DiCoInfo Exercise: Level I

Introduction

The DiCoInfo, developed by a team led by Marie-Claude L'Homme of the ÉCLECTIK research group at the Université de Montréal's *Observatoire de linguistique Sens-Texte* (OLST), is a specialized online dictionary of the Internet and information technology. Developed primarily in French, it describes a relatively restricted set of basic terms in these fields identified in corpora of mainly popularized and didactic texts. However, its nomenclature and the information in its entries are continually expanding.

The DiCoInfo was developed in the framework of lexico-semantic terminology (cf. L'Homme's *La terminologie : principes et techniques*, 2004), and draws on principles of Explanatory and Combinatorial Lexicology (ECL) (e.g. Mel'cuk, Clas and Polguère, *Introduction à la lexicologie explicative et combinatoire*, 1995). It also integrates some principles from a methodology developed as part of the FrameNet project (http://framenet.icsi.berkeley.edu/) for annotating contexts (i.e. labeling key elements of meaning in the example sentences provided in the entries).

Because of this perspective, the DiCoInfo describes terms not only using some familiar kinds of information (e.g. contexts, sources, synonyms) but also using some more unusual types of information (e.g. representations of structures in which terms appear, co-occurrents and collocates, related terms). In addition, the dictionary focuses mainly on single-word units that can be combined as needed to create more specific terms, rather than listing each individual combination of these items as a separate multi-word term, as conventional term banks (e.g. TERMIUM, the *Grand dictionnaire terminologique* (GDT)) generally do. Because of this perspective, the DiCoInfo has a higher proportion of entries for single-word terms and specialized adjectives and verbs than you might expect to find in term banks such as TERMIUM and the GDT (which tend to concentrate mostly on noun phrases).

The DiCoInfo was recently expanded to include a number of English terms and link them to the corresponding French entries; a Spanish component is also being developed.

You can learn more about DiCoInfo by consulting documentation available on its website (http://olst.ling.umontreal.ca/dicoinfo/ > **DOCUMENTATION**), on the Web page of the OLST (http://olst.ling.umontreal.ca/) and on Marie-Claude L'Homme's home page (http://www.ling.umontreal.ca/lhomme/).

Doing these exercises will help you learn to:

- Explore the DiCoInfo search interface,
- Browse through term lists,
- Search for terms,
- Interpret the DiCoInfo entries, and

- Compare the DiCoInfo entries to corresponding records in "traditional" term banks such as TERMIUM and the GDT.

Getting ready

1. Open the DiCoInfo:

    a. Open the Web browser Firefox using the shortcut on the desktop and type in the address http://olst.ling.umontreal.ca/dicoinfo/, or simply click on the hyperlink.

    NOTE: You can also use Internet Explorer as your browser, but some parts of the DiCoInfo entries will not be available or displayed properly. Using Firefox is best.

    b. Click on the **RECHERCHE** button to access the search page.

    NOTE: The DiCoInfo interface is currently only available in French, but will soon be developed in other languages.

Exploring the DiCoInfo search interface

1. You can type the term you are going to look for in the *Terme à rechercher* field.

2. You can set the search mode (*Mode de recherche*) to search for terms (**TERME**) or for lexical relations (**LIEN LEXICAL**) (i.e. to find entries in which the string you enter appears as a term related to the entry headword) by clicking the corresponding radio button.

3. You can set the language (*Langue*) to search in to **FRENCH**, **ENGLISH** or **BILINGUAL** (to search for the string in both languages) by clicking the corresponding radio button.

4. You can set the precision level of the search (*Niveau de precision*) to search for the exact term (**TERME EXACT**), for terms starting with the search string (**TERME COMMENÇANT PAR**) or for terms containing the search string (**TERME CONTENANT**) by clicking the corresponding radio button.

