Weighing in on End Weight

Stephanie Shih
Department of Linguistics
Stanford University

Jason Grafmiller
Department of Linguistics
Stanford University

Linguistic Society of America
85th Annual Meeting
Pittsburgh, Pennsylvania
January 9, 2011
The Principle of End Weight

• “Phrases are presented in order of increasing weight.” (Wasow 2002: 3; following Behagel 1909; Quirk et al. 1985)
  (1) peas and carrots > carrots and peas
  (2) the attitude of people who are really into classical music and feel that if it’s not seventy-five years old, it hasn’t stood the test of time >
  people who are really into classical music and feel that if it’s not seventy-five years old, it hasn’t stood the test of time’s attitude

• Facilitates planning, production, and parsing
• Cross-linguistic weight at peripheries
What is “weight”?  

**Syntax**

- Syntactic complexity: heavy constituents are structurally more complex.
  - Number of syntactic nodes (e.g., Hawkins 1994)

![Syntax Tree](image.png)
What is “weight”?  
Processing load

• Weight as structural integration cost: heavy constituents require more computational effort
  – Cost of relating an input into a projected structure depends on intervening computations
  – Dependency Locality Theory (Gibson 1998, 2000; Temperley 2007):
  – Each new referent (NP or finite verb) adds to integration cost
What is “weight”?  
Phonology

• Phonological complexity: Heavy constituents have complex prosodic properties
  – Number of primary stressed syllables (Anttila et al. 2010; following Selkirk 1984; Zec and Inkelas 1990)

• Phonological weight:
  – Number of syllables (Benor and Levy 2006; McDonald et al. 1993; a.o.)
What is “weight”?  
Word Count

• Many studies have used word count as proxy for other weight factors. (e.g., Wasow 2002; Szmrecsányi 2004; Bresnan and Ford 2010; a.o.)

• Correlated with many other measures
Which measure is appropriate?

• Most studies of syntactic alternations focus on syntactic/processing measures of weight

• Influence of phonological weight on syntax less understood

• Multiple weight measures rarely evaluated concurrently on the same data (cf., Szmrecsanyi 2004)
Present Study
The Data

• Two constructions in spoken American English
  (Switchboard Corpus, Godfrey & McDaniels 1992)

  (1) Genitive Alternation
  • ‘s -genitive ~ of genitive
  • e.g., the car’s wheel ~ the wheel of the car

  (2) Dative Alternation
  • double object construction ~ prepositional dative (to)
  • e.g., give the dog the bone ~ give the bone to the dog
Present Study

Weight measures investigated

• Syntactic nodes
• Referents (discourse new)
• Words
• Syllables
• Primary stressed syllables
Present Study
Analyses

• Simple and mixed effects regression modeling (Shih et al. 2009; Shih et al. submitted; Hinrichs & Szmrecsányi 2007; Bresnan et al. 2007; Bresnan & Ford 2010; a.o.)

  – 5 individual models using each weight predictor
  – Controlled for other known variables influencing construction choice
  – Model comparison using Akaike Information Criterion (Burnham & Anderson 2004)

• Variable comparison using Random Forests analysis (Strobl et al. 2009b)
  – Single model containing all predictors
Genitives
Fixed Effects Model

• 663 *of*-genitives + 460 *s*-genitives = 1123 total

• Predictors: Possessor animacy, final sibilancy, rhythm (Shih et al. 2009; submitted)

• Comparative weight (Bresnan & Ford 2010)

\[
\text{Comparative weight} = \log(\text{possessor weight}) - \log(\text{possessum weight})
\]

s-genitive favored

\[ - \]

0

\[ + \]

of-genitive favored

(*Referent counts were not log-transformed.*)
Genitives: results

Heavy Possessors favor *of*-gen

- Higher log odds value = higher *s*-genitive likelihood
- Lower log odds value = higher *of*-genitive likelihood

➢ As the number of words in the possessor increases relative to the number of words in the possessum, an *of*-genitive becomes more likely.

