

Language ideologies in social media: The case of Pastagate

Rachelle Vessey, Newcastle University

To appear in *Journal of Language and Politics*, 15 (2), 2016

Introduction

On February 19, 2013, inspectors from the *Office québécois de la langue française* (OQLF) sent a letter of warning to Massimo Lecas, owner of the Buonanotte Italian restaurant in Montreal, objecting to the use of Italian words on its menu. Lecas promptly reported the letter on the microblogging site Twitter, with a link to an Instagram image of the menu's linguistic offences, which included the word "pasta". The story was immediately picked up by a local journalist who retweeted (i.e. re-posted on Twitter) the photo and reported the story on his radio blog. The blog was then picked up by activist groups including *Put Back the Flag*, which shared the story, related stories, and memes tens of thousands of times over social media. Some individual examples were shared several thousand times. The story broke internationally and by February 26 had been chronicled in over 350 newspaper articles across 14 countries (Wyatt, 2013).

Social media have been credited with playing a major part in the proliferation of this story (e.g. Wyatt, 2013). Indeed, between February 20 and August 21, there were nearly 5000 instances of the coined term "pastagate" on Twitter that referred to this story; the hashtag #pastagate also trended online (Global Montreal, 2013). Reactions in social media predominantly expressed incredulity at the actions of the OQLF, often through mockery of the French language, Quebec, and then-Quebec Premier Pauline Marois. The impact of this social media deluge was immediate in Quebec: by March 8, the head of the OQLF had resigned, the OQLF warning to Buonanotte had been declared "overzealous", and on October 18 it was announced that the OQLF would be "modernising" its approach in order to deal with complaints more efficiently and in order to better support businesses and citizens (OQLF, 2013). In addition, the success of the so-called "Pastagate" affair has been linked to the provincial government's abandonment of the controversial Bill 14, which proposed amendments to Quebec's Charter of the French Language. In other words, the social media campaign appears to have had rather direct effects on language policy in Quebec.

While the OQLF conceded that it had been "overzealous", zeal also reverberated in the news and social media backlash. The international appeal of the story and its popularisation in social media suggest that the story touched a nerve internationally. Since the international response was largely unanimously disapproving, it may indicate that a loosely coherent set of beliefs and understandings about languages (i.e. language ideologies) permeate a range of different countries, cultures, and contexts. In addition, the fact that this level and style of language policing was so widely condemned in a globalised context suggests that minoritised nations may face challenges instituting language policy and linguistic autonomy even within territorial boundaries. In other words, Pastagate¹ highlights challenges not only in terms of language ideologies in transnational and multilingual social media, but also in terms of

language ideologies, language rights, and language policing that minoritised language groups may face more generally in a globalised world. To further flesh out these challenges, this paper addresses the following research questions:

1. How are French, English, and other languages represented in the social media that discuss Pastagate? What are the similarities and differences between these representations?
2. How do representations correlate with language medium (i.e. use of English, French, and other language)?
3. What is the meaning of “Pastagate”?

The paper proceeds as follows: first, a brief history of language policies and language rights in Canada is presented and the concept of language ideology is introduced. Next, the concept of superdiversity and its relevance to social media language is outlined. Section Three presents the data and methods, and then Section Four highlights the principal findings. These findings are then synthesised and discussed in Section Five.

Background, context, and theoretical assumptions

Although Pastagate is arguably the by-product of an era of mass communication in which news stories circulate widely and at an unprecedented rate, the “newsworthiness” of a Canadian language issue is by no means new (e.g. Vessey, 2013a). Language issues have been salient in Canada (and in Canadian media) throughout its history. When Britain won control over the majority of North America at the end of the Seven Years’ War, France withdrew and left behind an extensive French-speaking Canadian demographic; today, French speakers continue to comprise a significant proportion of the Canadian population. English became and remained the language of power, even in Quebec where the majority was French-speaking. By the mid-20th century, uprisings known as the “Quiet Revolution” brought the significance of the French-speaking population to the fore. New language laws were brought into force to address the linguistically-indexed inequality: the Charter of the French language (R.S.Q. c. C-11, s. 72) made French the official language of the Quebec government and the normal language of the workplace, schools, communication, commerce and business. The enforcement of these laws is largely the responsibility of the *Office québécois de la langue française*, which defines and conducts Quebec’s language policy and monitors the linguistic situation in the province. The federal Canadian government also stepped in, passing The Official Languages Act (R.S.C., 1985, c. 31 (4th Supp.)) in 1969; this made English and French the official languages of Canada. Canadian language policies thus address and are implemented in a country with a complex history and have never been problem-free (see e.g. Jedwab and Landry, 2011). Nevertheless, the policies become increasingly complicated with the rise of the Internet and in particular social media.

In 2011, the Standing Senate Committee on Official Languages undertook a study of Internet use and social media in Canada and found that the extent to which language rights are being “respected” is unclear (Senate Committee, 2012). The report observes that although the

Internet and new and social media present opportunities for the French language (e.g. web initiatives to promote French and those who speak it), the Internet remains English-dominant (Senate Committee, 2012, 102).ⁱⁱ Notably, the report stresses the need to expand the presence of French in the digital world (Senate Committee, 2012, 103). While recommendations focus primarily on the potential for action within the Government and civil service, the Committee also stresses the importance of the private sector and the significant role of financial incentives in encouraging growth (Senate Committee, 2012, 103). However, the extent to which any government can influence an online environment – even with financial incentives – remains unclear. Indeed, given the fact that the online Pastagate debate had direct impacts on language policy in Quebec, it is clear that the Committee’s report has raised pertinent questions about the viability of protecting language rights in online environments – and the role of the nation state in online language policing and language ideological debates.