    NOTE: The interface is currently designed primarily to search for single word forms. Even if you do a search for an exact term, if you enter two or more word forms in the search field, the results will include any or all of the forms entered. (That is, the search tool automatically uses an OR operator between the two strings. For example, if you enter *pavé* and *clavier* in the search field, you will find entries such as *mini-clavier*, *clavier*, *pavé*, and so on, rather than only terms containing both items.) This approach allows you to obtain the most inclusive results for multiple terms of interest to you, but of course means that you will need to evaluate the results to find the most appropriate entries for your search.

5. If you are not sure what term you want to search for, you can display an alphabetical list of all the terms listed in DiCoInfo by clicking on the **LISTE ALPHABÉTIQUE DES TERMES** link.

Browsing through term lists

1. In the search interface, click on the **LISTE ALPHABÉTIQUE DES TERMES** link.

2. Display the list of terms in French

    a. Click on the **TERMES FRANÇAIS** link to display the list of French terms in the DiCoInfo.

b. Click on a letter to display the list of terms starting with that letter. Try looking at the letter *C*.

c. What kinds of terms do you find in this list? What part-of-speech categories do they belong to? Are they single-word or multi-word terms, or both?

d. Click on a term to display the corresponding entry and take a quick look. (We will explore the entries in detail in a later section.)

e. Click on the **RETOUR À LA LISTE ALPHABÉTIQUE DES TERMES** link at the top of the page to go back to the list of terms.

3. Display the list of terms in English.

a. Click on **TERMES ANGLAIS**.

b. Click on a letter to display the list of terms starting with that letter. Try looking at the letter *C* in English as well.

NOTE: Since the DiCoInfo was originally created in French, it is currently more comprehensive in that language.

c. Click on a term to display the corresponding entry, and compare it to the French entry you examined.

d. Click on **RETOUR À LA LISTE ALPHABÉTIQUE DES TERMES** at the top of the page to go back to the list of terms.

Searching for terms

1. Click on the **RETOUR À LA RECHERCHE** link at the top left hand corner of the screen to return to the search interface.

Do a search for the term *zip* in English.

a. Type *zip* in the search field (*Terme à rechercher*).

b. Select English as the searching language. (*Langue*: **ANGLAIS**)

c. Set the level of precision (*Niveau de precision*) to exact term (**TERME EXACT**).

d. Click the **RECHERCHER** button.

e. The entry for the term that exactly matches your search, *zip*, will be displayed.

f. Scroll down to view the entry.

NOTE: After each search, the search settings return by default to search for terms in either French or English that begin with the string you enter in the search field. (This is intended to provide a relatively broad search to ensure that you don't miss too many results.) If you want to change these settings, you'll need to do so for each search.

2. Do a search for all entries containing *zip* in English and in French.

a. Scroll back up to the top of the page and type *zip* in the search bar (*Terme à rechercher*).

b. Select bilingual as the searching language. (*Langue:* **BILINGUE**)

c. Set the level of precision (*Niveau de precision*) to search for terms containing the search string (**TERME CONTENANT**).

d. Click on the **RECHERCHER** button.

e. Results will display a list of matching entries. You can click on a term in the hyperlinked list to jump to the corresponding entry, or you can scroll down through the entries to see all the results.

f. How are the results different from those of the previous search?

Do a search for the term *écrire* in French.

g. Type *écrire* in the search field (*Terme à rechercher*).

h. Select French as the searching language. (*Langue*: **FRANÇAIS**)

i. Set the level of precision (*Niveau de precision*) to exact term (**TERME EXACT**).

j. Click the **RECHERCHER** button.

k. Two entries for the term that exactly matches your search, *écrire1* and *écrire2*, will be displayed. Each of these represents a different sense of the verb *écrire*.

l. Scroll down to view the entries.