![Graph showing the relationship between log odds and word count](image.png)
Genitives: results

Individual Regression Analysis

• Nodes
  – $\beta = -1.234; z = -6.67; p < 0.000 (***)$

• Words
  – $\beta = -0.884; z = -5.50; p < 0.000 (***)$

• Referents
  – $\beta = -0.563; z = -3.71; p < 0.001 (**)$

• Primary Stresses
  – $\beta = -0.525; z = -3.44; p < 0.001 (**)$

• Syllables
  – $\beta = -0.412; z = -3.42; p < 0.001 (**)$
Genitives: results
High correlation of factors

- Syllable Count: $\rho = 0.62$
- Node Count: $\rho = 0.89$
- Stress Count: $\rho = 0.60$
- Referent Count: $\rho = 0.11$
## Genitives

### Model AICs and factor weights

<table>
<thead>
<tr>
<th></th>
<th>Nodes**</th>
<th>Words</th>
<th>Referents</th>
<th>Stresses</th>
<th>Syllables</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>809.962</td>
<td>821.277</td>
<td>836.889</td>
<td>841.002</td>
<td>841.416</td>
<td>851.218</td>
</tr>
<tr>
<td>(\Delta (AIC_m - AIC_{\text{min}}))</td>
<td>0.00</td>
<td>11.315</td>
<td>26.927</td>
<td>31.04</td>
<td>31.454</td>
<td>41.256</td>
</tr>
<tr>
<td>(w_m)</td>
<td>0.997</td>
<td>0.003</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

- Models with \(\Delta < 2\) have substantial support; \(\Delta > 10\) have no support.
- \(w_m\) = the probability that the model is the optimal one in the set (Burnham and Anderson 2006)
Genitives

Comparison of Models

![Bar Chart]

Weight measure(s) in model vs. Model AIC - Minimum AIC (in model set)

- Nodes
- Words
- Referents
- Stresses
- Syllables
- None
Datives

Mixed Effects Model

• 227 double objects + 183 prepositionals = 410 total

• Mixed effects model (Bresnan et al. 2007; Bresnan and Ford 2010)
  – Fixed effects: animacy of recipient, accessibility of recipient and theme, definiteness of recipient and theme
  – Random effects: Verb

Comparative weight = \log(\text{recipient weight}) - \log(\text{theme weight})
Datives: results
Individual Regression Analysis

• Nodes
  – $\beta = 1.312; z = 6.685; p < 0.000$ (***)

• Words
  – $\beta = 1.186; z = 6.877; p < 0.000$ (***)

• Primary Stresses
  – $\beta = 1.013; z = 6.304; p < 0.000$ (***)

• Syllables
  – $\beta = 1.040; z = 6.086; p < 0.000$ (***)

• Referents
  – $\beta = 0.207; z = 1.305; p = .19$
Datives: results
High correlation of factors

- **Syllable Count** \( \rho = 0.69 \)
- **Node Count** \( \rho = 0.94 \)
- **Stress Count** \( \rho = 0.83 \)
- **Referent Count** \( \rho = 0.35 \)
## Datives

### Model AICs and factor weights

<table>
<thead>
<tr>
<th></th>
<th>Words**</th>
<th>Nodes**</th>
<th>Stresses</th>
<th>Syllables</th>
<th>None</th>
<th>Referents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AIC</strong></td>
<td>397.77</td>
<td>398.58</td>
<td>409.81</td>
<td>414.32</td>
<td>456.58</td>
<td>456.90</td>
</tr>
<tr>
<td>Δ (AIC&lt;sub&gt;m&lt;/sub&gt; – AIC&lt;sub&gt;min&lt;/sub&gt;)</td>
<td>0.00</td>
<td>0.81</td>
<td>12.04</td>
<td>16.55</td>
<td>58.81</td>
<td>59.13</td>
</tr>
<tr>
<td>w&lt;sub&gt;m&lt;/sub&gt;</td>
<td>0.60</td>
<td>0.40</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

- Models with Δ < 2 have substantial support; Δ > 10 have no support
- w<sub>m</sub> = the probability that model is the optimal one in the set
Datives

Comparison of Models (AIC)

Model AIC - Minimum AIC
(in model set)

Weight measure(s) in model

- words
- nodes
- stresses
- syllables
- none
- referents
Random Forests

• Suited to datasets with complex interactions and highly correlated predictor variables (Strobl et al. 2008; 2009a; 2009b; a.o.)

