Variation in the English comparative adjective
Conflict across prosodic levels

Matthew E. Adams

Department of Linguistics
Stanford University
matthewa@stanford.edu

Exploring the Interfaces
Syntactic Interfaces Research Group
McGill University and UQAM
May 6–8, 2012
Synthetic versus analytic realization

(1)  a. **Synthetic, use of the suffix –er**
   i. A picometer is *smaller* than an nanometer.
   ii. Nico was *happier* than Lou.

   b. **Analytic, use of the independent word more**
   i. Merlin is *more cryptic* than Charlie.
   ii. Berlin is *more salubrious* than Weimar.
Dimensions of variation

- Adjectives of one or two syllables readily adopt –er (e.g., prouder, higher, happier)
- Often, however, these adjectives also accept more: more proud, more happy (Mondorf, 2009; Hilpert, 2008)
- Adjectives of three or more syllables are nearly categorically preferential to more: more intelligent, *intelligenter; cf. unhappier, more unhappy
The “cliff effect”

- Derived adjectives
  - *visc-ous*: more viscous, not *viscouser*
  - *cycl-ic*: more cyclic, not *cyclicker*
- Adjectives having undergone “inflectional” word formation processes: *puzzled, understanding*
- **Goal**: Ensure that the synthetic form is categorically blocked
The “frequency effect”

- Short but comparatively rare adjectives appear to prefer *more*
- Averaged across several corpora, *apt* occurs 95% of the time as *more apt* and only 5% of the time as *apter*
- Goes against length-based generalization for –er attachability: why is *more* more common here? What blocks the synthetic realization of the comparative?
- **Goal:** Ensure variable outputs for shorter adjectives, modulated (in part) by the frequency of the adjectives
Syntactic position further drives variation

- Further complicating the picture is the empirical observation (Lindquist, 2000; Mondorf, 2009; Adams, 2012) that syntactic position affects the rate of synthetic and analytic usage.
- Gross generalization: *more* is more tolerated in predicative environments, less so in attribute ones, all else being equal.
- In cases where only –*er* suffixation is impossible, *more* appears across all environments.
What is meant by syntactic environment?

- In this talk, a simple binary distinction between ATTRIBUTIVE and PREDICATIVE positions:

  (2) Attributive
  a. The *stockier/more stocky* linguist is on the grass.
  b. The *more intelligent* lunatic is on the grass.

  (3) Predicative
  a. The linguist is *more stocky/stockier* than the lunatic.
  b. The lunatic is *more intelligent* than the linguist.

- In study, approximately 50,000 tokens were extracted from the Corpus of Contemporary American English (coca, Davies 2008) for a variety of adjective types, lengths, frequencies, across these two syntactic environments.
Distribution of four adjectives by syntactic environment and morphosyntactic realization

<table>
<thead>
<tr>
<th>Adjective</th>
<th>Attr. –er</th>
<th>Attr. more</th>
<th>Pred. –er</th>
<th>Pred. more</th>
</tr>
</thead>
<tbody>
<tr>
<td>fast</td>
<td>4503</td>
<td>3</td>
<td>672</td>
<td>326</td>
</tr>
<tr>
<td>happy</td>
<td>348</td>
<td>175</td>
<td>265</td>
<td>272</td>
</tr>
<tr>
<td>apt</td>
<td>23</td>
<td>125</td>
<td>2</td>
<td>67</td>
</tr>
<tr>
<td>intelligent</td>
<td>0</td>
<td>54</td>
<td>0</td>
<td>43</td>
</tr>
</tbody>
</table>

Hand-corrected raw numbers from the spoken portions of the Corpus of Contemporary American English (COCA, Davies 2008), supplemented by hand-corrected Google counts
Initial observations and claims

- **Word-level variation:** faster, more fast; apter, more apt; more intelligent, *intelligenter

- **Phrase-level variation:** increased tolerance for more in predicative position, –er in attributive position

- **Conflict among levels:** apter is dis-preferred in comparison to more apt, but in attributive position, –er is more preferred
Plan of the rest of the talk

• Explain the proposed architecture of a stratally organized grammar with stochastic selection of outputs
• Demonstrate a simple word-level grammar that relates frequency to prosodic structure, and prosodic structure to suffixability
• Bring this grammar to the phrase level, showing how a word-level preference (synthetic form) may go against a phrase-level preference (analytic form)
• Discuss how comparative grammaticality in a stratally organized grammar leads to partial blocking in some instances, full blocking in others
Stratally organized grammar

