Economic and medical discourses under register variation in English and French languages: Terms and lexical bundles analysed through corpora for pedagogical purposes

Tiffany Jandrain
PhD Student
tiffany.jandrain@umons.ac.be

International Conference *Specialized Languages in Teaching and Translation: Theory and Practice*

17/11/2017, Adam Mickiewicz University, Poznań
Introduction

Problems of the transposition of “register, genre, style,...” in translations made by students...

... with still much work to be done in that field

So, what do we need?

→ A need to **characterise registers** in EN and in FR by **identifying patterns** of register features (based on Gledhill & Kübler 2016)

< recurrence - frequency
Introduction

For this study:

Phraseology in two different specialised languages associated to different registers
Outline

1. A theoretical overview
   - Register variation (RV), LSP and specialised discourse, popular discourse, Anglicisation, lexical bundles

   ➔ **Hypothesis:** To which extent are terms and lexical bundles of Eco. and Med. LSP marked by RV in specialised and popular texts in English and in French?

2. Methodological steps

3. Results and discussion

4. Conclusion
1. A theoretical overview

- Register variation
- LSP and specialised discourse
- Popular discourse
- Terms and lexical bundles
Register variation

A register

= “a variety of language that a language user considers appropriate to a specific situation”

(Baker 2011: 13)
Register variation

Register variation (RV)

= “functional or diaphasic variation determined by the situational context”

(Neumann 2014: 2)

→ Language is used in a particular way to fulfil communicational functions

(Biber & Conrad 2009)
Register variation

Register ≠ Genre

“a set of texts, spoken and written, which are institutionalized in so far as they are considered by a given speech community to be of the same type, for example the genre of political speeches or the genre of editorials”

(Baker 2011: 302)
LSP and specialised discourse

Definition:
"LSP is the language that is used to discuss specialized **fields of knowledge**. It is actually more accurate to talk about LSP in the **plural** (...) since different LSPs are used to describe different areas of specialized knowledge (...). Of course, there is some **overlap** between LGP [Language for General Purposes] and LSP. LSP typically contains a number of specialized **terms**; however, a lot of **general words** show up in LSP conversations. (...) An LSP may also have special ways of **combining** terms or of **arranging** information that differ from LGP."

(Bowker & Pearson 2002: 25-26; my emphases)
Specialised and popular discourses

SPE

POP

SPE

POP
Popular discourse

Defining features:

- Texts are written for a **heterogeneous public** (composed of specialists and non-specialists);
- Texts are written by **various authors** (journalists as well as members of other communities);
- Texts are **not expressed in usual terms**;
- Texts are published on **various supports and interactions**;
- Texts are **not peer-reviewed**.

(Delavigne 2003)
Popular discourse

Terminology  Popular texts

Phraseology? (Gledhill & Kübler 2016)

(Delavigne 2003)
Popular discourse

SPE discourse

Linguistic features

Specific expressive functions

Recontextualisation

Gotti (2014)
LSPs analysed in this study

Two LSPs:

- Economic LSP
- Medical LSP

→ Anglicisation

- EN as *lingua franca*
- EN flexibility and concision
- Supremacy of the USA in those disciplines

(Faure 2010, Soubrier 2016)
Anglicisation

Anglicism:

“an English word or phrase that is used in another language” (Cambridge Dictionary)

- Direct borrowings
- Semantics (~ false friends)

• Semantics, syntax, morphology, phonetics (Soubrier 2016)
Anglicisation?

- **Terminology**

  - Term, i.e. « word, phrase, symbol or formula expressing a concept specific to a field of use »

    (Pavel & Nolet 2001: 117; my translation)

- **Phraseology**

  - **Lexical bundles**

    (based on Grabowski 2013)
Lexical bundles

Lexical bundles =
“multi-word prefabricated expressions”

(Biber et al. 2005: 372)

"the most frequent recurring lexical sequences in a register"

(Biber et al. 1999 in Biber et al. 2005: 376)

”sequences of three or more words that occur frequently in natural discourse and constitute lexical building blocks used frequently by language users in different situational and communicative contexts (Biber et al. 1999: 990-991)”

(Grabowski 2013: 393)

3 primary functions: stance expressions, discourse organizers, referential expressions

(Biber et al. 1999 in Lefer & Grabar 2015)
Lexical bundles

Examples:

• Referential expressions: too much of
• Discourse organizers: in the same way
• Stance expressions: it is not surprising that

Lefer & Grabar 2015
Hypothesis

→ **Hypothesis:** To which extent are terms and lexical bundles of Eco. and Med. LSP marked by RV through anglicisation in specialised and popular texts in English and in French?
LSPs analysed here

Topics in this study:
- Economic LSP → property (real estate)
- Medical LSP → diabetes
2. Methodological steps

- Corpus Linguistics
- Corpus compilation
- Software
Methodology

A *corpus-driven* approach

Corpus

= “a large collection of *authentic* texts that have been gathered together in *electronic* form according to a *specific* set of *criteria*”

(Bowker & Pearson 2002: 9)

“texts (...) selected (...) in order to be used as a *representative* sample of a particular language or subset of that language”

