

A Practical Approach to Terminology Work

Author: José Gambin

E-mail: jgambin@abroadlink.com

¹Presentation delivered at the ATA Seminar: Translating for the Pharmaceutical Industry, January 24, 2004, San Juan, Puerto Rico.

1. INTRODUCTION

The current paper aims to be a practical proposal for terminology work based on practical experience having as scenario an individual translator or a small team of translators. Terminology is one of the most time-consuming tasks involved in a translation project. In order to face the challenge that terminology poses on translators it is necessary not only to know a set of appropriate tools but also to bear in mind strategies that will make possible it to deliver the job on-time and within the translator's time budget.

It is important to take into account that terminology search and management are two specialized tasks, which fully justify the role of translators as technical writers versus subject specialists. For a review of the importance of language resources in the translation profession see Yuste (2002). If we constrain ourselves to any particular sphere of knowledge (i.e., medicine), we cannot expect any specialist to know every aspect of his or her area of knowledge. Translators should be specialized, but the important objective is that they have the competence to investigate and deal with any text related to the field of their specialization due to their documentation and terminology search abilities.

This paper is aimed at those newcomers to the profession and at experienced translators willing to revise their terminology search strategies. The practical examples illustrated in this paper will use the linguistic combination English-Spanish.

2. TERMINOLOGY SEARCH TOOLS

Primary sources: Parallel texts and specialists.

A primary source (PS) is basically any kind of original oral or written text produced by a native speaker. Examples of primary sources are handbooks, newspapers, technical documentation, fiction books, websites, etc. The terminology work using primary sources will focus on finding context where source language (SL) terminology is used in the target language (TL) text.

Traditionally, hardcopy books have been the main resource of PSs. Nowadays electronic copies and mainly the Internet are switching the nature of PSs, and this switch is even more important for terminology work performed during a translation project. Note that dealing with electronic data makes possible the use of software applications to manage rapidly and automatically enormous batches of data

(i.e., automatic terminology extraction, quick term searches, etc.). As the goal of this paper is to be a practical guide for professional translators, we will consider the current biggest knowledge databases available at translators' fingertips, that is to say, Internet search engines. Please see section 5, *Using Internet Based Resources* for further information.

A versatile and important PS is the knowledge accumulated by field specialists. Any competitive translator

who aims to be a real specialist in a specific field should have a set of specialists related to his/her field of specialisation available for consultation. To achieve this goal may be seen as a sign of maturity as a specialized translator and, by no means, a stroke of luck; medical doctors, competent lawyers, engineers or any other knowledgeable highly-specialized professionals are generally all too busy and well paid to establish a mutually beneficial business relationship with translators. Perhaps translator should look for retired professional who may be suitable and willing to be the specialized support to translators. Besides, we need to evaluate and check the terminology that a particular professional may use; many times the economy principle of language makes professionals stay away of more formal and standardized terms in the written language, which are the most appropriate for translation.

Secondary sources: Dictionaries and glossaries (hard copy, electronic copy and the Internet)

Secondary sources are basically monolingual and bilingual dictionaries and glossaries. They can be found in their hardcopy, electronic or on-line versions. A comprehensive and quality-orientated secondary source is the dream of any translator. They can make terminology research really easy.

Nowadays, we find publishers releasing dictionaries in the above-mentioned different formats. For example, we can find the *Encyclopædia Britannica* on-line (www.britannica.com), on CD or in its traditional hardcopy version. As romantic as we may think hardcopy books are, the reality is that a term search in the electronic or on-

line version of a dictionary is much quicker, which as professionals we need to consider as a top relevant factor.

