Testing Canada's 'honour': Does Orthography Index Ideology?

Dustin Grue*

Recent studies in orthographic variation have assumed identity-driven motivations for spelling choice (Lipski 1975, Schieffelin & Doucet 1994, Sebba 2000), linking this motivation to national ideological positions. In the Canadian context, Heffernan et al. (2010) propose a method for representing the connection between national sentiment and orthography using quantitative measures and, using data extracted from the University of Alberta's student newspaper *The Gateway,* demonstrate a strong quantitative correlation between anti-Americanism and a decrease in the use of American spelling variants. This paper tests the ideology/orthography connection using Heffernan et al.'s method on data from the University of British Columbia's student newspaper *The Ubyssey* and finds an insignificant connection between ideology and orthography. However, this correlation appears to be indicated differently across different article genres. Observations are made on the methodological difficulty of establishing the orthography/ideology connection.

1 Introduction

In the latter part of the twentieth century, sociolinguistic conceptions of identity and ideology began to inform theories of orthography. For example, Lipski (1975) provides an overview of orthographic variation in Canada and argues that selecting between British and American spellings is a vehicle for linguistic nationalism. In his conclusion, Lipski summarizes: "That orthographic variation is used as an instrument of linguistic nationalism in Canada appears certain in many cases, although to what degree it is impossible to specify" (p. 46). Studies following Lipski (Schieffelin & Doucet 1994, Sebba 2000) would re-assert this connection, but this link remained qualified, its quantitative 'degree' unestablished. Heffernan, Borden, Earth and Yang (2010) (hereafter "Heffernan et al.") essentially take up Lipski's challenge, investigating nationalism and orthography in the Canadian context by proposing a method for representing the connection between spelling choice and national sentiment using quantitative measures. Specifically, they demonstrate a strong negative correlation between anti-Americanism expressed by Canadians and usage of American spelling variants, as opposed to British. Their paper argues that such short-term diachronic change is motivated by periods of increased anti-American sentiment, provoked mainly by the involvement of the United States in unpopular wars. That is, they hypothesize, for example, that during the Korean War in the early 1950s and the first Gulf War in the early 1990s, certain British spelling variants—such as *honour*—increased relative to their American counterparts, e.g. *honor*. Heffernan et al.'s data are extracted from The University of Alberta's student newspaper *The Gateway* and cover the years 1910 to 2006. The authors gauge anti-American sentiment from the text of this newspaper, establishing an 'Anti-American Index' by assigning scores from a 7-point Likert scale. The historical frequencies of these two measures—American orthographic variant count and Anti-American Index—are shown to be inversely correlated, with Pearson-r -0.715, p 0.001.

This paper tests the quantitative ideology/orthography connection and the claims of Heffernan et al. using similar methods on a different set of data from a different region of Canada. These data are taken from *The Ubyssey*, The University of British Columbia's student newspaper, published between 1920 and 2000 and consisting of an estimated 50 million words. Historically, British Columbia's orthography tends towards the British (along with Newfoundland and Ontario), while

* This paper is the revised product of research performed for Dr. Stefan Dollinger's graduate seminar, *Language and Identity: Canadian Constructions*, at the University of British Columbia. I would like to extend my thanks to Dr. Dollinger for his support and guidance of this work and also to the two anonymous reviewers and editors of *Strathy Student Working Papers on Canadian English* for their excellent comments.

Alberta's tends towards the American (with Manitoba and Saskatchewan) (Brinton and Fee 2001: 433); however, as will be shown, representing these tendencies is less clear than this common simplification.

Several improvements are offered to augment Heffernan et al.'s method. Establishing a correlation between orthography and ideology requires the isolation of multiple variables influencing linguistic motivations, whether these are language-internal (phrases preferring one variant over another) or language-external (variants preferred by certain geographic locations, etc.). The former has not yet been thoroughly investigated, but this paper finds that the latter—ideologically independent language-external motivations—is a factor in spelling choice. Further, although objective data are analyzed in this work, it must be recognized that the resulting statistics do not simply speak for themselves. Different statistics make different assumptions and tell different stories; that is, statistics have a rhetorical component. This paper identifies the assumptions and motivations behind varying representations and explains how these affect quantitative studies of identity. Finally, a practical, methodological modification is offered. In this study a 'tag' has been assigned to each article for which anti-American sentiment was evaluated to describe its 'type'. This measure recognizes that newspaper articles are comprised of many voices—through quotation, indirect reports, etc.—and that evaluating correlations by article-type can help strengthen the connection between speaker attitude and ideology by further classifying just who this speaker is. In addition, by adding to the method in this way, the method becomes more portable: we might more easily compare ideological position and orthographic change across different datasets.

The remainder of Section 1 provides background on the variation between British and American spellings in Canada, their history and origins, theories on orthographic variation and identity and a summary of Heffernan et al.'s work. Section 2 outlines the methodology for this paper. Section 3 explains the results of this study and contrasts these with Heffernan et al.'s findings, and Section 4 offers perspective on this in the context of Canadian English. Section 5 contains the conclusion.

1.1 Variation in Canada

> *[…] there are hundreds of words that have variant spellings in different parts of the English-speaking world, the principal cleavage being between the United Kingdom and the United States. Partly as a result of our historical links with Britain and our proximity to the United States, Canadian English has tended to waver between the forms used in these two countries, so that, to this day, there is no clearly established Canadian standard.* (Translation Bureau of Public Works and Government Services Canada 2012)

The above text comes from the introduction to the 'Spelling' section of the *Translation Bureau* of *Public Works and Government Services Canada's* website, providing a succinct, simplified summary of present-day orthographic variation in Canadian English. The verb "cleave" is apt, as the section following the introduction splits the page into two columns, listing British variants on the left and American variants on the right. The site lists seven orthographic variables, including a rule for digraphs: <ise>/<ize> (*civilize/civilize*), <re>/<er> (*fibre/fiber*), intermedial and terminal <l>/<ll> (*instil/instill*), <ce></se> (*defence/defense*), past tense <l>/<ll> (*traveled/travelled*) and <ae>/<e> (*encyclopaedia/encyclopedia*). The purpose of the site is functional and indeed presents a rather simplified guide to variation with a focus on "avoiding misspellings": not only is a highly limited selection of variables presented (which the site acknowledges), but variation is indicated as being substantially more clear-cut than in reality. For example, the site identifies verbs with single-l "including their derivatives" as British, but the participle of both *instil/instill* is, of course, *instilling*. However, this site ought not to be faulted for its simplification. Again, the scope is broad and the purpose is functional. Further, certain complications arise in the very study of orthographic variation—first, by a limitation in reliable data. Natural language in quantities large enough to be meaningful, from traditional sources like newspapers and other published media, has often gone through editorial revision, removing the 'natural' element to the selection of variants. (Data from social media, e.g. *Facebook* and *Twitter*, may prove to ameliorate this limitation.)

Moreover, what counts as orthographic variation? Every so often I see a word like *honourarium*, for example, which is a misspelling by most any convention and has little history as an orthographic variant—but what to do, here? Is this truly misspelling, or spelling variation? More interestingly, do forms like this arise only because of awareness of variation in the first place—an over-application of the 'rules' created by the American reaction to British forms? This is what Lipski (1975) terms "hypercorrection": the imposition or over-application of an orthographic rule. Such are the complexities of investigating variation, common to studies of linguistic innovation, generally.