Language ideologies are beliefs about language that become well-established in society to the extent that they end up being naturalised and taken for granted as “common sense” (Woolard, 1998). When beliefs about language surface in the context of pertinent social issues, they become manifested as “language ideological debates” (Blommaert, 1999). Language ideologies have been a particular focus in media studies (e.g. Johnson and Milani, 2010), and they have been studied in Canadian news media (e.g. Vessey, 2013a, b), but it remains unclear what role language ideologies have in transnational social media (cf. Androutsopoulos, 2013; Kelly Holmes and Milani, 2011). While more traditional media such as newspapers have been the mainstay of, for example, the ideology of the standard (i.e., beliefs in prescriptive, normative language use; Milroy, 2001), in superdiverse forums such as social media, the role of language ideologies is less clear. Social media normalise complex patterns of social relations and interactions across traditional boundaries; they also feature plurality, heterogeneity, and polycentricity in language use, communication, dissemination of information, and mediation of cultural practices (Leppänen, Kytölä, Jousmäki, Peuronen, and Westinen, 2014). Accordingly, social media provide contexts for discussions about language in which traditional offline categories such as “speech community”, “official language minority”, and indeed “nation state” have questionable relevance (Gee, 2005). In such diverse, boundary-free environments, language ideologies could play any number of different roles, especially because of the variety of contextual factors affecting the interpretation of online messages (Kelsey and Bennett, 2014). However, English continues to be a major player online and its role as a hegemonic medium of ideological transmission should not be underestimated. Thus, although social media have been argued to be “democratic” (see Small, 2011) in that they facilitate equal participation, their unregulated status means that there is little protection of minorities, including linguistic minorities. Thus, in online contexts it remains unclear whether and how beliefs about language are manifested, the role they play in communication across and through social media, and their implications for language policies. It is the objective of this paper to account for some the ways in which language ideologies feature in social media and their implications for language practices in the offline world of nation states. More specifically, this paper analyses a corpus of Tweets focusing on the Pastagate context, unearths some of the language ideologies embedded therein, and suggests some of the implications for language policy.

Data and methods

The “Pastagate corpus” under examination consists of 4795 Tweets and retweets (posts and repeated posts published on the microblogging site Twitter). Since Tweets are limited to 140 characters, the corpus of 4795 Tweets comprises only 123,853 words. The Tweets were collected based on their date of posting (February 20-August 21, 2013) and content: all publically-available Tweets containing the term “Pastagate” during this period were included. Hashtags were tagged so that they could be analysed in greater detail. Hashtags, which are lexical items marked by a # (hash), function on Twitter as metadata integrated into the linguistic structure of Tweets (Zappavigna, 2011, 4). The lexical item operates as a keyword that is searchable because of the hash; the inclusion of a salient, searchable item arguably contributes to a greater scope of meaning within a context of enforced brevity. Because hashtags are searchable, their use increases “loudness” and makes online conversations “more bondable” (Zappavigna, 2011, 18). Thus, even though they tend to be only singular lexical items, hashtags enable microbloggers to ambiently affiliate in co-present, impermanent communities (Zappavigna, 2011, 13).

This study adapts corpus-assisted discourse studies for data analysis (Baker, 2006). This approach combines corpus linguistics and discourse analysis in the study of large bodies of computer-readable texts called “corpora”. Corpus linguistic studies often begin with the study of frequency, because words that are repeated or avoided are understood to have particular functions within the society producing the texts. However, isolated words are not understood to be meaningful on their own; rather, meaning is understood to be achieved through the repeated use of words in fixed or semi-fixed phrases (e.g. “clusters” of words or collocation between individual items). With corpus linguistics, texts need not be read individually; instead, trends across texts can be established by reading concordance lines, which present an individual lexical item within its “co-text” across numerous texts (see e.g. Table 1). Thus, corpus linguistics can help discourse analysts contextualise their data and establish intertextuality and entextualisation (i.e. the process by which linguistic production is turned into a unit of text, decontextualized from its interactional setting, and metadiscursively recontextualised thus gaining new meaning). Discourse analysis can also help researchers to interpret the data with tools such as descriptive grammar, theorisations of communication and rhetoric, and sociopolitical and historical contextualisation. While discourse theory informs the research undertaken here, the focus is predominantly on the corpus methods and findings that emerged from that component of the analysis. In this study, WordSmith Tools (Version 6.0) (Scott, 2014) was used to analyse the Pastagate corpus in order to establish frequency and collocation trends and to explore the representations of languages and language ideologies more generally.

Findings

One of the most immediate findings to emerge from the data pertained to frequency. The most frequent lexical items (e.g. PASTAGATE, 5153; QUEBEC, 1779; LANGUAGE, 1260; OQLF 1003; MONTREAL, 792; PASTA, 559) refer to the most repeated themes – language issues in Montreal, Quebec, and the word “pasta”. However, non-content related words are

also frequent: several items (e.g. HTTP, COM, WWW, LY) suggest the extent to which websites – and in particular news websites featuring stories about Pastagate – figured in the Tweets. The frequency of HTTP (3575) suggests that at least 75% of all Tweets contain one website link.

Since URLs were not tagged in the corpus and punctuation was ignored, URL fragments occur as individual lexical items in the analysis. Thus, the most frequent clusters in the corpus tend to include these fragments (e.g. the cluster HTTP BIT LY consists of fragments containing <http://bit.ly>, a short URL redirection service). Other frequent clusters are fragments of headlines from these retweets of URL links. For example, all 170 instances of the cluster QUEBEC LANGUAGE OFFICE come from one of two verbatim retweeted headlines: “‘Pastagate’ prompts review at Quebec language office” (82) and “Head of Quebec language office resigns in wake of Pastagate” (87) (CBC, 2013a, b). The frequent cluster HTTP YOUTU BE also highlights the role of other media (in this case, YouTube videos) in the dissemination and popularisation of this story. Alternative media were used to mock and parody the Pastagate context. For example, one user posted parody videos of actors enforcing “French Only zones” in Montreal by informing passers-by – in English – of the pretend rules and regulations that would require them to use French only. These examples suggest the diversity of ways in which Twitter was used as a nucleus for affiliation with a shared condemnation of the enactment (and possibly the very existence) of Quebec’s language policy.

