

"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
Theory
Application

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

BNC spoken/demographic:
1901 speakers
~15 million words



BNC 1994:
1233 speakers
~4 million words

BNC 2014:
668 speakers
~11 million words

BNC SDA
500 speakers
~9 million words



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



1994

1233 speakers
3'942'768 words

speaker with highest
word count: ~ 70'000

speaker with lowest
word count: 1



2014

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:

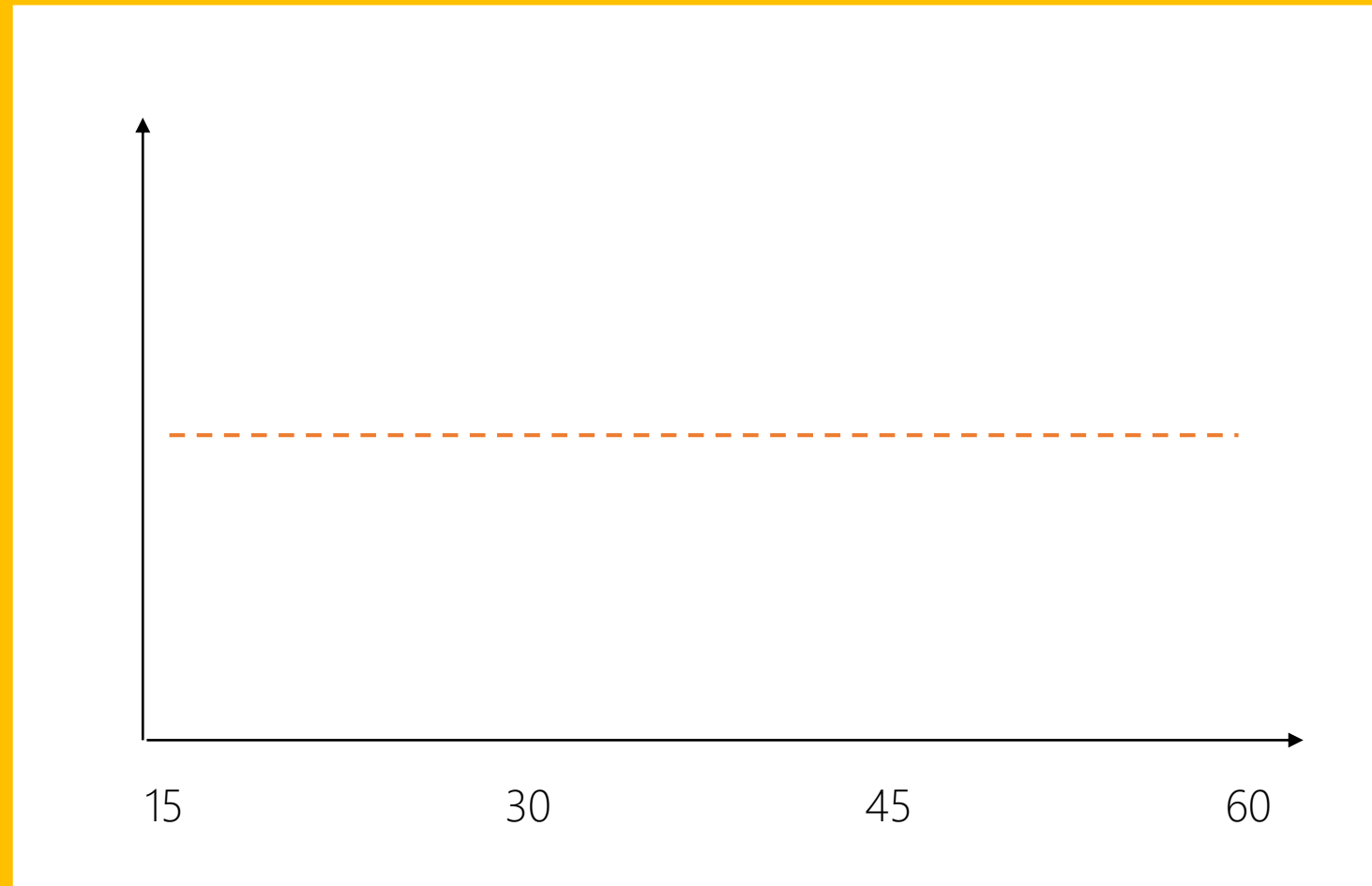
0-14	15-24
25-34	35-44
45-59	60+

Age distribution across **10**
groups:

0-10	11-18
19-29	30-39
40-49	50-59
60-69	70-79
80-89	90-99

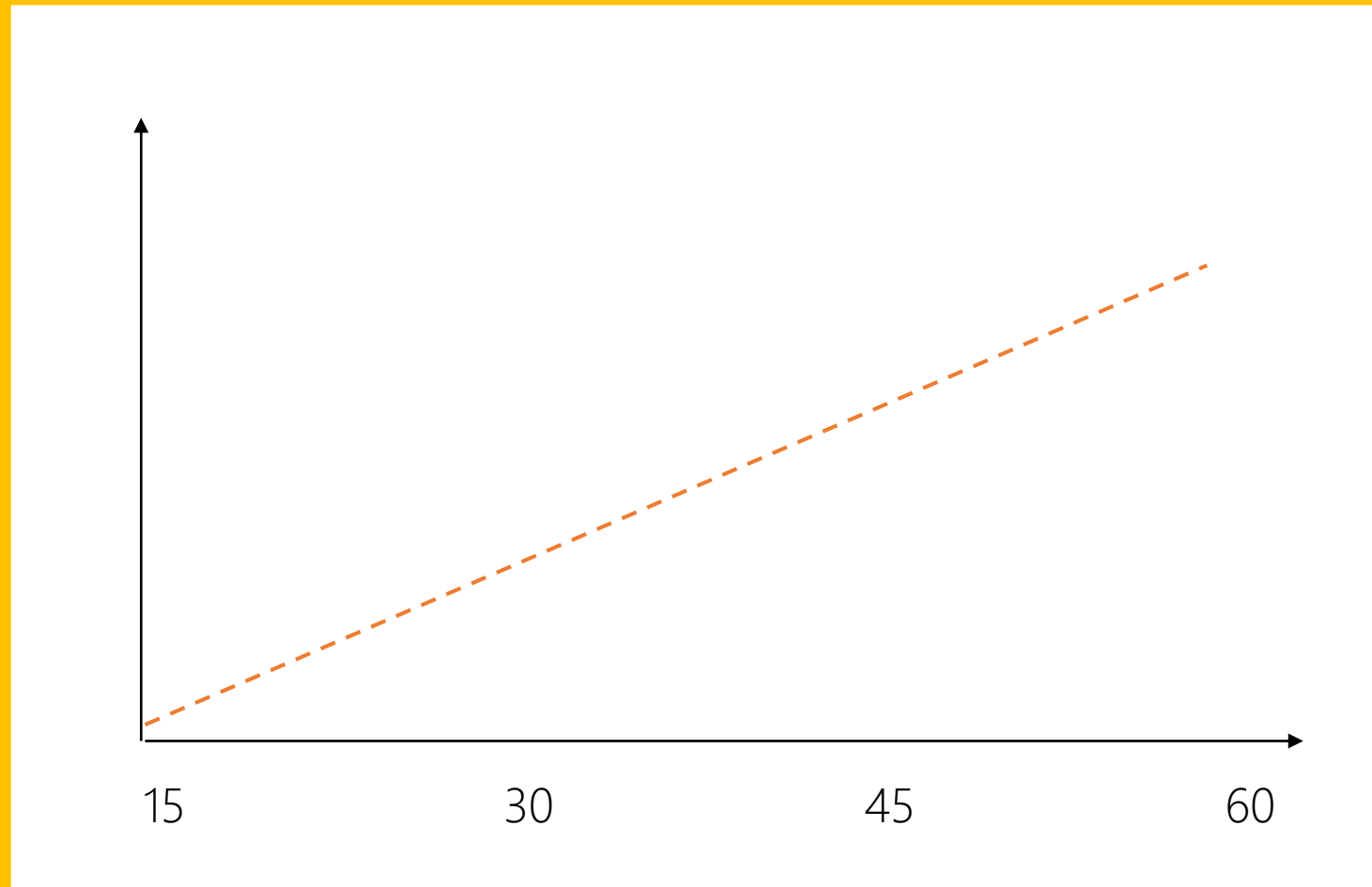
apparent time and possible interpretations