The first major study of Canadian orthography, Robert Ireland's *The Survey of Canadian Spelling* (1979, 1980a), overcame these challenges with mailed questionnaires and revealed the complex, contoured landscape of spelling variation in Canadian English. This survey project collected 3235 returns from students 15-17 years old, on 70 items in 12 categories of words. These results were reported in Ireland's (1979) dissertation, a condensed, article-length publication of his main findings (1980a), and informed his *Spelling Across the Curriculum: A Student Handbook* (1980b). In general, these findings demonstrate profound inter-provincial variability, show that certain provinces prefer a particular national orthography or mix of orthographies and therefore indicate that how one spells is at least partially dependent on the geographic area in which one was educated: for example, 87% of respondents from Ontario indicated *colour* as the correct spelling, while only 26% from Alberta selected this variant (Ireland 1980a: 66). Given these large inter-provincial differences, Ireland regards the idea of Canadian spelling as the mixture of orthographies, "a predictable pattern of American and British forms" (1980a: 80). What characterizes Canadian orthography (or orthographies), then, might not simply be a particular position between British and American 'poles' but the configuration of these variants into unique forms—and perhaps, as this paper investigates, used for particular purposes of identity. Indeed, if any generalization can be made about spelling differences between provinces, it is that the prairie provinces tend to adopt American variants while the east and west retain the British, according generally with European settlement patterns where 'old' provinces adopt the British and are less resistant to spelling change.[1] Brinton and Fee summarize Canadian spelling variability succinctly: "[s]pelling varies from province to province (with British Columbia, Newfoundland, and Ontario trending toward British spellings and Alberta, Manitoba, and Saskatchewan toward American ones)" (2001: 433).

T.K. Pratt (1993) calls such variability the "hobgoblin" of Canadian spelling and glosses the institutional motivations for such variability—the issue, as Pratt writes, of "the vexed question of standard English and its authorities" (p. 45). Pratt illustrates the often-conflicting demands imposed by different authorities—prescriptions in style guides, Canadian dictionaries and the pressures on editorial decisions—as well as the interesting story about how the Toronto newspaper the *Globe and Mail*, ostensibly national in its reach and scope, used American spellings for most of its history until abruptly shifting to the British in 1990.

However, these questions of variability only elicit—but do not answer—the most profound question of orthographic variation: if variation exists, how do speakers choose? If choice in spelling is available, why is one variant selected over the other? Which ones stay and which ones go away? As Pratt (1993) maintains, institutional factors by way of prescriptive/proscriptive rules are certainly one part of the answer but not the whole story. Ireland (1980a) alludes to socio-geographic conditions for orthographic choice: though respondents from the Atlantic provinces of Prince Edward Island and Newfoundland favoured <or> variants in general, the majority selected *harbour* as the correct spelling. The close proximity of these eastern provinces to the ocean and the trade with the rest of Canada are likely motivating factors—though the case is far from clear.

1.2 Noah Webster and North American orthographic variation

The origin of American variants and the resulting variability in North American orthographies can be largely attributed to dictionary author Noah Webster (Lipski 1975, Brinton and Fee 2001), whose 1828 *An American Dictionary of the English Language: The Origin, Orthography,*

[1] Newfoundland, of course, is technically the least 'old' province—joining the Confederation in 1949—but 'young,' here refers to European colonization.

Pronunciation, and Definitions of Words proscribes against British spellings and prescribes modern American variants, ostensibly for the purpose of providing more logical variants analogous to the orthographies of other words.[2] As Webster writes, "These words are here restored to their true analogous spelling, as recommended by Walker, Lowth, Perry, and others" (1848: n.p.). Here, Webster makes a case for "*the tendency of our language to greater simplicity*" (p. vi, emphasis in original), advocating for a consistent, generalizable system of orthographic rules—what Webster calls "completing analogies". For example, intermedial and terminal double-l <ll> are imposed for nouns such as *enroll* "to prevent a false pronunciation" (p. xxiii) and for consistency in their derivatives (e.g. *enrollment* and the participle *enrolling*). (Webster has no rule, however, standardizing the double versus single <l> between words like *skillful* and *skillfully*/*skillfulness*, etc.)

Webster prescribes seventeen orthographic rules to effect his linguistic vision, most of which are easily recognizable as contemporary salient British/American differences. For example:

1. "Terminations in *our* changed into *or*"
2. "Terminations of *re* changed into *er*"
3. "Distinctions between Verbs in *ize* and *ise*"
   (p. xxiii)

Interestingly, in addition to promoting orthographic consistency, preserving standard pronunciations and guiding the "simplification" of language, Webster also experimented with bringing orthographies into line with their etymological (and consequently national) origins—such examples are "*bridegoom* for *bridegroom*" and "*fether* for *feather*", which "the German critics highly applauded" (p. vii). To this end, while one linguistic national tradition was rejected another was embraced—a glimpse into the tug of war of nationalism embroiled within the ideology of orthography. However, this experiment in etymological alignment was largely a failure: as the editor of the 1848 edition explains, "however desirable these changes may be in themselves considered, as they do not relate to the general analogies of the language, and can not be duly appreciated by the body of the people, they will never be generally received" (p. vii).

Indeed, this experiment was likewise a long process. Webster's first foray into American nationalist linguistics occurred decades before his orthographic reforms in his dictionaries. In 1783, Webster began publishing a series of books initially titled *A Grammatical Institute of the English Language*, of which his famed blue-backed speller was a part. His goal with the publication of these books was, as Warfel (1966) writes, "to develop a simple system of elementary education by which the speech and language of the United States would be rendered uniform, moral and religious truths would be propagated, and a love of country would be developed" (p. 60). But however radical these books may have been pedagogically, his experiments with language and nationalism continued for some time, and his orthographic reform did not come until years later with the publication of his dictionaries.

To say, then, that Webster's changes were not entirely consistent would be too superficial—more accurate to say that they were designed, over the span of his career, to be more rhetorical and more publicly palatable than strictly logical. Part of this palatability had to do with the consequences of changing historical associations, or national nuances carried with changing orthographies. For example, Webster considered both <or> and <our> forms of *Saviour* acceptable. This was because the term linked a religious heritage that Webster would not contravene: "The *u* will be speedily omitted in all words of this class, unless, from the sacredness of its associations, it be retained in *Saviour*" (p. vi, emphasis in original). Orthography, therefore, certainly exhibits extra-linguistic sentiment linked to tradition. The question is how speakers might employ this sentiment, this "sacredness of associations", to construct identities.

Webster's act of simplifying the language resonated with, and was adopted by, certain populations in North America (or was at least enforced through institutions such as school, government, etc.) until the orthography in the United States was distinct from that realized in Canada—where British variants remain common. And though Webster's work was a significant

[2] This is a simplification. A search in *Google Book's Ngram Viewer*, which represents data from 1800 to 2000, shows that *color* certainly did exist in North American writing before Webster's 1828 dictionary.

factor spurring innovation, these prescribed forms were not the definitive determining factor in fixing orthography. First, as Robert Ireland (1980a) reports of Donald Emery's (1974) work, variation does exist in American dictionaries: the hegemony of Webster's work is nowhere near total. As well, Canada did not react to this linguistic 'threat' passively. The diffusion of perceived American English into Canada was met with strong resistance from Britain: according to J.K. Chambers, a "massive migration" took place in the early to mid-1800s "intended by the British governors of Canada as a kind of counter-Loyalist wave" (2004: 226). The effect of the migration, claims Chambers, was a new "heady anti-American spirit that became a marker of Canadian imperialism for the next century" (p. 227), creating "the Canadian double standard in many matters of spelling and pronunciation" (p. 230). Canada, caught between two linguistic superpowers, developed non-standard orthographies as a result.