Interpreting entries in the DiCoInfo

1. Look at the information provided in the default view of the entries for *écrire*:

a. The term is followed by a number, which is used to number senses in case more than one meaning is identified for a form. You'll see that there are two senses identified for this verb: *écrire1* and *écrire2*. Can you immediately see the difference between these two meanings? If not, exploring the entries should help you to discover how they are different.

b. The terms are also accompanied by their actantial structure (S*tructure actancielle*). This describes the term's actants (other terms that are "linked" by their meanings to the central term), and how they are connected.

i. In this case, we see that the meaning of *écrire1* involves three other entities: an agent who writes something, a patient that is written, and a destination, the thing on which the patient is written.

ii. For *écrire2*, we also find three actants. The first two are similar to those of *écrire1*, but the third is not a destination but rather the means of doing the writing. Does this help you to start to see what the difference between the two senses may be? Why or why not?

iii. The description also identifies (in curly brackets) one or more terms that are often used to name these actants: for *écrire1*, for example, the agent who does the writing may be a *programme* or a *processeur*, the patient that is written may be *données*, and the destination to which the data is written may be the *mémoire* or a *support de stockage*.

iv. For *écrire2*, the agent may be a *programmeur*, the patient a *programme*, and the means a *langage*. Does this information help you to see the difference between the two senses? Why or why not?

2. However, this is not all the information available about the actants. Click on the **RÉALISATIONS LINGUISTIQUES DES ACTANTS** link to see other possible actants for these two terms.

a. How might this information be useful for you as a translator or writer?

NOTE: These are <u>semantic</u> actants, integral parts of the meaning of the term. That is, it would be impossible to explain the meaning of the term *écrire* in either of its senses without referring, for example, to an agent who does the writing, or the patient that is written.

3. Under the list of actants, a definition of the term is provided. This definition shows how these different actants can be combined to describe the meaning of the term. Is it easier to see how the actants fit into the terms' meaning in this kind of structure? How does it compare to other definitions you have seen?

NOTE: Definitions are available for some, but not all, of the entries in the DiCoInfo. They are gradually being added.

4. To view one or more contexts in which the term is used, click the **CONTEXTS** link.

   a. Do these contexts help you to better understand the terms? If so, how? Can you identify the actants in any of the contexts? Are all of the actants always present at the "surface" level (i.e. represented by words or other expressions in the sentence)?

NOTE: While it would be impossible to explain the meaning of the term _écrire_ without referring to its semantic actants, these actants are not always expressed at the surface level every time the term is used. For example, we can use the term _écrire_ without specifically mentioning who or what is doing the writing.

5. Click the **LIENS LEXICAUX** link to view the list of terms identified as being related in various ways to the term _écrire_. A two-column table appears.

   a. The right-hand column lists the related terms. Terms that are described in the DiCoInfo are hyperlinked to the appropriate entry. As in all DiCoInfo entries, in case more than one entry exists for a form, the sense of each item is numbered.

   b. What kinds of terms are identified as related to _écrire_? To what part of speech categories do they belong?

   c. The left-hand column describes the relationship between the main term (_écrire_) and the related terms. This is a fairly specific description of the relationship.

   d. What kinds of relationships are identified? Is it easy to understand the explanations? How might knowing about these relationships be useful for you as a translator or writer?

6. To see a description of the relationship between these related terms at a more general level (e.g. that could apply to several different pairs of terms), click on the **RÔLES ACTANCIELS** link above the table. A third column appears in the table.

   a. How do these descriptions compare to the default explanations?

   b. When/for what purposes do you think having the two types of explanations might be useful? (If you wish, consult the entries for a few more terms to get a better idea of the kinds of variations you may observe.)

7. To see an even more generally applicable description of the relationship between these related terms, click on the **FONCTIONS LEXICALES** link above the table. A fourth column appears in the table.

NOTE: The codes that appear represent Lexical Functions (LFs), which are used in Explanatory and Combinatorial Lexicology (cf. Mel'cuk, Clas and Polguère, 1995). LFs reflect a classification of recurrent types of relationships between lexical units. The general idea behind this classification is that similar relationships between various pairs of lexical units can be observed in a wide range of languages, and observing parallels can help us to better study how particular kinds of meaning are expressed and to make links between different instances to assist in expressing these meanings idiomatically. We will not go into Lexical Functions in detail here, other than to note that the choice and classification of relationships between terms represented in the DiCoInfo are based in part on these LFs.