stable feature



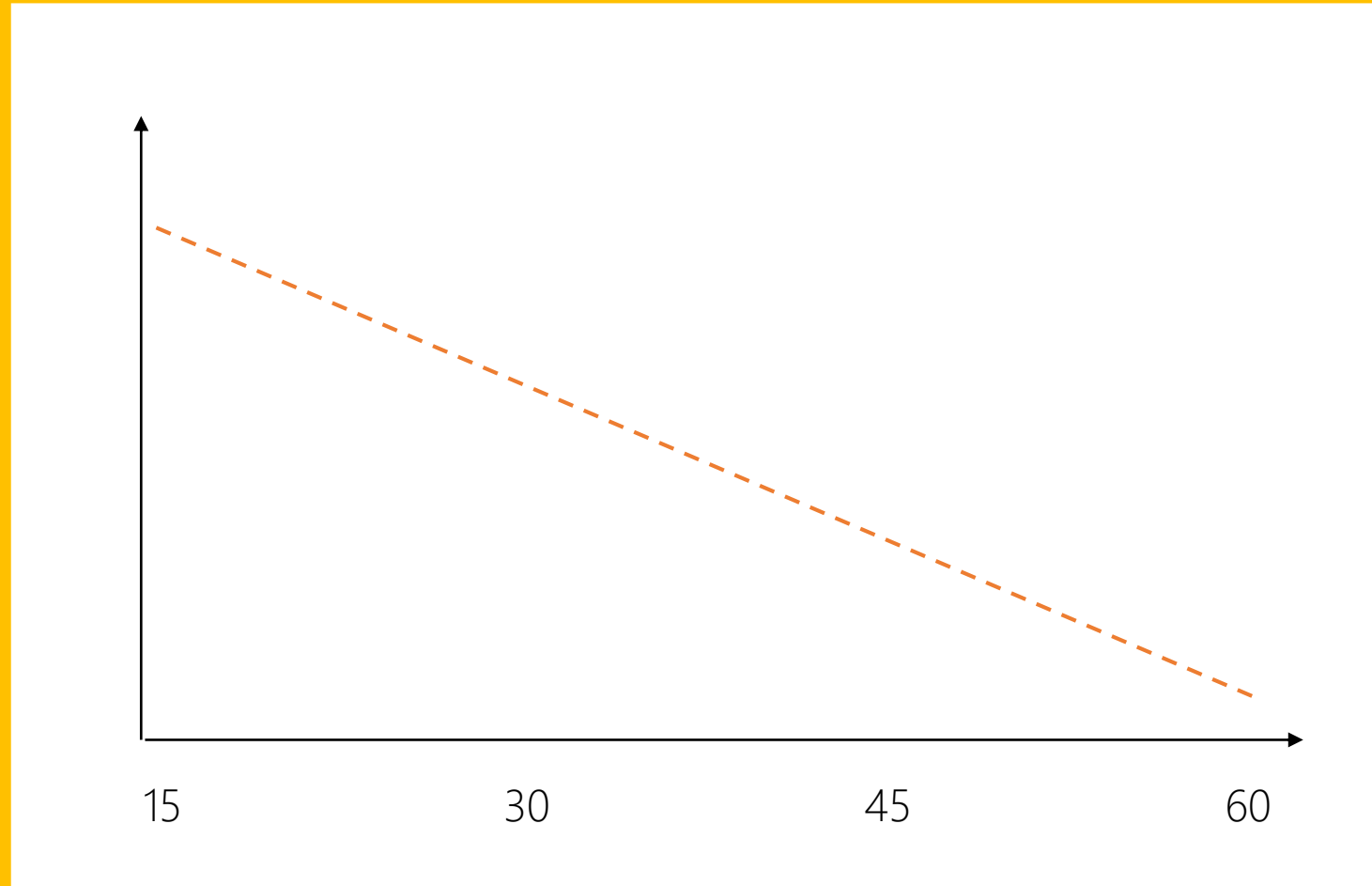
apparent time and possible interpretations

outgoing feature



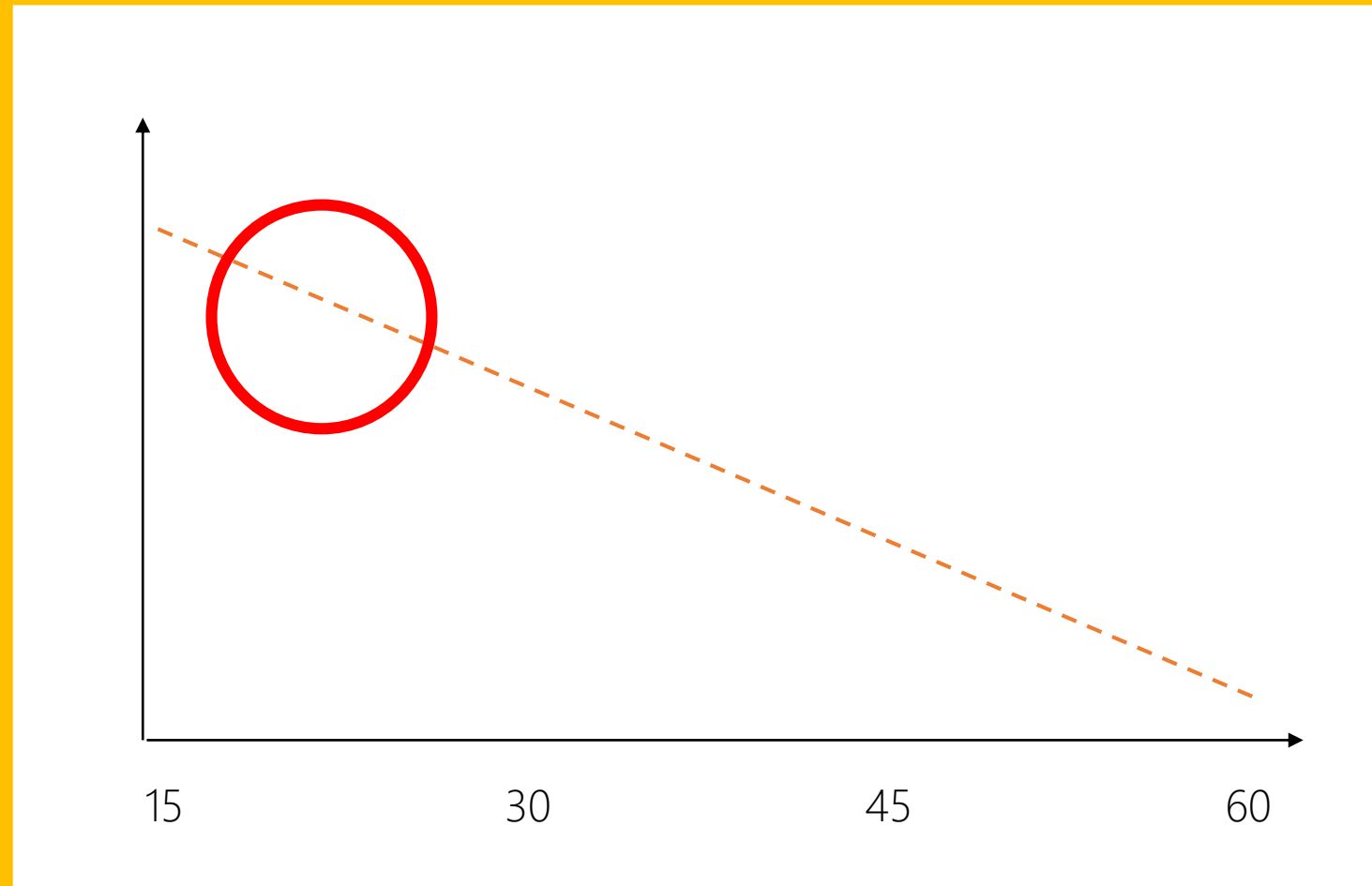
apparent time and possible interpretations

incoming feature



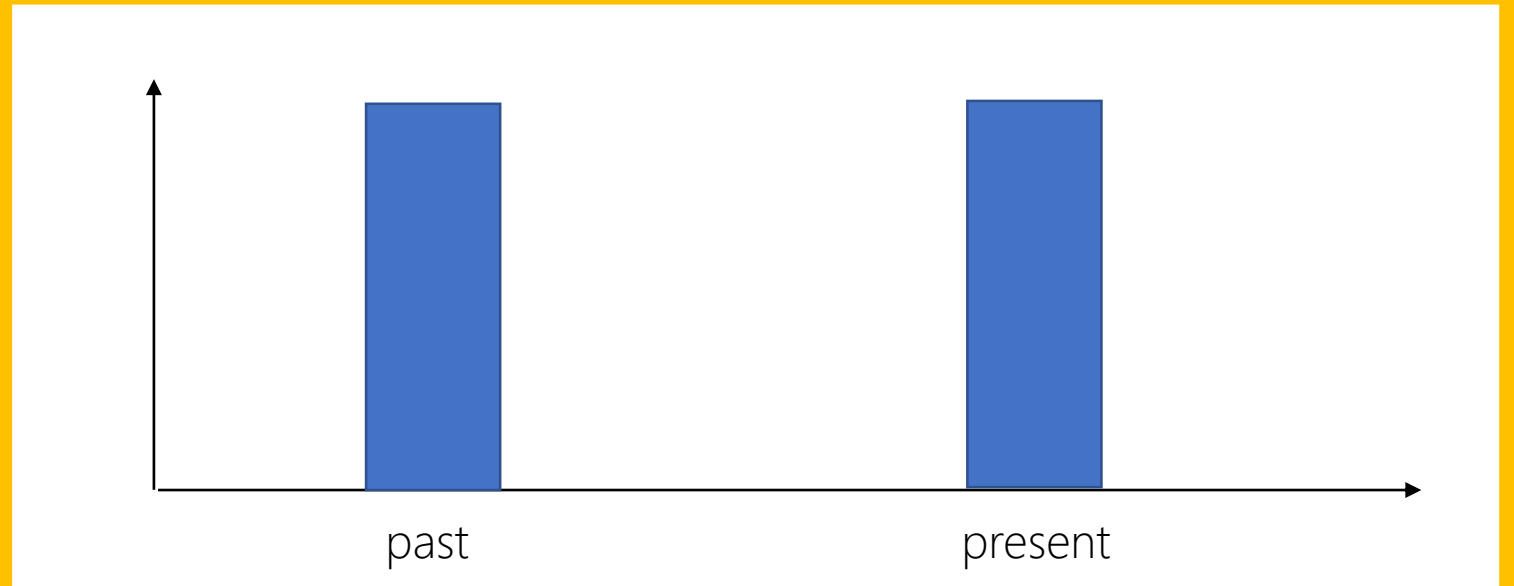
apparent time and possible interpretations

... or maybe
age grading?