Another notable finding from the frequency lists was the presence of the French language as a medium of some Tweets. English was the predominant medium within the corpus; nevertheless, French functional words occur frequently (e.g. DE, 1066; LE, 897; LA, 840) and the frequent cluster HTTP WWW LAPRESSE (166) refers to a French online news source (*La Presse*). Other frequent clusters such as DE LA LANGUE (93) and PRÉSIDENTE DE L (89) attest to the presence of the French medium within this dataset. These instances suggest that the term “Pastagate” resonates within the French-speaking community; further investigation indicates that French users engage with the term and contribute to and contest its evolving meaning.

Certainly, many instances of the term “Pastagate” occur in French Tweets because they are used in news stories being retweeted (e.g. 52 retweets of “‘Pastagate’: la présidente de l’OQLF démissionne”, Teisceira-Lessard, 2013). However, numerous original French-medium Tweets contain “Pastagate”, and some of these express concern over what is seen to be an overreaction to the issue. For instance, one user laments that people so quickly jump to the conclusion that the OQLF should be abolished; she implies that its role is important but is often forgotten (*c’est dommage que certains crient à l’abolition de l’OQLF. On oublie facilement sa raison d’être...*). Another poster expresses frustration that the Buonotte restaurant is profiting from the OQLF’s actions, and calls for the restaurant owners to send their profits to the government (*Les proprio. devraient faire un chèque au gvt*).

There are many who express disgust with Pastagate, the actions of the OQLF, and Quebec's language laws more generally. In Example 1, the Tweeter questions the real source of the scandal: is the problem a lack of judgement or intelligence on the part of the OQLF or on the part of the law?

Example 1

Pastagate: Est-ce qu'on parle de manque de jugement/intelligence de l'OQLF ou d'une loi qui à la base manque de jugement/intelligence?

Some Tweeters are more virulent, describing the OQLF as having an attitude of "zealots" and the recent practices as "linguistic Nazism". In Example 2, the Tweeter explains that while she "adores" her language and wants to protect it, the "overzealous" actions of the OQLF are embarrassing and inappropriate.

Example 2

J'adore ma langue et veux bien la défendre, mais je trouve embarrassant et carrément déplacé cet excès de zèle de l'OQLF!!! #OQLF #pastagate

Condemnation of former OQLF president Louise Marchand was particularly fierce. Many Twitter users took advantage of the opportunity to play on words with the term "noodle" (*nouille*). This term can serve as an adjective, meaning "dumb", or as slang for "penis" or "prick". For instance, in Example 3, the Tweeter explains that "the #PQ [Parti Québécois] prefers seeing the word #noodle instead of #pasta [because] it better reflects its image".

Example 3

C vrai que le #PQ aime mieux voir le mot #Nouille au lieu de #Pasta ça reflète mieux leur image #Pastagate

Nevertheless, there was some support voiced for the former president through retweets of an article in which Marchand is described as the "victim of Pastagate" (see Example 4).

Example 4

Victime du "pastagate" Présidence de l'OQLF: Louise Marchand quitte son poste | Le Devoir [website]ⁱⁱⁱ via @[username]

Thus, although English was the dominant medium of this dataset, the French language was also used to both condemn and defend the actions of the OQLF.

These preliminary findings suggest that raw frequency findings, which are often a useful starting point in corpus linguistic research (Baker, 2006, 47), can prove less advantageous within a corpus of Twitter data. Indeed, frequency was also not particularly useful here because some of the core lexical items of interest to this study (e.g. English/*anglais*, French/*français*) were not among the most frequent words. The words ENGLISH and ANGLAIS/E/S occurred 123 and 45 times, respectively, and the words FRENCH and FRANÇAIS/E/S occurred 277 and 141 times, respectively. The word FRENCH ranked 50 on

the wordlist of most frequent words and only combined do the wordforms FRANÇAIS/E/S and FRANCAIS/E/S rank 98 on the wordlist.

Although FRANÇAIS/E/S occurs as a standalone hashtag (i.e., with no grammatical function) in three instances and within websites and usernames in fifteen instances (e.g. *@Impératif Français*), it predominantly exists in nominal and adjectival forms. FRANÇAIS/E/S occurs as an adjective in 66 instances, but 56% (37) of these occur in the semi-fixed phrase *Office [québécois/e] de la langue [français/e/français/e]*. Most (40%, or 15) instances of this semi-fixed phrase occur in the standard format (*Office québécois de la langue française*), but there are twelve references to the *Office de la langue française*. There are also eight erroneous uses of gender, accent, and capitalisation (e.g. *québécoise* instead of *québécois*), so this semi-fixed phrase does not emerge as a frequent cluster. References to *Office [québécois/e] de la langue [français/e/français/e]* aside, only 21% (29 occurrences) of all instances of FRANÇAIS/E/S occur as adjectives. Most (45%) adjectival uses of FRANÇAIS/E/S refer to the French language (*la langue française*, 13). In particular, there are five specific references to protecting, defending, and the “combat” for the French language, as in the following examples (see Examples 5-7).

Example 5

Pastagate : jusqu’où va la **défense de la langue française**... (En anglais !)
@[username]: [website] by @[username]

Example 6

Pastagate, ou de la nécessité de **reconcevoir le combat pour la langue française** au Québec [website]

Example 7

C’est ok de protéger la langue française, mais là les zélés on se calme ...!! -
@[username] digs into “pastagate”

Notably, these examples take issue with the actions of the OQLF while at the same time recognising the necessity of protecting French. For instance, Example 5 questions how far (*jusqu’où*) action should go to protect the French language, but does not suggest that French is not in need of protection. Example 6 retweets a news story about “the need to rethink the combat for the French language” (*la nécessité de reconcevoir le combat pour la langue française*), which also presupposes that the French language requires a battle. Example 7 argues that while the “overzealous [OQLF agents] need to calm down” (*là les zélés on se calme*), it is “ok” to protect the French language.

