MY ALTERNATION, MY RULES
INVESTIGATING SYNTACTIC VARIATION IN INDIVIDUAL ENGLISHES

Lauren Fonteyn

Leiden University Centre for Linguistics (LUCL)
“The changes in language fulfill themselves in the individual, partly through his spontaneous activity, …, and partly through the influence which each individual receives from others.” (Paul 1891: 5)

Growing body of historical (socio-)linguistic studies devote more attention to **individual language use** (e.g., Pratt & Denison 2000 [but: Van Bergen 2013]; Bergs 2005; Raumolin-Brunberg 2009) and **how that feeds into population-level change** (e.g., Nevalainen *et al.* 2011; Nevalainen & Raumolin-Brunberg 2016; Baxter & Croft 2016; Hundt *et al.* 2017; Petré 2017; Petré & Van de Velde 2018).
Why is this important?

(1) Increasing evidence that individuals are quite different in L1-attainment (Dabrowska 2012), and differences in individual usage exist to the extent that they can be used to identify speakers (Barlow 2013, Nini 2018).

(2) If the population is indeed very heterogeneous, looking at the generalized trend alone makes you run the risk of:
   - Missing patterns (Tagliamonte & Baayen 2012)
   - Mistakenly identifying patterns (Fonteyn 2017)
<table>
<thead>
<tr>
<th><strong>Labov</strong> (2001): “the individual speaker can only be understood as the product of a unique social history”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eckert</strong> (2008; 2012): individual language is <em>bricolage</em>, in which “speakers make social-semiotic moves, reinterpreting variables, combining and recombining them”</td>
</tr>
<tr>
<td><strong>Coulthard</strong> (2004): “whereas in principle any speaker/writer can use any word at any time, speakers in fact tend to make typical and individuating co-selections of preferred words.”</td>
</tr>
</tbody>
</table>

**phonological – lexical variation**
Syntax (and syntactic change) is underrepresented:

“Over the last forty years, more sociolinguistic work has been done on morphology than on syntax and much more on phonology than on morphology and syntax put together”. (Nevalainen et al. 2011 in reference to Labov 2006: 380-403)

“Nevalainen and colleagues (2011), for instance, discuss a number of instances of variation in Early Modern English, but most of these are more morphophonological than syntactic in nature.” (Petré & Van de Velde 2018: 2)
THE RISE OF THE VERBAL GERUND
THE GERUND ‘ALTERNATION’

- Deverbal nominalization in -ing

- Two types:
  1. **nominal gerund**: … the dishonour of Gods Name should affect us more then *the shedding of the warmest blood in our veins* (John Flavell, 1668)
  2. **verbal gerund**: he also made an end of … ø *Shedding ø the Blood of Rams, Lambs, Heifers, Goats and other Creatures* for the Sins and Transgressions of Men (George Fox, 1686)
HISTORICAL DEVELOPMENT

Old English
Gerund is an abstract deverbal noun, with nominal syntactic features (NG) (e.g. *by writing of a letter*).

Middle English
Gerund was re-analysed as part of the verb system and acquired the ability to govern a *direct object* (c. 1250 - e.g. *by writing a letter*).

Modern English
Gradual spread of the verbalized gerund (Fanego 2004) to new syntactic environments.
HISTORICAL DEVELOPMENT

- Competition between two (or more) forms in the language system has either one of two outcomes (Traugott & Trousdale 2013: 18): substitution or retention (development towards division of labour);

- Historical process described as substitution of NG by VG in Middle and Modern English (De Smet 2008; Nevalainen et al. 2011);

- Gradual reorganisation of the English ‘ing-form network’ – the functionally hybrid gerund splits into a nominal and a clausal component (Fonteyn 2019).
Nevalainen et al.
- To capture slow, long-term processes of change, for which **real-time data** are required.
- Individual language users’ participation in six morphological and syntactic changes in English as they unfold in real time over three centuries.
- Study ranged from **morpheme replacements** (e.g. ye > you) to replacement in **more abstract structural patterns** (e.g. Ving (of), negation)

Results
- degree to which language users participated in variability seen in mean depended on (i) the type of language change analyzed, (ii) the stage of development of the change, and (iii) the rate of diffusion of the process over time.
In the graphs, the x-axis represents the proportion of the new form, and the y-axis represents the total number of uses. The figures show the distribution of the frequency of forms over different periods.