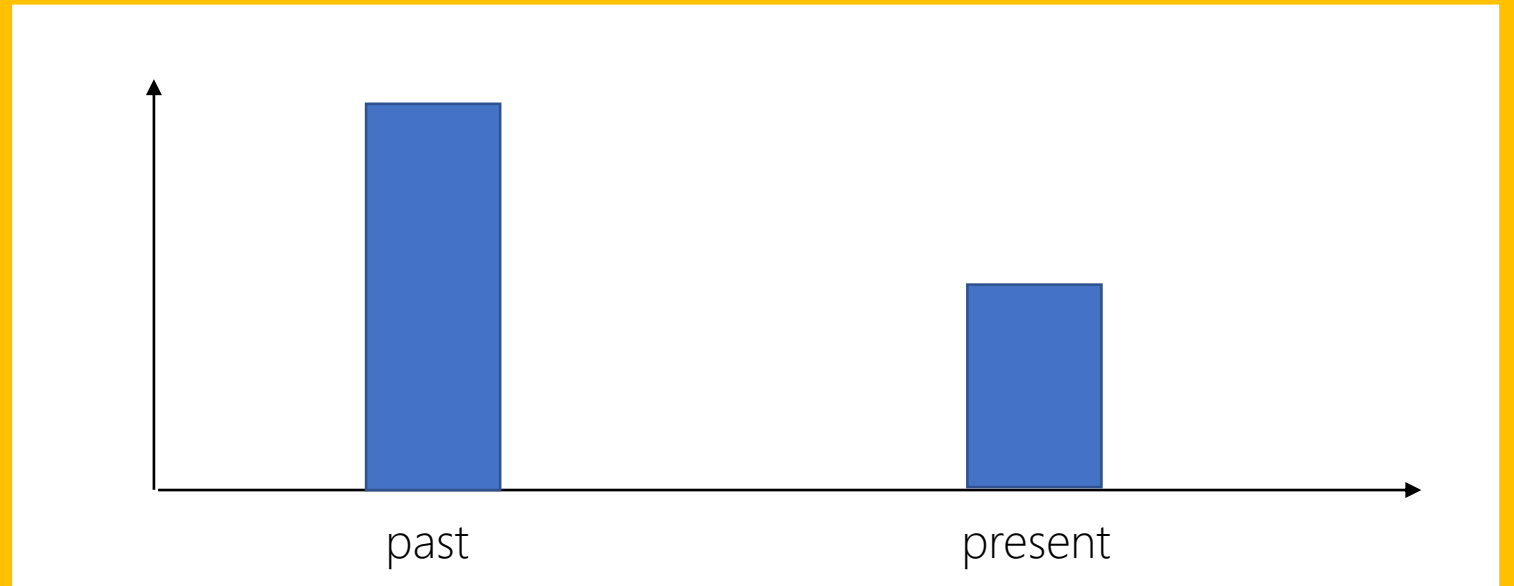
real time and possible interpretations

stable feature



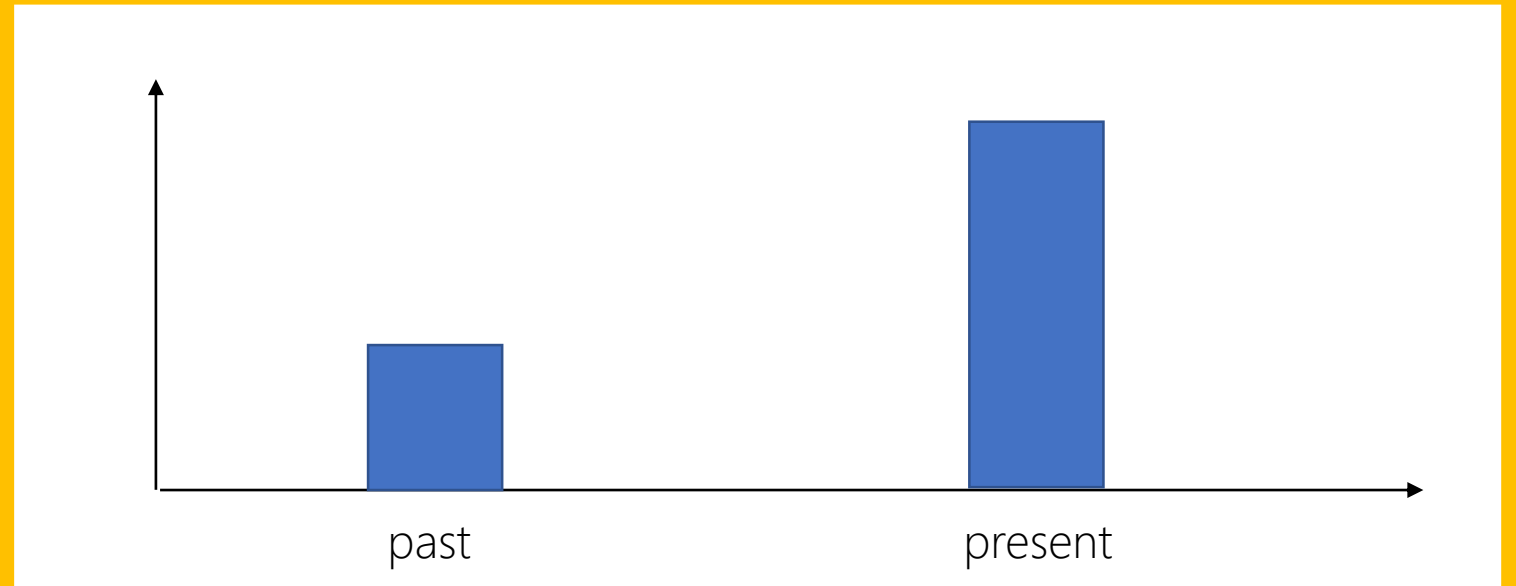
real time and possible interpretations

outgoing feature



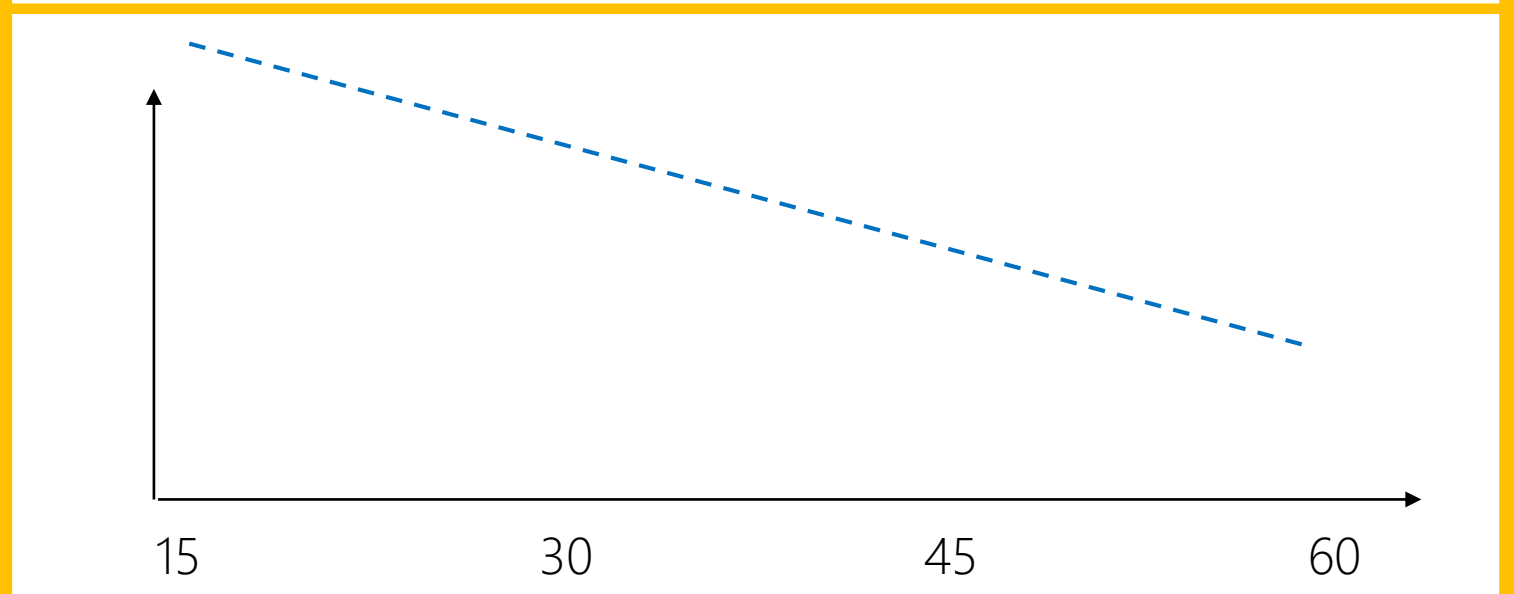
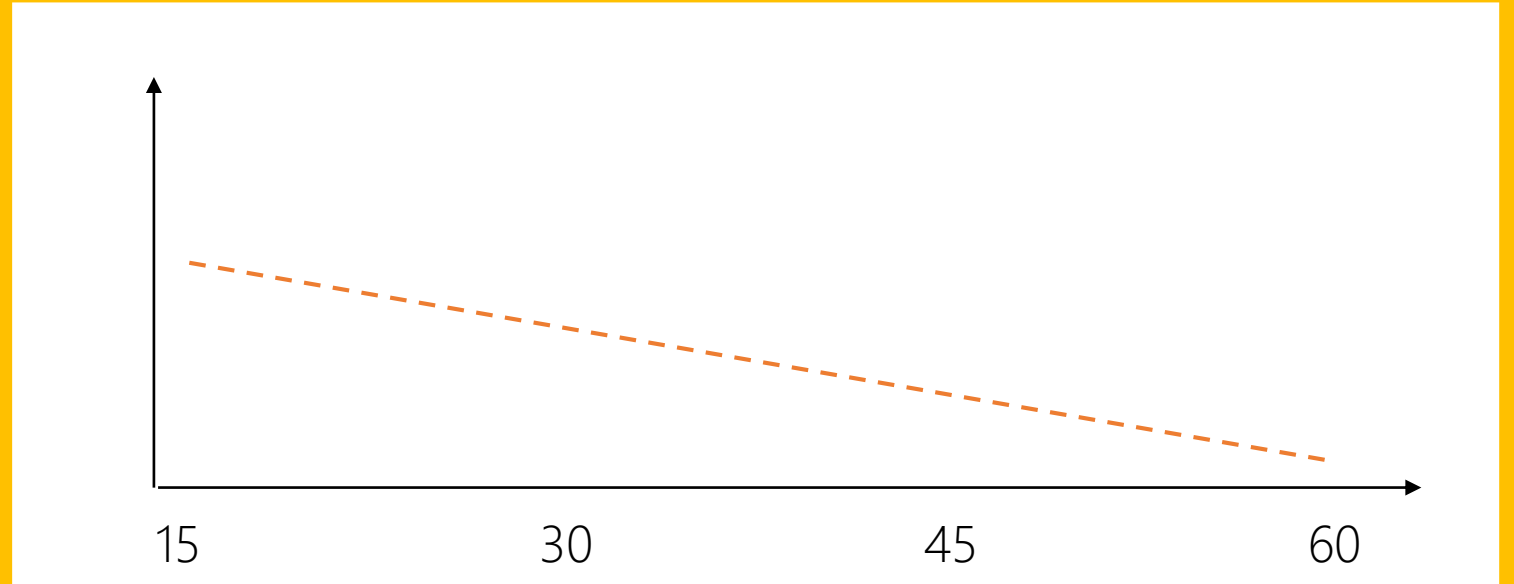
real time and possible interpretations