FRANÇAIS/E/S also occurs as a noun in 57 instances, of which 13 (23%) refer to a person or people from France. All other nominal uses refer specifically to French as a language. Notably, a salient trend is for Tweets to discuss things that exist or occur “in French” (13), or things being done *to* or *for* French (13) (combined, 46% of occurrences). Notably, there are nine instances in which French is represented as a cause or something in need of defence and

protection (*défendre [...] le français; protéger le français*). Similar to the examples listed above, none of these examples suggests that French should not be protected. Example 8 argues that Law 101 is meant to “protect us from the proliferation of English” (*nous protéger de la prolifération de l’anglais*), but explains that Italian does not threaten French. Thus, it seems clear that “we” are in need of protection. Similarly, Example 9 argues that Pastagate does not help “the French cause” (*la cause du français*), once again suggesting that French requires protection. Example 10 clearly advocates the need for French protection by exclaiming “Defend French yes! Defend French stupidly, no!”.

Example 8

@[username] La loi 101 vise à **nous protéger de la prolifération de l’anglais;** l’italien ne menace pas le français. @[username] #pastagate

Example 9

Pasta ou comment se mettre les pieds dans les plats!Ça n’aide absolument pas **la cause du français** au Quebec, amis anglos furieux! #pastagate

Example 10

Défendre le français oui! Défendre connement le français, non! #pastagate à #Montréal

Another trend – albeit less salient – is for French literacy to be criticised (see Examples 11-13).

Example 11

@[username] @[username] Je déplore que #OQLF fasse la guerre à l’anglais ou #pastagate mais rien contre **le mauvais français & fautes**

Example 12

#ecoeuredepayer pour des zélés de l’#OQLF quand nos jeunes ne savent même pas **écrire et parler correctement le français** #pastagate

Example 13

Quand #OQLF s’attaquera-t-il aux vraies problèmes, comme celui de **la qualité du français de nos jeunes à l’école?** #pastagate #polqc

These instances condemn the OQLF for not addressing the diminishing quality of French in Quebec and allude to wider concerns about the standard of French in Quebec, a topic that has been debated for many decades (if not longer – see Bouchard, 2002).

Finally, it is notable that while there are frequent references to the French language, there are few references to what French *is* or *does*. For example, there are no references to why French should be protected, spoken, or used more generally. Furthermore, there are few references to *using* French (e.g. speaking, learning, teaching). In other words, the French language is predominantly represented as a passive recipient of action or as a circumstantial adjunct contextualising other action taking place.

When the French language is discussed in English Tweets (e.g. FRENCH, 277), some of the most frequent clusters pertain to retweeted content. For example, the 76 references to FRENCH ONLY ZONE pertain to retweets of the aforementioned YouTube videos. Also, the 18 references to QUEBEC'S FRENCH HARDLINERS pertain to retweets of a *Globe and Mail* article, as in Example 14.

Example 14

'Pastagate' reveals the hypocrisy of Quebec's French hardliners /via @[username]

Retweets aside, FRENCH occurs as an adjective in 75 (43%) instances, and 20 of these refer to the OQLF (e.g. "French language office"). There are three references to the Charter of the French Language, and notably six references to French language endangerment (e.g. protect/promote/supporting/save the French language). These examples do not presuppose that French should be protected, as was the case with FRANÇAIS/E/S. For example, Example 15 declares that Quebec's promotion of French is discriminatory and creates apartheid in Canada.

Example 15

APARTHEID IN CANADA? Yes. Quebec and all levels of government **promote French, even to the point of discrimination**

Similarly, Example 16 argues that there is a "fine line" between protecting French and "whitewashing multiculturalism", suggesting that such protectionism is xenophobic and potentially racist.

Example 16

@[username] I love Québec (avec tout mon coeur), but there's a fine line between **protecting French** + whitewashing multiculturalism. #Pastagate

Example 17 uses scare quotes to argue that "protecting" French is simply a guise for discrimination.

Example 17

#Quebec needs a #harlemshake to rid off the #oqlf and discrimination disguised as **"protecting" French** #pastagate [website]

In a similar vein, there are also nine references to "French language police", a popular label in English-speaking Canada used to negatively evaluate OQLF agents investigating complaints. POLICE, which occurs 293 times in the corpus, collocates with LANGUAGE in 265 cases, but notably collocates with LANGUE in only nine cases.

FRENCH also occurs as a noun in 71 cases and 17 (24%) of these discuss things taking place in French. However, unlike the uses of FRANÇAIS/E/S, 20 (28%) references also discuss speaking and learning French (see Table 1).

olf and Pauline Marois got to DjokerNole who's speaking French at CoupeRogers bill1
 13 Mar A Marie Antoinette moment in Quebec? Can't speak French? Walk! no metro for
 bec ministers to speak to Canadian counterparts only in French <http://natpo.st/Z2to>
 ues. Like pastagate. And people who don't care to learn French. <http://instagram.co>
 e DJ_20_1003 Apr je veux ce t-shirt Keep Calm and Learn French #Pastagate ? Pfff. p

Table 1: Selected concordance lines discussing speaking and learning French

Thus, English Tweeters place more emphasis on using (i.e. speaking, learning) languages than French Tweeters.

Notably, ENGLISH and ANGLAIS/E/S occur considerably less frequently (123, 45, respectively) than FRENCH and FRANÇAIS/E/S (277, 141, respectively). These terms are also used rather differently, too. In French, *anglais* occurs most often as a noun (84% or 38 instances), indexing things taking place “in English” (*en anglais*). The most salient trend in nominal uses (32% or 12 occurrences) is to refer to the original complaint about the Buonanotte restaurant. It eventually transpired that the OQLF inspectors had misunderstood the original complaint that an individual had been given an English menu (with Italian) and not a French menu (Bourgault-Côté, 2013). When the inspectors investigated the claim, they were given a French menu, where the inspectors noted the extent of Italian. Nevertheless, in the original complaint, the problem was not seen to be the use of Italian, but rather the use of English (see Examples 18 and 19).

