A form- and corpus- based approach to understanding Aspectual be in African American English

Charlie Farrington
Janneke Van Hofwegen

LSA Annual Meeting
Portland, OR; January 8, 2015
Revisiting a classic variable

Aspectual be

• Found widely across many varieties of African American English (AAE).
• Marks habituality (Labov et al. 1968) / iterativity (Fasold 1972) / recurring eventualities (Green 2000)

Examples:

Sometimes they be working.
She usually be here.
Principle of Accountability

“For the section of speech being examined all occurrences of a given variant are noted, and where it has been possible to define the variables as a closed set of variants, all non-occurrences in the relevant environments” (Labov 1982, p. 30).
Quantifying Aspect

• Aspectual constructions are notably difficult to tabulate (Walker (ed.) 2010)
• With Aspectual be in sociolinguistics…
  • Raw counts are often employed (Wolfram 1969)
  • Number of tokens per hour used (Rickford 1999)
  • Quantification within the copula system (Bailey & Maynor 1987)
A classic variable in the 21st century

**BIG DATA:** Large-scale, transcribed speech corpora become more and more widely available. Linguists should use them!

Problem:
- Often transcribed by machines or naïve humans (i.e., not linguists!)

But…

If we’re dealing with an overt, invariant form, we can largely trust these sorts of text corpora.
Frank Porter Graham Corpus  
AAE, Durham/Chapel Hill, NC

<table>
<thead>
<tr>
<th>Time Point</th>
<th>Age</th>
<th>Interaction Types</th>
<th>Speech samples*</th>
<th>Speakers</th>
<th>C-Units**</th>
</tr>
</thead>
<tbody>
<tr>
<td>G6</td>
<td>11</td>
<td>Peer interaction; Adult formal/informal interview</td>
<td>74</td>
<td>122</td>
<td>15801</td>
</tr>
<tr>
<td>G8</td>
<td>13</td>
<td>Peer interaction</td>
<td>67</td>
<td>113</td>
<td>14025</td>
</tr>
<tr>
<td>G10</td>
<td>15</td>
<td>Peer interaction</td>
<td>66</td>
<td>99</td>
<td>15711</td>
</tr>
<tr>
<td>PHS</td>
<td>19-20</td>
<td>Sociolinguistic interview</td>
<td>64</td>
<td>64</td>
<td>10866</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td></td>
<td></td>
<td>211</td>
<td>398***</td>
<td>56403</td>
</tr>
</tbody>
</table>

* Speech samples are interactions between two speakers at G6 through G10, and one speaker (and an interviewer) at PHS.

** Communication Units (a.k.a. C-units) an independent clause plus its modifiers (c.f. Loban, 1976)

*** 170 unique speakers, most are recorded at multiple time points
Raw Counts for IBE across Age/Grade*

Wolfram and Van Hofwegen (2012)
(See also Van Hofwegen and Wolfram 2010)

*31 FPG longitudinal speakers
Construction Grammar

• Traditionally in sociolinguistics, Aspectual *be* is talked about as a function-based form.

• Construction grammar (or a purely functional/structural approach) allows us to think of Aspectual *be* as a form/function composite, without assuming an underlying form.

• Since Aspectual *be* is learned alongside other MAE *be* forms, this approach allows us to treat them together in an analysis.
Construction Grammar

• The traditional approach puts a lot of burden on the analyst to recognize and understand the AAE forms.

• Here, we rely on structural coding of both MAE and AAE forms, then sorting through tokens based on previous studies of Aspectual *be*.
This study

Utilizing a large text corpus of AAE we can:

1) Show how a construction-based (structural) approach informatively quantifies Aspectual be usage.

2) Examine a relatively new innovation (aspectual be like) vis-à-vis other Aspectual be forms.