incoming feature



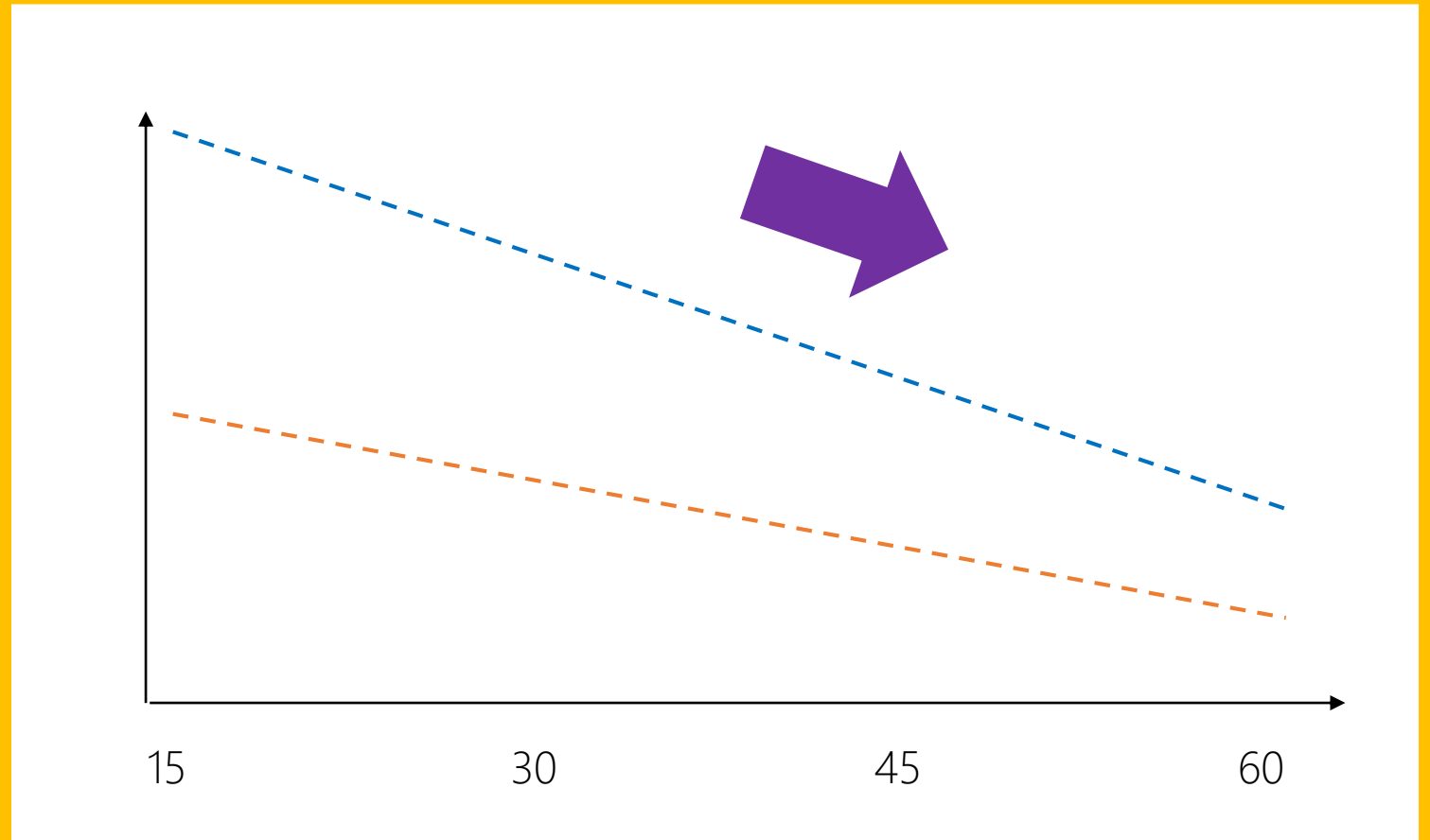
Combination of apparent and real time

ongoing change



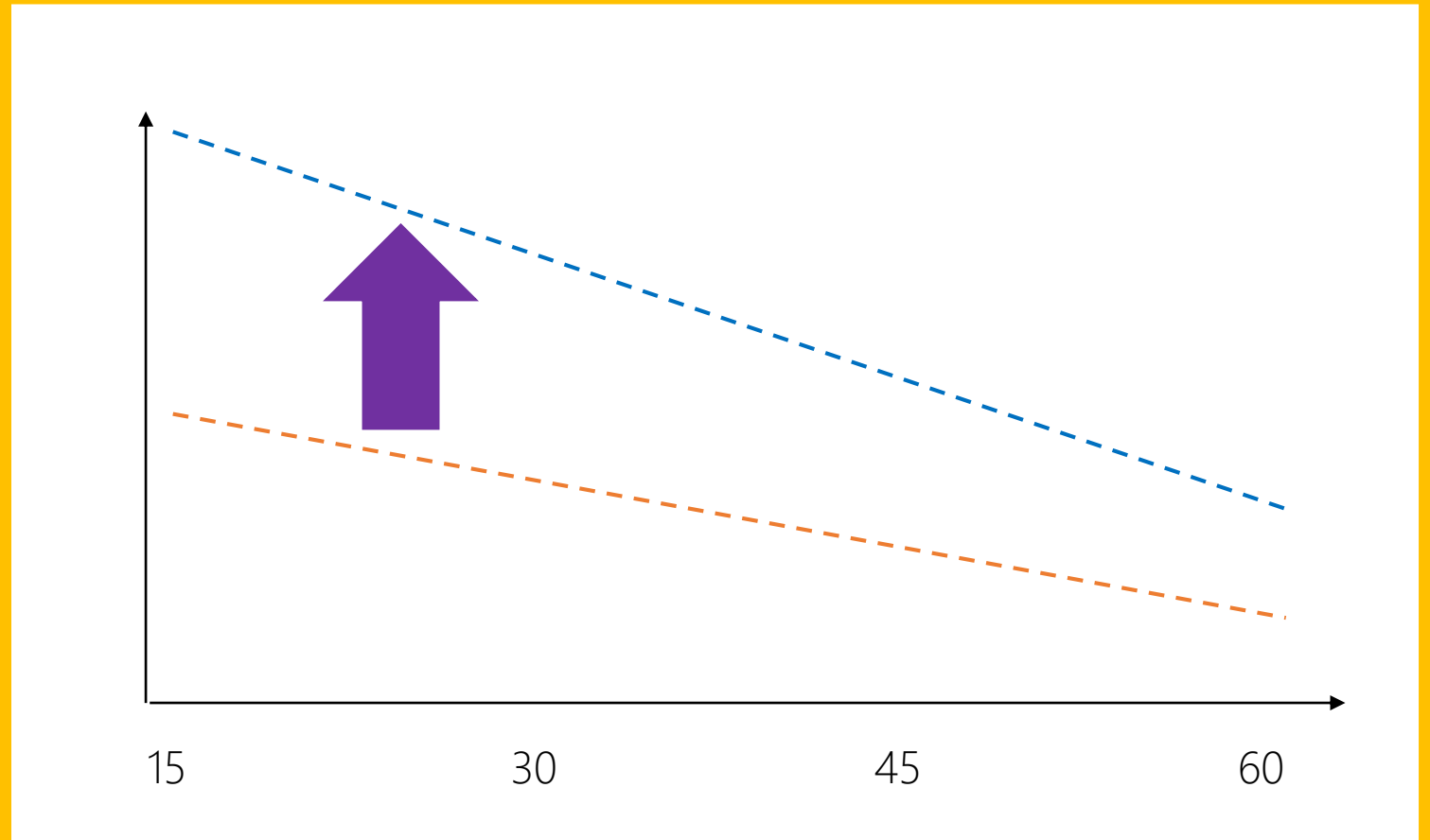