Example 18

‘Pastagate’ (français au Québec) Une manipulation. La plainte portait sur l’anglais,
 pas l’italien.. [website]

Example 19

le mensonge est la vérité! Le menu est entièrement en anglais >>> FAUX! menu
 italien & francais OU italien & anglais #pastagate #ORWELL

In the rare cases where *anglais* occurs as an adjective, it is used to describe English menus, Canadians (*Canadiens anglais*), songs (*chansons anglaises*), terms (*terme/s anglais*), media (*Médias anglais*), and the “English minority” (*la minorité anglaise*). In contrast, *English* occurs equally as a noun and as an adjective, and it is predominantly discussed with reference to its uses (speaking, making, teaching/learning), and *prevention* of uses (ban, forbid, disallow), as in Examples 20 and 21.

Example 20

In Quebec it is officially prohibited to speak English [website] #pastagate

Example 21

[website] Raw vid of complete asshole screaming at person for speaking english in
 public in Montreal #pastagate #qcpoli

Notably, *English* is also used as an adjective to describe various items ranging from programs (11 occurrences) to muffins (one occurrence).

The difference between the uses of ANGLAIS/E/S and ENGLISH seem to suggest that French Tweets were focused on establishing the facts about the original complaint (i.e. about English

on the menu) whereas the English Tweets pertain to wider issues about banning English in other contexts. Furthermore, the English-medium Tweets seem to focus more on using (e.g. speaking, learning, writing) English and French rather than simply denoting things taking place in these languages. In contrast, there is hardly any reference to language use in the French Tweets, which in some instances suggest that the OQLF’s efforts would be better spent improving the quality of French. Also, while many French Tweets presumed that French is threatened and requires protection, in English Tweets this kind of protectionism is criticised.

The final step of the analysis was to examine hashtags. The most frequent hashtag in the corpus is #PASTAGATE and with 2829 instances it is five times more frequent than any other hashtag. However, other hashtags are also used. Since raw frequencies proved to be somewhat problematic (see discussion above), these hashtags were not analysed according to their frequency but rather according to the following categories of emergent themes: (1) news outlets, (2) geography, (3) politics, (4) topic-specific items, (5) humour, (6) negative evaluation, and (7) hashtags linking the story to other contexts.

The first category of hashtags pertained to news outlets or news sources, and not just Canadian news sources, but also American, British, English and French news outlets (Table 2).

Hashtag	Frequency
#CNN	49
#NEWS	42
#SUNNEWSNETWORK	40
#CJAD	28
#NBC	19
#CBC	17
#ABC	16
#BBCWORLD	13
#BAZZOTV	9
#BBCNEWS	8

Table 2: 10 most frequent news hashtags

Notably, most of these hashtags occurred in Tweets that contained numerous different news hashtags. It seems that Twitter users were hashtagging multiple news outlets in order to flag the story to the international community and to international news outlets more specifically, as in Example 23.

Example 22

Quebec at it again: [website] #OQLF #BILL14 #CNNI #CNN #BBCNews
 #BBCWorld #Montreal #Canada #NBC #ABC #pastagate #nationalpost

Another category consisted of hashtags pertaining to geography (see Table 3). While Montreal is the most frequently cited location, other Canadian locations are mentioned and

international sites, too. These instances seem to suggest an appeal to the international community or the international interest in this story.

Hashtag	Frequency
#MONTREAL	281
#QUEBEC	211
#CANADA	57
#MONTRÉAL	11
#MTL	9
#TORONTO	7
#514 [Montreal area code]	6
#CANADIENS	5
#QUÉBEC	5
#FRANCE	4

Table 3: 10 most frequent geography hashtags

A third large category of hashtags pertains to politics (see Table 4), including references to specific political parties such as the elected minority government of Quebec, the Parti Québécois (#PQ, 323), and opposing parties such as the Parti Libéral du Québec (#PLQ, 33).

Hashtag	Frequency
#OQLF	555
#PQ	323
#POLQC	233
#ASSNAT	179
#BILL14	145
#QCPOLI	134
#CDNPOLI	78
#BILL101	70
#OLF	67
#LOI101	65

Table 4: 10 most frequent politics hashtags

Hashtags also refer to individual politicians (e.g. #MAROIS, 34) and to Canadian (e.g. #CDNPOLI, 78, #CANPOLI, 3, #POLCAN, 5) and Quebec politics (#QCPOLI, 134, #POLQC, 233). Notably, different hashtags are used in English in French: while #CDNPOLI and #QCPOLI tend to be used in English to tag Canadian and Quebec politics, #POLCAN and #POLQC tend to be used to refer to the same topics in French. While these are user-generated preferences rather than fixed rules, they are preferences adhered to, for example, by the Prime Minister of Canada in his verified English and French Twitter accounts (compare the use of #cdnpoli for the account @pmharper and #polcan for the account @premierministre). Notably, these hashtags only occur together in four instances within the corpus (see Table 5).

```
#polqc #assnat #pq #pastagate #qcpoli
e=8870035 ... #pastagate #polqc #qcpoli #pq
- http://bit.ly/YcCXSf #polqc #qcpoli #cdnpoli #pastagate #OQL
- http://bit.ly/YcCXSf #polqc #qcpoli #cdnpoli #pastagate #OQL
```

Table 5

Similarly, #BILL14 and #LOI14 are English and French hashtags used refer to the same language policy document and #BILL101 and #LOI101 refer to Quebec’s Charter of the French Language. Again, only in very rare cases do these occur as collocates. For example, #BILL14 occurs 145 times and collocates with #BILL101 in 40 (28%) occurrences, but only collocates with #LOI101 and #LOI14 in fewer than 5% of instances (8 and 2, respectively). Similarly, #LOI101 occurs 70 times but only collocates with #BILL101 in 2 (3%) instances. While these numbers cannot reveal anything in and of themselves, the collocation patterns suggest that little effort is being made to appeal to both English and French Twitter users at the same time. Instead, in most cases it seems that the exclusive use of monolingualism in hashtagging suggests an affiliation with *either* English Twitter users *or* French Twitter users. In other words, the use of these distinct hashtags suggests that different language groups are not only participating within the debate, they are also appealing to distinct communities of ambient affiliation (Zappavigna, 2011). This divide further enhances the potential for a communication gap between the two groups – what is commonly known as the “two solitudes” in Canada (see e.g. Heller, 1999).