1.3 Ideology and spelling

Thus far, rule-based and socio-geographic explanations for spelling variation have been largely piecemeal and incidental and fail to explain orthographic variation in terms of a system or theory. Recently, the basis of ideology, and the concomitant link between orthographic choice and national identity, has been proposed to account for such selection. Sebba (2000), for example, argues "that orthographies are shaped less by the phonological facts of the language concerned than by social and cultural factors in the context where the orthography is used" (p. 926). Indeed, only on the surface did Webster attempt to shape the phonological (and orthographic) facts to which Americans would adhere. Rather, he tried to do the reverse—sculpt the language to shape culture (respecting that the actual process is quite less linear than this). Sebba's view assumes that orthography is not a passive reflection of a linguistic system but that "orthography itself is a social practice" (p. 926). Furthermore, as social practice, orthographies can take shape to index one's position within that practice and come to be associated with a certain prestige or disposition. As John Lipski (1975) writes, specifically regarding orthography and ideology, "a speaker may choose to employ his [sic] language as a vehicle of nationalistic sentiments" (p. 37). Heffernan et al. advance these claims, giving qualitative and quantitative evidence to Lipski's (1975) assertion in particular, by operationalizing this theory in a historical study and establishing the correlation between anti-American sentiment and increased use of British spelling variants.

Heffernan et al. test two hypotheses regarding orthography and ideology: first, that the relative frequency of British variants increases in times of increased anti-American sentiment, in turn motivated by the United States' involvement in unpopular wars; second, that such short-term change eventually results in long-term change—in other words, Canadians would increasingly adopt British spelling variants as the norm. Heffernan et al.'s findings support the first hypothesis but not the second: they find a striking, statistically significant negative correlation between American variant usage and anti-Americanism (Pearson-r -0.715, p 0.001), but the split between American/British variant use is essentially the same at the beginning of their data (the year 1920) as at the end (the year 2006). Heffernan et al. conclude their paper by speculating on the degree to which orthography is used, as they say, "strategically" (p. 19) and question whether similar patterns will be found in other regions of Canada: "One outstanding question is the degree to which these results can be extended beyond our source population, that is, university students who contributed articles to *The Gateway*" (p. 18). This paper responds to this question by analyzing data from The University of British Columbia's *The Ubyssey*. The question is interesting, not the least because, as Ireland (1979, 1980a) found, Alberta—the geographic source of Heffernan et al's data—manifests very different orthographic choices than does British Columbia, the source of the present study.

1.4 Heffernan et al. (2010)

To probe the link between orthography and ideology, Heffernan et al. engage in a two-part process: 1) they plot the aggregate American/British frequencies of 15 variables across 96 years of newspaper data, and 2) they determine the correlation of these aggregate frequencies with an index of anti-American sentiment. Table 1 lists their tested variables as well as the frequencies of their respective variants in *The Gateway*.

American Variants			British Variants		
	#	%		#	%
honor	1961	68%	*honour*	938	32%
neighbor	366	53%	*neighbour*	327	47%
humor	1244	47%	*humour*	1426	53%
gray	864	33%	*grey*	1739	67%
program	6266	92%	*programme*	518	8%
defense	2597	42%	*defence*	3539	58%
offense	898	40%	*offence*	1375	60%
enrollment	890	48%	*enrolment*	982	52%
jewelry	431	69%	*jewellery*	192	31%
marvelous	353	50%	*marvellous*	355	50%
centered	408	54%	*centred*	348	46%
kilometers	73	41%	*kilometres*	104	59%
judgment	931	67%	*judgement*	456	33%
labeled	129	22%	*labelled*	467	78%
fulfill	572	77%	*fulfil*	168	23%
Totals	**17983**	**58%**		**12934**	**42%**

Table 1. Variant count and relative frequencies in *The Gateway* (Heffernan et al. 2010: 8).

What is immediately striking from Heffernan et al.'s selection of variables is that <color>/<colour> is absent. On the one hand this makes sense, since *colour* has an odd history,[3] but on the other this decision itself is odd—especially in light of their own recognition of this variable's saliency: "In contrast to words such as *fulfill*, a very small set of words, particularly *colour* and *center*, seem to do more than their share of ideological work" (p. 18). However, while Heffernan et al. do not elaborate specifically on all of their individual variable selections, they do specify criteria for the inclusion of variables, including: 1) a minimum frequency of 500 tokens, 2) non-categorical preference in the dataset of one variant over another and 3) that modern American English shows preference for one variant over the other, which they confirm using counts from the *Corpus of Contemporary American English* (Davies 2008). Exceptions were made for <kilometers>/<kilometres> and <program>/<programme>—defying rules 1 and 2, respectively—to test their hypotheses with variants of low frequency and categorical preference.

In addition to the three criteria discussed, a tacit fourth requirement also exists—one that applies to the dataset as a whole. This requirement is that the data reflect ideology 'purely', in that spelling ought to be free from editorial oversight since prescriptive rules might skew variant selection one way or the other: this is to say, spelling ought to be allowed to display variation 'naturally'. As Heffernan et al. write, this is a very difficult thing to prove, though they do show editorial independence by finding variation within single articles: as they report, "both spelling variants of the word *honor* (<honor>/<honour>) occur within the same article 31 times between the years 1914 and 2001" (p. 7).

The second step in Heffernan et al.'s method involves calculating the correlation between variant-use and anti-American sentiment. This itself is a three-step process. First, the authors located articles from each year that contained the tokens *american* or *united states*. Second, the order of these articles was randomized and examined: three articles were selected from each year that made some statement about Americans as a people or the United States as a country. Additionally, selected articles had to come from different months of the year. Third, each author rated each article for pro or anti-American sentiment on a 7-point Likert scale, where 4 was a neutral rating (no perceived sentiment), and scores above and below 4 indicated anti-American

[3] For the sake of brevity I cannot enter into a full explanation, but the historical frequency of *colour* is unusual because, in *The Ubyssey* data, it increases dramatically (both in token count and count relative to *color*) in the 1990s. A brief investigation of this trend—which defies almost all other variable trends—shows that *colour* increased in usage in 'racialized' phrases, by which I mean such phrases as *person of colour*, *student of colour*, *woman of colour*, etc. There is also a strong economic correlation, in that *colour* increases in usage due to advertisements for *colour televisions*, *colour printers*, *colour monitors*, etc.

and pro-American sentiment, respectively. To determine correlation, scores were averaged per five-year period and compared with average percent American spelling variant usage during the same periods. To determine the correlation, Pearson's *r* was used, which measures the dependence between two variables—in this case, anti-American sentiment and American variant use. The negative correlation Heffernan et al. find is quite high: a Pearson-r value of -.715 *p* 0.001, indicating that when anti-American sentiment is high, American variant usage is low. However, since the relative use of American variants is the same at the end of their period of data as at the beginning, they fail to confirm their second hypothesis—that such short-term diachronic changes (i.e. a decline followed by a recovery) would eventually motivate long-term change (overall decline.)

2 Methodology

My method follows Heffernan et al.'s, with some modifications. First, I checked for variation in *The Ubyssey* to verify that editorial orthographic oversight is likely minimal. Indeed, 322 British variants co-occur in a short proximity with American variants, fairly evenly across all years of publication. In fact, in an article from 1984 on that year's Progressive Conservative national political campaign, party leader Brian Mulroney is reported as 'saying' *honour* and then *honor* in the same sentence. Figure 1 reproduces this excerpt.

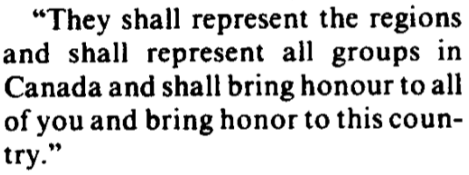

Figure 1: 'honour' and 'honor' in *The Ubyssey* ("PC Leader Promises New Prosperity", July 31, 1984: 8).