   Example: “My momma be like ‘Clean your room!’” compared to “My ears be itchin’”
Form-based structural approach

1. Search for the lexical word *be* in a corpus.
2. Code all *be* construction types (both MAE (Gildea 2011) and AAE constructions).
Construction Types (N = 2855)

Auxiliary + be (n=1068)
  “It would be so funny” (1001, g8)

Semi-Auxiliary + be (n=649)
  “They useta be fussin” (1072, g8)

Imperative be (n=40)
  “Now you be quiet!” (K280, g6)

Infinitive be (n=221)
  “And I’m happy to be different” (1019, PHS)

Aspectual be (n=744)
  “He still be trying to call me” (1061, G8)
Structural Analysis

A construction-based (structural) approach informatively quantifies Aspectual *be* usage:

1. Corpus measures (frequency/spread) (Torgersen et al. 2011)
2. By all *be* constructions (Farrington & McLarty 2014)

Gives us a *measure* to incorporate in regression-based variation analyses.
Aspectual *be* Coding

Following Dayton (1996), *be* tokens are coded into the following groups:

- Be + verb-ing
- Be + adv.prep
- Be + adj
- Be + noun phrase
- Be + past.participle
- Other (be done, etc.)
- Quotative “be like”
## Aspectual *be* by construction type

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>be+verb-ing</em></td>
<td>395</td>
<td>53</td>
<td>647</td>
<td>49</td>
</tr>
<tr>
<td><em>be.like</em></td>
<td>143</td>
<td>19</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><em>be+adv./prep.</em></td>
<td>87</td>
<td>12</td>
<td>311</td>
<td>24</td>
</tr>
<tr>
<td><em>be+adjective</em></td>
<td>64</td>
<td>9</td>
<td>173</td>
<td>13</td>
</tr>
<tr>
<td><em>be+NP</em></td>
<td>37</td>
<td>5</td>
<td>68</td>
<td>9</td>
</tr>
<tr>
<td><em>be+past part.</em></td>
<td>8</td>
<td>1</td>
<td>113</td>
<td>5</td>
</tr>
<tr>
<td><em>other</em></td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>&lt;1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>742</td>
<td>100%</td>
<td>1320</td>
<td>100%</td>
</tr>
</tbody>
</table>
Innovative Invariant be. like

“My momma be like ‘Clean your room!’”

Cukor-Avila (2012): “When be like first diffused into Springville, young AAVE speakers adapted it into the AAVE tense/aspect system by…encoding invariant be…to mark present habitual be like quotatives” (p. 630).

• ~48% of 558 be like tokens were invariant/habitual.
Innovative Invariant be. like

“My momma be like ‘Clean your room!’”

Kohn and Franz (2009): “Although infrequent, the pairing of invariant be with quotative be like occurs across ethnicities and field sites” (p. 257)

• 4-7% of be like tokens for African American and Latino speakers were invariant.
Aspectual *be like* in a text corpus

Hypothesis:
Invariant *be like* = Aspectual *be like*

Challenges of a text corpus:
“I will not be attempting to tell you what the [speakers] are attempting to convey, only what their sentences are syntactically and semantically capable of meaning to the best of my knowledge of the linguistic system they are using” (Fasold 2011).
Aspectual *be like* in a text corpus

Semantics of Aspectual *be* (Green 2000):
“There is no certain number of times greater than the number one that an eventuality expressed by a predicate has to have occurred to be used in Aspectual *be*-type constructions” (p. 5).

**Iterativity**
(e.g., Fasold 1972, Green 2000, Dayton 1996)
Coding for **Iterativity** in a text corpus

Fasold (2011): “Is it possible to interpret this line as containing [Aspectual be], given my understanding of Green’s analysis?”
Coding for **Iterativity** in a text corpus

Coding:

1. **Explicitly Iterative** – presence of frequency-of-occurrence adverbs (cf. Crystal 1966) and *when* clauses. (How often? How long?)

Examples:

It seems like *every time* we go (to like) to a hotel or a motel them dudes be getting mad with us. (1976, G8)

It *usually* just be Kendall and my sister fighting. (1086, G6)

And *when* it's time for me to tell her say something she be like yeah yeah. (1020, G8)
Coding for **Iterativity** in a text corpus

Coding:


Examples:

You and your brother and sister *be fighting*? (P364, G8)

People *be having* like nicknames and all that. (2011, G10)

I *be scared* [of ordering things online]. (1072, G10)

I know but some of them mugs [boys’ faces] *be ugly*. (K283, G8)
Coding for **Iterativity** in a text corpus

Coding:
3. **Not Iterative** – iterative reading is impossible

Examples:
And he be male, 19 year[s]. (1058, G10)