Nevalainen et al. (2011)
Nevalainen et al.

Results
- The results indicate that in the mid-range phase of a change in progress abstract structural patterns are more apt to be used variably than morphological changes.
- The more protracted the process, the more individuals have variable usage.
- These two phenomena combine to produce the extreme cases of a rapid pronoun change with proportionately fewer in-between language users (ye > you) and a protracted syntactic change with proportionately fewer progressive and conservative individuals (i.e. verbalization of gerund).
Baxter and Croft (2016)

- “The largest standard deviation value observed, .26, is just below the threshold of .29, meaning that there is a significant fraction of mixed-usage speakers throughout the change.”

- “We see indeed that the speakers are mostly in the middle of the range in the intervals 1580 to 1599 and 1600 to 1619, indicating that this change is a different pattern from the previous two.”

(2016: 165-167)
Variable users?

- Verbalization of the gerund is a **slow, gradual** change of an abstract pattern, with a **significant fraction** of mixed-usage speakers throughout the change (population is less polarized than with faster and more ‘salient’ changes).
- Suggestion that **speakers will behave more like the population level mean** in such cases.

**However:**
- Only looks at overall proportions of old versus new variant for each individual – not at whether (and how) individual conditions variation.
- Not every 50-50 distribution is the same…

(Figure taken from Fonteyn (2017))
Frank

breakfast – lunch – snack – dinner

50% hot meals

Laura

breakfast – lunch – snack – dinner

50% hot meals
“Some subtypes of the new construction [i.e. VG] became possible before others, their generalization being largely governed by two different linguistic hierarchies” (Fanego 2004: 50)

HISTORICAL DEVELOPMENT

- hierarchy of relative ‘nominality’
  1. $\emptyset$ (by) eating the forbidden fruit
  2. poss (by) Adam’s eating the forbidden fruit
  3. the (by) the / an unadvised eating the forbidden fruit

- grammatical relations hierarchy
  1. prep - by my not doing it
  2. object - It does not excuse my not doing it
  3. subject - my not doing it may be laid upon the account of my weariness

- verb type (Maekelberghe 2017; Fonteyn forthc.)
  1. light - giving of thanks
  2. state - the having of a sword
RESEARCH QUESTIONS

- Investigate variable importance of factors involved in the diachronic verbalization of the English gerund (NG > VG)

- Investigate the issue in aggregated vs. individual data:
  - “How consistently do speakers in a community converge on the same constraint effects on linguistic variables?” (Guy 2015)
  - Is there substantial inter-speaker variation, or, in other words, do we observe ‘individual conditioning’ of the existing variation?
  - If so, where/how does that individuality reveal itself in the individual’s linguistic behaviour?
DATA AND ANALYSIS
CORPUS

**Early Modern Multiloquent Authors** (EMMA; Petré *et al.* 2018)
- Sample of 50 of the most prolific English writers born in the 17th century (mostly belonged to the London-based elite)
- 5 generations

**In this study:**
- 14 randomly selected speakers, born between 1600 and 1635 (2 generations)
- Focus on prose and letters
- 27,000+ *ing*-forms (including lexicalised *ing*-forms)
- 7,127 nominal and verbal gerunds
## CORPUS - RAW