Input

Stem-Level Morphology ↔ Stem-Level Phonology

Word-Level Morphology → Word-Level Phonology

Phrase-Level Morphosyntax → Postlexical Phonology

Output

Matthew E. Adams  Variation in the English comparative adjective
Stochastic outputs at each level

- **Input**: *fast+er* is *input* to word level
- Stochastic selection of outputs: in the vast majority of cases, *faster* is the selected output; in a small minority of cases, *Deg* is not morphologically realized (*fast* is therefore the output)
- Most of the inputs the phrase level receives for the comparative of *fast* are therefore *faster*. A small set of inputs are *Deg+fast*. The analytic morpheme *more* is inserted to express the comparative degree (as the word level is now closed off to further morphophonological manipulation)
Constraining the input set

- Stratal organization allows candidate set to be constrained at each level. *faster* and *more fast* do not compete at the word level.
- Rather, *faster* competes at word level with a potential output form that fails to attach the suffix (*fast*): realization of comparative meaning is “pushed” to the next level.
- The essence of *COMPARATIVE GRAMMATICALITY* in stratal OT: Richness of the Base is by design restricted. At the word level *more fast* is not a viable candidate because it is a multi-word construct; it is able to compete only with *faster* at the phrase level, and only in some cases.
Word level

- General claim: –er suffixation is possible only when the suffix can be incorporated into the primary stressed foot
- As a word-attaching suffix (Bermúdez-Otero, 2010), it notably can’t be combined with other word-attaching suffixes, like inflectional –ed and –ing; this follows from the non-cyclicity of the word level
- In infrequent adjectives, the prosodic representation is more detailed, and suffixation is consequently blocked in cases where visibility conditions are not met (to be discussed in more detail)
Phrase level

• General claim: *more* is avoided in attributive position because of the **Stress-to-Stress Principle** (Anttila et al., 2010), which discourages multiple lexical stresses in environments that cannot receive phrasal stress.

• While historically *more* has gained in precedence across the language (Kytö and Romaine, 1997), the one area that is slowest to take over is with *short* adjectives in *attributive* position.
Cohering suffixes and the prosodic foot

- *–er* is a cohering suffix (Booij, 1995; Raffelsiefen, 1999); it needs to fit inside some prosodic unit.
- Proposal: *–er* must occupy the weak position in a prosodic foot that carries main stress: *(fás.ter),* *(hígh.er)*, but not *ma(lí.ciou)ser* (rendered extrametrical via FT-BIN) or *in(tél.li)(gèn.ter)* (in a secondary stressed foot).
- An alignment constraint can enforce this:
- **ALIGN(Suffix, R, Σ, R),** where Σ is a prosodic foot, indicated by parentheses.
The dimensions of variation  
Overview of the stratal stochastic model  
Word-level variation  
Phrase-level variation  
References

What can –er attach to?  
The beta distribution  
Prosodic representation

<table>
<thead>
<tr>
<th>FT-BIN</th>
<th>ALIGN-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>(fást) +er → (fás.ter)</td>
<td></td>
</tr>
<tr>
<td>ma(lí.cious) +er → ma(lí.ciou.ser)</td>
<td>!*</td>
</tr>
<tr>
<td>ma(lí.cious) +er → ma(lí.ciou)ser</td>
<td>!</td>
</tr>
<tr>
<td>ma(lí.cious) +er → ma(lí.cious)</td>
<td>!</td>
</tr>
</tbody>
</table>

Matthew E. Adams  
Variation in the English comparative adjective
But what about infrequent, short adjectives like *apt*?

- Based on constraints so far, the output form (*áp.ter*) should be perfectly acceptable, but usage data show that *more apt* is far more frequent than *apter*.
- Short answer: Frequency matters (Hilpert, 2008)
- Longer answer: The beta distribution describes the frequency profile of different adjectives; different areas under the curve correspond to different prosodic representations, which in turn affect the suffixability of the adjective
- The longer answer helps explain *why* frequency should affect word edges, including morphological processes like suffixation
The beta distribution

- Coetzee (2009) applies the beta distribution to explaining how frequency interacts with English $t/d$-deletion: the more rare a word, the more likely it is to preserve $t$ or $d$ in coda position at the word edge.
- The beta distribution has desirable properties for handling frequency: its $x$ axis ranges from 0 to 1, appropriate for probabilities, and by manipulating two parameters, $\alpha$ and $\beta$, the shape of the line changes.
- In current study I graphed the beta distributions of 465 adjectives, using how much they diverged from the mean frequency of adjectives in the corpus as a basis for calculating the $\alpha$ and $\beta$ parameters, which determine the shape of the curve.
Some sample beta distributions
The dimensions of variation
Overview of the stratal stochastic model

**Word-level variation**

**Phrase-level variation**

References

What can –er attach to?