Second, using the same set of spelling variables as Heffernan et al. (Table 1), I plotted percent American usage per five-year period in *The Ubyssey*, from 1920 to 2000. This was a straightforward task, though there are discrepancies in the methods of aggregation Heffernan et al. use, which could be problematic for different data. There are two basic methods for aggregating data: all of the counts of each variant per five-year period can be summed together (i.e. an aggregate of all American variants versus British variants, regardless of the specific variable), or the proportion of each American/British variant for each variable can be calculated, and then these proportions can be averaged together. Table 1 uses the former method, but the calculations Heffernan et al. use for historical correlation employ the latter: Heffernan et al. write, "[p]ercent American variant was then calculated for each word by half-decade and averaged across the fifteen words" (p. 10). It is hard to tell which method is better; each carries different assumptions. Using an aggregate count of all variants makes the assumption that ideological weight is also approximated by raw usage—though results using this method can be heavily skewed with high frequency counts. On the other hand, taking the average of the proportion of variables assumes that every variable exhibits the same ideological weight, such that a 5:1 American/British variant ratio of a highly salient item would average out with a 1:5 American/British variant ratio of a less salient item—this might be complicated with a variable such as *gray*/*grey*, where variation is restricted (and therefore ideological weight is likely very low). Confounding the matter further is the unknown degree to which frequency and saliency are also correlated. Given that the 'best' way to represent this figure is likely not a 'one-or-the-other' type answer but a mixture of both—and also likely different for each variable—I will follow Heffernan et al. and calculate historical frequencies using averages.

Further innovation included performing a simple collocation analysis on the synchronic data, identifying 1L (one token left) and 1R (one token right) collocates of each variant. Because this research depends on language-external motivations for variant choice—national sentiment—language-internal motivations ought to be considered, including variants occurring in fixed

phrases. That is, perhaps variants are lexically or simply conventionally motivated: they are assumed to mean differently or historically have developed to form language formulas. Ideally, more sophisticated methods would be used to account for this, but a simple collocation search of the synchronic data might identify abnormalities.

Following the three steps for assigning an anti-American rating to each article, I 1) extracted the 3985 issues containing *american* or *united states* from the total 4792 issues, 2) selected 276 for rating and 3) rated each selection. The only problem I encountered here was not being able to find any articles referring to Americans from the years 1918 or 1919—consequently, the five-year period 1915-1920 is excluded from this study. The averaged index rating I obtained for *The Ubyssey* was 4.82, while Heffernan et al. report a 4.63 average for *The Gateway*. By way of testing this method, I also determined how 'long' it took to perform step 2. That is, I recorded how many issues for each year I went through to find three articles making claims about the United States. I tested this not only because it seemed difficult for some years (particularly the early years) and very easy for others, but also because it provides another dimension to the research: do greater sentiments of anti-Americanism come in a greater amount, or vice-versa? To determine how 'long' it took to perform this step, I calculated the percentage of each year's issues of *The Ubyssey* I evaluated (and passed by) before finding three that fit the above criteria.

My method of rating for historical anti-American sentiment differed from Heffernan et al.'s since I did this alone, whereas each author in Heffernan et al. rated each article independently and then averaged their scores. However, although I rated the articles alone I also rated them twice, with an appreciable amount of time (~8 weeks) between ratings. Since Heffernan et al. report that, "the ratings tended to coincide with each other with very little disagreement (mean intraclass correlation of .967)" (p. 15), I can be somewhat confident in the ability to apply consistent ratings: the correlation between my two sets of ratings is Pearson-r of 0.932.

In addition to rating anti-American sentiment, I assigned to each article an article 'type'—this followed from the observation that student newspapers are heterogeneous assemblages of genres and represent disparate voices performing different actions. For example, a sentiment of anti-Americanism could be expressed in a news article or in a sports article, where, in reporting on sports, nationalism can be rather hyperbolic, essentially the rhetorical 'spice' given to the genre.[4] And though it falls to the researcher to decide where the sentiment lies on the 7-point Likert scale, it remains that different types of articles arrange voices in varying ways and likewise assume different audiences. For example, news articles are almost exclusively an assemblage of voices (interviewed subjects, others' reports, etc.). Articles on special guest speakers at the university (of which there are many) would likewise report another's—this guest's—speech, while an opinion piece more directly frames the sentiment as that of the author. Further, there are often anti-American jokes in *The Ubyssey*, so even if it is decided that, rhetorically, the purpose of the joke is indeed to express negative sentiment, the question remains, 'from where does this sentiment emanate?'. Practically, labeling each article-type for evaluation answers the important question, 'for whom is this author speaking?'—important, since assigning these sentiments is meant to characterize historical periods, and so a post-analysis review of article-types might shed light on the robustness of this characterization. Article-type labels were generally taken from the labeled section in which they were found, though some consistency was enforced by changing labels, recognizing historical variability in the newspaper's form (in other words, in some cases a 'news' article would be more appropriately labeled 'opinion', etc.). These class labels, with counts, are given in Table 2. The labels should be descriptive in themselves, and rationale for assignment is given above. The type 'speaker' is assigned for articles that report the words of a guest speaker at the university or in the community.

[4] This is especially true for *The Ubyssey*, since University of British Columbia sports teams often played cross-border games, especially in the university's early history.

Type	Count
News	76
Opinion	66
Speaker	45
Sports	22
Letter to the editor	19
Total	**228**

Table 2. Article-types and counts in *The Ubyssey* for evaluated articles between 1920 and 2000.

Finally, I considered problems inherent to Optical Character Recognition (OCR) on which technology this research relies. Problems innate to OCR are well known, where the technology can misread a token or interpret one word or words as others. This problem warrants particular attention in historical investigations because the documents under investigation are generated across a span of time (in this case 92 years of *The Ubyssey*): in this light OCR accuracy is problematic not because it misreads terms but because it might misread different terms at different times. This inconsistent inaccuracy is due to historical trends in typesetting, degradation of hard copies, etc., which is problematic because it might not only introduce error but also introduce different errors at different times. However, though early texts might be assumed to attest higher rates of error, I was surprised to find a relative historical homogeneity of error-rates across all years and documents. Therefore, this is not assumed to affect results.

3 Results

3.1 Synchronic data

Table 3 represents synchronic data from this study, as well as from Heffernan et al.'s study.

American Variants					**British Variants**				
	The Gateway		***The Ubyssey***			***The Gateway***		***The Ubyssey***	
	#	**%**	**#**	**%**		**#**	**%**	**#**	**%**
honor	1961	68%	2276	65%	*honour*	938	32%	1251	35%
neighbor	366	53%	238	49%	*neighbour*	327	47%	252	51%
humor	1244	47%	1418	50%	*humour*	1426	53%	1442	50%
gray	864	33%	1555	17%	*grey*	1739	67%	7366	83%
program	6266	92%	24749	90%	*programme*	518	8%	2840	10%
defense	2597	42%	2772	39%	*defence*	3539	58%	4285	61%
offense	898	40%	769	43%	*offence*	1375	60%	1040	57%
enrollment	890	48%	1034	33%	*enrolment*	982	52%	2086	67%
jewelry	431	69%	281	49%	*jewellery*	192	31%	291	51%
marvelous	353	50%	150	39%	*marvellous*	355	50%	230	61%
centered	408	54%	244	42%	*centred*	348	46%	342	58%
kilometers	73	41%	36	19%	*kilometres*	104	59%	155	81%
judgment	931	67%	860	66%	*judgement*	456	33%	451	34%
labeled	129	22%	135	20%	*labelled*	467	78%	551	80%
fulfill	572	77%	650	81%	*fulfil*	168	23%	156	19%
Totals	**17983**	**58%**	**37167**	**62%**		**12934**	**42%**	**22738**	**38%**

Table 3. Variant count and relative frequencies in *The Gateway* (Heffernan et al. 2010: 8) and in *The Ubyssey*.