(ibid.: 10)
Methodology

Corpus linguistics
= a methodological approach to study language use

(Bowker & Pearson 2002: 9)

→ To study language as a social phenomenon, based on actual data
(Robert de Beaugrande)

→ Functionalist approach (J.R. Firth, M.A.K. Halliday, J. Sinclair)
(Zanettin 2012)
Corpora

1. ECO – SPE – EN
2. ECO – SPE – FR
3. ECO – POP – EN
4. ECO – POP – FR
5. MED – SPE – EN
6. MED – SPE – FR
7. MED – POP – EN
8. MED – POP – FR
Corpora

8 corpora

ECO

SPE

POP

MED

SPE

POP

EN

FR

EN

FR

EN

FR

EN

FR
Corpus compilation: ECO sources

1. ECO – SPE – EN


2. ECO – SPE – FR


3. ECO – POP – EN


4. ECO – POP – FR

Le Monde, L’Express, Le Figaro, Libération, La Tribune, Forbes, Le Parisien, Pap.fr, L’OBS, MagImmo (French of France)
Corpus compilation: MED sources

5. MED – SPE – EN


6. MED – SPE – FR


7. MED – POP – EN


8. MED – POP – FR

Le Figaro, L’Express, OMS, Association Belge des Diabétiques, Fédération Française des Diabétiques, Top Santé, e-Santé, Le Monde, Futura Sciences, Pour la Science (→ French of France)
Corpus compilation: criteria

<table>
<thead>
<tr>
<th>SPE Corpora</th>
<th>POP Corpora</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 extracts of research articles (500 words)/corpus</td>
<td>30 extracts of press articles (500 words)/corpus</td>
</tr>
<tr>
<td>Originals (EN – FR)</td>
<td>Originals (EN – FR)</td>
</tr>
<tr>
<td>Written by specialists</td>
<td>Written by natives</td>
</tr>
<tr>
<td>In Am/Br EN or Fr/Be French</td>
<td>In Br EN or Fr French</td>
</tr>
<tr>
<td>From 2007 to 2017</td>
<td>From 2007 to 2017</td>
</tr>
<tr>
<td>From 10 sources (3 x 10)</td>
<td>From 10 sources (3 x 10)</td>
</tr>
</tbody>
</table>

= a total of 240 texts (120,000 words)

Criteria based on Bowker & Pearson 2002
Software AntConc©

By Laurence Anthony

http://www.laurenceanthony.net/software.html
Methodological steps

LB $\rightarrow$ 4-word sequences (Grabowski 2013)
appearing min. 4 times
in min. 4 different texts
Software AntConc©

By Laurence Anthony

http://www.laurenceanthony.net/software.html
3. Results and discussion
Results
Results

At first glance,
MED > ECO

In MED: FR > EN in POP
FR = EN in SPE

In ECO: FR = EN in POP
FR > EN in SPE

MED seems more formulaic (POP > SPE)
Results

But the majority of 4-word sequences are related to the **topics** of the texts:

*le diabète de type, de sucre dans le, glucose dans le sang, une activité physique régulière, action de l’insuline, of type 2 diabetes, risk of type 2, of people with diabetes, type 2 diabetes in, their blood sugar levels, linked to being overweight, ...

*Prix de l’immobilier, marché de l’immobilier, le marché de l’, de la crise financière, la baisse des prix, ...

Except **ECO – SPE – EN:**

*At the same time, for the first time, the first time in*
## Results

<table>
<thead>
<tr>
<th></th>
<th>Lexical Bundles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MED EN</strong></td>
<td><em>has been shown to, is one of the, studies have shown that, a major cause of</em></td>
</tr>
<tr>
<td></td>
<td><em>at the university of, people in the UK, the number of people, published in the journal</em></td>
</tr>
<tr>
<td><strong>MED FR</strong></td>
<td><em>Au cours de la</em></td>
</tr>
<tr>
<td></td>
<td><em>Qu’est-ce que, est-ce que le, au cours de la, l’action de l’</em></td>
</tr>
<tr>
<td><strong>ECO EN</strong></td>
<td>*/</td>
</tr>
<tr>
<td></td>
<td><em>At the same time, for the first time, the first time in</em></td>
</tr>
<tr>
<td><strong>ECO FR</strong></td>
<td><em>De plus en plus</em></td>
</tr>
<tr>
<td></td>
<td><em>Ce n’est pas, pour la première fois</em></td>
</tr>
</tbody>
</table>
Results

Tiffany Jandrain
Economic and medical discourses under register variation
Conclusion
Conclusion

- An exploratory analysis of LB in two different LSPs; those small corpora showed that MED seems more formulaic than ECO, but for the sequences with one or more terms, not for the LB (the opposite!)
- Apparently no anglicisation for those LB (with some shared structures)
- Limits of the study
- Register variation
Conclusion

For future research:

• A bigger corpus
• Anglicisation with more LB, with collocates
• LB: referential expressions especially
• A comparison between LB and collocates
• A comparison between LSPs of medicine and between LSPs of economics
• A comparison between originals and translations
References


Thank you for your attention!

Tiffany JANDRAIN
tiffany.jandrain@umons.ac.be
UMONS, Belgium