**The beta distribution**

Prosodic representation

Matthew E. Adams

Variation in the English comparative adjective
Driving suffixation via satisfaction of a total order of \( \text{Max} \) constraints

- To drive suffixability, the input to the word level is assumed to include the base of the positive adjective (e.g., fast or apt), concatenated to \(-er\)
- A series of \( \text{Max}(\text{Suffix}) \) constraints are proposed (in the simplified picture, just two)
- The probability (\(x\)-axis) line is bisected by a vertical line, which partitions the probability mass under the curve into two regions
- Each region corresponds to a particular \( \text{Max} \) constraint, supplied with its own subscript (e.g., \( \text{Max}_1, \text{Max}_2 \))
The probability that a particular adjective will fall into one Max region or another directly follows from its frequency, which is used to determine the parameters that give the beta line its shape.

- In *apt*, a low-frequency adjective, most of the probability mass is located in the left partition, which is labeled $\text{Max(Suffix)}_1$.
- In *fast* or *high*, high-frequency adjectives, most of the probability mass is located in the right partition, labeled $\text{Max(Suffix)}_2$.
- The Max constraints prohibit deletion of suffixal material.
By calculating the region under the curve bounded by the midsection line, the probability that an adjective will fall into either Max region can be determined.

At time of selection, the adjective is indexed with the region it fell under: apt has a strong probability of being selected from the region labeled Max(Suffix)\(_1\).

If –er is not present in the candidate form, the Max constraint matching that subscript will be violated:

<table>
<thead>
<tr>
<th>apt(_1)+er</th>
<th>Max(Suffix)(_2)</th>
<th>Max(Suffix)(_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>apt(_1)</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>apter(_1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
• The number of times *apt₂* will be selected is much smaller, as given by the beta distribution, but it is a marginal possibility:

<table>
<thead>
<tr>
<th>apt₂+er</th>
<th>Max(Suffix)₂</th>
<th>Max(Suffix)₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>apt₂</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>apter₂</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• The interaction of *markedness* constraints gives variable outputs for low- and high-frequency adjectives

• When a markedness constraint is situated between the two *Max* constraints, sometimes the suffixed form will win; sometimes it will not (it depends on the shape of the beta line, which gives the probability that a particular indexed form will be selected)
• Assume a general markedness constraint exists that militates against suffixation:

• The frequent case: *apt* is the preferred word-level output, even though it violates the relevant \text{Max} constraint:

<table>
<thead>
<tr>
<th>apt_{1}+er</th>
<th>\text{Max(Suffix)_2}</th>
<th>Markedness</th>
<th>\text{Max(Suffix)_1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{/apt}_{1}</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>\text{apter}_{1}</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• The far less frequent situation: *apter* is selected

<table>
<thead>
<tr>
<th>apt_{2}+er</th>
<th>\text{Max(Suffix)_2}</th>
<th>Markedness</th>
<th>\text{Max(Suffix)_1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{apt}_{2}</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\text{apter}_{2}</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
• Based on the reverse shape for high-frequency adjectives, the opposite case obtains

• The frequent case: *higher* is the preferred word-level output, even though it violates some relevant markedness constraint:

<table>
<thead>
<tr>
<th>high(_2)+er</th>
<th>(\text{Max} (\text{Suffix})_2)</th>
<th>Markedness</th>
<th>(\text{Max} (\text{Suffix})_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>high(_2)</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☞ higher(_2)</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• The far less frequent situation: *high* is selected

<table>
<thead>
<tr>
<th>high(_1)+er</th>
<th>(\text{Max} (\text{Suffix})_2)</th>
<th>Markedness</th>
<th>(\text{Max} (\text{Suffix})_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>☞ high(_1)</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>higher(_1)</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

• By dividing the area under the beta curve into two regions, corresponding to two $\text{Max(Suffix)}$ constraints and sandwiching a markedness constraint in between the two faithfulness constraints, *variable outputs* can be obtained.

• Sometimes the word-level output includes the comparative morpheme; sometimes it doesn’t.