Either using a primary or secondary source, it is important to consider their reliability. We need to stress here that printed publications will normally rank higher when considering the reliability of the source. This is not a fixed criterion as there are excellent resources available in electronic format and, at the same time, there are lousy printed publications which may be the result of a bad translation work or poor quality terminology compilation. Furthermore, we increasingly find a great deal of data reproduced in electronic and hardcopy versions. Basic criteria to evaluate the reliability of a source are:

- ✦ Reputation of the author (well-know, specialist, reliable)
- ✦ Type of publication (specialized magazine, thesis, original technical documentation...)
- ✦ Whether the source is an original or a translation
- ✦ Publisher's prestige (public institutions, universities...).

The following chart presents Harris' CARS Checklist (Harris, 1997), which may help us to evaluate deeper resources found on the Web

Credibility	Trustworthy source, author's credentials, evidence of quality control, known or respected authority, organizational support. Goal: an authoritative source, a source that supplies some good evidence that allows you to trust it.
Accuracy	Up to date, factual, detailed, exact, comprehensive, audience and purpose reflect intentions of completeness and accuracy. Goal: a source that is correct today (not yesterday), a source that gives the whole truth.
Reasonableness	Fair, balanced, objective, reasoned, no conflict of interest, absence of fallacies or slanted tone. Goal: a source that engages the subject thoughtfully and reasonably, concerned with the truth.
Support	Listed sources, contact information, available corroboration, claims supported, documentation supplied. Goal: a source that provides convincing evidence for the claims made, a source you can triangulate (find at least two other sources that support it).

3. A MODEL OF TERMINOLOGY MANAGEMENT

When approaching terminology work from a professional point of view, there is a somewhat complicated balance between time and quality. As professionals seeking to maintain, broaden or improve our customer portfolio, we can only deliver quality but, at the same time, we need to deliver that competitive quality within a reasonable and profitable timeframe. In order to achieve this goal, there are some points we need to take into account when facing terminology work in any translation project:

- † **Whether developing a glossary is worth the effort**
- † **Automatic terminology extraction versus traditional human extraction**
- † **Setting priorities**

- † **Strategies to overcome terminology uncertainty**
- † **Retrieving terminology**

3.1 Whether developing a glossary is worth the effort

Depending on the nature of the translation project and our customer, we can estimate that the time invested in creating a glossary will pay off in future assignments or within the same project, not only regarding terminology search time but also in quality through consistency and accuracy. As a rule of thumb, if we got a long-term relationship with our client or are seeking it, the creation of a glossary will be a key factor to render a quality and profitable job. On the other hand, if the translation is going to be done by a team of translators, which will mean that we are dealing with a large project, the creation of a glossary will be highly advisable as to save the

time requested for the project by avoiding redundant term searches and to provide a higher consistency.

The creation of a glossary will be equally important if we are thinking of translating using a Machine Translation system. Most of these applications will allow us to create a customary dictionary that will be used during translation. As described by Kübler (2002), this method may produce good quality results.

3.2 Automatic Terminology Extraction versus Traditional Human Extraction

Once we have decided to create a glossary, there are at least three different approaches to the compilation of SL list of terms: automatic terminology extraction, human terminology extraction before or human terminology extraction during translation.

Automatic terminology extraction

This is a technology which is a good candidate to become a standard in the industry. Automatic terminology extraction is based on the principles of corpus linguistics, that is to say, the study of language based on large text databases (or corpora) through software applications. This makes possible a statistical approach to language through the study of word frequencies and concordance patterns, among other variables.

Terminology extraction applications generally use word frequencies to propose a list of candidate words. The principle is easy: if a term is repeated several times through a text, it may be a term specific to that particular field. After this automatic selection of potential terms, the translator

needs to go over it and select the terms he considers to be relevant terms to be included in the glossary.

This technology is far from providing exact results. For example, pertinent terms may have a low frequency and therefore will not be included in the candidate list. The point is that for large projects this could be an important time saver. A commercial application using this basic principle is for example *TerminologyExtractor* or *Trados Term Extract*. The latter features a function for exclusion terms (for example, terms already stored in a MultiTerm database). *Termight*, another extraction terminology tool, uses taggers and a set of syntactic patterns defined by regular expressions to identify candidate terms and expressions. It also features an interface showing term concordance to help the user to decide if a word should be considered as a term.