<table>
<thead>
<tr>
<th></th>
<th>NG</th>
<th>VG</th>
<th></th>
<th>NG</th>
<th>VG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heylyn, Peter</td>
<td>344</td>
<td>401</td>
<td>Bunyan, John</td>
<td>401</td>
<td>374</td>
</tr>
<tr>
<td>Fuller, Thomas</td>
<td>172</td>
<td>271</td>
<td>Dryden, John</td>
<td>90</td>
<td>320</td>
</tr>
<tr>
<td>Milton, John</td>
<td>235</td>
<td>339</td>
<td>Flavell, John</td>
<td>58</td>
<td>165</td>
</tr>
<tr>
<td>Prynne, William</td>
<td>496</td>
<td>558</td>
<td>Fox, George</td>
<td>213</td>
<td>382</td>
</tr>
<tr>
<td>Taylor, Jeremy</td>
<td>102</td>
<td>380</td>
<td>Pierce, Thomas</td>
<td>91</td>
<td>295</td>
</tr>
<tr>
<td>Boyle, Robert</td>
<td>79</td>
<td>515</td>
<td>Swinnock, George</td>
<td>55</td>
<td>262</td>
</tr>
<tr>
<td>Boyle, Roger</td>
<td>79</td>
<td>192</td>
<td>Tillotson, John</td>
<td>82</td>
<td>176</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2497</strong></td>
<td><strong>4630</strong></td>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>NG</td>
<td>VG</td>
<td>Name</td>
<td>NG</td>
<td>VG</td>
</tr>
<tr>
<td>----------------</td>
<td>-----</td>
<td>-----</td>
<td>-----------------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Heylyn, Peter</td>
<td>46.17%</td>
<td>53.83%</td>
<td>Bunyan, John</td>
<td>51.74%</td>
<td>48.26%</td>
</tr>
<tr>
<td>Fuller, Thomas</td>
<td>38.83%</td>
<td>61.17%</td>
<td>Dryden, John</td>
<td>21.95%</td>
<td>78.05%</td>
</tr>
<tr>
<td>Milton, John</td>
<td>40.84%</td>
<td>59.16%</td>
<td>Flavell, John</td>
<td>26.01%</td>
<td>73.99%</td>
</tr>
<tr>
<td>Prynne, William</td>
<td>47.06%</td>
<td>52.94%</td>
<td>Fox, George</td>
<td>35.80%</td>
<td>64.20%</td>
</tr>
<tr>
<td>Taylor, Jeremy</td>
<td>21.16%</td>
<td>78.84%</td>
<td>Pierce, Thomas</td>
<td>23.58%</td>
<td>76.42%</td>
</tr>
<tr>
<td>Boyle, Robert</td>
<td>13.30%</td>
<td>86.70%</td>
<td>Swinnock, George</td>
<td>17.35%</td>
<td>82.64%</td>
</tr>
<tr>
<td>Boyle, Roger</td>
<td>29.15%</td>
<td>70.85%</td>
<td>Tillotson, John</td>
<td>31.78%</td>
<td>68.22%</td>
</tr>
<tr>
<td><strong>Total (AVG)</strong></td>
<td><strong>31.8%</strong></td>
<td><strong>68.2%</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Graphs showing the percentage of VG_tot and VG_G1 and VG_G2 over the decades from 1620s to 1690s.
<table>
<thead>
<tr>
<th>Name</th>
<th>R²</th>
<th>Name</th>
<th>MAE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pierce, Thomas</td>
<td>0.02854212</td>
<td>Boyle, Robert</td>
<td>16.482</td>
</tr>
<tr>
<td>Milton, John</td>
<td>0.04123702</td>
<td>Bunyan, John</td>
<td>16.34</td>
</tr>
<tr>
<td>Fuller, Thomas</td>
<td>0.04522199</td>
<td>Pierce, Thomas</td>
<td>15.435</td>
</tr>
<tr>
<td>Boyle, Robert</td>
<td>0.06929551</td>
<td>Fuller, Thomas</td>
<td>10.685</td>
</tr>
<tr>
<td>Fox, George</td>
<td>0.7123062</td>
<td>Heylyn, Peter</td>
<td>10.42</td>
</tr>
<tr>
<td>Bunyan, John</td>
<td>0.748092</td>
<td>Milton, John</td>
<td>7.4525</td>
</tr>
<tr>
<td>Heylyn, Peter</td>
<td>0.8357582</td>
<td>Prynne, William</td>
<td>4.272</td>
</tr>
<tr>
<td>Prynne, William</td>
<td>0.8785116</td>
<td>Fox, George</td>
<td>3.39</td>
</tr>
</tbody>
</table>
Tagliamonte & Baayen (2012)

Mixed effect logistic regression model
- Individual contributors (if contributions > 1) added as random effect:
  - “If the individual is not considered as a predictor in the model and the individuals in the data use a variant with widely diverging individual probabilities (...) we may miss out on the opportunity to better understand the data, and to explain more of the variation.” (2012: 8)
  - “It is only when the individual is brought into the model (...) that it becomes possible to correct for this systematic prediction error” (2012: 9)
METHODOLOGY