LFs may not be available for all DiCoInfo entries.

8. In the bottom right corner of the entry, codes indicate the author(s) who contributed to the entry and the date it was created or modified. In addition, the *Statut* code indicates the completeness of the entry. The lower the code, the more complete the entry. For example, the records for *écrire* are Status 0, which means that they are complete, approved and finalized.

9. The English terms *write1* and *write2* have been identified as equivalents of *écrire1* and *écrire2*. Click on the **ÉCRIRE1** and **ÉCRIRE2** links to view the English entries.

   a. The English entries are laid out similarly to the French entries for *écrire*. Click on the links to view the complete entries. (As indicated by their Status 2, they are still in progress and so don't contain quite as much information as the French entries.)

   b. We see here that the equivalents of the two senses of *écrire* have the same form. How important do you think it is, then, to have two separate entries for these terms? Why do you think so?

   c. Do you see any parallels in the related terms for *écrire* and *write*? How might these be useful for a writer or translator?

10. Do an exact search for the French term *signet*.

11. Look at the information provided in the entry. There are a few differences from the entries for *écrire*.

    a. The synonym *marque-pages* has been identified and is displayed in a *Synonyme(s)* field.

    b. The status of this entry is 1 instead of 0 for *écrire* or 2 for *write*. In your opinion, what work might remain to be done on this record to make it Status 0? (If you want to find out if you are right, you can consult p. 12 in the user's guide at http://olst.ling.umontreal.ca/dicoinfo/manuel-DiCoInfo.pdf.)

12. Click the **LIENS LEXICAUX** link to view the list of terms identified as being related in various ways to the term *signet*.

    a. What kinds of terms are identified as related to *signet*? To what part of speech categories do they belong?

    b. What kinds of relationships are identified? Is it easy to understand the explanations? How might knowing about these relationships be useful for you as a translator or writer?

    c. For *signet*, the entry indicates a synonym and a near-synonym (*quasi-synonyme*), in different parts of the entry. Why do you think this distinction is made? Why might it be important for a translator or writer?

13. The English term *bookmark* (in the first of two entries for this form) has been identified as the equivalent of the term *signet*. Click on the **BOOKMARK1** link to view the English entry.

    a. The entry is laid out similarly to the French entry for *signet* (although it has a different status — 2 instead of 1 — and so does not contain quite as much information). Click on the links to view the complete entry.

    b. Do you see any parallels in the related terms for *signet* and *bookmark*? How might these be useful for a writer or translator? How does having them represented in this way help you to evaluate them (e.g. as compared to how these are represented in conventional term banks or bilingual dictionaries)?

14. In addition to the entry for *bookmark1*, we also see a second entry *bookmark2*.

a. How is this second entry different from the one for *bookmark1*? Are these terms related? How? How can you tell? How might knowing about this relationship be useful for you as a writer or translator?

Comparing entries in the DiCoInfo to similar entries in TERMIUM and the GDT

1. Do a search for French terms containing the term *écran*. Look at the entries for the three senses of *écran* (alone).

   a. Can you tell the difference between the senses?

   b. What parts of the entries help to clarify the difference between the senses? What parts help you to see how they are related?

   c. Do you think it is important for a translator or writer to recognize the difference between these senses? Why or why not? What about for a terminologist?

2. Look up *écran* in TERMIUM and in the *Grand dictionnaire terminologique*.

   NOTE: For instructions and tips on how to search effectively in TERMIUM and the GDT, you can consult the CERTT tutorials *TERMIUM Exercise, Level I* and *Exercices avec le Grand dictionnaire terminologique* on the CERTT site (ACCESS BY TOOL NAME > TERMIUM / GRAND DICTIONNAIRE TERMINOLOGIQUE).

   a. Do you find records in these term banks for the term? For which senses (if any)?

   b. How does the description of the term in the term banks differ from that in the DiCoInfo? How is it similar?

   c. Do you have an opinion about whether one of the two formats would be more useful than the other for a translator or writer? If so, which one, for what kinds of uses, and why?