• Recursive partitioning method:
  – Random subsamples of data, each fit with a single classification tree.
  – Randomly restricted set of predictor variables to select from in each split.

• Detects contributions and behavior of predictor variables otherwise masked by competitors.
Random Forests

Conditional Variable Importance and Model Parameters

• Conditional Variable Importance
  – Permutation Accuracy: the difference in model accuracy before and after randomly permuting the values of a given independent variable, averaged over all trees in the forest. (Strobl et al. 2009b)
  – Ranks the importance of independent variables.

• Model parameters:
  – Genitives: $ntree = 2000; mtry = 3$
  – Datives: $ntree = 8000; mtry = 3$

• Model stability verified on two random seeds.
Genitives | Random Forests

Variable Importance

Predictors to the right of dashed vertical line are significant.
Genitives | Random Forests

Variable Importance

Variable Importance in Genitives (animacy not shown)
Predictors to the right of dashed vertical line are significant.

- Syntactic nodes
- Referents
- Primary stresses
- Words
- Syllables
### Genitives

**AIC vs. Random Forests**

<table>
<thead>
<tr>
<th>Genitives</th>
<th>AIC</th>
<th>Random Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Syntactic nodes</strong> &gt;</td>
<td><strong>Syntactic nodes</strong> &gt;</td>
</tr>
<tr>
<td></td>
<td><strong>Words</strong> &gt;</td>
<td><strong>Referents</strong> &gt;</td>
</tr>
<tr>
<td></td>
<td><strong>Referents</strong> &gt;</td>
<td><strong>Primary stresses</strong> &gt;</td>
</tr>
<tr>
<td></td>
<td><strong>Primary stresses</strong> &gt;</td>
<td><strong>Words</strong> &gt;</td>
</tr>
<tr>
<td></td>
<td><strong>Syllables</strong></td>
<td><strong>Syllables</strong></td>
</tr>
</tbody>
</table>
Datives | Random Forests

Variable Importance

- Animacy of Recipient
- Word Count
- Givenness of Theme
- Node Count
- Referent Count
- Primary Stress Count
- Definiteness of Theme
- Givenness of Recipient
- Definiteness of Recipient
- Syllable Count

Variable Importance in Datives
Predictors to the right of dashed vertical line are significant.
## Summary

### AIC vs. Random Forests

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>Random Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genitives</strong></td>
<td><strong>Syntactic nodes &gt;</strong></td>
<td><strong>Syntactic nodes &gt;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Words &gt;</strong></td>
<td><strong>Referents &gt;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Referents &gt;</strong></td>
<td><strong>Primary stresses &gt;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Primary stresses &gt;</strong></td>
<td><strong>Words &gt;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Syllables</strong></td>
<td><strong>Syllables</strong></td>
</tr>
<tr>
<td><strong>Datives</strong></td>
<td><strong>Words &gt;</strong></td>
<td><strong>Words &gt;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Syntactic nodes &gt;</strong></td>
<td><strong>Syntactic nodes &gt;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Primary stresses &gt;</strong></td>
<td><strong>Referents &gt;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Syllables&gt;</strong></td>
<td><strong>Primary stresses &gt;</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Referents</strong></td>
<td><strong>Syllables</strong></td>
</tr>
</tbody>
</table>
Discussion

Syntactic Complexity

- Number of syntactic nodes = best *individual* predictor of end weight in English genitive and dative construction choice.

