Henri Kauhanen and George Walkden
University of Manchester and University of Konstanz

When is a constant rate truly constant? A Monte Carlo power analysis of the logistic operationalization of constant rate effects

The Constant Rate Hypothesis (CRH) maintains that when a change targets multiple linguistic contexts simultaneously, the rate of change is the same across contexts (Kroch, 1989). Assuming that linguistic changes are S-shaped (see Denison, 2003; Blythe and Croft, 2012; Nevalainen, 2015), most scholars working on Constant Rate Effects (CREs) have followed Kroch (1989) in operationalizing this hypothesis using the simple logistic

$$p_c(t) = \frac{1}{1+e^{-s_c(t-k_c)}}. \tag{1}$$

Here $p_c(t)$ is the frequency of the innovation in context c at time t, s_c gives the rate of change, and k_c serves to translate the entire curve along the time axis. The Constant Rate Hypothesis is the statement that $s_c = s$ for some unique s, for all contexts c. Variation in the k_c parameter, on the other hand, is allowed.

Since its formulation 30 years ago, a number of studies have sought to establish the CRH by fitting (1) to various historical datasets (e.g. Kroch, 1989; Santorini, 1993; Pintzuk, 1995; Kallel, 2007; Wallage, 2008; Fruehwald, Gress-Wright, and Wallenberg, 2013; Wallage, 2013; Zimmermann, 2017). This pursuit has, however, been criticized on two grounds. Firstly, it has been pointed out that the currently available statistical methodology casts the CRH as the null, not the alternative, hypothesis (Paolillo, 2011). This has the unwelcome consequence that the rate at which the methods report false positives is unknown, a problem which is compounded by the typically small size of the datasets encountered in historical linguistic work. Secondly, it has been argued that operationalizing the CRH on the logistic (1) is ill-founded, as the constancy of rates of change across contexts is not actually being derived from any underlying model of speaker–speaker interactions (Kauhanen & Walkden, 2018).

In this contribution, we provide a systematic *a priori* power analysis of the logistic operationalization of the CRH using a Monte Carlo set-up. We consider two models: (A) one in which the s_c (rate of change) parameter is allowed to vary across contexts, and (B) one in which it is tied to a constant value: $s_c \equiv s$. The two models were used to generate synthetic data in a binomial experiment in which V tokens of the innovative linguistic form were drawn for each of H time points across the length of the logistic, for L contexts. The two models were then cross-fitted to each other's data using a logistic regression model involving a time–context interaction term (cf. Fruehwald et al., 2013); whenever the interaction term proved statistically significant at the $\alpha = 0.05$ level, model A (no CRE) was selected, otherwise model B (CRE) was selected. The procedure was repeated 1,000 times with each model as the generator and a confusion matrix established. Effect size was estimated as

$$E = 1 - \frac{|\min_c(s_c)|}{|\max_c(s_c)|}. \tag{2}$$

In other words, the closer the shallowest and steepest slopes s_c are to each other, the smaller the effect of s_c on the probability of the new linguistic form; E ranges from 0 to 1.

The Monte Carlo simulations indicate that the type II error rate β of the method may be approximated to a good degree by a combination of the effect size E and an overall data resolution quantity R, which is simply expressed as the product of the three data size parameters: $R = VHL$. The dependence of β on E and R is

$$\beta = \frac{1-\alpha}{1+\left(\frac{E}{11.60R^{-0.45}}\right)^{4.44}} \tag{3}$$

for a type I error rate α (see Figure 1).

Equation (3) can be used to estimate the data size required to reach a desired level of statistical power when modelling CREs using the logistic operationalization. It can also be used to evaluate to what extent previous studies have managed to reach reasonable standards of statistical rigour. To this end, we compiled a database of 27 studies from existing literature on CREs, and measured the data resolution ($R = VHL$) and effect size (E) quantities for each study. Of the 27 studies, 9 reach a type II error rate of 0.05 or less (a power of 0.95 or more), but nearly half of the studies (13) attest an error rate in excess of 0.5. The significance of this finding for the methodology of quantitative corpus-based diachronic linguistics will be discussed. We will also comment on whether the adoption of alternative models of the CRE (Kauhanen & Walkden, 2018) can be expected to ameliorate the situation.