Combination of apparent and real time

ongoing change



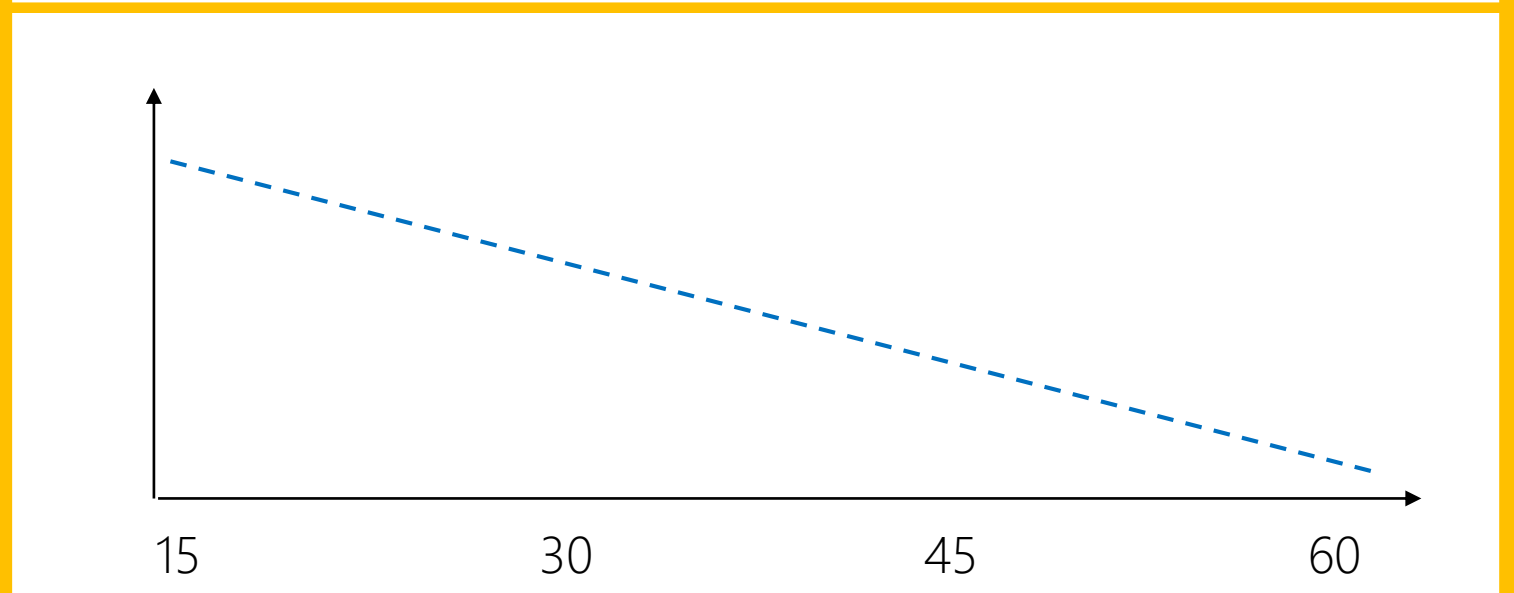
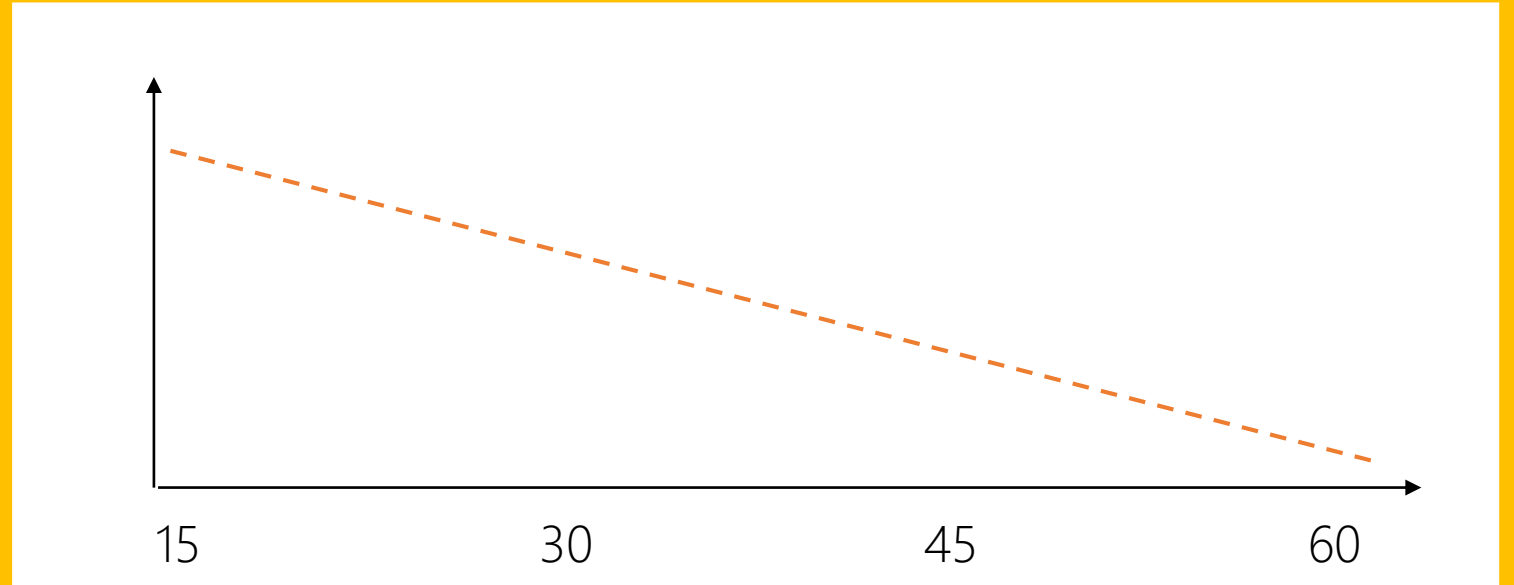
real time and possible interpretations