- Is “weight” purely syntactic?
  - English binomial ordering studies: number of syllables affect ordering of nouns in binomial pairs. (Wright et al. 2005; cf., McDonald et al. 1993; Benor & Levy 2006)

- At a higher-level domain (i.e., genitives, datives), syntactic complexity is the most salient manifestation of “weight.”
Discussion

Word count as a proxy

- Methodologically, the number of words—though not perfect—can act as a sufficient proxy for syntactic complexity and ‘weight’.

- Dative construction choice:
  - Syntactic nodes and words are the best measures in comparison to the other measures tested.

- Genitive construction choice:
  - AIC: words are second best, though not great.
  - Random forest: not the most important measure
Discussion

Referents and DLT

• In comparison, referents are not the best measures of weight.
  = Gibson (1998; 2000): Non-given and definite nouns and verbs

• What can contribute to integration costs? (Temperley 2006)
  e.g., the green ball
  Gibson: x = 1 new referent
  alternatively: x x = 2 new referents

➢ Redefinition of “referents” -> content words?
Discussion
Phonological complexity and weight

• Stresses and syllables rank low as good measures of weight for genitive and dative construction choice.

• Prosodic theory of end weight (=number of primary stresses) is not entirely syntax-independent.
  – phonological words ≈ content words

• Do possible phonetic correlates of weight or complexity play into end weight effects?
  – e.g., duration, complexity of segments, syllable weight or complexity of syllable structure (e.g., Benor & Levy 2006)
Future directions

Weight Beyond English

• How do measures of weight generalize beyond English?

• Is there a better proxy for cross-linguistic syntactic complexity?
  – Morphological complexity and weight?
Conclusion

• Two statistical methods resistant to collinearity:
  – AIC model comparison and selection
  – Random forest conditional variable importance

• Two alternations in spoken American English:
  – Genitives | Datives

• Tested syntactic, processing, and phonological measures of “weight.”
  – Syntactic nodes (syntactic complexity)
  – Referents (Dependency Locality Theory)
  – Words
  – Primary stresses (phonological complexity)
  – Syllables (phonological weight)
Conclusion

- Syntactic-based measures contribute most to weight-driven alternations in higher-level constituent ordering
  - (though perhaps heavily theory dependent)

- Methodologically, the number of words can be an appropriate and sufficient proxy for (syntactic) complexity and weight.

- “Weight” effects cannot be reduced to a single dimension.
Thank you!

Thank you to Arto Anttila for the initial push and subsequent support on this project. Many thanks must go to Joan Bresnan for invaluable advice, discussion, and patience. Acknowledgements also to Matthew Adams, Susanne Gahl, Sharon Inkelas, Victor Kuperman, Beth Levin, Robin Melnick, Daphne Theijssen, Benedikt Szmarcsányi, Tom Wasow, Christoph Wolk, the audience at the Development of Syntactic Alternations workshop, and the members of Quorum (UC Berkeley) for further discussion, support, skepticism, questions, and statistical aid.

The authors’ names are listed in reverse alphabetical order so as to satisfy the Principle of End Weight.

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.

Contact:

Stephanie Shih
stephsus@stanford.edu

Jason Grafmiller
jasong1@stanford.edu

Slides available online: http://stanford.edu/~stephsus/ShihGrafmillerLSA2011.pdf
Selected References


Bresnan, Joan; Anna Cueni; Tatiana Nikitina; and R. Harald Baayen. 2007. Predicting the Dative Alternation. in G. Bouma,; I. Kraemer; and J. Zwarts (ed). *Cognitive Foundations of Interpretation*. Royal Netherlands Academy of Science. 69-94.


Shih, Stephanie; Jason Grafmiller; Richard Futrell; and Joan Bresnan. 2009. Rhythm’s role in genitive and dative construction choice in spoken English. Paper presented at the 31st annual meeting of the Linguistics Association of Germany (DGfS). University of Osnabrück, Germany. 4 March 2009.

Shih, Stephanie; Jason Grafmiller; Richard Futrell; and Joan Bresnan. submitted. Rhythm’s role in predicting genitive construction choice in spoken English.


