Contact-induced lexical variation in Quebec English: an accountable description

Filip Miletic

To cite this version:
Filip Miletic. Contact-induced lexical variation in Quebec English: an accountable description. 22èmes Rencontres des jeunes chercheurs en Sciences du Langage (RJC 2019), Jun 2019, Paris, France. hal-02280295

HAL Id: hal-02280295
https://hal.science/hal-02280295
Submitted on 6 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Contact-induced lexical variation in Quebec English: an accountable description
Filip Miletic  •  CLE-ERSS: CNRS & University of Toulouse  •  filip.miletic@univ-tlse2.fr

Introduction
QUEBEC ENGLISH (QCE)
• A regional variety of Canadian English (CanE) spoken by a minority of Quebecers;
• Ongoing contact with Quebec French (QF);
• French borrowings and semantic shifts: a secondary phenomenon (e.g. Poplack et al. 2006) or a defining characteristic (e.g. Fee 2008, Bobberg 2012)?

STUDY OF LEXICAL VARIATION
• Lexical variables are routinely examined in dialectology, butch challenging to systematically analyze in corpora;
• It is difficult to identify all possible lexical variants: some propose more generic analyses (e.g. Poplack 1993);
• Corpus-based studies of QCE lexicon are often imprecise (e.g. only reporting raw frequencies of one variant).

Results

IMPOSED DIRECT LEXICAL TRANSFER
• Terms referring to Quebec administration, whose only legally valid version is in French;
• 6 (out of 13) variables fully attested in both cities, while another 6 present only French variants.

Table 1. Lexical variable CEGEP vs. junior college.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Montreal</th>
<th>Toronto</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEGEP</td>
<td>287%</td>
<td>11%</td>
</tr>
<tr>
<td>Junior college</td>
<td>1%</td>
<td>31%</td>
</tr>
<tr>
<td>Total</td>
<td>288%</td>
<td>42%</td>
</tr>
</tbody>
</table>

CULTURALLY-BOUND LEXICAL TRANSFER
• Terms with referents typical of Quebec, with no single established English equivalent;
• 10 (out of 25) variables fully attested in both cities;
• Most lexical choices are geographically shared, with a general preference for French borrowings.

Table 2. Lexical variable francophone vs. parapraphrase (French-speaking, French speaker, speaker of French).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Montreal</th>
<th>Toronto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francophone</td>
<td>198%</td>
<td>14%</td>
</tr>
<tr>
<td>Paraphrase</td>
<td>29%</td>
<td>25%</td>
</tr>
<tr>
<td>Total</td>
<td>227%</td>
<td>170%</td>
</tr>
</tbody>
</table>

ELECTIVE DIRECT LEXICAL TRANSFER
• Terms of French origin for which established English equivalents exist;
• 17 (out of 29) variables fully attested in both cities;
• Significant geographic differences in 8 cases, with French items relatively more frequent in Montreal.

Table 3. Lexical variable cause (populaire) vs. credit union.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Montreal</th>
<th>Toronto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cause (populaire)</td>
<td>11%</td>
<td>8%</td>
</tr>
<tr>
<td>Credit union</td>
<td>5%</td>
<td>92%</td>
</tr>
<tr>
<td>Total</td>
<td>16%</td>
<td>100%</td>
</tr>
</tbody>
</table>

CALQUE
• Translations of French lexical items into English;
• 2 (out of 4) variables fully attested in both cities;
• Limited data, although some geographic differences can be observed.

Table 4. Lexical variable cultural community.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Montreal</th>
<th>Toronto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultural</td>
<td>35%</td>
<td>35%</td>
</tr>
<tr>
<td>Immigrant</td>
<td>13%</td>
<td>65%</td>
</tr>
<tr>
<td>Total</td>
<td>48%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Methodology

RESEARCH AIMS
We examine known contact-induced lexical variables in Montreal (contact) and Toronto (no contact), focusing on:
• their degree of sociolinguistic integration, as reflected by use in formal contexts (newspapers);
• their geographic diffusion outside of Quebec;
• differences between distinct types of contact influence.

VARIABLES
• 100 variables manually defined from previous research (e.g. Bobberg 2012, Fee 1991, 2008, Grant-Russel 1999);
• Variable a set of synonymous lexical variants, at least one of which is typical of QCE and one of CanE;
• The relative frequencies of all variants are examined as per the principle of accountability (Labov 1969).

DATA
News of the Web corpus (online newspaper articles): CanE section, 900m words.

ANALYSIS
• Look up 1,000 most recent occurrences for each variable (all variants combined);
• Extract those from Montreal and Toronto;
• Manually exclude homonymy / polysemy;
• Get relative frequencies for all variants / cities.

Ongoing work: semantic shifts

The observed sparsity of semantic shifts is contradicted by previous work on QCE (e.g. Fee 1991, 2008, Bobberg 2012) and by psycholinguistic evidence of semantic interference in bilinguals (e.g. Romaine 1995).

Possible explanations: corpus type (e.g. semantic shifts may be more stigmatized than borrowings in writing); limits of human annotation.

Ongoing PhD research: distributitional semantic models (computational representations of meaning) applied to the detection of semantic shifts in QCE.

A large dataset of geolocalized tweets (in preparation) will be used to train the distributional models.

A sociolinguistic study will be conducted in Quebec in order to evaluate the computational analysis.

Conclusions

48 out of 100 variables were fully attested (i.e. all variants were found in both cities):
• the presence of these contact-induced variants suggests they are sociolinguistically integrated;
• the absence of other variants may reflect the formality of the written context.
Most attested contact-related variants are used in both Montreal and Toronto:
• their relative frequency tends to be higher in Montreal;
• they appear to be typical of Quebec, but not exclusive to it.

Distinct usage patterns linked to different types of contact influence:
• the differences are generally not clear-cut;
• exception: culturally-bound lexical transfer, with a generally geographically shared preference for items deriving from contact.

Issues such as semantic shifts may benefit from computational methods. Ongoing work will shed more light on this matter.

References