ongoing change



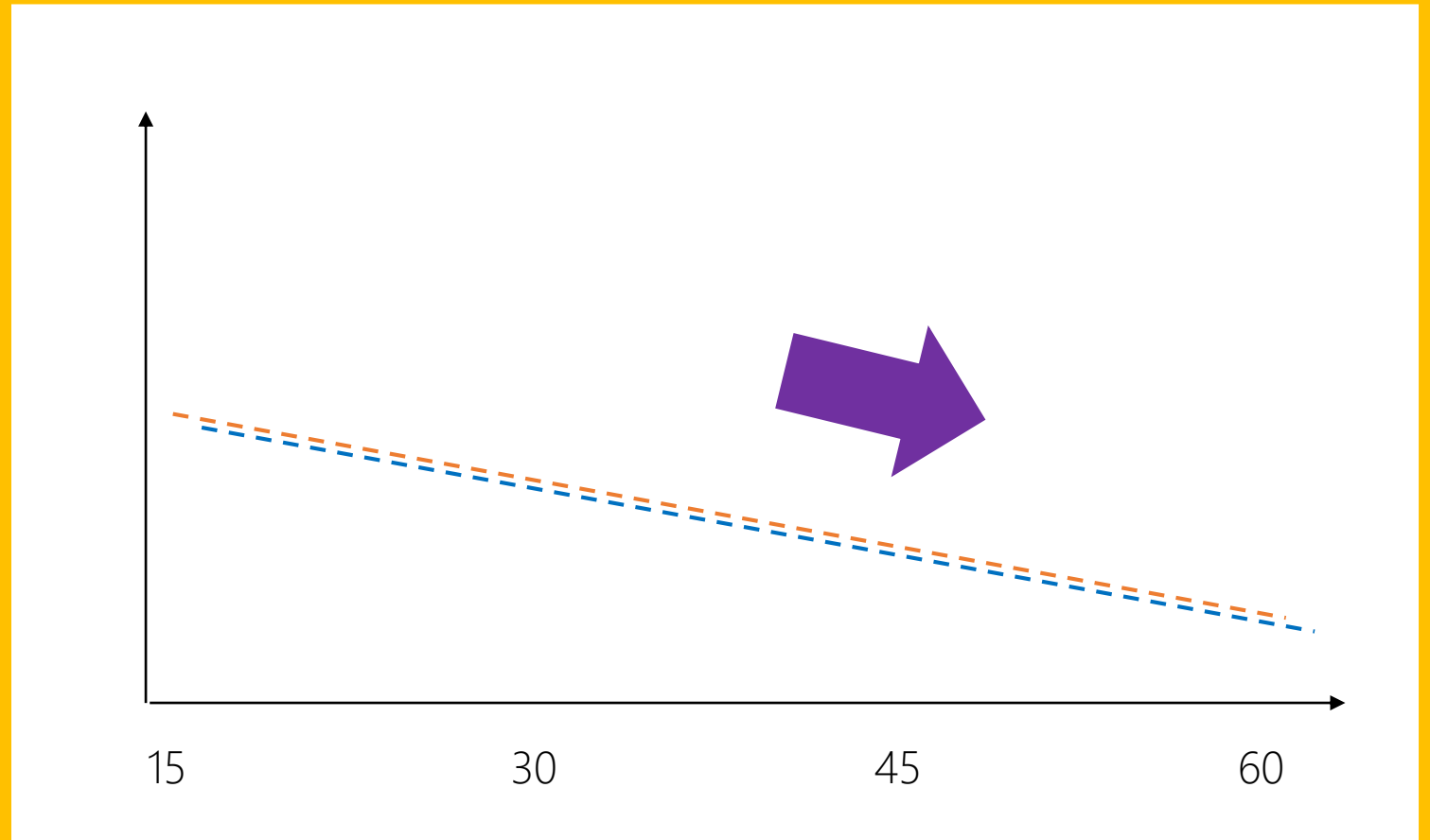
Combination of apparent and real time

age grading

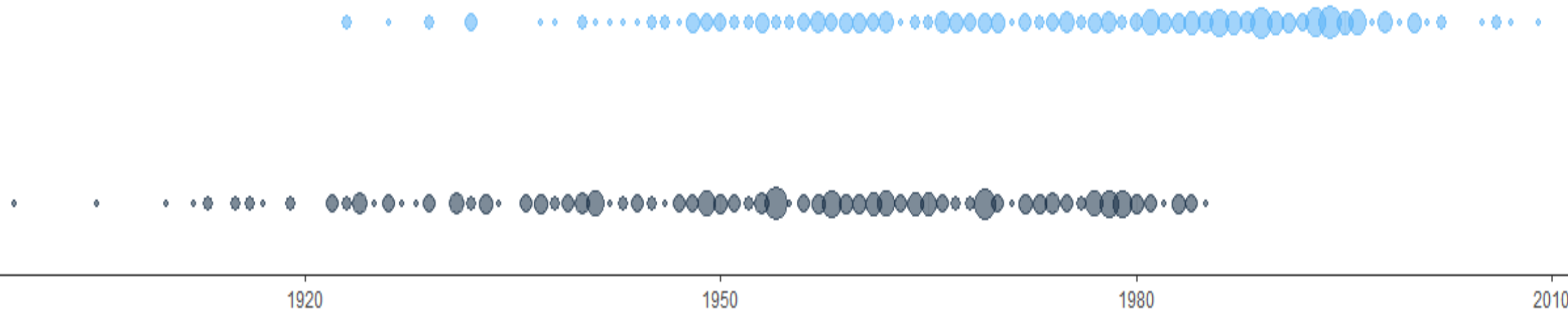


real time and possible interpretations

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)



What can time tell us, then?

application

Adjective intensification in Spoken British English: the past 20 years

They're *very* nice.

She's *so* thoughtless.

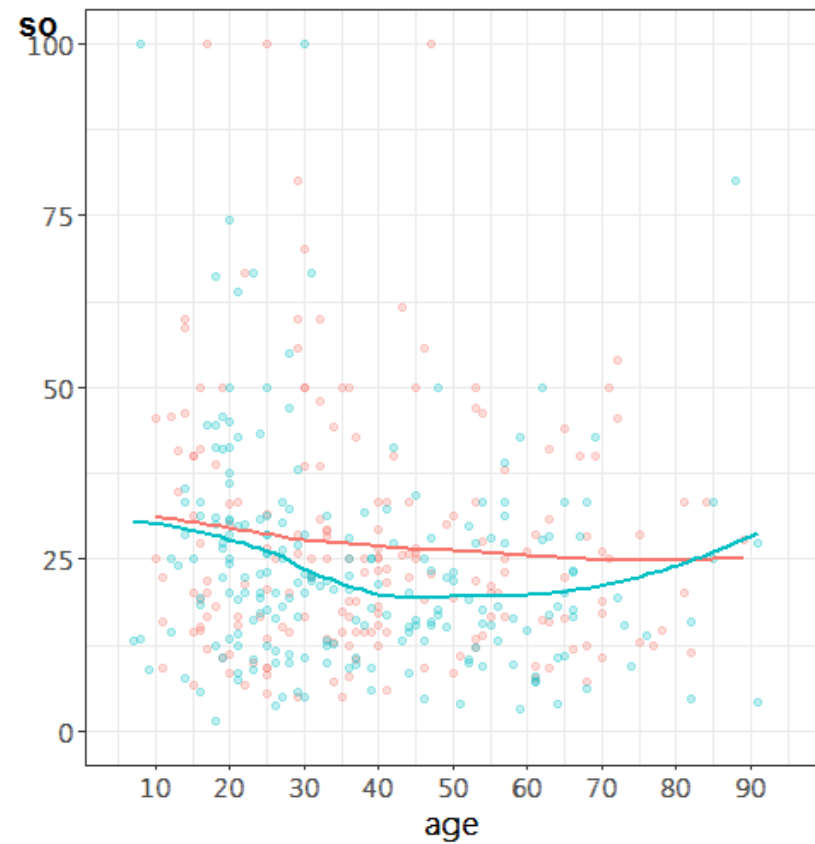
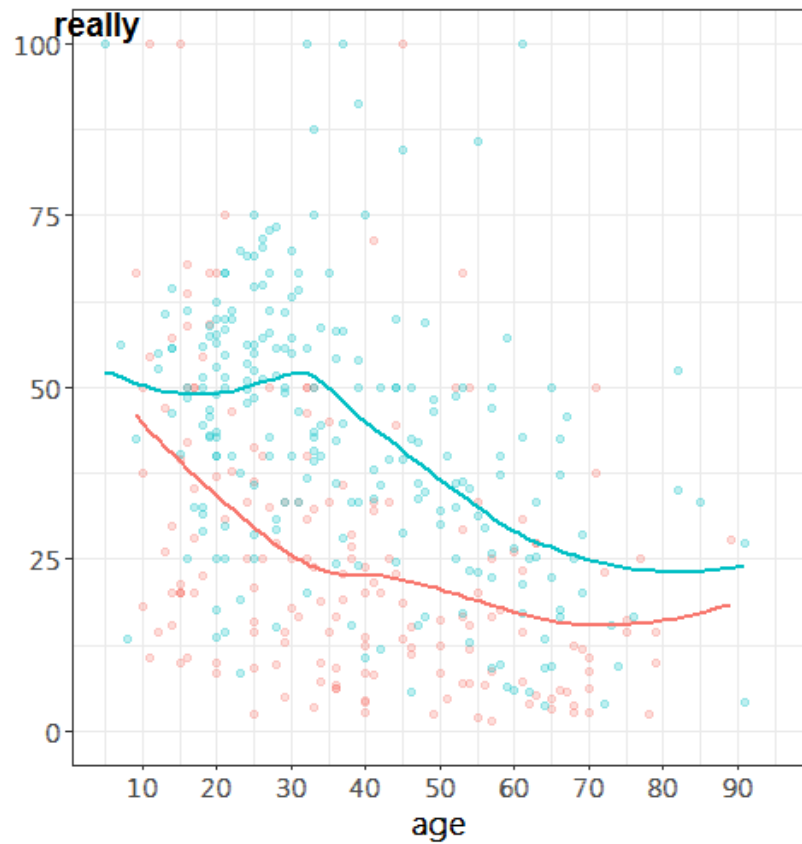
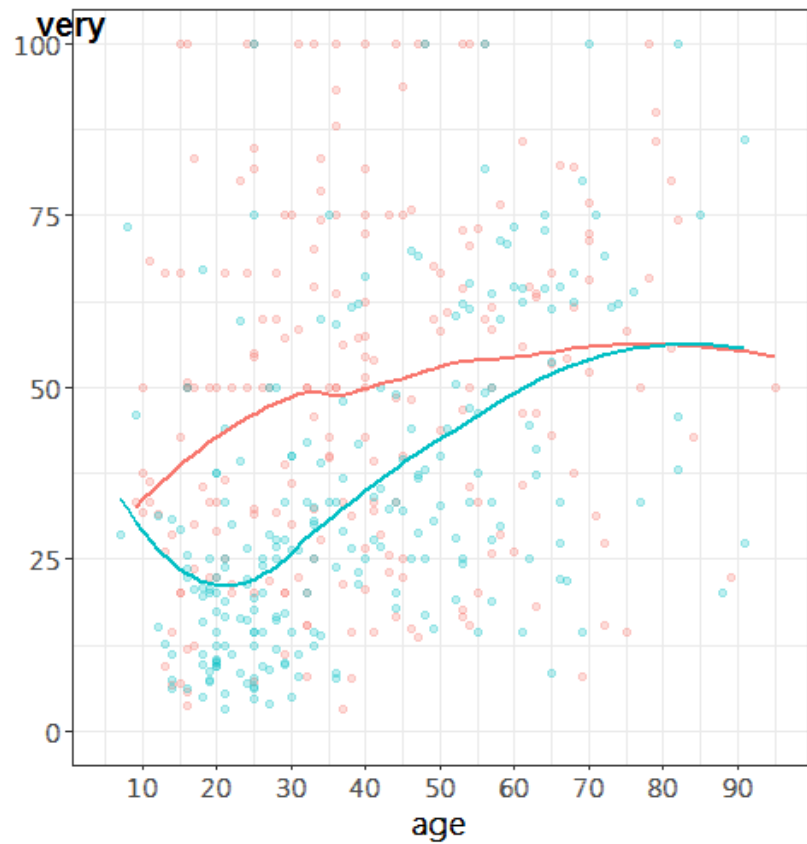
That's *really* cheap.