Table 3, on first glance, defies expectations. Though Ireland (1979, 1980a) and Brinton and Fee (2001) agree that British Columbian spelling differs from Albertan spelling—in that the former manifests a larger percentage of British variants than the latter—in Table 3 we find the reverse, a

higher proportion of American variants in *The Ubyssey* than in *The Gateway*: 62% American variants in *The Ubyssey* compared with 58% in *The Gateway*. Initially, this strikes as especially odd because every American variant ratio is lower in *The Ubyssey* than in *The Gateway* (and considerably so), save for four variables: *program, humour, offence,* and *fulfill*. The culprit turns out to be the variable <program>/<programme>: with an American variant *program* count of 24749, it makes up almost two-thirds of the entire combined American variant count in *The Ubyssey*, heavily skewing the data and highlighting the pitfalls of using raw frequency counts outlined above. Using average frequencies, which Heffernan et al. only use for their diachronic data, re-aligns the proportions: *The Gateway* exhibits 54% American variants and *The Ubyssey* 47%. Indeed, though the average frequency of all American variants except *program* in *The Ubyssey* is 887, the frequency of *program* is 27.9 times this, and it is unclear why. A search for *program*'s collocates in *The Ubyssey* reveals predictable results, since *program* refers, of course, most often to a university or extracurricular program. Examples include *work study program* (60 instances), *women's studies program* (56 instances) and *the athletic program* (50 instances). This does not reveal why *program* would occur more often in *The Ubyssey*, however, and may simply be due to differences in the papers' purposes.

In other respects, the synchronic American and British spelling variant frequencies between newspapers are remarkably similar, in relative and in absolute terms. (Not knowing the total wordcount of *The Gateway* dataset, I am unable to determine if the variables themselves differ in frequency between newspapers.) <or> and <our> variable frequencies are largely similar, with the exception of *honor*, which *The Gateway* prefers. The use of the variants <se>/<ce> and double <ll> are likewise similar between papers. Slight preference is shown in *The Ubyssey* for <re> *centred* and strong preference for *kilometres* over *kilometers*. Both of these differ in proportion between *The Gateway* and *The Ubyssey*: the American variant <er> *kilometers* accounts for 41% of the variable in *The Gateway* but for only 19% in *The Ubyssey*.

Results of the collocation search likewise largely show similarities between the papers but with one key difference—revealing a large concern. In Table 3, we observe a slight difference between *The Gateway* and *The Ubyssey* in the frequency of *grey* versus *gray*. Not only is the proportional usage of *grey* 16% higher in *The Ubyssey* than in *The Gateway*, the raw occurrence is over 4 times higher. The collocation search of synchronic data helps to explain this. The most frequent 1L (left) and 1R (right) collocations for all 15 American and British variants are listed in Table 4.

American Variants			**Canadian Variants**		
Variant	**l-collocate**	**r-collocate**	**Variant**	**l-collocate**	**r-collocate**
honor	the honor	honor of	*honour*	the honour	honour of
neighbor	a neighbor	neighbor to	*neighbour*	your neighbour	neighbor to
humor	of humor	humor and	*humour*	of humour	humour and
gray	stuart gray	gray and	*grey*	point grey	grey.
program	the program	program.	*programme*	the programme	programme of
defense	the defense	defense of	*defence*	the defence	defence of
offense	the offense	offense.	*offence*	the offence	offence.
enrollment	the enrollment	enrollment of	*enrolment*	the enrolment	enrolment in
jewelry	and jewelry	jewelry and	*jewellery*	and jewellery	jewellery.
marvelous	a marvelous	marvelous.	*marvellous*	a marvellous	marvelous in
centered	is centered	centered around	*centred*	is centred	centred on
kilometers	square kilometers	kilometers of	*kilometres*	two kilometres	kilometres of
judgment	the judgment	judgment on	*judgement*	the judgement	judgement of
labeled	be labeled	labeled as	*labelled*	be labelled	labelled as
fulfill	to fulfill	fulfill the	*fulfil*	to fulfil	fulfill the

Table 4. Most frequently occurring 1L and 1R collocates of 15 American and British spelling variants in *The Ubyssey*.

The most frequent 1L collocate of *grey* is *point*, and the geographic area on which the University of British Columbia lies is called "West Point Grey". Of the 7366 instances of *grey* almost 44% of those, 3265, constitute *point grey*. Since *grey* in this case is part of a proper noun, certainly less choice exists for variation (though there is some: 15 occurrences of *point gray* versus 3265 of *point grey* in *The Ubyssey*), which poses a problem for the study. Moreover, we also find high occurrences of proper nouns in the American variant *gray*. As Table 4 shows, the most common 1L collocate of *gray* is *stuart*: Stuart Gray was *The Ubyssey* editor from 1967 to 1968. Indeed, many other proper nouns occur in the most frequent collocates of *gray*: the second highest occurring collocation is *John Gray*—other *Gray*-names with high frequencies include *Stu Gray, Bob Gray, Stephanie Gray, Allan Gray, Herb Gray, Jerry Gray, Bruce Gray, Jack Gray, Miss Gray, Wally Gray, Mabel Gray, James Gray,* etc. The sheer abundance of proper nouns with *gray* or *grey* as a component part suggests that it should be dropped from the data. Doing so to *The Ubyssey* changes the proportion of its American variants to 46%. Changes in results from *The Gateway* are unclear. To my knowledge, there is no "Point Grey" in Edmonton, Alberta (the University of Alberta's location), though I am certain that there are *Gray* surnames. This is a concern that should be resolved in future work.

For the most part, collocates of each remaining American and British variant are the same: for example, the most frequent 1L and 1R collocates of both *honor* and *honour* are *the* and *of*, respectively. Even where they do differ, they often take the same form, as in noun + prep. for *judgment on* and *judgement of*, or realize collocates similar in sense as in *centered around* and *centred on*. Indeed, it would be wise to perform this collocation check to ensure that candidate variants are realized outside of restricted formulas: a collocation search can be performed and the variable (or a portion of these counts) discarded if variants are found to largely comprise longer fixed phrases or prefer particular lexical items.[5]

Finally, in the synchronic data represented in Table 3 we see variation in preference for variants between datasets. Not every American variant in *The Ubyssey* is selected 47% of the time, of course: a range is exhibited (from a low of 17% for *gray* to a high of 90% for *program*) and this variation is higher in *The Ubyssey* than in *The Gateway*. The means of the frequencies are given above, and the standard deviations of variant frequency for *The Gateway* and for *The Ubyssey* are 18% and 22%, respectively. Perhaps ironically, this does not suggest that there is greater variation in *The Ubyssey*, but less. Greater deviation suggests that a stronger categorical preference is expressed for each variable.

3.2 Diachronic data

Figures 2 and 3 depict historical aggregate American variant frequencies for *The Ubyssey* and for *The Gateway*, respectively. Without exact values it is difficult to compare figures rigorously, but the trends depicted in each figure appear similar and mostly in the second part of the century: both feature increasing American variant usage through the 1970s and a decrease in the 1980s, with a slight increase before a large decrease in 1990, highlighted by another uptick in the years following. Trends from the first half of the century are perhaps similar on average but feature different peaks and valleys—for example, the 1950 increase in *The Ubyssey* and sharp decrease in *The Gateway*. The most striking dissimilarity in the data from *The Ubyssey* is the modest, short-lived increase of American variants in 1965 that is absent from *The Gateway*.

