"It doesn't stop, it never, never stops, er, it doesn't stop evolving"

Observing Spoken British English of the past 20 years through apparent and real-time evidence

Susan Reichelt
@susanreichelt1

The research has been supported by the ESRC grant no. EP/P001559/1.
Data
Project: The British National Corpus (BNC) as a sociolinguistic dataset: Exploring individual and social variation

Funding: ESRC grant no. EP/P001559/1.

Team: Vaclav Brezina (PI), Dana Gablasova (Co-I), Tony McEnery (Co-I), Miriam Meyerhoff (Co-I), Susan Reichelt (RA)
• Focus on teaching SLX and CL

• Focus on new research investigating social factor age and language change

• Focus on methodology and new ways of analysing variation
• Focus on new research investigating social factor age and language change
<table>
<thead>
<tr>
<th>Subset</th>
<th>Speakers</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNC spoken/demographic</td>
<td>1901</td>
<td>~15 million</td>
</tr>
<tr>
<td>BNC 1994</td>
<td>1233</td>
<td>~4 million</td>
</tr>
<tr>
<td>BNC 2014</td>
<td>668</td>
<td>~11 million</td>
</tr>
<tr>
<td>BNC SDA</td>
<td>500</td>
<td>~9 million</td>
</tr>
</tbody>
</table>

subsets for the SDA project
subset of the subset “BNCeption”
“(…) for most speakers we do not have combined information about sex, age and social class (…). This should not mean that we cannot use the BNC to investigate sociolinguistic variation, but we should be clear about any shortfalls in terms of representativeness, particularly when we start splitting the corpus up into finer slices.”

Baker 2010: 40
1233 speakers
3’942’768 words

speaker with highest
word count: ~ 70’000

speaker with lowest
word count: 1

668 speakers
10’982’869 words

speaker with highest word
count: ~ 351’000

speaker with lowest word
count: 18
**1994**

region:

27 regional codes without clear geographical boundaries or detail on what regional background entails for the individual speaker.

---

**2014**

region:

Distinctions between birthplace, current location (and duration of stay) and perceived accent. Coding follows four levels, from broad (UK, non-UK) to narrow (town).
we follow system by Gerwin (2014), who adjusted the region system in the old BNC to compare to other data sets
Social grades:

AB (upper) middle class
C1 lower middle class
C2 skilled working class
D working class
E not working (unknown)

National Statistics Socio-economic Classification, or 'NS-SEC':

1.1 Employers in large organisations, higher managerial occupations
1.2 Higher professional occupations
2 Lower professional and higher technical occupations, higher supervisory occupations
3 Intermediate occupations
4 Employers in small organisations
5 Lower supervisory occupations, lower technical occupations
6 Semi-routine occupations
7 Routine occupations
8 Never worked and long-term unemployed
N/C Full-time students
### Age distribution across 6 groups:

<table>
<thead>
<tr>
<th>1994</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-14</td>
<td>15-24</td>
</tr>
<tr>
<td>25-34</td>
<td>35-44</td>
</tr>
<tr>
<td>45-59</td>
<td>60+</td>
</tr>
</tbody>
</table>

### Age distribution across 10 groups:

<table>
<thead>
<tr>
<th>1994</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10</td>
<td>11-18</td>
</tr>
<tr>
<td>19-29</td>
<td>30-39</td>
</tr>
<tr>
<td>40-49</td>
<td>50-59</td>
</tr>
<tr>
<td>60-69</td>
<td>70-79</td>
</tr>
<tr>
<td>80-89</td>
<td>90-99</td>
</tr>
</tbody>
</table>
Theory
Tracking language change as it happens is, according to Chambers (1995:147), “the most striking single accomplishment of contemporary linguistics”.

“(....) giving the analysis of variation the status of an in vivo study of historical change” (Eckert 2012: 89)
apparent time & real time
Time ...
it's more like a big ball of wibbly wobbly... timey wimey... stuff.

(Doctor Who, *Blink*, 2007)
apparent time and possible interpretations

stable feature
apparent time and possible interpretations

outgoing feature
apparent time and possible interpretations

incoming feature
apparent time and possible interpretations

... or maybe age grading?
real time and possible interpretations

stable feature
real time and possible interpretations

outgoing feature
real time and possible interpretations

incoming feature
Combination of apparent and real time

ongoing change
Combination of apparent and real time ongoing change
real time and possible interpretations

ongoing change
Combination of apparent and real time age grading
real time and possible interpretations

age grading
“In the best of circumstances, of course, researchers will be able to combine apparent-time data with real-time evidence, with the relative strengths of one approach offsetting the weaknesses of the other” (Bailey, 2008:330)
application
Adjective intensification in Spoken British English: the past 20 years
They’re very nice.
She’s so thoughtless.
That’s really cheap.

Included: all items that amplified a following adjective

The variable set includes a great number of variants with the most common intensifiers (very, really, so) representing approx. 90% of the results.
Data extraction & coding
very

really

so

CORPUS  BNC1994  BNC2014
What does the BNC subset study offer in terms of new insights?

A side-by-side investigation of apparent time and real time (trend) which adds detail to our interpretations of language change and who is involved in what type of change.
Thank you
References:


