikt Szmezcanyi and Lars Hinrichs

This material is based in part upon work supported by the National Science Foundation under Grant Number IIS-0624345 to Stanford University for the research project “The Dynamics of Probabilistic Grammar” (PI Joan Bresnan). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References:

Anttila, Arto. 2008. Phonological constraints on constituent ordering. In C.B. Chang and H.J. Haynie (ed). *Proceedings of the 26th West Coast Conference on Formal Linguistics*. 51-59.

Ariel, Mira. 1990. Accessing noun phrase antecedents. London: Routledge.

Baayen, R.H., Davidson, D.J. and Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390–412.

Behagel, O. 1909 Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen*, 25/110.

References:

Joan Bresnan, Anna Cueni, Tatiana Nikitina, and Harald Baayen. 2007. "Predicting the Dative Alternation." In *Cognitive Foundations of Interpretation*, ed. by G. Boume, I. Kraemer, and J. Zwarts. Amsterdam: Royal Netherlands Academy of Science, pp. 69–94.

Marie-Catherine de Marneffe, Bill MacCartney and Christopher D. Manning. 2006. Generating Typed Dependency Parses from Phrase Structure Parses. In *LREC 2006*.

Marie-Catherine de Marneffe and Christopher D. Manning. 2008. The Stanford typed dependencies representation. In *COLING 2008 Workshop on Cross-framework and Cross-domain Parser Evaluation*.

References:

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, W. ONeil (ed). Image, language, brain. Cambridge, MA: MIT Press. 95-126.

Harrell, Frank E. 2001. Regression Modeling Strategies . Springer.

Hawkins, John A. 1994. A Performance Theory of Order and Constituency. Cambridge: Cambridge University Press. Hawkins, John A. 2004. Efficiency and Complexity in Grammars. Oxford: Oxford University Press.

References:

Lars Hinrichs and Benedikt Szmrecsányi. 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), pp. 335-378.

Dan Klein and Christopher D. Manning. 2003. "Accurate Unlexicalized Parsing." *ACL 2003*, pp. 423-430.

Anthony Kroch and Cathy Small. 1978. *Grammatical Ideology and its Effect on Speech*. *Linguistic Variation: Models and Methods*, D. Sankoff, ed., Academic Press.

Osselton, N. 1988. *Thematic genitives. An Historic Tongue: studies in English Linguistics in Memory of Barbara Strang* ed. by G. Nixon & J. Honey. London: Routledge.

References:

Stephanie Shih, Jason Grafmiller, Richard Futrell, and Joan Bresnan. (2009) "Rhythm's role in predicting genitive and dative alternation choice in spoken English." Presentation at DGfS: Rhythm beyond the word. Osnabruck, Germany.

Szmrecsányi, Benedikt. 2004. On operationalizing syntactic complexity. In Purnelle, G., C. Fairon, and A. Dister (ed) *Le poids des mots: Proceedings of the 7th International Conference on Textual Data Statistical Analysis*. Louvain-la-Neuve. Presses universitaires de Louvain. 2:1032-1039.

Szmrecsányi, B. 2005 Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistics Theory*, 1, 113150.

References:

Wasow, Thomas Postverbal Behavior. CSLI Publications. 2002.