• Two important questions:
  • What do the indexes indicate beyond being a notational device (an unspecified “lexical class” in Coetzee’s work)?
  • What are the markedness constraints that militate against suffixation?
What are the markedness constraints?

- The first markedness constraint has already been introduced: \( \text{ALIGN}(\text{Suffix}, R, \Sigma, R) \)
- To drive categoricity, this constraint is undominated (it dominates all of the \( \text{Max} \) constraints)
- Another relevant markedness constraint: \( \text{Onset} \)

<table>
<thead>
<tr>
<th>high(_2) + er</th>
<th>( \text{Max}(\text{Suffix})_2 )</th>
<th>( \text{Onset} )</th>
<th>( \text{Max}(\text{Suffix})_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>high(_2)</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \bowtie ) higher(_2)</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>high(_1) + er</th>
<th>( \text{Max}(\text{Suffix})_2 )</th>
<th>( \text{Onset} )</th>
<th>( \text{Max}(\text{Suffix})_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bowtie ) high(_1)</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>higher(_1)</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Matthew E. Adams  Variation in the English comparative adjective
What do the indexes mean?

• Rather than the nebulous idea of a “lexical class,” I propose that the indexes represent different prosodic representations.
• Thus, low- and high-frequency adjectives have potentially different representations.
• In general, low-frequency adjectives will be more carefully prosodified, respecting conditions on syllable weight; high-frequency adjectives will have “greedy” prosodifications that are less detailed and potentially violate coda weight.
The dimensions of variation
Overview of the stratal stochastic model
Word-level variation
Phrase-level variation
References

What can –er attach to?
The beta distribution
Prosodic representation

Matthew E. Adams
Variation in the English comparative adjective
• When *apt* falls under the part of the beta curve labeled with the 1 subscript, it is part of a class of analytically stored adjectives (Bermúdez-Otero, 2006): a detailed prosodic representation is present.

• In the unlikely event *apt* falls under the other part of the curve, it will have a more sloppy prosodification, as exemplified by *fast*.

• *fast* violates a $^*\mu\mu\mu$ (an anti-trimoracity constraint)
- Another relevant markedness constraint respects Structure Preservation (Borowsky, 1993)
- \textit{*Resyllabify-EM}: Do not resyllabify a consonant marked as extrametrical in the input
- The important upshot is that \textit{apter} will violate this constraint (\textit{faster} will not because only \textit{apt} has an extrametrical \textit{t} in the input when it is drawn from region indexed as \textit{1})

\begin{tabular}{|c|c|c|c|}
\hline
\textit{(fast)}\textsubscript{2}+er & \textit{Max(Suffix)}\textsubscript{2} & \textit{*Resyll-EM} & \textit{Max(Suffix)}\textsubscript{1} \\
\hline
\textit{(fast)}\textsubscript{2} & & \textbf{!} & \\
\hline
\textit{(fas.ter)}\textsubscript{2} & & \textbf{*} & \\
\hline
\end{tabular}

\begin{tabular}{|c|c|c|c|}
\hline
\textit{(ap)t}\textsubscript{1}+er & \textit{Max(Suffix)}\textsubscript{2} & \textit{*Resyll-EM} & \textit{Max(Suffix)}\textsubscript{1} \\
\hline
\textit{(ap)t}\textsubscript{1} & & & \textbf{*} \\
\hline
\textit{(ap.ter)}\textsubscript{1} & & \textbf{!} & \\
\hline
\end{tabular}

Matthew E. Adams  
Variation in the English comparative adjective
Stochasticity and prosodic representation

- Frequency does not directly affect deletion processes; rather, prosodic representation provides structure that is more or less prone to deletion.
- “Frequency matters,” but it’s unsuccessful by itself in explaining why right word edges and codas are more susceptible to deletion than onsets or stressed nuclei.
- A prosodic/representational solution is more than just a set of diacritics: it provides an explanation for why codas, and not onsets, are deleted.
Inputs

- The inputs to the phrasal level are the stochastically selected outputs form the word level.
- In a case like *fast*, the majority of outputs will be *faster*; hence the majority of inputs evaluated for the comparative degree of *fast* at the phrase level will have *faster* as an input.
- A minority of inputs will be *fast* (phonetically unrealized morpheme).
- The opposite case holds for an adjective like *apt*: most inputs will not have –er attached to them.
- The categorical cases are handled by the fact that an input with –er never is an input (because the word level never generates forms like *intelligenter*).
The Sentence Accent Assignment Rule
(Gussenhoven, 1984)