[5] Here, discretion ought to be exercised by the researcher, where data realizing non-Zipfian distributions, particular variants preferring certain lexical items or data otherwise contradicting expectations might be given closer scrutiny.

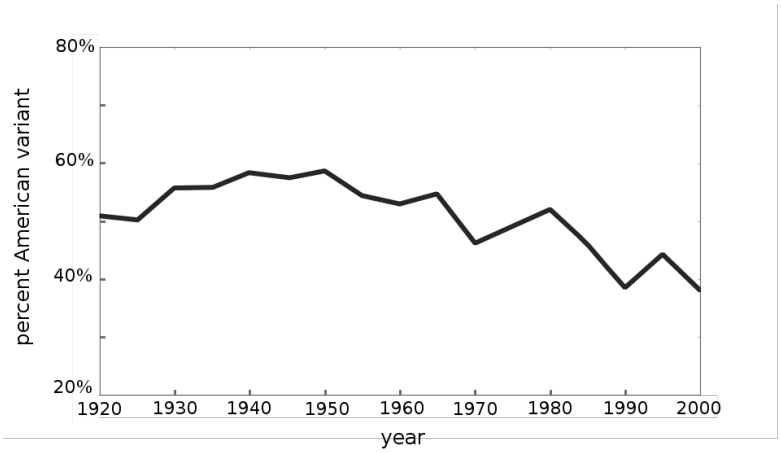

Figure 2. Aggregate American variants in *The Ubyssey*.

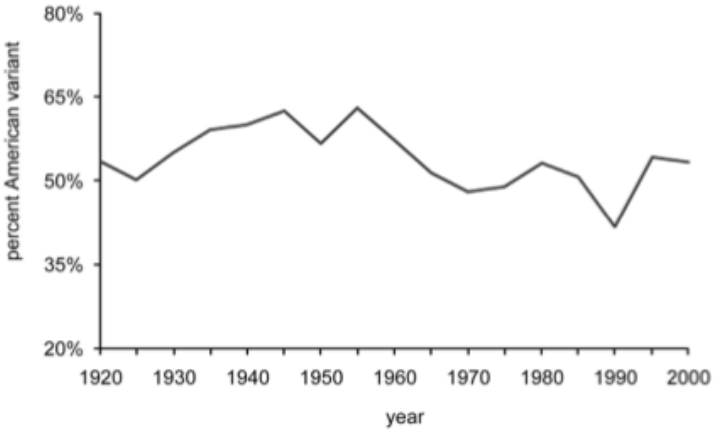

Figure 3. Aggregate American variants in *The Gateway* (Heffernan et al. 2010: 10).

Figure 4 plots the obtained Anti-American Index for data from *The Ubyssey*, and Figure 5 displays this index from Heffernan et al.

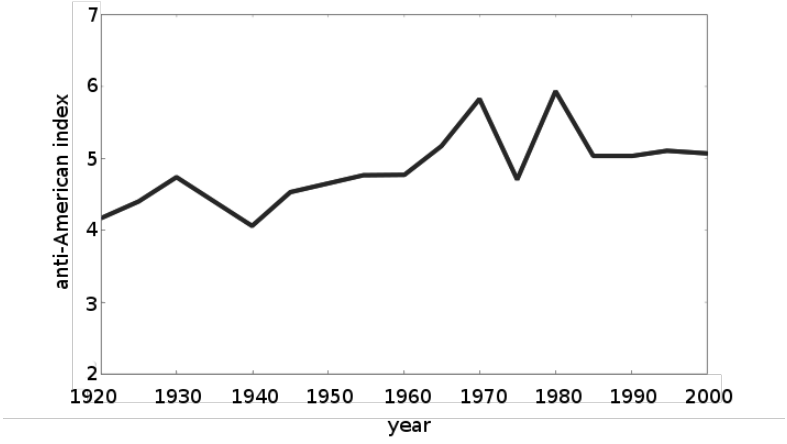

Figure 4. Anti-American Index for *The Ubyssey*, per five-year period.

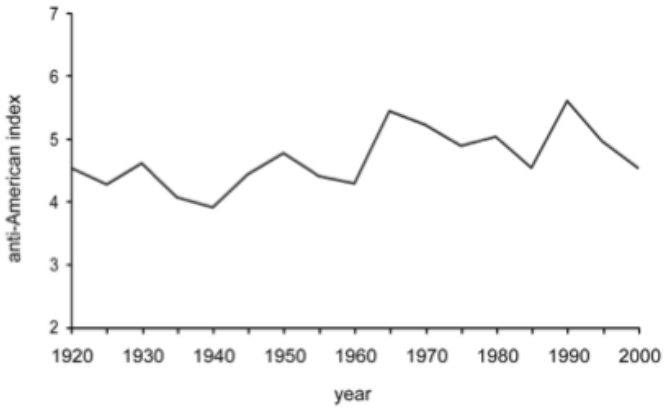

Figure 5. Anti-American Index for *The Gateway*, per five-year period (Heffernan et al. 2010: 15).

The Anti-American Index per five-year period for *The Ubyssey* appears relatively similar to Heffernan et al.'s for *The Gateway*: both feature 'peaks' in the years 1930 and 1980 and a 'valley' in 1940. Specifically, since Heffernan et al. theorize that the United States' involvement in unpopular wars motivates orthographic change, we might look to specific points in time for comparison. These are given as follows: 1940 or 1945 for World War II (where they theorize we would find low negative sentiment), 1950 for the Korean war (high anti-American sentiment), 1965 or 1970 for the Vietnam war (high) and 1990 for the first Gulf War (high). In the intervening periods we might find lower anti-American sentiment. Figure 4 shows an increase (instead of the expected decrease) through 1940 and 1945, but this continued increase from 1950 through 1970 is expected (attributed to the Vietnam War). The greatest discordance between figures occurs in 1965, where instead of a decrease as in Heffernan et al.'s plot there is a sharp increase in anti-American sentiment detected in *The Ubyssey*.

Figure 6 plots how long it took to find articles expressing American sentiment, while Figure 7 contains a scatterplot correlating this measure with anti-American sentiment. Ostensibly, this measure gauges the historical strengths of national sentiment, as detected in *The Ubyssey*, across the twentieth century. I do find the relationship between article-selection time and Anti-American Index informative, too: the correlation is fair, with a Pearson-r value of -.581, indicating that articles took less effort to find in years also ranked as exhibiting more anti-Americanism. The meaning of this is unclear, but several possibilities exist. Perhaps sentiments of anti-Americanism correlate with greater amounts of nationalism in general. Further, this could simply be a function of the dataset as a whole exhibiting slightly more anti-American sentiment. Finally, perhaps either in the mind of the analyst or in the corpus, the prototypical comments about Americans in *The Ubyssey* is anti-American.

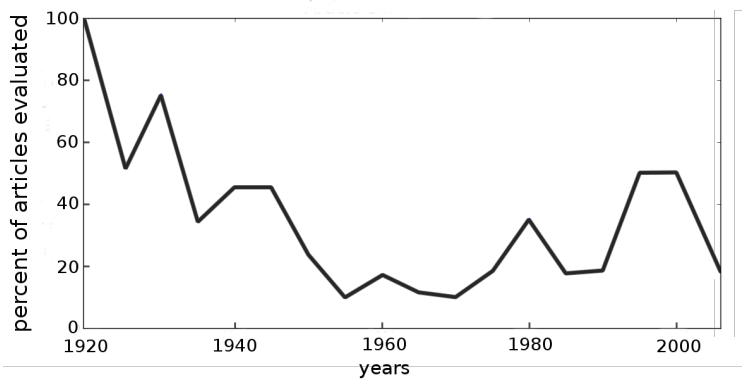

Figure 6. Percent of articles evaluated per 5-year period in *The Ubyssey* in order to locate three containing American sentiment.

