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
  - US: Brown
  - UK: LOB

- 1990’s
  - US: Frown
  - UK: F-LOB
Background: The Dative Alternation

We apply a comparable methodology to a new case study: the dative alternation

(1) a. Who gave that watch to you? (NP PP)
    theme  recipient

b. Who gave you that watch? (NP NP)
    recipient  theme
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
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?
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.
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).
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
This would give the hydrogen atom a slight charge-excess.
Developed Python script which ran over the parsed output and extracted the structures which had both the relevant grammatical relations (*iobj*; *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 purply”.
  - Instances with (unambiguous) spatial goals
  - Non-alternating idioms
### Evaluation of the Automated Extraction: False Positives

<table>
<thead>
<tr>
<th></th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>PP</th>
<th>Total</th>
<th>After Filtering</th>
<th>% retained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown</td>
<td>854</td>
<td>666</td>
<td>1520</td>
<td></td>
<td></td>
<td>819</td>
<td>53%</td>
</tr>
<tr>
<td>Frown</td>
<td>992</td>
<td>773</td>
<td>1765</td>
<td></td>
<td></td>
<td>759</td>
<td>43%</td>
</tr>
<tr>
<td>LOB</td>
<td>805</td>
<td>759</td>
<td>1564</td>
<td></td>
<td></td>
<td>765</td>
<td>49%</td>
</tr>
<tr>
<td>FLOB</td>
<td>834</td>
<td>1076</td>
<td>1910</td>
<td></td>
<td></td>
<td>771</td>
<td>40%</td>
</tr>
</tbody>
</table>
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.
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
Length in words is a convenient proxy for syntactic complexity (Szremscányi 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

Scott Grimm and Joan Bresnan
Spatiotemporal variation in the dative alternation: a study of four corpora of British and American English
Distribution: length of recipient arguments

Scott Grimm and Joan Bresnan
Spatiotemporal variation in the dative alternation: a study of four corpora of British and American English
Annnotation: 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

Scott Grimm and Joan Bresnan
Spatiotemporal variation in the dative alternation: a study of four corpora of British and American English
Distribution: recipient pronominality

Scott Grimm and Joan Bresnan
Spatiotemporal variation in the dative alternation: a study of four corpora of British and American English
"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).”
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

Scott Grimm and Joan Bresnan

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

Scott Grimm and Joan Bresnan
Spatiotemporal variation in the dative alternation: a study of four corpora of British and American English
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 $\alpha$ 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.

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

where \( X \) is the model matrix of independent variables \([x_1, x_2, \ldots]\) and \( \beta \)s are their coefficients.
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).
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

<table>
<thead>
<tr>
<th>Factor</th>
<th>Odds</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log of Recipient Length</td>
<td>6.67</td>
<td>0.000</td>
</tr>
<tr>
<td>Log of Theme Length</td>
<td>0.15</td>
<td>0.000</td>
</tr>
<tr>
<td>Recipient Pronoun</td>
<td>0.09</td>
<td>0.000</td>
</tr>
<tr>
<td>90’s Corpus</td>
<td>0.63</td>
<td>0.032</td>
</tr>
<tr>
<td>Theme Pronoun</td>
<td>3.48</td>
<td>0.059</td>
</tr>
<tr>
<td>Recipient Thematicity</td>
<td>0.98</td>
<td>0.052</td>
</tr>
</tbody>
</table>
Model Results: US corpora interaction effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Odds</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theme Pronoun * 90’s Corpus</td>
<td>10.42</td>
<td>0.014</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Factor</th>
<th>Odds</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log of Recipient Length</td>
<td>7.00</td>
<td>0.000</td>
</tr>
<tr>
<td>Log of Theme Length</td>
<td>0.21</td>
<td>0.000</td>
</tr>
<tr>
<td>Recipient Pronoun</td>
<td>0.03</td>
<td>0.000</td>
</tr>
<tr>
<td>Theme Pronoun</td>
<td>8.18</td>
<td>0.000</td>
</tr>
<tr>
<td>90’s Corpus</td>
<td>0.44</td>
<td>0.001</td>
</tr>
<tr>
<td>Theme Thematicity</td>
<td>1.15</td>
<td>0.002</td>
</tr>
<tr>
<td>Recipient Thematicity</td>
<td>0.90</td>
<td>0.030</td>
</tr>
</tbody>
</table>
Model Results: UK corpora interaction effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Odds</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recipient Pronoun * 90’s Corpus</td>
<td>6.12</td>
<td>0.000</td>
</tr>
<tr>
<td>Theme Thematicity * 90’s Corpus</td>
<td>0.86</td>
<td>0.001</td>
</tr>
<tr>
<td>Recipient Thematicity * 90’s Corpus</td>
<td>10.42</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Stronger preference for \( V \ Pron \ NP \) in the FLOB corpus

- Consistent with increasing harmonic alignment
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.
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)
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
The C statistic (discriminatory accuracy) for the models was quite high .96
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 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)
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 Szmrecsanyi 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:


References:


References:


References:


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.


References: