Ness-less-ness
Zero-derived adjectival nominals in Internet forum data

Marisa Brook and Emily Blamire
(University of Toronto)
Open A World of Possible

Something special happens when a child finds the right book.

Hear What Kids Say
We have something of a novel variable:

Variable realization of noun-forming suffixes that attach to adjectives.

-ness/ -ity  Ø

We started spotting zero forms all over the place and wanted to investigate.
In advertising...

[Images of advertisements]

Open A World of Possible
Something special happens when a child finds the right book.

Hear What Kids Say

YOU JUST GRABBED A WHOLE BUNCH OF HEALTHY

Cyclable bag is made from 100% recycled paper.
Introduction

In pop music...

And commentary...
In the bookstore…
And all over the Internet.

4. Dorky Music Gets Injected With Awesome

13 People Who Turned Their Canadian Up To 100

SARCASM IS A BODY’S NATURAL DEFENSE AGAINST STUPID
We are not the first!

Whitman (2009), Zwicky (2009, 2010), Francis (2013), and Lighter (2013a, 2013b) have all made note of these forms and analyzed them as zero-derived nominals.

Zwicky (2009) reports having only found it in “of” PPs.

However it has been found in other context such as “Creating Amazing” (Francis, 2013), and (Lighter, 2013a, 2013b) makes a connection to “I’m bringing sexy back”.

Lighter (2013c) suggests zero-derived nouns are now so common that the pattern is syntactic rather than lexical.
However, this is also not completely novel in English!

Nouns of moral judgement

– Good vs Evil vs Bad
  • Good exists in everyone
  • Axis of evil
  • The root of all evil
  • The bad outweighs the good

Nouns of flavor

– Sour and Sweet
  • This recipe has too much sour in it
  • If you want to up the sweet, add some honey to your cookies
Research questions:

1. Is this a change in progress?

2. Is the variation productive enough (across words and contexts) to be syntactic rather than lexical?

3. Will null suffixes appear more often in PPs (particularly ‘of’) than other phrases?

4. Does the presence of modifiers (determiners/adjectives) before the noun make the overt suffix less necessary (leading to increased rates of the null form)?
Data
Colloquial pattern $\rightarrow$ needed casual language data.

Occurring a lot in writing $\rightarrow$ focused on written data.

Low-frequency phenomenon $\rightarrow$ needed a lot of words.

The Internet

Open access forum data
Claridge 2009:87-88:

- “different from other forms of CMC in various respects”
- “dialogic…and conversational in style”
- “completely public”
- “asynchronous”
- “a wide range of topics…an open public-opinion platform”
But where in the dark, depths of the Internets would make a good starting point?

Forums for the webcomic **xkcd**.

Fans of xkcd tend to be:

- Tech-savvy/nerdy
- Young
- Playful/witty

The comic itself often touches on language and creativity...

and the language of the fans often reflects this.

Methodology
Methodology

We used a Python web-scrapping script to get us **300,000 public posts** from the xkcd forum

- “Individual comic threads” Subforum
  - “Time” Thread
- Threads in the “News and Articles” Subforum

We converted the spreadsheets into .txt files and used AntConc (Anthony, 2012) to search for nominalized adjectives.

We ran a fixed effects logistic regression using Goldvarb X (Sankoff et al., 2005)
We focused on…

awesome(∅/ness)
stupid(∅/ity/ness)
crazy(∅/-ness)
adorable(∅/-ness)
cute(∅/-ness)
dumb(∅/-ness)

Not a lot of competition between –ness and –ity, we treated the variation as a binary dependent variable: presence or absence of a suffix.
For variable context we only used syntactic positions where nouns were licensed

We excluded:

• Ambiguous contexts
  • Bare complement of copula: *This is awesome*
  • “Compounds” where the token ≠ head: *cute overload*
  • Single word exclamations: *Awesome!*
  • Implicit “people”: *Home of the stupid*
• Definitions: *That is the definition of “awesome”*
• Fossilized phrases with no variation: *human stupidity*
• Poems, songs, and translations
• DP/NP fragments: *EEEEEEEEEEEEeeeee the cuteness!*
Methodology

We coded for:

- Suffix (or lack thereof)
- Year of post
- Username
- Smallest containing phrase (CP, VP, PP, ConjP)
- Phrase head
- Presence or absence of determiner
- Presence or absence of adjective
In the end:

- **987 tokens** (293 Null & 694 Suffix) from **596 posters**
  - 8 in ConjP
  - 109 in CP
  - 583 in PP
  - 287 in VP

- 392 had determiners
- 232 had modifying adjectives
Results
<table>
<thead>
<tr>
<th>Word</th>
<th>Total</th>
<th>Null</th>
<th>Suffix</th>
<th>% Null</th>
</tr>
</thead>
<tbody>
<tr>
<td>adorable</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>75%</td>
</tr>
<tr>
<td>awesome</td>
<td>367</td>
<td>144</td>
<td>223</td>
<td>39%</td>
</tr>
<tr>
<td>crazy</td>
<td>84</td>
<td>41</td>
<td>43</td>
<td>49%</td>
</tr>
<tr>
<td>cute</td>
<td>29</td>
<td>7</td>
<td>22</td>
<td>24%</td>
</tr>
<tr>
<td>dumb</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>25%</td>
</tr>
<tr>
<td>sad</td>
<td>138</td>
<td>10</td>
<td>128</td>
<td>7%</td>
</tr>
<tr>
<td>smug</td>
<td>57</td>
<td>9</td>
<td>48</td>
<td>16%</td>
</tr>
<tr>
<td>stupid</td>
<td>304</td>
<td>78</td>
<td>226</td>
<td>26%</td>
</tr>
<tr>
<td></td>
<td>987</td>
<td>293</td>
<td>694</td>
<td>30%</td>
</tr>
</tbody>
</table>
Null Suffix Rate Over Time

- Awesome stupid: 18%
- Crazy adorable: 94%
- Smug: 132%
- Dumb: 89%
- Cute: 87%
- Null suffix rate over time from 2006 to 2015.
Variation occurred both between and within Posters

Null Suffix Rate by Poster

“…Or, to be skeptical, the group whose craziness is most known. Everyone else just hides their crazy.” –Vaniver, May 10, 2007
Results

Null Suffix Rate by Smallest Containing Phrase

PP: 583
CP: 109
VP: 287
ConjP: 8
Null Suffix Rate by Absence/Presence of Modifiers

- Determiner: Absent (595), Present (392)
- Adjective: Absent (755), Present (232)
## Results

<table>
<thead>
<tr>
<th>Corrected mean (null suffix)</th>
<th>0.26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N</td>
<td>971</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lexical adjective</th>
<th>FW</th>
<th>%</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>crazy</strong></td>
<td>0.76</td>
<td>48.8</td>
<td>84</td>
</tr>
<tr>
<td><strong>awesome</strong></td>
<td>0.64</td>
<td>39.5</td>
<td>365</td>
</tr>
<tr>
<td><strong>stupid</strong></td>
<td>0.48</td>
<td>25.8</td>
<td>302</td>
</tr>
<tr>
<td><strong>cute</strong></td>
<td>0.44</td>
<td>24.1</td>
<td>29</td>
</tr>
<tr>
<td><strong>smug</strong></td>
<td>0.32</td>
<td>16.1</td>
<td>56</td>
</tr>
<tr>
<td><strong>sad</strong></td>
<td>0.14</td>
<td>6.7</td>
<td>135</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>62</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Determiner</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Absent</strong></td>
<td>0.66</td>
<td>37.7</td>
<td>584</td>
</tr>
<tr>
<td><strong>Present</strong></td>
<td>0.40</td>
<td>17.6</td>
<td>387</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adjective</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Absent</strong></td>
<td>0.53</td>
<td>30.7</td>
<td>748</td>
</tr>
<tr>
<td><strong>Present</strong></td>
<td>0.41</td>
<td>26.0</td>
<td>223</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fixed-effects logistic regression in Goldvarb X (Sankoff et al. 2005)

Application value: null suffix.

Time and containing phrase were not significant.
<table>
<thead>
<tr>
<th>preposition</th>
<th>awesome</th>
<th>stupid</th>
<th>crazy</th>
</tr>
</thead>
<tbody>
<tr>
<td>of</td>
<td>2007</td>
<td>2007</td>
<td><strong>2006</strong></td>
</tr>
<tr>
<td>with</td>
<td><strong>2006</strong></td>
<td>2009</td>
<td>2012</td>
</tr>
<tr>
<td>for</td>
<td>2007</td>
<td>2014</td>
<td>2013</td>
</tr>
<tr>
<td>to</td>
<td>2013</td>
<td>2013</td>
<td></td>
</tr>
<tr>
<td>about</td>
<td></td>
<td></td>
<td><strong>2013</strong></td>
</tr>
<tr>
<td>in</td>
<td></td>
<td></td>
<td>2009</td>
</tr>
<tr>
<td>on</td>
<td>2008</td>
<td></td>
<td><strong>2006</strong></td>
</tr>
<tr>
<td>by</td>
<td>2013</td>
<td></td>
<td><strong>2013</strong></td>
</tr>
<tr>
<td>from</td>
<td>2011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>at</td>
<td>2013</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Null Suffix Rate in PPs by Phrase Head

- of: 311
- with: 54
- for: 48
- to: 42
- about: 19
- in: 15
- on: 14
- by: 13
- from: 10
- at: 8
- because of: 8

(Percentages are not shown in the image.)
Discussion
Revisiting our research questions…

Is this a change in progress?

If this is a new change, it is already fairly stable in our data and has been since 2007
- Null suffix rate varies between 20-40 percent
- Preliminary data suggest some increase for certain words
- No interaction of year with any other factor
Productive enough (across words and contexts) to be a syntactic change?

May have been spreading from word to word, but even in 2006 half of the nouns we examined were already variable.

**Different overall rates** of null form for each word

Some words have low Ns (*adorable*), but some do not (*sad*). 

Preliminary result from running the more frequent words individually: some role of idiosyncrasy in significant conditioning factors and effect sizes.
More null suffixes with PPs (particularly ‘of’)?

No evidence that the null form spread from PPs to CPs/VPs. Not even a trend in this direction.

No interaction of word and containing phrase.

Within the tokens introduced by PPs, ‘of’ does seem to be more conducive to the null form than most other PPs as per Zwicky (2010).
Modifiers leading to increased null rates?

**Significant effect in the opposite direction.**

Why does the null form appear less often with determiners and/or adjectives? What does this tell us?

Extra contextual information **not required** to recognize nouns

Is the null form spreading through the simplest syntactic context (without modifiers) first?
- Difficult to reconcile with overall stable rates of the null form.

Do overt modifiers plus null set up parsing confusion?
- When people get to the end of the null form NP there is more information to reparse if a determiner and or adjective is present
To recap:

1. Change in progress?  **Not in our data**

2. Syntactic rather than lexical?  **Debatable. Seems to be heading there**

3. Null suffixes favour PPs (and of)?  **No to PPs, but yes to ‘of’**

4. Modifiers leads to increased null rates?  **Nope!**
Thanks!

Matt Hunt Gardner
Ron Smyth
Naomi Nagy
Dmitry Vlasov
U of T Language Variation and Change Group
NWAV 44 organizers

…and all of you!

marisa.brook@mail.utoronto.ca    emily.blamire@mail.utoronto.ca