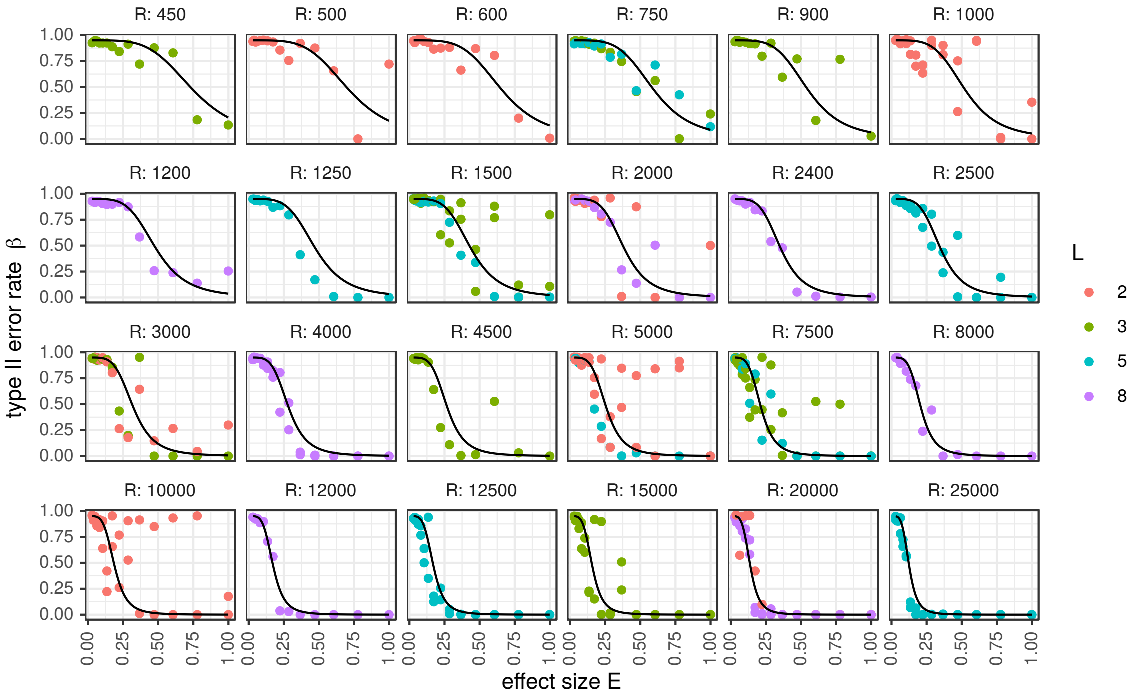

Figure 1. Type II error rate β of the logistic operationalization of CREs as a function of effect size E, overall data resolution R and number of contexts L, for type I error rate $\alpha = 0.05$. Points from simulations, curves from equation (3).

References

Blythe, R. & Croft, W. (2012). S-curves and the mechanisms of propagation in language change. *Language*, *88*(2), 269–304. Retrieved from https://www.jstor.org/stable/23251832

Denison, D. (2003). Log(ist)ic and simplistic s-curves. In R. Hickey (Ed.), *Motives for language change* (pp. 54–70). Cambridge: Cambridge University Press.

Fruehwald, J., Gress-Wright, J., & Wallenberg, J. C. (2013). Phonological rule change: The constant rate effect. In S. Kan, C. Moore-Cantwell, & R. Staubs (Eds.), *NELS 40: Proceedings of the 40th Annual Meeting of the North East Linguistic Society* (Vol. 1, pp. 219–240). GLSA Publications. Retrieved from https://www.research.ed.ac.uk/portal/files/14416788/Fruewald_Gress_Wright_Wallenberg_Phonological_Rule_Change.pdf

Kallel, A. (2007). The loss of negative concord in Standard English: Internal factors. *Language Variation and Change*, *19*(1), 27–49. doi:10.1017/S0954394507070019

Kauhanen, H. & Walkden, G. (2018). Deriving the Constant Rate Effect. *Natural Language & Linguistic Theory*, *36*(2), 483–521. doi:10.1007/s11049-017-9380-1

Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, *1*(3), 199–244. doi:10.1017/S0954394500000168

Nevalainen, T. (2015). Descriptive adequacy of the s-curve model in diachronic studies of language change. In C. Sanchez-Stockhammer (Ed.), *Can we predict linguistic change?* (Vol. 16). Studies in Variation, Contacts and Change in English. Helsinki: VARIENG. Retrieved from http://www.helsinki.fi/varieng/series/volumes/16/nevalainen/

Paolillo, J. C. (2011). Independence claims in linguistics. *Language Variation and Change*, *23*, 257–274.

Pintzuk, S. (1995). Variation and change in Old English clause structure. *Language Variation and Change*, *7*(2), 229–260. doi:10.1017/S0954394500001009

Santorini, B. (1993). The rate of phrase structure change in the history of Yiddish. *Language Variation and Change*, *5*(3), 257–283. doi:10.1017/S0954394500001502

Wallage, P. (2008). Jespersen's Cycle in Middle English: Parametric variation and grammatical competition. *Lingua*, *118*(5), 643–674. doi:10.1016/j.lingua.2007.09.001

Wallage, P. (2013). Functional differentiation and grammatical competition in the English Jespersen Cycle. *Journal of Historical Syntax*, *2*(1), 1–25. doi:10.18148/hs/2013.v2i1.4

Zimmermann, R. (2017). *Formal and quantitative approaches to the study of syntactic change: Three case studies from the history of English* (Doctoral dissertation, University of Geneva). Retrieved from http://www.old-engli.sh/myres%20documents/Zimmermann_R.(2017)PhDDissertation.pdf