She be about eighteen then? (2020, G10)

His page be easy to make. (1025, G10)

He looks like his name be Tony, no Marcus. (2037, G10)

I be writing [at this very moment]. (1061, G10)
## Results

All FPG Aspectual *be* tokens (no *be.like*) (N = 589)

<table>
<thead>
<tr>
<th>Iterativity Code</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Explicitly Iterative</td>
<td>Plausibly Iterative</td>
<td>Not Iterative</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td><strong>% Total</strong></td>
<td>% Total</td>
<td>% Total</td>
<td>% Total</td>
</tr>
<tr>
<td>177</td>
<td>30.1</td>
<td>362</td>
<td>61.5</td>
</tr>
<tr>
<td>50</td>
<td>8.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Results

All FPG Aspectual *be* tokens (no *be.like*) \((N = 539)\)

<table>
<thead>
<tr>
<th>Iterativity Code</th>
<th>Explicitly Iterative</th>
<th>Plausibly Iterative</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>% Total</td>
<td>N</td>
</tr>
<tr>
<td>177</td>
<td>32.8</td>
<td>362</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Explicitly Iterative</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasold (1972)</td>
<td>22.5</td>
</tr>
<tr>
<td>Wolfram (1969)</td>
<td>35.3</td>
</tr>
</tbody>
</table>
## Results

### All FPG `be.like` tokens (N = 136)

<table>
<thead>
<tr>
<th>Iterativity Code</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicitly Iterative</td>
<td>N</td>
<td>% Total</td>
</tr>
<tr>
<td></td>
<td>53</td>
<td>44.9</td>
</tr>
<tr>
<td>Plausibly Iterative</td>
<td>N</td>
<td>% Total</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>55.1</td>
</tr>
</tbody>
</table>

### All FPG Aspectual `be` tokens (no `be.like`) (N = 539)

<table>
<thead>
<tr>
<th>Iterativity Code</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicitly Iterative</td>
<td>N</td>
<td>% Total</td>
</tr>
<tr>
<td></td>
<td>177</td>
<td>32.8</td>
</tr>
<tr>
<td>Plausibly Iterative</td>
<td>N</td>
<td>% Total</td>
</tr>
<tr>
<td></td>
<td>362</td>
<td>67.2</td>
</tr>
</tbody>
</table>
## Results

### All FPG Aspectual *be* tokens (N = 651)

<table>
<thead>
<tr>
<th>BE Construction Type</th>
<th>Iterativity Code</th>
<th>N</th>
<th>% Total of Type</th>
<th>N</th>
<th>% Total of Type</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Explicitly Iterative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plausibly Iterative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>be.like</td>
<td></td>
<td>53</td>
<td>44.9</td>
<td>65</td>
<td>55.1</td>
<td>118</td>
</tr>
<tr>
<td>be+verb-ing</td>
<td></td>
<td>111</td>
<td>29.7</td>
<td>263</td>
<td>70.3</td>
<td>374</td>
</tr>
<tr>
<td>be+adv./prep.</td>
<td></td>
<td>31</td>
<td>40.3</td>
<td>46</td>
<td>59.7</td>
<td>77</td>
</tr>
<tr>
<td>be+adjective</td>
<td></td>
<td>18</td>
<td>37.5</td>
<td>30</td>
<td>62.5</td>
<td>48</td>
</tr>
<tr>
<td>be+NP</td>
<td></td>
<td>11</td>
<td>42.3</td>
<td>15</td>
<td>57.7</td>
<td>26</td>
</tr>
<tr>
<td>be+past part.</td>
<td></td>
<td>4</td>
<td>50.0</td>
<td>4</td>
<td>50.0</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>228</td>
<td></td>
<td>423</td>
<td></td>
<td>651</td>
</tr>
</tbody>
</table>
Conclusion

Construction-based (structural) approaches to be usage enable us to utilize text corpora of AAE

• Aspectual be usage for incorporation into larger quantitative analyses.

• Investigate a new innovation vis-à-vis other Aspectual be forms.
References

(Labov 1982, p. 30)
This research was made possible by:

National Science Foundation Grants
BCS-0544744
BCS-0843865