A production bias model of the Constant Rate Effect

Henri Kauhanen & George Walkden
The University of Manchester
henri.kauhanen@manchester.ac.uk, george.walkden@manchester.ac.uk

Kroch (1989) advanced the hypothesis that when two grammatical options compete across a number of linguistic contexts and one replaces the other over time, the rate of change will be the same in all contexts. To date, this hypothesis has been studied in a number of languages and data sets (Kroch 1989; Santorini 1993; Taylor 1994; Pintzuk 1995; Postma 2010; Wallage 2013) and has accumulated enough support for it to be referred to as the Constant Rate Effect, or CRE (e.g. Pintzuk 2003).

CREs provide a fresh perspective on causation in syntactic change: evidence of a CRE is evidence against the view (e.g. Bailey 1973) that linguistic innovations adapt to linguistic contexts based on their functionality; instead, patterns of use observed in historical data are to be thought of as reflexes of more abstract, underlying grammatical changes. Despite the wealth of empirical studies that over the years have sought to establish CREs in historical data, this central intuition of Kroch (1989) has, however, never been explicated formally in a detailed model of change that takes both grammatical competition and contextual effects into account.

What is more, certain doubts have recently been raised concerning the standard way of detecting CREs in corpus data, which is to fit a number of independent logistic curves, one per each context of interest. Firstly, (1) Wallenberg (2015) and Willis (2015) show that, using this method, CREs can be empirically demonstrated in situations where they cannot be taken to support underlying grammatical unity: across languages and across geographical areas, respectively. On the other hand, (2) customary research practice in diachronic syntax has long acknowledged that fitting a number of independent logistic curves to a set of contexts leaves variation in the time dimension entirely unexplained: it would, in principle, be possible to establish a CRE across two contexts where the change goes to completion in one before it even takes off in another. Together, problems (1) and (2) imply that the standard operationalization of CREs is not sufficient for assuming that a single underlying change has occurred.

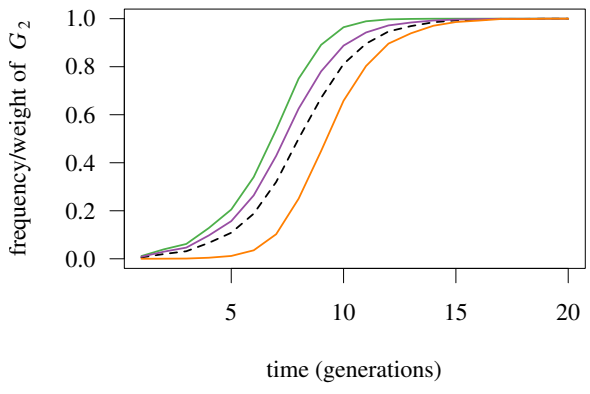

Fig. 1. CRE in a computer simulation.

In this talk, we aim to overcome these problems and to shed light on the nature of causation in language change by introducing a model of the CRE that is more tightly constrained, and therefore makes stronger (more restricted) empirical predictions than the traditional formulation. Starting with Yang's (2000) mathematical model of grammar competition, we augment the model with production biases across an arbitrary number of linguistic contexts. We show that this extension of Yang's framework naturally gives rise to the CRE in computer simulations (Fig. 1). Crucially, however, it is a theorem of the model that the time separation possible between any two contexts of one underlying grammatical change has a finite upper bound which is inversely proportional to the rate of the underlying change. This time separation theorem overcomes problem (2) identified above, and invites us

to reconsider a number of data sets in which CREs have previously been studied using the independent logistics operationalization.

For this purpose, we introduce a novel curve-fitting algorithm based on nonlinear least squares regression (Bates & Watts 1988). We investigate the model in the light of historical data by focussing on a number of changes for which a CRE has previously been established using the method of independent logistics. We show that the fit of our model to these data is no worse than a fit made using the traditional method (Fig. 2, top). Crucially, however, our model implies a maximal time separation for each change, which we also test, finding that the empirically observed time separations fall within the range prescribed by our model (Fig. 2, bottom).

