German morphosyntactic change is consistent with an optimal encoding hypothesis

David M. Howcroft  
*Universität des Saarlandes*

Cynthia A. Johnson  
*Universiteit Gent*

Rory Turnbull  
*École Normale Supérieure*

cynthia.amy.johnson@ugent.be  
rory.turnbull@ens.fr

howcroft@coli.uni-saarland.de

The assumption that language approximates an optimal code for human communication has implications for diachronic linguistics, predicting that the disruptions due to language change result in pressure to re-optimize the code. We examine these implications in the context of morphosyntactic change between Middle High German (MHG) and New High German (NHG). The adjectival paradigm of MHG was bipartite (strong, weak) but developed into a tripartite system in NHG (strong, weak, mixed). The choice of adjective form depends on the morphosyntactic context; e.g., the weak form is used with the definite article, but after an indefinite article the strong form is used in MHG and the mixed form in NHG. Introducing a new paradigm looks like an increase in systemic complexity, but the novel distinctions re-optimize the code such that overall complexity need not increase in successive stages of the language.

We examine the relative entropy of the adjectival systems in MHG and NHG, using the predictability of adjective forms given their preceding context and the predictability of the context forms given the adjective form. These measures allow us to examine the balance of information within the noun phrase before and after the shift to the tripartite system, showing how the distribution of information across the words of the noun phrase changes while situating our findings in terms of the optimality of the resulting code.

Our results provide information theoretic support for claiming that strong forms ‘compensate’ for the lack of information from articles (synchronously; cf. Durrell 2002) and, furthermore, suggest that the morphosyntactic changes observed between MHG and NHG, far from being a process of increasing complexity, are consistent with an optimal encoding hypothesis.

**References:**  