Included: all items that amplified a following adjective

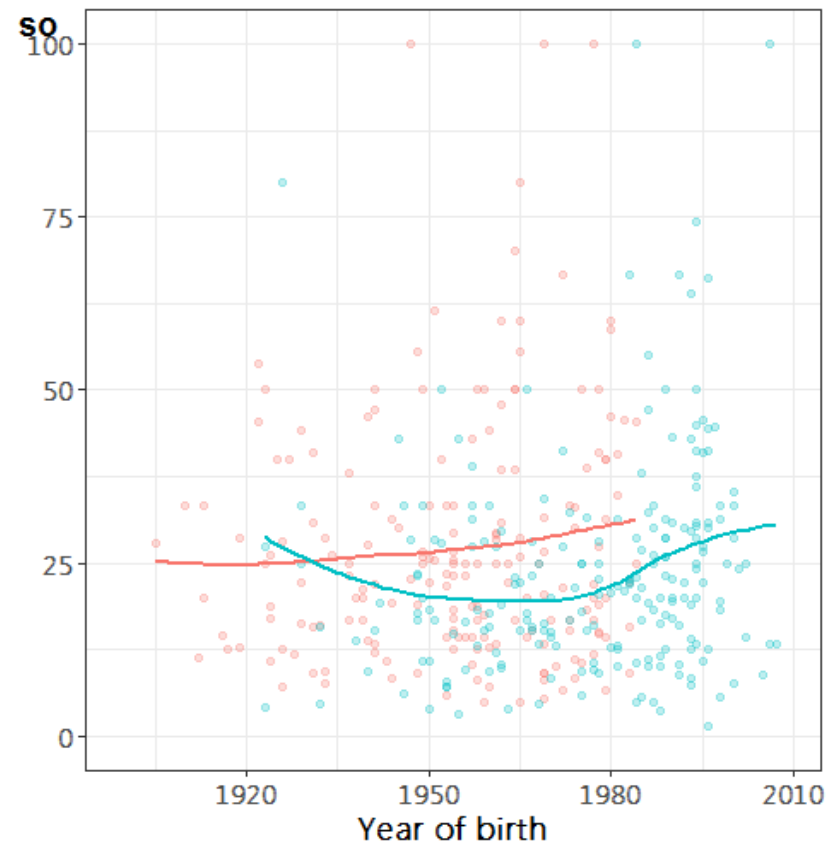
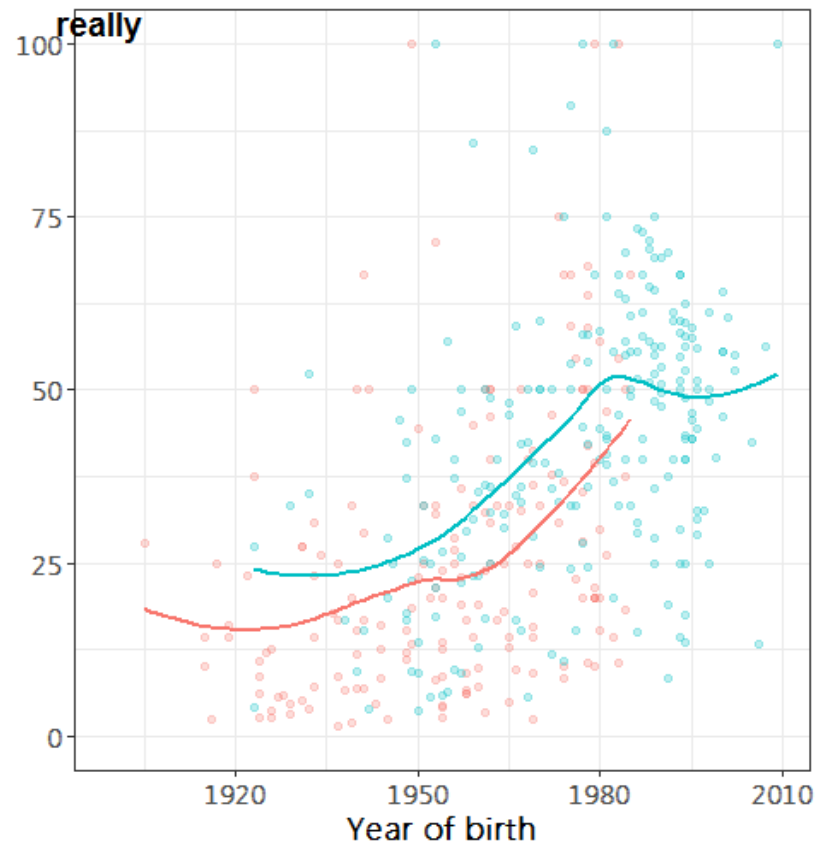
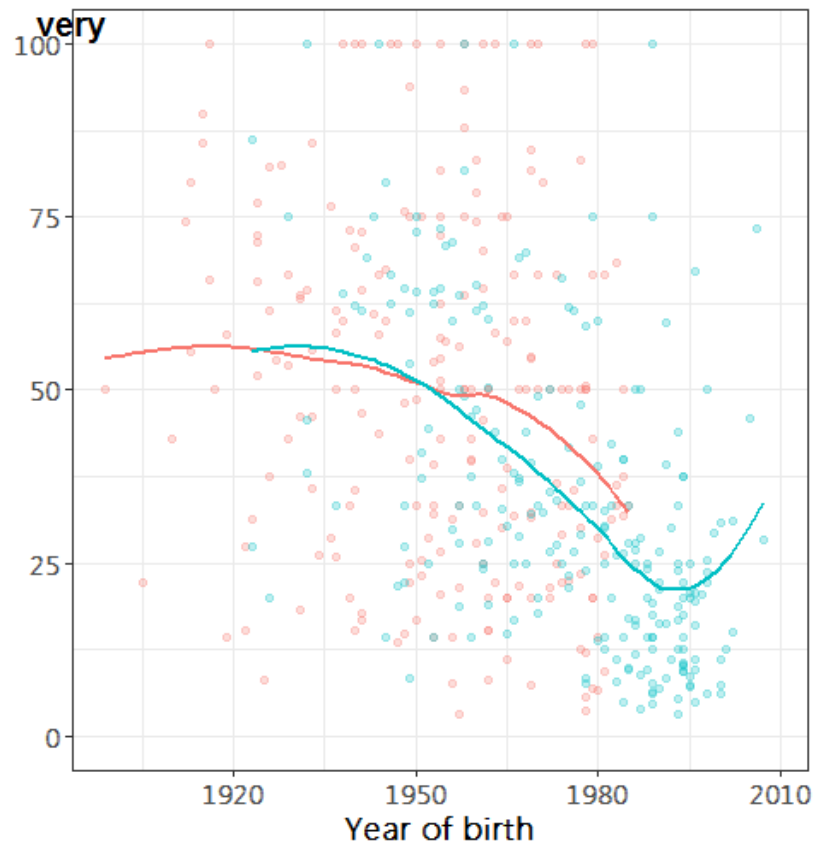
The variable set includes a great number of variants with the most common intensifiers (*very, really, so*) representing approx. 90% of the results.