A fourth category of hashtags pertains specifically to the Pastagate controversy (see Table 6).

Word	Hashtag frequency	Raw frequency	% of raw frequency occurring as hashtag
#PASTAGATE	2829	5153	55
#PASTA	91	654	14
#LANGUAGE	54	1266	4
#BUONANOTTE	29	248	12
#COFFEEGATE	29	29	100
#SPOONGATE	26	28	93
#FRENCH	16	337	5
#CAFFEGATE	12	14	86
#TURBANGATE	10	12	83
#ENGLISH	9	151	6

Table 6: 10 most frequent topic-specific hashtags

Notably, most words that directly index the topic – e.g. *pasta*, *language*, *French*, *English*, *Italian* – are used less often as hashtags than more playful words or neologisms. For example, although PASTA occurs 654 times in the corpus, only 91 (14%) occur as hashtags. Similarly, although the words LANGUAGE and ITALIAN occur 1266 and 339 times, respectively, only 4% and 2% of these instances occur as hashtags. In contrast, spinoffs of Pastagate – which included reports of OQLF investigations into coffee cups inscribed with *caffè* and spoons at a frozen yogurt venue with English catchphrases such as “this is my mix” – took on names such as *caffegate* and *spoongate* that only ever occur as hashtags within this corpus. Indeed, the items that occur predominantly as hashtags rather than as regular functioning

words in the corpus include *coffegate*, *spoonagate*, *caffegate*, *turbagate*, *dildogate*, *spoonscandal*, *steakgate*, *redialgate*, *strudelgate* – all items pertaining to stories related to the premise of Pastagate. Since these hashtags are all searchable, they function as a simple method of extending the scope of Pastagate to new and developing domains. In contrast, lexical items pertaining specifically to the actual Pastagate topic (e.g. language, French, English, Italian) occur proportionally less often as hashtags. This suggests that the facts surrounding Pastagate were less salient than the more evaluative, sensationalist descriptors and labels that occurred proportionally more often as hashtags.

Another category of hashtags pertains to the humorous appeal of the story (Table 7).

Hashtag	Frequency
#COMEDY	49
#FUNNY	42
#LOL	10
#HUMOR	5
#HUMOUR	2
#LMFAO	1

Table 7: Humour hashtags

References to “comedy”, “funny”, “lol”, “humor/humour”, and “LMFAO” all suggest the extent to which a story about such language policing aroused incredulity. However, other readers did not evaluate the story as humorous and instead highlighted negativity and embarrassment with hashtags such as #helpme, #idiots, #laughingstock, #cretins, #honte (“shame”) (see Table 8).

Hashtag	Frequency
#HELPME	8
#HONTE	8
#NOFRENCHZONE	6
#RACIST	5
#QUEBECBASHING	4
#DISCRIMINATION	3
#EQUALITY	3
#FRANCOFASCISTS	3
#ENOUGHISENOUGH	2
#ETHNOCIDE	2
#ZÉLOTE	1

Table 8: Negative evaluation hashtags

More extreme negativity was expressed through the use of hashtags such as #racist (5), #francofascists (3), #ethnocide (2), and #anglophobes (1). These hashtags are used exclusively in English Tweets and express a more extreme negative evaluation of Pastagate; they naturalise a link between linguistic protectionism (and language policy enforcement) and varying levels of xenophobia. Hashtags such as #zélote are used by French Tweeters who also make links between Pastagate and extremism. However, the (English) hashtag

#quebecbashing only occurs in Tweets that are otherwise entirely written in French. The label “Quebec bashing” tends to be used to highlight English Canada’s negative representation of events in Quebec – an issue discussed at length, for example, in the work of Barbeau (2013). Two of these four hashtags retweet the story that the original complaint against the Buonanotte restaurant did not pertain to the use of Italian on the menu, but rather the use of English on the menu.

Notably, the story about the original complaint seems to have been only reported in French, and not picked up in English. When English-language Tweets refer to this story, the original complainant is described as a “[s]erial language complainer” and someone who “hates English” (see Examples 23 and 24).

Example 23

Oh good lord. RT @[username]: OQLF Buonanotte complainant says doesn't hate Italian, just hates English. [website] #pastagate

Example 24

(FR) Serial language complainer behind #pastagate says he complained about menu in Eng, not Italian. [website] #qcpoli #oqlf

However, these descriptors were not used in the retweeted article (Bourgault-Côté, 2013) and English hashtags are used to propagate this misrepresentation of the story to the wider English-speaking Twittersphere. In other words, even when the English Tweeters do pick up the factual context of the original complaint, they misrepresent and propagate the misrepresentations to a wider public.

The final category of hashtags pertains to links made between Pastagate and other contentious or negative stories in and about Quebec (Table 9). These include the hashtags #CEIC and #CORRUPTION, which refer to the Charbonneau Commission, a public inquiry into corruption in the management of public construction contracts in Quebec. This Commission is irrelevant to Pastagate but the linkage serves to negatively evaluate Quebec more generally. Similarly, the hashtag #MAGNOTTA refers to Luca Magnotta, the Montrealer accused of (among other things) killing and dismembering an international student. This individual is also irrelevant to Pastagate but serves to perpetuate and extend the negativity to wider and more serious issues.

Hashtag	Frequency
#CEIC	24
#CORRUPTION	2
#MAGNOTTA	2

Table 9: Hashtags links Pastagate to other stories

Discussion and conclusion

To synthesise the findings, let us return to the research questions posed at the beginning of this paper.

