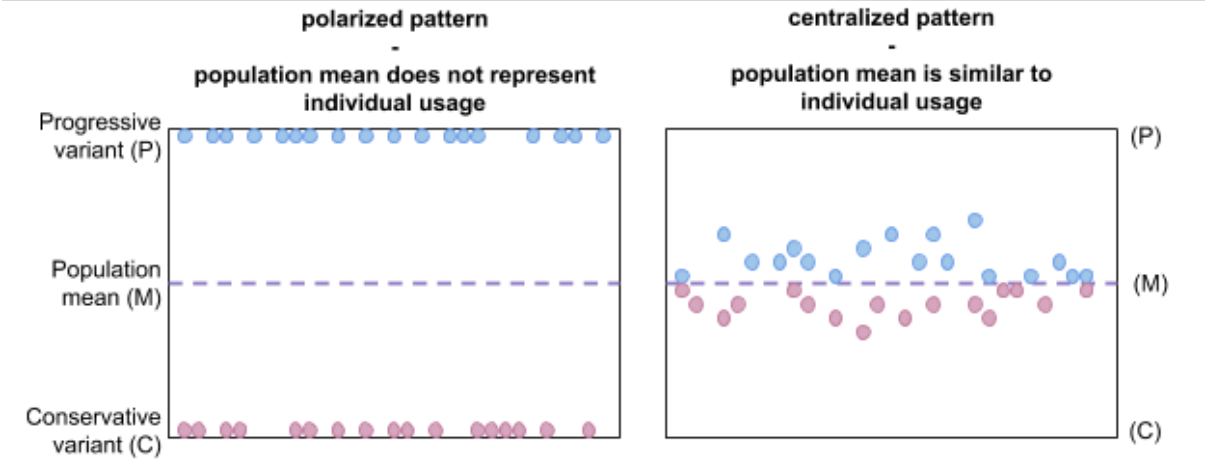
***Doing (of) your own thing*: a quantitative analysis of syntactic variation and idiolect in Late Modern English**

Despite the longevity of interest in the role individuals play in language change, it appears that historical (corpus) linguistics has predominantly focussed on studying change in the 'grammar-generated variation' in aggregated, population-level data. However, in recent years, scholars have been placing more explicit emphasis on the relation between the linguistic behaviour of individuals and the changes we observe in such population-level language (see, among many others, Nevalainen *et al.* 2011; Croft & Baxter 2016; Hundt *et al.* 2017; Petré & Van de Velde 2018). One notable example is a study by Nevalainen et al. (2011), who use real-time historical corpus data from the Early Modern English period to document how individual language users 'participate' in different types of morpho-syntactic change. From their study, Nevalainen et al. (2011) concluded that, when confronted with the fact there are "different ways to say the same thing" (Labov 1972:188), individual language users are more consistently fully progressive or conservative with simple morpheme replacement processes, and are more likely to actually participate in the variability that is observable in population-level language if the process is protracted and involves a change in an abstract structural pattern. In other words, if two variant forms are used at a 50-50 proportion in an aggregate data pool, the behaviour of individuals tends to be more centralized (rather than polarized) when the two variants concern abstract syntactic patterns in competition.



Given such figures, it is tempting to conclude that slow, gradual changes of abstract patterns are essentially changes whereby the behaviour of individual language users aligns with the mean observed in aggregate language. However, such conclusions cannot and should not be drawn, as such figures simply represent how often individuals opt for a progressive or conservative variant, and thus, they do not reveal anything regarding the potential differences in 'grammatical rules' these individuals apply to condition the observed variation.

The aim of this study, then, is to investigate (i) whether individuals who use 'alternate ways of saying the same thing' employ shared or idiosyncratic rules to condition that variation, and (ii) at what level of specificity such possible idiosyncrasies emerge. The specific variation pair under scrutiny is illustrated in (1):

(1) a. *...the dishonour of Gods Name should affect us more then* [*the shedding **of** the warmest blood in our veins*] (John Flavell, 1668)

b. *he... made an end of...* [*Shedding **ø** the Blood of Rams, Lambs, Heifers, Goats and other Creatures*] (George Fox, 1686)

The structures in (1a)—the nominal gerund [NG]—and (1b)—the verbal gerund [VG]—illustrate two types of *ing*-nominals that were used interchangeably during the Modern English period (see, among many others: Fanego 2004, Nevalainen et al. 2011, De Smet 2013, Fonteyn 2019). The study presents a quantitative analysis of 15,000 NGs and VGs (taken from the EMMA corpus (Petré *et al.* 2017)) found in the writings of 19 authors born in three subsequent generations (between 1599 and 1640), all of whom have all been proven to be connected to the same social circles. The data set has been examined by means of two complementary statistical models that have recently been added to the variationist toolkit: Conditional Inference Trees and Random Forests (Tagliamonte & Baayen 2012). Each observation in the data set has been coded for language-internal (e.g. *determiner, function in clause, verb type*) and external factors (e.g. *individual, age, generation, genre*).

The results of the Random Forest analysis indicate that individuality is an important predictor of the variation, which trumps the predictive power of some of the proposed language-internal factors as well as the higher-order sociolinguistic factors such as the age or generation of the author. Furthermore, the complementary Conditional Inference Tree reveals that, while all authors do share some 'rules', the grammatical choices of certain (subsets of) authors significantly deviate from those of others at lower levels of the tree. These findings suggest that (i) syntactic variation may be conditioned by means of a (limited set of) grammatical rules that are shared by all individuals in a population, but (ii) at the same time, where grammatical rules interact, much of the observed variation seems to be associated with the individual.

Baxter, G. & W. Croft. 2016. Modeling language change across the lifespan: Individual trajectories in community change. *Language Variation and Change* 28.2, 129-173.

De Smet, H. 2013. *Spreading patterns: Diffusional change in the English system of complementation.* Oxford: Oxford University Press.

Fanego, T. 2004. On reanalysis and actualization in syntactic change: The rise and development of English verbal gerund. *Diachronica* 21.1, 5–55.

Hundt, M., Mollin, S. & S. E. Pfenninger. 2017. *The Changing English Language: Psycholinguistic Perspectives*. Cambridge: CUP.

Labov, W. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Nevalainen, T., H. Raumolin-Brunberg & H. Mannila. 2011. The diffusion of language change in real time: Progressive and conservative individuals and the time depth of change. *Language Variation and Change* 23.1, 1–43.

Petré, P., L. Anthonissen, S. Budts, E. Manjavacas, W. Standing, E.-L. Silva & O. Strik. 2017. *Early-Modern Multiloquent Authors (EMMA)*. Antwerp: Linguistics Dept, (www.helsinki.fi/varieng/CoRD/corpora/EMMA).

Petré, P. 2017. The extravagant progressive: an experimental corpus study on the history of emphatic [*be* V*ing*]. *English Language and Linguistics* 21.2, 227-250.

Petré, P. & F. Van de Velde. 2018. The real-time dynamics of the individual and the community in grammaticalization. *Language* 94.4.

Tagliamonte, S. & H. Baayen. 2012. Models, forests and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24.2, 135-178.