<table>
<thead>
<tr>
<th>Attributive syntax</th>
<th>Predicative syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>VP</td>
</tr>
<tr>
<td>A</td>
<td>V</td>
</tr>
<tr>
<td>NP</td>
<td>AP</td>
</tr>
</tbody>
</table>

- Functor-argument relationship and prosodic asymmetry (Wagner, 2005):
- When a projecting element A precedes its complement B, a sequence of two prosodic domains that are on a par is derived: A B. The last domain provides the “nuclear stress”
- If A selects for the NP (Gazdar and Pullum, 1981) in attributive position, then the NP receives nuclear stress
The Stress-to-Stress Principle (Anttila et al., 2010)

• The idea of the principle is that a lexical stress wants to receive phrasal stress
• If a lexical stress sits in a position that does not receive phrasal stress, this is considered prosodically less well-formed
• Prediction: Analytic forms in which multiple elements carry lexical stress prefer the rightmost (or leftmost) edge, depending on the parametrization of nuclear stress in the language
• Upshot for the analytic comparative: the degree marker *more*, when combined with an adjective, carries lexical stress but when it’s in attributive position, it cannot receive phrasal stress.

• Recall that the outputs at the word level are variable from stochastic selection: depending on the adjective, its frequency, and its structural properties, it either outputs a higher proportion of suffixed forms or simply the positive form with an abstract degree marker (*faster* or *fast*+\text{Deg})
Expressiveness and Faith

- Because the phrase level is the last “chance,” as it were, elements that have not received phonetic form must do so: Expressiveness (Kiparsky, 2005) is undominated at this level
- Expressiveness is satisfied either by –er or by more, which can only be inserted at the phrase level
- Faith is a general faithfulness constraint that militates against changing the form of the input form of the comparative adjective
Attributive position

<table>
<thead>
<tr>
<th>the faster car</th>
<th>Expr</th>
<th>S-to-S</th>
<th>Faith</th>
</tr>
</thead>
<tbody>
<tr>
<td>☞ the (fàster) cár</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>the (mòre fàst) cár</td>
<td></td>
<td>**!</td>
<td>*!</td>
</tr>
<tr>
<td>the (Deg fàst) cár</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>the Deg apt game</th>
<th>Expr</th>
<th>S-to-S</th>
<th>Faith</th>
</tr>
</thead>
<tbody>
<tr>
<td>☞ the (mòre àpt) gáme</td>
<td></td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>the (Deg àpt) gáme</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Matthew E. Adams
Variation in the English comparative adjective
Predicative position

<table>
<thead>
<tr>
<th>is faster</th>
<th>Expr</th>
<th>S-to-S</th>
<th>Faith</th>
</tr>
</thead>
<tbody>
<tr>
<td>❅ is (fäster)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is (mòre fást)</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>is (DEG fást)</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>is DEG fast</th>
<th>Expr</th>
<th>S-to-S</th>
<th>Faith</th>
</tr>
</thead>
<tbody>
<tr>
<td>❅ is (mòre fást)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is (DEG fást)</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Selection of outputs at phrasal level

- Is deterministic in the sense that the ranking of the constraints, given an input, uniquely determines an output.
- The STRESS-TO-STRESS PRINCIPLE only matters in attributive position.
- However, this depends on the relative number of inputs that are suffixed or not suffixed: in the fast case, most of the inputs are faster (suffixed); hence, most of the outputs are predicted to be faster.
- With apt, the opposite holds: most of the inputs are unsuffixed, so most of the outputs are realized with more.
Comparative grammaticality and conflict

- The success of the analysis depends on ultimately here on what the candidate set is.
- At the phrasal level, when the input is unsuffixed, it simply does not compete with an suffixed form; there is no syntactic operation available at the phrasal level that allows the DEG morpheme to lower onto the adjective — this is sometimes that would have needed to have been accomplished at the word level.
The stratal organization of the grammar allows for this “non-comparison” to take place, which allows the analytic form to win even in a syntactic position that disfavors it: the attributive position.

Conflict across levels is resolved via procedural (stratal organization) and representational methods (prosodic structure that ultimately determines the relative number of suffixed and unsuffixed outputs from the word level to the input of the phrase level).
• The word level appears to influence how the variability in phrasal syntax operates
• In a separate regression-based study, not presented here, the word level is represented as an independent variable and turns up significant and predictive: this may provide support for the ideas (1) that lower level phonological and prosodic processes can influence some syntactic operations and (2) for a stratally organized grammar
References I


References II