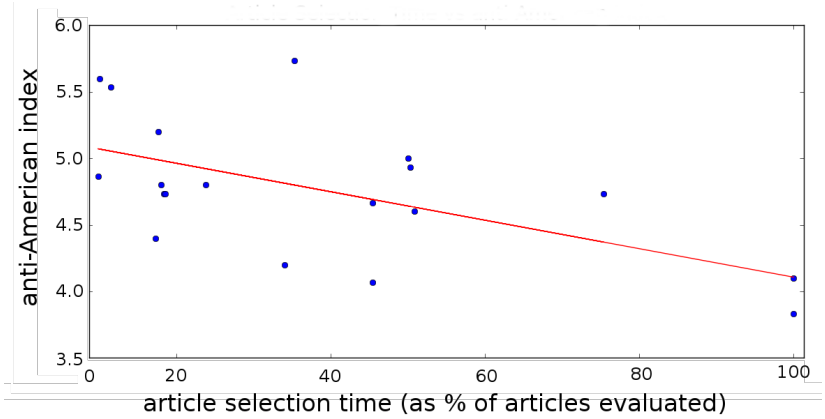

Figure 7. Article selection time per five-year period and correlation with Anti-American Index, for *The Ubyssey*.

The calculated correlation between American variant usage and anti-American sentiment in *The Ubyssey* is marginal. The Pearson-r score for 14 variables (excluding <gray>) is -0.434 with a *p* value of 0.064. This correlation is represented by the scatterplot in Figure 8. Therefore, while a correlation between orthography and national sentiment was detected by Heffernan et al. (2010) in *The Gateway*, this connection cannot be demonstrated in *The Ubyssey*.

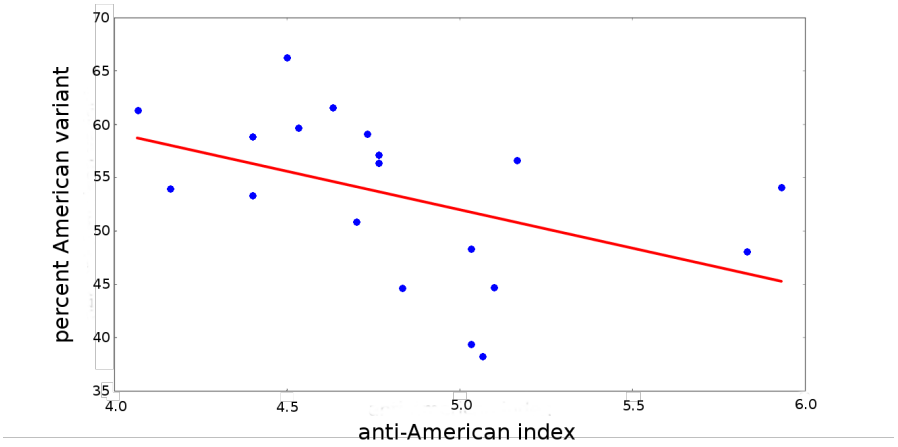

Figure 8. Scatterplot correlation with Anti-American Index for 14 variables, for *The Ubyssey*.

However, as Heffernan et al. note, not all variants have the same histories. Some are more closely correlated with anti-Americanism than others; some are simply more salient as identity markers, and some have historical variability in their social meanings. They write, "This inconsistency in historical identification for [double <ll> variables] suggests that the correlation between their American spelling variant usage and anti-American sentiment may be weaker for words such as *fulfill*" (p. 16). To discover these histories, to further interrogate the nationalism/orthography connection in *The Ubyssey* data, I calculated each variable's individual correlation to the Anti-American Index (Table 5). In addition, to discover how removing variables with inconsistent historical identification would affect correlations, I calculated the correlation of only eight variables' frequencies to the index (the original 15, minus <program>/<programme> because of the discrepancy in frequency noted above, <gray> / <grey> and the five double <ll> variables). This correlation becomes positive, however, at +0.445, *p* 0.06. Table 5 might explain this, listing all 15 variables under consideration rank-ordered by their correlation with the Anti-American Index. Interestingly, the double <ll>/single <l> variables are distributed throughout the

list: while <enrollment>/<enrolment> exhibits the second greatest negative correlation, <fulfill>/<fulfil> exhibits the greatest positive correlation, with the other variables in between. <or>/<our> variables are clustered together, which seems intuitive, but they exhibit almost non-existent correlation. If nothing else, if one accepts the premise that orthography does index ideology, Table 5 demonstrates that an array of variables is required to establish this correlation since correlation does not reduce to single terms.

Variables		**Pearson-r, *p***
kilometers	kilometres	-0.699, 0.001
enrollment	enrolment	-0.659, 0.002
labeled	labelled	-0.371, 0.118
defense	defence	-0.246, 0.309
offense	offence	-0.240, 0.322
centered	centred	-0.220, 0.365
humor	humour	-0.163, 0.505
honor	honour	-0.148, 0.545
neighbor	neighbour	-0.118, 0.63
marvelous	marvellous	+0.153, 0.532
judgment	judgement	+0.184, 0.45
jewelry	jewellery	+0.254, 0.294
gray	grey	+0.424, 0.07
program	programme	+0.445, 0.06
fulfill	fulfil	+0.481, 0.037

Table 5. Correlation of 15 variables with anti-American ratings, from *The Ubyssey*.

Expanding upon Heffernan et al.'s method, I added a 'tag' to each article rated that assigns its type. The motivation for this tag is the thoroughly heterogeneous makeup of a newspaper's articles: since a news article tends to report the speech of others while opinion pieces are presented as the author's own words, I captured these differences to see how they might affect correlations with different orthographies. Indeed, newspapers are composed of many different kinds of articles, which are further composed of many different voices, and these multiple ways of conveying sentiment (national or otherwise) ought to be recognized: it seems to be the case that an article about sports, for example, is simply doing something different than an article reporting on politics.

I propose that one benefit of tagging for article-type in newspapers might be greater ease of comparison. If, for example, we can show that in one newspaper orthographic variation correlates with sentiments in articles of the same type, we might be better able to show overall similarity with other newspapers. In this way, too, we might not need to show outright correlation between orthography and ideology (i.e. as high or higher than Heffernan et al.'s) but correlation trending in the same way. Table 6 contains the Pearson-r correlation values of orthographic variation and anti-American sentiment in the top three indicated types. Indeed, such a trend seems to be present. Respecting that only three categories can be reliably evaluated, strength of correlation is greatest for 'speaker' second greatest for 'news' and least for 'opinion', seeming to vary with the tacit attribution of sentiment described above.

Type	**Correlation**	**Count**
Speaker	-0.572	45
News	-0.523	76
Opinion	-0.445	66
Sports	---	22
Letter to the editor	---	19
Total	**---**	**228**

Table 6. Correlation of variation by article-type in *The Ubyssey*.

4 Discussion

Synchronically, *The Ubyssey* indicates a preference for British orthography but only by deviating from Heffernan et al.'s method of comparison. Historically, this preference is increasing. This historical trajectory contrasts with Heffernan et al.'s findings in *The Gateway*, where a negligible to slight preference—depending on the method of calculation—for American orthography is manifest, and this preference does not change over time. Only marginal correlation was found between variant use and anti-American sentiment in *The Ubyssey*, with Pearson-r -0.434, *p* 0.064. Heffernan et al. show a strong correlation in their data, with Pearson-r -0.715, *p* 0.001.