Interested in knowing how important individuals are and what they do:

**Conditional inference tree** (binary splits until no longer justified)
- Clearly visualize interactions and how multiple predictors operate in tandem
- `ctree_ger <- ctree(gerund ~ author + det + func + verb_type + genre + generation, data=df)`

**Random forest** (1000 trees)
- Large number of conditional inference trees that contribute a vote to final model
- `cforest_ger <- cforest(gerund ~ author + det + func + verb_type + genre + generation, data=df, controls=cforest_control(mtry = 6, ntree=1000))`
<table>
<thead>
<tr>
<th>Determiner</th>
<th>Function</th>
<th>Verb Type</th>
<th>Genre</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BARE</strong></td>
<td><strong>BY, IN, FOR, OF, TEMP, ...</strong>&lt;br&gt;by onely torturing of men&lt;br&gt;in the destroying of the ...&lt;br&gt;after his blaspheming Shakespeare.</td>
<td><strong>LEX</strong>&lt;br&gt;... whilst others make them groan, by abusing them to sin, and subjecting them to their lusts.</td>
<td><strong>PROSE</strong></td>
</tr>
<tr>
<td><strong>POSS</strong></td>
<td><strong>THE</strong>&lt;br&gt;The seeing of our Friends in Heaven will ...&lt;br&gt;<strong>A</strong>&lt;br&gt;a cry will be among you, and a wishing you had never been born&lt;br&gt;<strong>NO</strong>&lt;br&gt;... no reverencing of images&lt;br&gt;<strong>DEM</strong>&lt;br&gt;This forgetting of the God that saves us ...</td>
<td><strong>‘LIGHT’</strong>&lt;br&gt;He is accus'd of Malevolence, and of taking Actions in the worst Sence&lt;br&gt;... that prayers, and supplication, and giving of thanks be made for all men</td>
<td><strong>LETTERS</strong></td>
</tr>
<tr>
<td><strong>THE</strong></td>
<td><strong>SUBJECT COMPLEMENT</strong>&lt;br&gt;... that there should be christening of children&lt;br&gt;It is not the giving out of mercy that troubles him, but ...</td>
<td><strong>HAVE</strong>&lt;br&gt;there is more required to make a good Scholler, then onely the having of many bookes</td>
<td></td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>SUBJECT</strong>&lt;br&gt;The laying down of life did abundantly proclaim his love</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RANDOM FOREST
Model residuals

F1-score 0.81
Accuracy 0.87

Variable importance

Det 0.28
Author 0.05
Func 0.02
Generation 0.008
verb_type 0.006
Genre 0.002
CONDITIONAL INFERENCE TREE
Model residuals

C 0.92
F1-score 0.80
Accuracy 0.87

Binary splits

(1) determiner
{a, the, dem} vs. {bare, no, poss}

(15) Within {bare, no, poss}
> split: function
{prepositions} vs. {other +temp}
(2) author
{J Taylor, Rob Boyle, W Prynne} vs. {others}

(10) verb type
{have, light} vs. {lexical}

(12) Func
{direct object, Temp prep} vs. {others}
Main results:

• All authors engage in mixed usage (// Nevalainen et al. 2011; Baxter & Croft 2016), but the relationship between their usage profiles and the population-mean is still complex:

• Individual trajectories (lifespan developments) can deviate from population trend;
• Variable importance of author ranks second, below determiner, but above function, generation, verb type and genre.
• Indications that different individuals condition the variation differently (following one or two initial, higher-order splits).
CONCLUSIONS

- The ‘relative nominality’ of the context (i.e., type of determiner; cf. Fanego 2004) is the most important predictor of the variation, consistently returning as the factor that most ‘effectively’ explains NG vs. VG usage.
- But further grammatical conditioning depend on individual:
  - This indicates ‘shared constraints’, but more chances of idiosyncrasies at lower levels of the tree (when a number of features have been ‘co-selected’);
  - Speakers do not only ‘adjust’ overall rates of use of specific variants but also ‘adjust’ the contextual constraints on a variable process (cf. Guy 2015).
CONCLUSIONS

- The ‘relative nominality’ of the context (i.e., type of determiner; cf. Fanego 2004) is the most important predictor of the variation, consistently returning as the factor that most ‘effectively’ explains NG vs. VG usage.
- But further grammatical conditioning depend on individual:
  - In line with the idea that different speakers have ‘different grammars’ because they each have unique linguistic ‘histories’ (Dabrowska 2012; Raumolin-Brunberg 2009);
  - If the factors conditioning variation differ from individual to individual, the aggregate data set contains ‘contradictory’ information, leading to ‘information loss’ (cf. Tagliamonte & Baayen 2012).
Future research?

