

Identifying local grammars in a corpus of female Irish migrant correspondence

The post-famine period from the 1850s to the 1920s was a time that saw a significant increase in female migration from Ireland to North America. A small glimpse into the lives of these women – their preoccupations, feelings, perceptions and beliefs – can be found in the letters they wrote home to their families. This essay uses traditional historical sciences methods and digital humanities to explore letter-writing practices within the context of nineteenth-century mass migration to America. Specifically, it uses a mixed methods approach to analyse the Lough family correspondence, extracted from a larger collection of Irish migrant correspondence collated by Professor Kerby Miller (University of Missouri) in the 1970s. The Lough collection contains 99 letters written by four sisters (Elizabeth, Alice, Annie and Julia), who left Ireland for America in the 1870s and 1880s. The essay begins with a qualitative study of the letters to manually identify topics and emotions within the content. These topics and emotions are then annotated, extracted and analysed using corpus and computational methods (including *Wmatrix* (Rayson, 2009) and *Sketch Engine* (Kilgarriff and Kosem, 2012)) to identify local grammars (specific words, phrases and structures in the language) for each of the twenty-four topics. Four topics, in particular, will be examined in detail: Ireland/America (*‘I think you two ought to be very comfortable there and us here working hard’*); reunion (*‘I hope I shall meet you once more in life and have a happy time again’*); homesickness (*‘I was heartbroken the other night I dreamed you was dead and I could not see you and you never left any message for me so I woke up crying’*); and recollections (*‘I have thought of Mother very much all through May I remember the prayers we used to say during May’*). Finally, the essay will draw on these local grammars to explore the possibility of semi-automating the process of topic identification in a reference corpus of nineteenth century migrant correspondence (around 500 letters by male and female authors from a range of socio-economic backgrounds) to examine whether the language used by the Lough sisters to express, for instance, feelings of homesickness, is typical – or not – of other migrant letter writers from the same period.

Kilgarriff, A., & Kosem, I. (2012). Corpus Tools for Lexicographers. In S. Granger, & M. Paquot (Eds.), *Electronic Lexicography* (31–56). New York, NY: Oxford University Press. Available from: <http://www.sketchengine.co.uk>.

Rayson, P. (2009). *Wmatrix*. Lancaster University. Available from: <http://ucrel.lancs.ac.uk/wmatrix/>.