The first two questions asked how English, French and other languages are represented within the corpus of Twitter data and if representations differ according to language medium. Findings showed that English and French Tweets tended to represent languages rather differently. While ANGLAIS/E/S was used to describe the menu in the original context of the Buonanotte complaint, ENGLISH was used to represent language as a tool to be used (or prevented from being used) within wider contexts. Similarly, FRENCH was used to represent the language as a tool for use, rather than labelling things as “being” or “existing in” French. These examples suggested evidence of instrumental language ideologies underpinning the English-medium Tweets. If a language has “instrumental” value, then it is seen as a tool that enables individuals to achieve specific goals (Ager, 2001, 2-10). Garvin (1993, 51) cites English-speaking countries as examples of places where the instrumental attachment to language dominates.

Rather than focusing on language use, references to FRANÇAIS/E/S represented the French language as needing protection – even if the actions of the OQLF were sometimes highlighted as embarrassing, unnecessary, or “overzealous”. These examples suggest evidence of ideologies of language endangerment, which presume that some languages require protection (and, indeed, that it is possible to protect them) because of ecological factors that threaten their existence (Heller and Duchêne, 2007, 4). Not only were ideologies of language endangerment not found in English Tweets, the concept of language endangerment was linked to discrimination. Finally, a small number of French Tweets suggested that attention should be paid to the quality of French in Quebec rather than the use of other languages. These examples could suggest evidence of standard language ideologies, which presuppose that there are “good” and “correct” ways of speaking that must be adhered to whereas “poor” and “incorrect” ways of speaking should be eliminated (see Milroy, 2001). No standard language ideologies were found in the English Tweets.

This leads us to the final research question, which asked about the meaning of “Pastagate”. Pastagate seems to index affiliation with (1) Quebec-related interests, (2) a value system concerning language(s), and/or (3) an interest in novelty news stories. Notably, the term also indexes an anglocentric understanding of language issues in that it is predominantly used in English Tweets that miss, misrepresent, or misunderstand the original complaint that sparked Pastagate and these Tweets advocated English interests (e.g. the abandonment of Bill 14). “Pastagate” impacted language policy, and more than a year later the term is still frequently used to index anglocentric perspectives on Quebec news and events (see e.g. www.pastagate.com). What began as a story about one menu in one restaurant turned into a way of opposing language policy more generally, and Quebec politics, too. This seems to have been achieved by linking the Pastagate context to related contexts (e.g. #caffegate) and unrelated contexts, too (e.g. #CEIC).

Certainly, there are limitations to this study. Only a small corpus of data was examined and it was not charted diachronically to establish change over time. Also, Twitter metadata were not examined and user details were not available, so it is impossible to know if English Tweets come from “English speakers” and French Tweets from “French speakers” – or if these

categories have any meaning online. Accordingly, findings from these Tweets cannot in any way be generalised from these communities of “ambient affiliation” to offline language communities. Additionally, not all URLs contained within Tweets were investigated; this kind of “hypertext” can be considered an important form of intertextuality in the discourse analysis of Web 2.0. In terms of methodology, although frequency counts revealed some findings, frequencies were affected by the high concentration of retweeting, the nature of the multilingual data, and abbreviations and misspellings. These all impacted the analysis to the extent that salience based on frequency was rather difficult to establish. In other words, frequency was less useful here than in other datasets that contain less verbatim repetition (e.g. retweets), less multilingualism, and more “standard” language. Since superdiverse online environments are increasingly part of offline daily life, the ability for corpus linguistics to address these challenges will be essential in short order.

Limitations notwithstanding, Pastagate could have implications for other minorities in a globalised context. Given that the Pastagate social media campaign used a global language (English) to impact language policy on the ground, it seems that linguistic representations in social media may have implications for other contexts. Although the Canadian Senate Committee on Official Languages acknowledged that social media pose new challenges for language rights, it failed to note the consequences of language ideologies online. A flawed “benign neglect” approach to language (i.e. the belief that individuals’ language choices should be unrestricted and ungoverned, Kymlicka and Patten, 2003, 10) seems to underpin English Tweets’ disparagement of language protectionism and prioritisation of language use. Thus, it would seem that social media complicate and replicate offline language ideological systems and power hierarchies (Kelsey and Bennett, 2014; Page, 2012, 181-2; Vessey, 2013a, b). So long as English predominates in online spaces, ideologies of instrumental language that predominate in English-speaking countries (Garvin, 1993) may be imported online, where language policies and the protection of minorities have little relevance and “democratic” (see Small, 2011) participation in the negotiation of meanings (Zappavigna, 2011) is contingent on the use of the dominant language – English. Although French is an international language, French perspectives on this issue were still marginalised within this dataset and the English-dominant representation of Pastagate arguably impacted upon Quebec language policy. English-medium representations could therefore have an even greater impact on “smaller” languages. Thus, this paper contends that dominant, transnational language ideologies in online forums can have rather direct implications for linguistic minorities, their rights, and the policies protecting them in the off-line world of nation-states.

