Jean-Marie Lessard
Senior Advisor, Department of Justice Canada
Francophonie, Justice in Official Languages and Legal Dualism
Editor and Co-author of the
Legal Dictionary of Property in Canada (LDPC)
with Anne Des Ormeaux
jlessard@justice.gc.ca

Conference: Using Corpora in Contrastive and Translation Studies (UCCTS 2010)
Edge Hill University, St Helens Road, Ormskirk, Lancashire, L39 4QP, UK

Bilingual and Bisystemic Legal Dictionary (1)

Using judicial bilingual and bisystemic corpora
- Canadian federal law in private law matters is expressed with equal authority in both official languages and must reflect the Civil Law and Common Law systems. Since both the French and English carry equal authority, both versions should be taken into account for a full understanding of the law.

Promoting Canada’s unique contribution to legal dualism
- … to the scientific community and reflect the advances made in all areas of expertise that are relevant to its activities.
- Attending UCCTS conference and presenting the Dictionary is part of the Legal Dualism Team’s mandate.
More than 400 copies of the Dictionary were sent to specialists in Canada, England, United States and France for feedback.

We still have a few copies of the Dictionary in print version for those who would like to examine it and share their comments with us.

Is it really a dictionary?

This is a question we could debate. We consider it to be a dictionary prototype, and our nomenclature is a working hypothesis that current and future corpora will enable us to verify.

This is the first of a four volume dictionary, and the corpus is in its initial phase.
Ownership as a phenomenon

**Property:** There is no property without an owner and there is no significance in owning a thing with no value.

Any thing that can be appropriated due to its use or exchange value, whose ownership is opposable to a third party and whose possession is prescriptible by law. They are objects of legal commerce. (LDPC)

**Knowledge:** To learn about the phenomenon, the relationship itself and the terms of this relationship must be known: the protagonist is the functional subject of the relationship.

**Perspective:** Categorial knowledge of legal objects is inseparable from the juridical act or fact. Each legal object is known as it is appropriated, enforced and prescribed.

**Legal dualism:** Exists as it appears in the qualitative opposition between the DC and CL systems in the Canadian legal framework.

Moodle as a collaboration platform

**Community of interest**
Anything we can learn on the use of bilingual and aligned corpora provides us with insights and drives our efforts.

**The collaboration platform**
The aim is to foster interaction between Canadian researchers, but it is open to everyone interested in legal dualism issues.

**Moodle** was selected because of its strong popularity among teaching professionals and its growing use at universities around the world.
Building corpora for the Dictionary (1)

Selecting the situation of communication

- The Dictionary’s corpus **consists exclusively of decisions from Canada’s appellate courts**.
- The Dictionary’s corpus is **not a reference corpus** for studying the language of the legal corpus on property.
- It is a **specialized language corpus** related to a specific area of knowledge: Canadian Property Law.

Building corpora for the Dictionary (2)

Selecting the texts for the corpus

- (1) the **open-access corpus** (or available corpus) includes all the decisions of the selected courts (79727 texts)
- (2) the **total corpus** is a subset of the open-access corpus (or available corpus) and consists of texts selected for the purpose of analyzing the area of property law knowledge (20627 texts) and
- (3) the **bilingual corpus** is the section of the total corpus that incorporates all bilingual decisions (4232 texts).
Distribution of texts, types and tokens

<table>
<thead>
<tr>
<th>Court</th>
<th>Period</th>
<th>Available Corpus (AC)</th>
<th>Total Corpus (TC)</th>
<th>% Total Corpus from Available Corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supreme Court</td>
<td>1980-2009</td>
<td>2563</td>
<td>1700</td>
<td>68%</td>
</tr>
<tr>
<td>Federal Court of Appeal</td>
<td>1980-2009</td>
<td>10343</td>
<td>1878</td>
<td>18%</td>
</tr>
<tr>
<td>Court of Appeal of New Brunswick</td>
<td>1980-2009</td>
<td>7070</td>
<td>762</td>
<td>11%</td>
</tr>
<tr>
<td>Court of Appeal of Quebec</td>
<td>1990-2009</td>
<td>31742</td>
<td>6447</td>
<td>11%</td>
</tr>
<tr>
<td>Court of Appeal for Ontario</td>
<td>1980-2006</td>
<td>9555</td>
<td>2168</td>
<td>23%</td>
</tr>
<tr>
<td>Manitoba Court of Appeal</td>
<td>1980-2009</td>
<td>1249</td>
<td>512</td>
<td>73%</td>
</tr>
<tr>
<td>British Columbia Court of Appeal</td>
<td>1990-2009</td>
<td>17072</td>
<td>7073</td>
<td>45%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>79727</strong></td>
<td><strong>20827</strong></td>
<td><strong>28%</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Court</th>
<th>Period</th>
<th>Bilingual Corpus (BC)</th>
<th>Bilingual Corpus Words</th>
<th>Bilingual Corpus Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supreme Court</td>
<td>1980-2009</td>
<td>1776</td>
<td>120031</td>
<td>39643969</td>
</tr>
<tr>
<td>Federal Court of Appeal</td>
<td>1980-2009</td>
<td>1724</td>
<td>37967</td>
<td>16736553</td>
</tr>
<tr>
<td>Court of Appeal of New Brunswick</td>
<td>1980-2009</td>
<td>959</td>
<td>18733</td>
<td>5801797</td>
</tr>
<tr>
<td>Court of Appeal of Quebec</td>
<td>1980-2009</td>
<td>51</td>
<td>26953</td>
<td>978769</td>
</tr>
<tr>
<td>Court of Appeal for Ontario</td>
<td>1980-2009</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Manitoba Court of Appeal</td>
<td>1980-2009</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>British Columbia Court of Appeal</td>
<td>1980-2009</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>4232</strong></td>
<td><strong>157810</strong></td>
<td><strong>62035190</strong></td>
</tr>
</tbody>
</table>

Corpus as a testing environment

The corpus is seen as an environment to evaluate the definitions as describing attested senses.