Heffernan et al. confirm their first hypothesis, that anti-American sentiment motivates the use of fewer American variants, but not their second—that such short-term change results in the long-term reduction of the use of American orthography. Interestingly, the opposite is found in *The Ubyssey*. A strong ideology/orthography connection is not observed, but diachronic change is: British variant use increases over the twentieth century. Perhaps long-term change inhibits short-term, in that an early reaction to American orthographies motivates the increasing preference of British variants which ultimately stifles variation. A basic analysis—by comparing relative variant use with the ideology index for the first and last half of the century—however, suggests that this is not the case: the variant use/ideology correlation from 1920 to 1960 is lower than the correlation found from 1960 to 2000. A more general observation is that the more, and increasingly, stalwart orthography observed in *The Ubyssey* might suppress smaller-scale changes. Finally, it was observed from the synchronic data in Table 3 that there are greater differences in the relative use of each variant in *The Ubyssey* than in *The Gateway*, which suggests stronger preferences for particular orthographic variants in *The Ubyssey*: though only by degree, British Columbians exhibit more clear selectivity in their orthography. Perhaps, then, this stalwart orthography is characterized by both long-term historical change and this greater preferentiality.

Ideology/orthography in different classes of documents showed varying correlations, and for the most frequently occurring type this correlation was quite high: for 'speaker', with Pearson-r of -0.572, it was the highest correlation observed so far. Correlation was lower for 'news' and low for 'opinion' at -0.445. Interestingly, then, sentiment that correlates with variant use tends to come from speakers not necessarily affiliated with the paper—those reported *through* the paper, rather than authors *of* the paper. Beyond a methodological addition that might make this work more translatable, as noted above, it is unclear what this means: are these opinions not the opinions of the paper? Likely they are, of course: by way of inclusion and tacit endorsement, they must be. Moreover, they perhaps better characterize prevailing attitudes, since the very presence of a speaker of note (or speakers, as in news articles) often informs, motivates and encapsulates opinion. It should be noted, though, that since taking this step segments the full dataset, it reduces the sample size with the consequent attenuation of reliability. (Scores could not be reliably calculated for article-types with low occurrence in the data: there were too many periods for which data would have been absent.)

Central to this type of analysis is the method of representing frequencies—what I have been calling 'averaged frequencies' and 'aggregate counts'. Each carries a different assumption—that each variable is weighted ideologically equal and that usage reflects ideological weight, respectively. Both are used in Heffernan et al.—aggregate counts for synchronic data and averaged frequencies for diachronic data, though it is not clear that this motivates each method's use or if these methods are fit for the purpose. A consistent method should be adopted in this regard.

Variables should be tested for variation. Heffernan et al. are careful to determine this at the corpus-level—for example by establishing that editorial oversight does not suppress mixed orthographies—but it does not appear that this is confirmed for each variable. A basic collocation analysis of variables in *The Ubyssey* revealed *grey* as an especially high frequency token because it occurs in the name of the geographic area on which the University of British Columbia lies, Point Grey. Since this stifles variability, it was determined that *grey* was an unusable variant. The American variant *gray*, however, did not escape censure: this token also often comprised proper nouns—names of individuals—to such a high degree that variation could not be assumed. On this note, it would be interesting to expand the scope of this search and find how the method might be refined by incorporating tokens' patterned occurrence into the analysis. At some level, we must

remember that these words are actual words, used in contexts for purposes, and that this contextual use might also motivate a preference for particular orthographies.

5 Conclusion

Although orthographic variation in *The Ubyssey* was found to corroborate earlier regional studies on American/British variation (Ireland 1979, Ireland 1980a, Brinton and Fee 2001) in that spelling in British Columbia favours—and continues to favour—the British, it was not found to be motivated by ideological position to the same extent as that found in *The Gateway* by Heffernan et al. (2010). As I have suggested, these findings might be related. Though variation is certainly ongoing, perhaps the orthographic system in British Columbia is more entrenched than that in Alberta and is becoming more entrenched over the long-term and hence more resistant to short-term change. The changes might still be manifest (indeed, this paper does not disprove the ideology/orthography hypothesis) but are more slight and less salient.

Methodological questions persist in establishing this ideology/orthography connection, pertaining to reliable ways of establishing and representing variation influenced by nationalism. Ideally, a comprehensive investigation would take into account context, perhaps by determining to what extent variant selection is motivated by the surrounding lexicon, if this motivating lexical context is patterned and how potential patterns might more finely index attitudes. Given the source, especially, methodological improvements will need to continue in order to effectively wrangle great quantities of unruly Canadian English data—Pratt's (1993) hobgoblin of Canadian English writ large.

References

Brinton, Laurel and Margery Fee. 2001. Canadian English. In *The Cambridge History of the English Language*, vol 6, ed. John Algeo, p. 422-439. Cambridge: Cambridge UP.

Chambers, J. K. 2004. 'Canadian dainty': The rise and decline of Briticisms in Canada. In *Legacies of Colonial English: Studies in Transported Dialects*, ed. Raymond Hickey, p. 224-241. Cambridge: Cambridge UP.

Davies, Mark. 2008-. *The Corpus of Contemporary American English: 450 million words, 1990-present*. Online: http://corpus.byu.edu/coca/.

Emery, Donald W. 1974. *Variant Spellings in Modern American Dictionaries*. Urbana: National Council of Teachers of English.

Heffernan, Kevin, Alison J. Borden, Alexandra C. Erath and Julie-Lynn Yang. 2010. Preserving Canada's 'honour': Ideology and diachronic change in Canadian spelling variants. *Written Language and Literacy* 13.1, 1-23.

Ireland, Robert J. 1979. *Canadian Spelling: An Empirical and Historical Survey of Selected Words*. Toronto: York University dissertation.

Ireland, Robert J. 1980a. Canadian spelling: How much British? How much American? *English Quarterly* 12.4, 64-80.

Ireland, Robert J. 1980b. *Spelling Across the Curriculum: A Student Handbook*. Toronto: Gage.

Lipski, John. 1975. Orthographic variation and linguistic nationalism. *La Monda Lingvo-Problemo* 6, 37-48.

Pratt, T. K. 1993. The hobgoblin of Canadian English spelling. In *Focus on Canada*, ed. Sandra Clarke, p. 45–64. Amsterdam: Benjamins.

Schieffelin, Bambi B. and Rachel C. Doucet. 1994. The 'real' Haitian Creole: Ideology, metalinguistics and orthographic choice. *American Ethnologist* 21, 176–200.

Sebba, Mark. 2000. Orthography and ideology: Issues in Sranan spelling. *Linguistics* 38, 925–948.

Translation Bureau, Public Works and Government Services Canada. 2012. Spelling. 15 January 2014. Online: http://www.btb.termiumplus.gc.ca/tcdnstyl-chap?lang=eng&lettr=chapsect3&info0=3.

The Ubyssey. 1918-2010. The Ubyssey Publications Society. Online: http://ubcpubs.library.ubc.ca/?db=ubyssey

Warfel, Harry R. 1966. *Noah Webster, Schoolmaster to America*. New York: Octagon Books.

Webster, Noah. 1846. *A Dictionary of the English Language; Abridged from the American Dictionary*. New York: Huntington and Savage.

Department of English
University of British Columbia
Vancouver, BC V6T 1Z1
dgrue@alumni.ubc.ca