We therefore show that a more constrained, theoretically motivated model of the CRE can fit historical data no worse than a less constrained one, and that it also generates new empirical predictions, also in line with the data, in the form of the time separation theorem. To complement these results, we investigate a number of pseudo-CREs – data sets that appear to exhibit a CRE if probed using the traditional method of independent logistics but that plausibly cannot due to unassailable a priori grounds (see problem (1), above). We show that here, when quantified by the residual error of the regressions, our model gives consistently worse fits than the traditional method, as desired (Fig. 2, top).

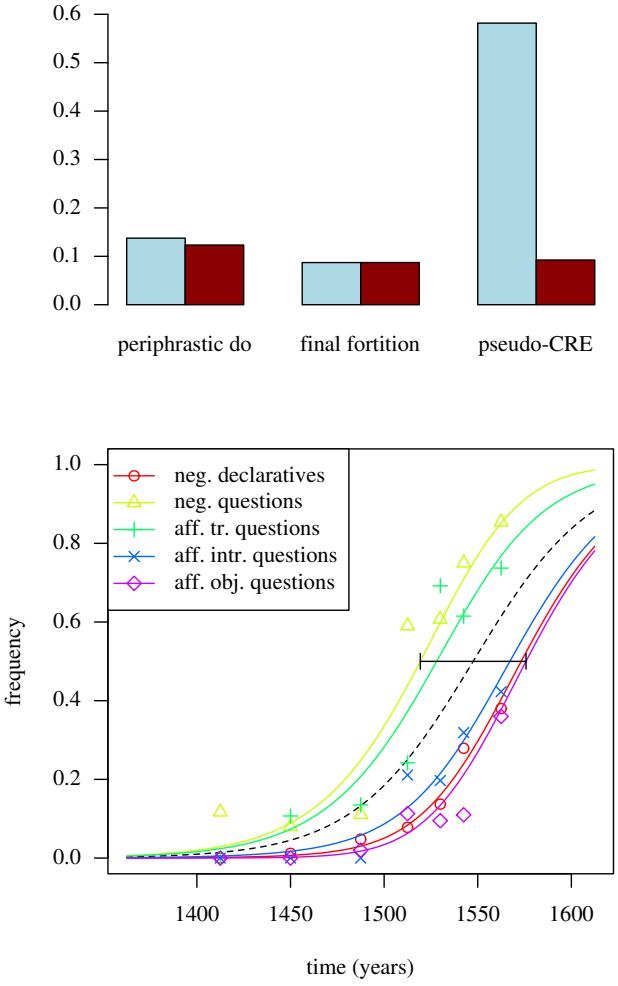

Fig. 2. *Top:* Errors of fits of our model (blue) and the standard model (red) to two CREs, rise of periphrastic *do* in English (Kroch 1989) and loss of final fortition in Early New High German (Fruehwald, Gress-Wright & Wallenberg 2009). *Bottom:* Fit of our model to the data on English *do*-support. The horizontal bar gives the maximal time separation between contexts licensed by the model.

Finally, we discuss a number of additional predictions the model makes about change in the presence of contextual biases. In brief, we show that in this extended model Yang's (2000: 239) Fundamental Theorem of Language Change ceases to hold, so that a distributional difference in the proportion of sentences parsed by two competing grammars is neither a sufficient nor a necessary condition of change on its own: the production biases induce a bifurcation in the parameter space of the model, and whether an innovatory grammatical option overtakes a conventional one comes to depend on a nonlinear interaction of grammar advantages (as defined in Yang 2000) and the magnitude and direction of the production biases. Conducting a full bifurcation analysis of the two-grammar case of the extended model, we work out the exact mathematical form of this dependence, and discuss its implications for population-level modelling of language change.