No description tags were added to the files. The texts are identified only by their filenames. The aligned files are HTML with UTF-8 character encoding.

**Filenames consist of four parts**

- (1) the year of the decision
- (2) the court of appeal of origin (for Supreme Court only)
- (3) the number assigned to the decision and
- (4) the direction of the alignment

For example, 1980_1_SCR_122_FR-EN indicates that, in 1980, the Supreme Court of Canada, rendered decision number 122c, and the alignment direction was French first.
The structure of the judgements

DtSearch search engine software and Terminotix’s LogiTerm alignment software were selected.

Source files indexed with dtSearch retain their format and alignment in the search results.

Catchwords, Summary, Cases Cited, Statutes and Regulations Cited, etc. were excluded from indexing.

The abridged Conclusion part of the judgement was indexed.
Searching with dtSearch (1)

Composing dtSearch queries

The dtSearch search engine can handle very complex queries.

**Definitional markers** the use of definitional markers such as “is a”

**Frequency analysis and lemmatization of headwords occurrences**

A different process based on frequency analysis and conjunction of bilingual forms of dictionary entries, as well as forms from the “semantic” part of articles.

The results of the examination of simple forms are displayed under the lemma, indicating the frequency of each of the forms encountered.
Concept of maximum frequency

- The maximum frequency of a compound word is defined by the number of occurrences of its least frequent form. The same applies to bilingual equivalents.

- The number of occurrences of bilingual headwords, used together in a bilingual equivalent, cannot exceed the frequency of the least frequent headword or lemma.

- By combining the frequency of English and French headword forms, the maximum frequency of equivalents is easily calculated.
Searching with dtSearch (3)

Lack of bilingual contexts

- In practice, verifying acceptations involves finding contexts that best illustrate the concept described in the dictionary entry. If we cannot find a perfect example of how the concept is used, we select the excerpts most likely to clearly illustrate its meaning.

- The lack of an equivalent in the other language results in observing a large number of variants. In those cases we must use the unilingual part of the corpus to find an excerpt containing a possible equivalent.

- Translation is sometimes necessary for an excerpt to appear in the dictionary.

Definition and ontological consistency

Building definitions

- A short list of dictionaries were systematically consulted. They were used to produce computerized records of the semantic features of their definitions.

- The process involved documenting the semantic features that would enable us to provide sufficient evidence of the consistency, as a whole, of our dictionary’s definitions.

- Each definition is a “condition nécessaire et suffisante” leading to ontological consistency of property law.
### Property Rights

**Cat.** (Civ. Law) proprietary rights; (DC) right

**Incl.** (Civ. Law) property interest, security interest

**Prot.** (DC) Title (SL) Title

**Graph.** (SL) Sl

<table>
<thead>
<tr>
<th>Source:</th>
<th>Design Services Ltd. v. Cantral (2009) 3 S.C.R. 888</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>Clause</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>Property rights are a property interest of the owner in, or a security interest on, the property or a security interest in the property.</td>
</tr>
<tr>
<td>(2)</td>
<td>Approval of an agreement is subject to the approval of the owner of the property.</td>
</tr>
</tbody>
</table>

### Action

**Cat.** (Civ. Law) action; (DC) action against an insurer

**Incl.** (Civ. Law) personal action, personal action

**Prot.** (Civ. Law) holder (SL) holder (SL)

### Graph.** (Civ. Law) **Corpus**

<table>
<thead>
<tr>
<th>Source:</th>
<th>Design Services Ltd. v. Cantral (2009) 3 S.C.R. 888</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>Clause</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>Action is a legal proceeding to recover personal property in the context of an obligation or a security interest.</td>
</tr>
<tr>
<td>(2)</td>
<td>Approval of an agreement is subject to the approval of the owner of the property.</td>
</tr>
</tbody>
</table>
Moving forward

Next steps...

- We plan to increase the number of entries to over 1000, as each volume was initially supposed to contain about 250 bilingual and bisystemic entries.

Internet version and complementary studies

- Publishing the Dictionary online allowed us to add more than 900 graphics to illustrate the conceptual relationships described in the Dictionary.

- The search for definitional contexts led us to compile the definitions cited in decisions. Interestingly, the definitions cited are primarily from general language dictionaries.

![Ontology graphs and Etymology](image-url)
Conference: Using Corpora in Contrastive and Translation Studies (UCCTS 2010)

Edge Hill University, St Helens Road, Ormskirk, Lancashire, L39 4QP, UK