

A statistical analysis of the effects of censorship on the letters of Australian soldiers in World War I

Ashley Dennis-Henderson, Matthew Roughan, Jonathan Tuke

During World War I, letters were an important way for Australian soldiers to keep in touch with those back home. Before the war, Australian literacy levels had increased due to compulsory schooling, meaning that large numbers of letters were sent home daily (Longo 2018: 4-5). In early 1917, estimates suggest an average of 558,000 letters sent by Australian's overseas were processed daily (Longo 2018: 8). However, Australian mail was censored (War Precautions Act 1914). Censorship aimed to prevent soldiers sending information which, if intercepted by the enemy, could derail the war effort, such as troop numbers, locations, and movements. Furthermore, soldiers could not include information which could upset or alarm people back home. Many soldiers also practiced self-censorship to avoid worrying those back home (Longo 2018: 12-14).

Some historians, such as Jean Norton Cru and Frédéric Rousseau, believe that due to censorship, war letters do not contain useful information (Hanna 2003: 1340). However, other historians, such as Martha Hanna, Martyn Lyons, Alistair Thomson, and David Taylor, have argued that letters can provide useful insights into soldiers' experiences (Hanna 2014, 2, 6-7, 10; Hanna 2003 1340-1341; Lyons 2008, 189-191; Thomson 2006, 6, 13-14; Taylor 2013, 70). It is noted that many soldiers attempted to circumvent military censorship whilst others wrote frank letters home hoping it would not be censored. While it is unclear how many letters were officially censored, a 1917 report estimated that only 5% of Australian mail being sent home went through official censorship (Cain, 1983: 109).

This paper aims to investigate how censorship affected the contents of letters sent home by Australian soldiers in World War I through a statistical analysis comparing war letters to war diaries. Many historians consider diaries as a more personal form of writing, allowing more opinions and emotions. Furthermore, unless sent posted, diaries did not go through official censorship, and hence could contain more military information. This belief is not completely accurate as some diarists did not include personal thoughts and opinions, some sent diaries home, and some wrote their diary for others to read. However, comparing letters to diaries can still provide useful insights into the overall effects of censorship. The letters and diaries used in this study come from the State Library of New South Wales.

The first element of these documents that is considered are the locations mentioned within them, as official censorship aimed to exclude military information such as locations. Therefore, we investigate the effects of censorship by determining if there is a statistically significant difference in the number of locations mentioned within letters and diaries. To do this, locations are extracted using named entity recognition (NER) and a Chi-squared test is performed. Another important consideration is whether there is a significant difference in the locations that were mentioned. This is determined by comparing heat maps of the locations mentioned in letters and diaries.

The second element of the letters and diaries that is considered is the sentiment of the text. It could be expected that diaries would have a more negative overall sentiment than letters. This study uses a dictionary-based method to calculate sentiment, using the SentiWordNet dictionary (Reagan et al. 2017: 5). A t-test is then performed to determine if there is a statistically significant difference in the true average sentiment of all Australian World War I letters and diaries.

Preliminary analysis of these methods has been performed on a set of 70 diaries and 774 letters. Initially, the 100 most frequent words in the letters and diaries were considered. It is noted that diaries more frequently use words related to fighting and troop movements, whereas letters use more words related to family and correspondence. A Chi-squared test of the number of location words showed with 95% confidence that the number of location words in the text is dependent on whether it is a letter or a diary, with letters containing more location words. Whilst this is somewhat surprising given censorship restrictions, when considering which locations are mentioned, it is noted that letters contain more locations in Australian and the United Kingdom, whilst diaries contain more locations in France and the Middle East. This could be due to letters including the receivers address, but it also suggests that soldiers wrote more about their overseas service in their diaries. A t-test shows with 95% confidence that the true average sentiment of these letters and diaries is not equal. However, it suggests that the difference is between 0.029 and 0.035, which is very small.

References

- Cain, Frank (1983): *The Origins of Political Surveillance in Australia*. Sydney: Angus & Robertson.
- Hanna, Martha (2003): "A Republic of Letters: The Epistolary Tradition in France during World War I," *The American Historical Review* 108, no. 5: 1338-61, <https://doi.org/10.1086/529969>.
- Hanna, Martha (2014): "War Letters: Communication between Front and Home Front," *International Encyclopedia of the First World War*, 2014, https://encyclopedia.1914-1918-online.net/article/war_letters_communication_between_front_and_home_front.
- Longo, Don (2018): "Yorke Peninsula, the Great War and Soldiers' Voices," in *Pens and Bayonets: Letters from the Front by soldiers of Yorke Peninsula, South Australia, during the Great War*. Adelaide, Wakefield Press: 1-31.
- Lyons, Martyn (2008): *Reading Culture and Writing Practices in Nineteenth-Century France*. Toronto: University of Toronto Press. <https://ebookcentral.proquest.com/lib/adelaide/detail.action?docID=4672659>.
- Reagan, Andrew et al. (2017): "Sentiment analysis methods for understanding large-scale texts: a case for using continuum-scored words and word shift graphs," *EPI Data Science* 6, no. 28, <https://doi.org/10.1140/epids/s13688-017-0121-9>.
- Taylor, David (2013): *Memory, Narrative and the Great War: Rifleman Patrick MacGill and the Construction of Wartime Experience*. Liverpool: Liverpool University Press.
- Thomson, Alistair (2006): "Anzac Stories: Using Personal Testimony in War History," *War & Society* 25, no. 2: 1-21, <https://www.tandfonline.com/doi/abs/10.1179/072924706791601900>.
- War Precautions Act 1914, <https://www.legislation.gov.au/Details/C1914A00010>.
- Picture Credit**
- Marshall, Henry: "Kensington to Cairo and from Cairo to Gallipoli: album of photographs, 1914-1915," *State Library of New South Wales collection of World War I pictorial material*, PXA 1861, <https://collection.sl.nsw.gov.au/record/Ydmdmgd9>.