References

- Ager, Dennis. 2001. *Motivation in Language Planning and Language Policy*. Clevedon: Multilingual Matters.
- Androutsopoulos, Jannis. 2013. Participatory Culture and Metalinguistic Discourse. In *Discourse 2.0: Language and New Media*, ed. by Deborah Tannen and Anna M. Trester, 47-71. Washington, D.C.: Georgetown University Press.
- Baker, Paul. 2006. *Using Corpora in Discourse Analysis*. London & New York: Continuum.
- Barbeau, Genviève. B. 2013. *Le dossier Maclean's et le Québec bashing : Analyse sociodiscursive d'une affaire médiatique controversée*. Unpublished PhD Thesis. Université Laval.
- Blommaert, Jan. 1999. The Debate is Open. In *Language Ideological Debates*, ed. by Jan Blommaert, 1-38. Berlin: Mouton de Gruyter.
- Bourgault-Côté, Guillaume. 2013. Buonanotte: la plainte originale ne portait pas sur l'italien. *Le Devoir*. 02.03.2013. <http://www.ledevoir.com/politique/quebec/372332/buonanotte-la-plainte-originale-ne-portait-pas-sur-l-italien>.
- CBC (Canadian Broadcasting Corporation). 2013a. Head of Quebec Language Office Resigns in Wake of Pastagate. 03.08.2013. Available: <http://www.cbc.ca/news/canada/montreal/head-of-quebec-language-office-resigns-in-wake-of-pastagate-1.1308461>.
- CBC (Canadian Broadcasting Corporation). 2013b. 'Pastagate' Prompts Review at Quebec Language Office. 26.02.2013. Available: <http://www.cbc.ca/news/canada/montreal/pastagate-prompts-review-at-quebec-language-office-1.1305047>.
- Crystal, David. 2011. *Internet Linguistics: a Student Guide*. London: Routledge.
- Garvin, Paul. 1993. A Conceptual Framework for the Study of Language Standardization. *International Journal of the Sociology of Language* 100/101: 37-54.
- Gee, James P. 2005. Semiotic Social Spaces and Affinity Spaces: From *The Age of Mythology* to *Today's Schools*. In *Beyond Communities of Practice*, ed. by David Barton and Karen Tusting, 214-232. Cambridge: CUP.
- Global Montreal. 2013. Now Trending: #Pastagate. *Storify.com*. Available: http://storify.com/global_montreal/now-trending-pastagate.
- Heller, Monica. 1999. Heated Language in a Cold Climate. In *Language Ideological Debates*, ed. by Jan Blommaert, 143-170. Berlin: Mouton de Gruyter.
- Heller, Monica. and Duchêne, Alexandre. 2007. Discourses of Endangerment: Sociolinguistics, Globalization and Social Order. In *Discourses of Endangerment: Ideology and Interest in the Defence of Languages*, ed. by Alexandre Duchêne and Monica Heller, 1-13. London: Continuum.
- Jedwab, Jack and Landry, Rodrigue (eds). 2011. *Life After Forty/ Après quarante ans: Official Languages Policy in Canada/ Les politiques de langue officielle au Canada*. Montreal: McGill-Queen's University Press.
- Johnson, Sally and Milani, Tommaso (eds). 2010. *Language Ideologies and Media Discourse: Texts, Practices, Politics*. London: Continuum.

- Kelly-Holmes, Helen and Milani, Tommaso (eds). 2011 Thematising Multilingualism in the Media. *Journal of Language and Politics* 10 (4): 467-489.
- Kelsey, Darren and Bennett, Lucy. 2014. Discipline and resistance on social media: Discourse, power and context in the Paul Chambers 'Twitter Joke Trial'. *Discourse, Context & Media* 3 (1): 37-45.
- Kymlicka, Will and Patten, Alan. 2003. Language Rights and Political Theory. *Annual Review of Applied Linguistics* 23: 3-21.
- Milroy, James. 2001. Language Ideologies and the Consequences of Standardization. *Journal of Sociolinguistics* 5/4: 530-555.
- Leppänen, Sirpa Kytölä, Samu, Jousmäki, Henna, Peuronen, Saija and Westinen, Elina. 2013. Entextualization and Resemiotization as Resources for Identification in Social Media. In *The Language of Social Media*, ed. by Phillip Seargeant and Caroline Tagg, 112-138. Houndsmills: Palgrave Macmillan.
- Office québécois de la langue française. 2013. L'Office québécois de la langue française modernise ses pratiques. Available: http://www.oqlf.gouv.qc.ca/office/communiqués/2013/20131018_modernisation.html
- Page, Ruth. 2012. The Linguistics of Self-Branding and Micro-celebrity in Twitter: The Role of Hashtags. *Discourse & Communication* 6 (2): 181-201.
- Scott, Mike. 2014. *WordSmith Tools, Version 6*. Liverpool: Lexical Analysis Software.
- Small, Tamara A. 2011. What the Hashtag? A Content Analysis of Canadian Politics on Twitter. *Information, Communication & Society* 14 (6): 872-895.
- Teisceira-Lessard, Philippe. 2013. "Pastagate": La présidente de l'OQLF démissionne. 08.03.2013. Available: <http://www.lapresse.ca/actualites/politique/politique-quebecoise/201303/08/01-4629080-pastagate-la-presidente-de-loqlf-demissionne.php>.
- Vessey, Rachelle 2013a. *Language Ideologies and Discourses of National Identity in Canadian Newspapers: A Cross-Linguistic Corpus-Assisted Discourse Study*. Unpublished PhD thesis. Queen Mary, University of London.
- Vessey, Rachelle 2013b. Too Much French? Not Enough French? The Vancouver Olympics and a very Canadian language ideological debate. *Multilingua*, 32 (5), 659-682.
- Woolard, Kathryn A. 1998. Introduction: Language Ideology as a Field of Inquiry. In *Language Ideologies. Practice and Theory*, ed. by Bambi B. Schieffelin, Kathryn A. Woolard and Paul V. Kroskrity, 3-50. Oxford: Oxford University Press.
- Wyatt, Nelson. 2013. Pastagate: Quebec to Review Language Law Violation Policies after PR Mess. *Canadian Press*, 02/26/2013. Available: <http://news.nationalpost.com/2013/02/26/quebecs-pastagate-pr-nightmare-story-gets-60-times-more-coverage-outside-province-than-marais-investment-trip/>.
- Zappavigna, Michele. 2011. Ambient Affiliation: A Linguistic Perspective on Twitter. *New Media & Society* 13 (5): 708-806.

ⁱ Although the use of the term "pastagate" arguably perpetuates its social meaning, it is used here for brevity and coherence with the larger context.

ⁱⁱ Other research had found that this trend is changing (see e.g. Crystal, 2011: 78-91).

ⁱⁱⁱ Since many websites are fragmented by the process of scraping Tweets from the web, no websites are included in examples in this paper.