Data extraction & coding





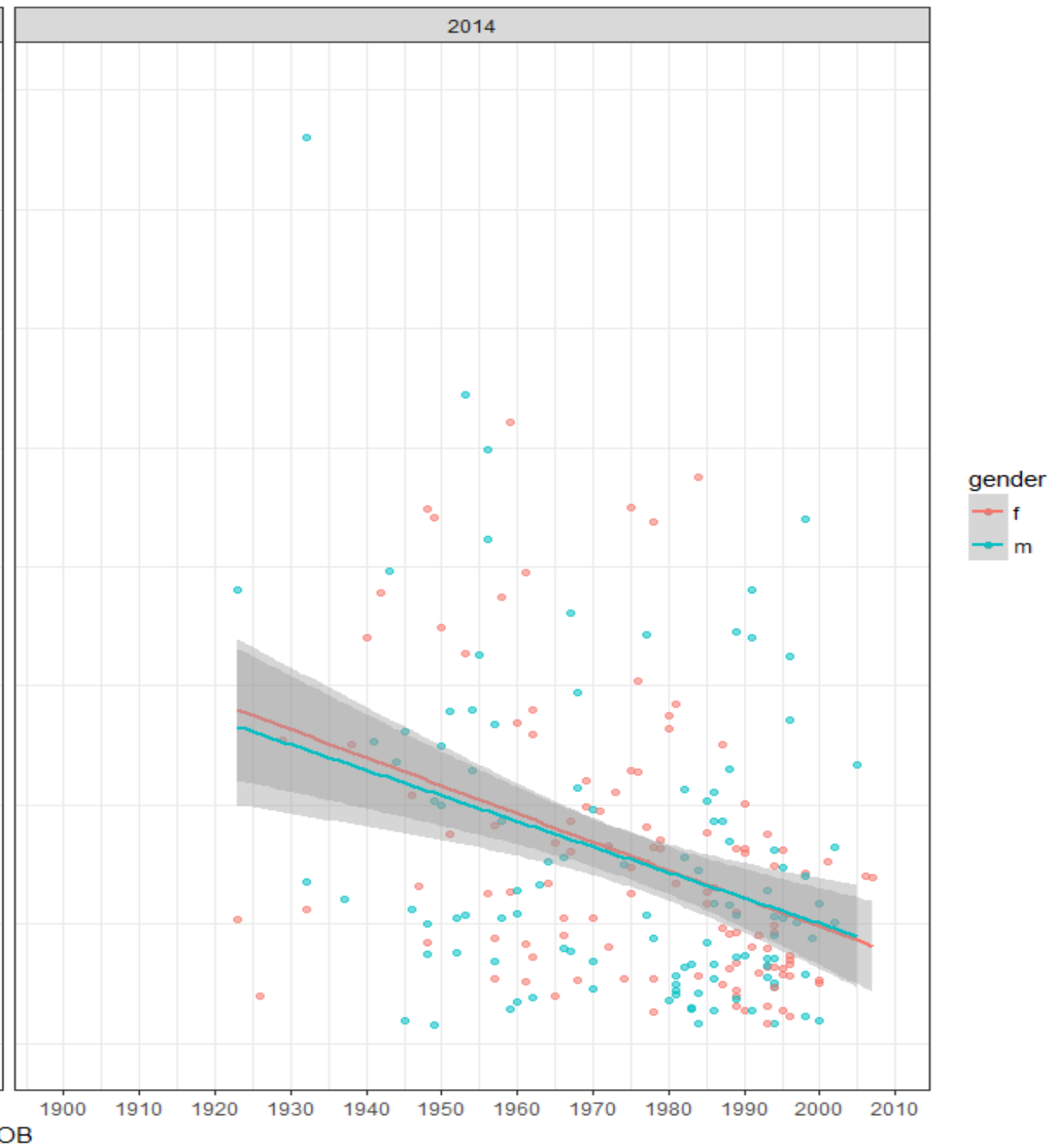
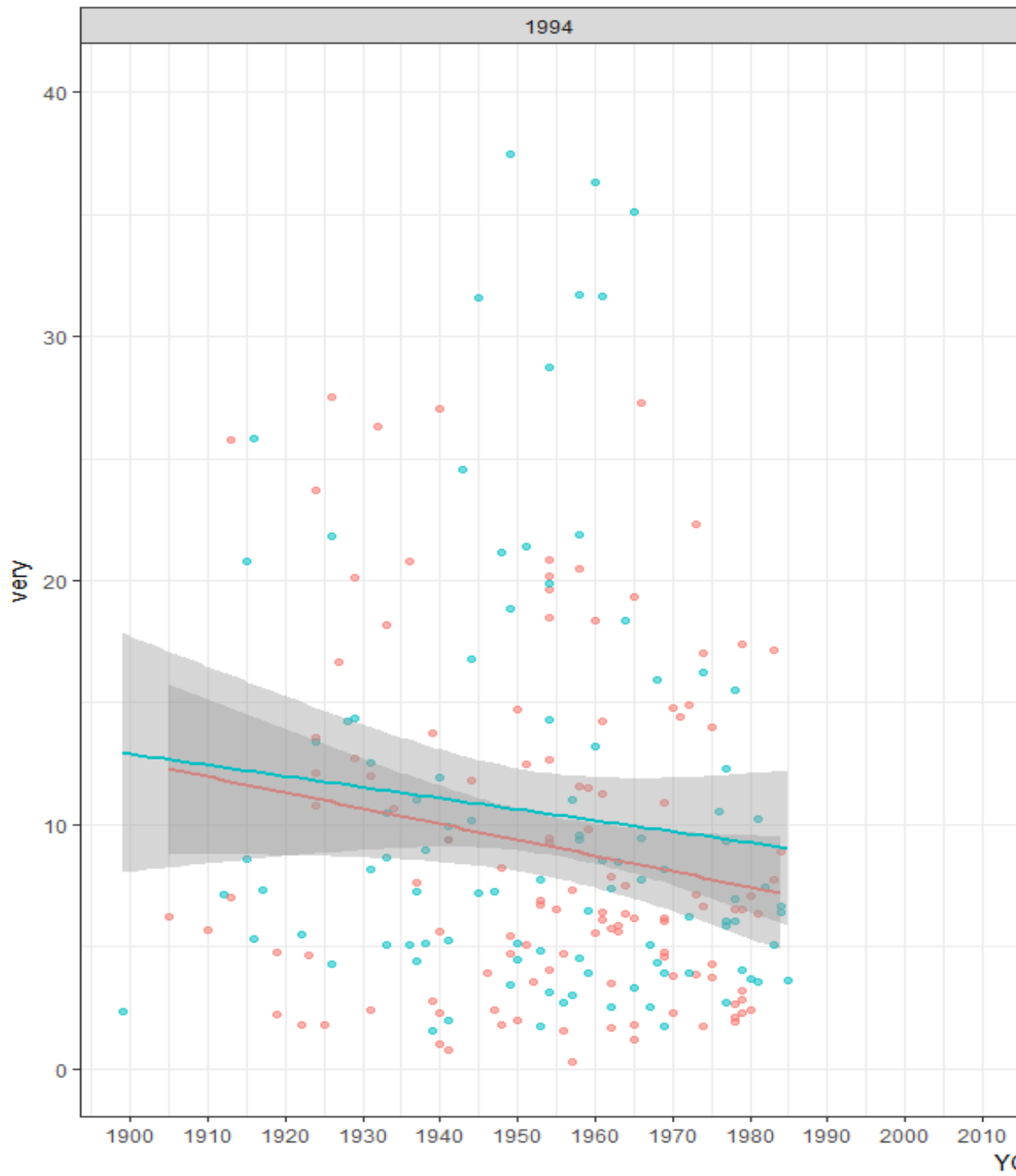
CORPUS — BNC1994 — BNC2014



CORPUS — BNC1994 — BNC2014

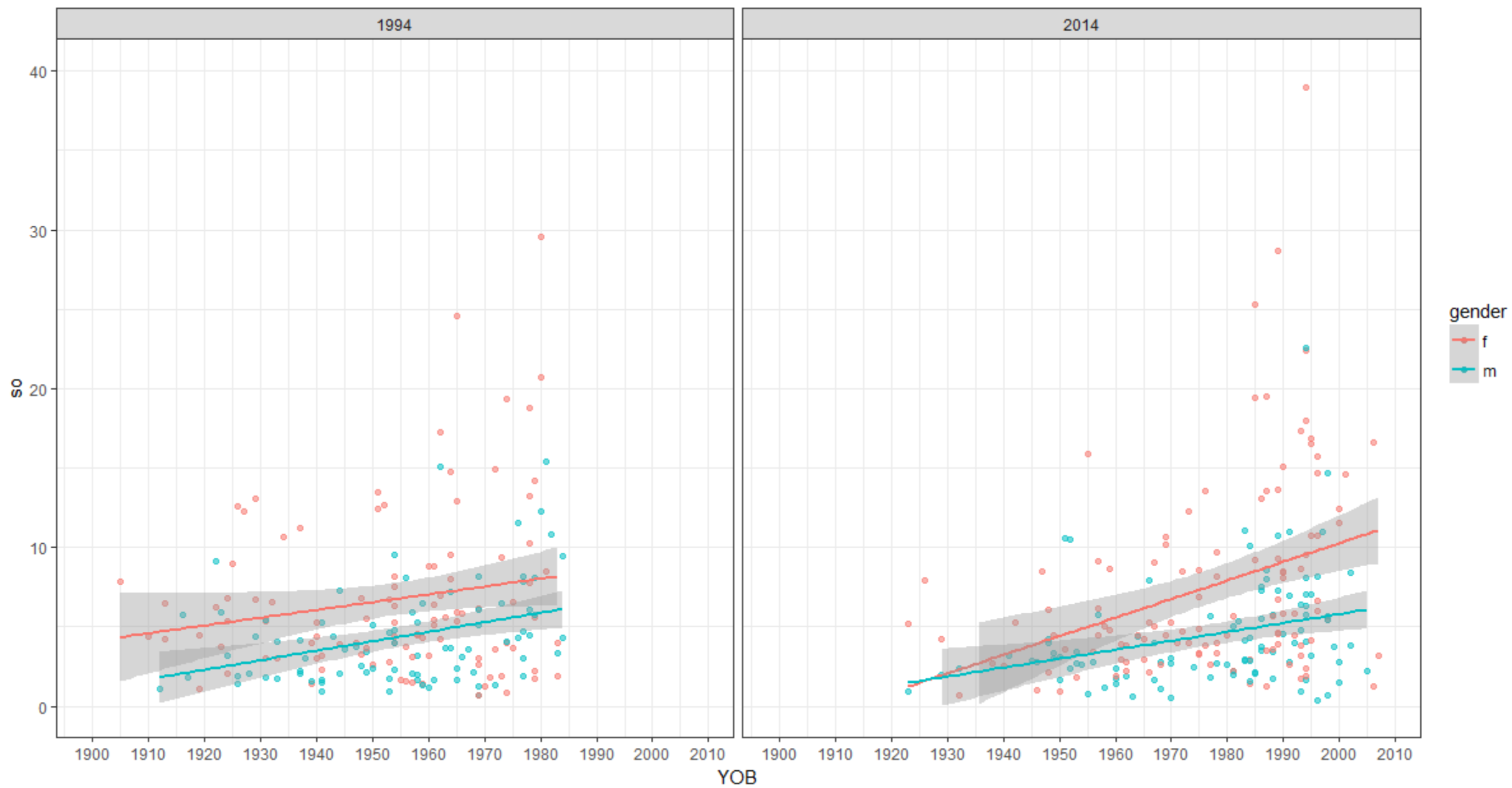
What does the BNC subset study offer in terms of new insights?

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.



gender
f
m





Thank you