3. Do a search for French terms containing *pavé* in the DiCoInfo and in TERMIUM and/or the GDT. (In the latter two, concentrate on the records relating to computers / information technology. Consult the tutorials indicated above for tips on how to design effective searches if need be.)

   a. What sense(s) of the term is/are present? Are the senses of the term in these resources the same? How can you tell?

4. Do a search for French terms containing *numérique* in the DiCoInfo and in TERMIUM and/or the GDT, again concentrating on the field of computers / information technology.

   a. What sense(s) of the term is/are present? Are the senses of the term in these resources the same? How can you tell?

5. Find the entries (if any) for the French term *pavé numérique* in the DiCoInfo and in TERMIUM and/or the GDT.

   a. Which of these resources have an entry for this term? Which do not?

   b. What kind of relationship is present between the two components of this term? Can you paraphrase this relationship?

6. Do a search in the related terms section of the DiCoInfo entries for French terms containing *numérique*.

   a. Enter *numérique* in the *Terme à chercher* field.

    b.   In *Mode de recherche*, choose the **LIENS LEXICAUX** option.

    c.   In the *Langue* section, choose the **FRANÇAIS** option.

    d.   In the *Niveau de précision* section, choose the **TERME CONTENANT** option.

    e.   Click the **RECHERCHER** button to start the search.

7. Look at the search results. What kinds of terms has *numérique* been linked to? In what ways? Can you find a link between *pavé* and *numérique*?

NOTE: **This exercise helps to demonstrate the difference in the approaches used in more conventional term banks (e.g. the GDT, TERMIUM) and the DiCoInfo. While the former tend to store multi-word terms (e.g. *pavé numérique*) as independent units, in the DiCoInfo the terms are broken down and stored as components that can be combined in regular patterns to produce multi-word terms. (For example, the sense of *pavé* and the sense of *numérique* occur in various multi-word terms with little change in their meaning. Rather than storing each of these terms separately, the DiCoInfo stores the components and shows how they can be combined in regular ways to form multi-word terms such as *pavé numérique*, but also, for example, *pavé alphabétique* and *touche numérique*.)**

The exception to this policy is when the meaning of a multi-word term is different from or more than the sum of the meanings of its parts. (This is referred to as *non-compositional* meaning.) In these cases (e.g. *disque dur*, which has a richer meaning than the combination of the meanings of *disque* and *dur* alone can produce), the DiCoInfo does store multi-word terms.

To learn more, take a look at the list of entries in the DiCoInfo and the entries for some of the multi-word terms it includes (as well as their components, if they appear in the DiCoInfo). Try to determine by analyzing the entries why these terms have been stored as multi-word units rather than as individual components.

8. What do you think of this difference in approach for multi-word terms? What are the advantages of each approach, for the translator/writer and for the terminologist? Which do you ultimately prefer, for what purposes, and why?

Wrapping up

1. Close the web browser window.

Questions for reflection

1. As you did these exercises, what did you notice about how the DiCoInfo works?

2. What could the DiCoInfo help you to do? In what kind of situation?

3. What criteria can be used to evaluate specialized online dictionaries such as the DiCoInfo?

4. How does the DiCoInfo compare to other online dictionaries and term banks you have used? To other specialized dictionaries?

5. What are some of the advantages and disadvantages of using the DiCoInfo to search for information about terms, compared to using other terminology resources?

6. What did you like best about the DiCoInfo?

7. What did you find less useful?

8. Do you have any suggestions for ways to improve the DiCoInfo? If so, what are they, and why will they help to make it more useful?

Tutorial developed by Baris Bilgen and Elizabeth Marshman