- Weak Variable importance of Generation
  - Worth adding more generations?
- Link individual trajectories to lifespan (Raumolin-Brunberg 2009)
- Individual ranks high, but more modest than, for instance, in Tagliamonte & Baayen (2012):
  - More homogeneous set of speakers (male, educated, etc.)?
  - Different type of change (duration/speed, ‘salience’)?
  - Investigate was/were and other types of changes in same corpus?
Multifactorial Prediction and Deviation Analysis with Random Forests (MUPDARF)

- Train Random Forest Model on ‘baseline’ (i.e. population) and use it to predict ‘target’ (i.e. one individual)
- If the model predicts correctly, deviation score is 0.
- If the model fails, a positive or negative deviation score is assigned depending on whether target choice was NG or VG.
- Gives a more detailed view on who deviates in which contexts.
good fit, small deviations

less optimal fit, larger deviations
THANK YOU.

Lauren Fonteyn – l.fonteyn@hum.leidenuniv.nl

Leiden University Centre for Linguistics
REFERENCES

<table>
<thead>
<tr>
<th>Name</th>
<th>NG</th>
<th>VG</th>
<th>Name</th>
<th>NG</th>
<th>VG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heylyn, Peter</td>
<td>32.78</td>
<td>38.21</td>
<td>Bunyan, John</td>
<td>77.25</td>
<td>72.05</td>
</tr>
<tr>
<td>Fuller, Thomas</td>
<td>62.07</td>
<td>97.79</td>
<td>Dryden, John</td>
<td>79.49</td>
<td><strong>282.65</strong></td>
</tr>
<tr>
<td>Milton, John</td>
<td><strong>115.35</strong></td>
<td>166.4</td>
<td>Flavell, John</td>
<td>17.74</td>
<td>50.46</td>
</tr>
<tr>
<td>Prynne, William</td>
<td>61.88</td>
<td>69.61</td>
<td>Fox, George</td>
<td>60.76</td>
<td>108.98</td>
</tr>
<tr>
<td>Taylor, Jeremy</td>
<td>38.18</td>
<td>142.22</td>
<td>Pierce, Thomas</td>
<td>34.08</td>
<td>110.47</td>
</tr>
<tr>
<td>Boyle, Robert</td>
<td>54.48</td>
<td><strong>355.13</strong></td>
<td>Swinnock, George</td>
<td>22.64</td>
<td>108.27</td>
</tr>
<tr>
<td>Boyle, Roger</td>
<td>62.61</td>
<td>152.16</td>
<td>Tillotson, John</td>
<td>63.7</td>
<td>136.71</td>
</tr>
<tr>
<td><strong>Total (AVG)</strong></td>
<td>55.93</td>
<td>135.08</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RELATED CONSTRUCTIONS

- Gerunds are ‘linked’ to present participle.
  - *Iudas’s religion lay much in the bag, but his soul is now burning in hell* (Bunyan, 1676)
  - *The stench of whose burning flesh (offensive to others) was a perfume to him* (Fuller, 1646)
  - … *his fair Behaviour in the War, Not plundering Towns, nor burning Villages*; (Dryden, 1694)

- Given the increasing overlap in form and function between them, it seems likely that “the different constructions have influenced and reinforced each other in various ways” (Killie & Swan 2009: 359; van de Pol & Petré 2015; Fonteyn & van de Pol 2016).

- Correlation between ‘pres part + argument’ and relative frequency of VG?
No significant correlations.
Positive correlation frequency (p < 0.05):
- Adverbial
- Attributive
- Participle general
Frequency NG

Positive correlation frequency (p < 0.05):
- Adverbial
- Attributive
- Participle general