This approach should be seriously considered if we are dealing with large projects as it will help to rapidly deliver a glossary to a translator team. The shortcoming of having some terms left out of the glossary may be managed by including those terms during translation. In any case, these terms are likely not to have a high impact due to their low frequency.

Human terminology extraction before translation.

This is the current practice and consists of reading the document to be translated, focusing on tables, figures, indexes, tables of contents and other key text elements in order to establish which terms are good candidates (Dagan and Church, 1994). As happens with automatic extraction, this

approach may be defective due to human errors, so we can also expect to have some terms left out of the glossary. The advantage of this traditional approach is that the translator will have available the context at any moment, which may be a key factor for deciding about the appropriateness of a word to be added to the glossary. Note that some automatic terminology applications offer practical solutions to make the context available. A drawback to this method is that it is very time-consuming compared to the automatic approach. Despite this, we can assume that this method will be valid for relatively small projects as automatic terminology extraction is expected to be more efficient when dealing with large projects.

A practical tip to minimize the time required to create a monolingual glossary when working in a word processor, is to create a macro (so called in *MS Word*) which automatically copies and pastes the selected word or expression into a table created in another file. This way we avoid the distracting and time-wasting task of coping and pasting or typing entry by entry.

Either with this approach or using some of the automatic terminology extraction tools existing on the market, we will need to set strategies to leave out of our term list those terms that we already have in an existing term database. A possible procedure would be to look up any individual term in the database. A less time-consuming way is to copy and paste the candidate terms highlighted from a table into another table containing all the database terms. If we sort this table alphabetically, we could easily discern redundant terms.

Human extraction terminology during translation.

Another possible option is to introduce terms in the glossary at the same time we are performing the translation. I have not found an empirical study comparing translation word ratios to evaluate productivity including terminology when creating the glossary before or during the translation. We can assume that when creating bilingual glossaries before translating, the difficulties to render some solutions may be greater or more time consuming as the translator will have generally a more superfluous understanding of the term, but, when translating, we will have less interruptions due to terminological searches which can help to concentrate better and to produce more coherent translations.

On the other hand, the creation of glossaries during translation will help us to produce directly bilingual glossaries with a deeper knowledge of the context. While using many of the commercially available memory translation tools (i.e., *Start Transit*, *SDLX* or the latest *Trados* version), it will be relatively easy to add a new term to an existing term database, checking simultaneously if that term has been already collected. Note that even using the automatic terminology extraction or the human extraction before translation strategies, the use of this ongoing term extraction feature is highly probable. In any case, the use of human terminology extraction during translation is discouraged if we are dealing with a large project involving a team of translators, as we cannot guarantee the consistency of the terms and this will need to be thoroughly checked afterwards.

3.3. Setting Priorities

Ideally, we will search a list of terms to solve all the terminology work until we are fully documented and satisfied with the results. This means we can do a four-hour search for a single term, depending on the efficiency of our search and the available resources (we may need to visit some libraries or wait to contact a reliable specialist). But normally the job is for yesterday, so setting search priorities is a valid way to get the best possible results within a limited timeframe. Here there are some guidelines which may help to judge which terms should be considered the most important:

- ‡ Terms appearing in an index, glossary of contents, titles and other relevant text elements.
- ‡ Repetitive terms
- ‡ Terms we reckon to be established and fixed formulas in the TL in an industry, regardless the frequency or the place where they appear

For example, it seems rather obvious that a product name in an instruction manual will have maximum priority. The chances are that such term will appear in the title of the publications and that the frequency of occurrence in the text will be rather high; it is also highly probable that the term has a standard translation solution in the TL.

On the other hand, when dealing with terminology, we can find terms that are not consistent in their use in TL within a specific field of knowledge. When facing those terms (and recognising them requires sometimes an experienced translator's intuition), we can translate them with a solution which

delivers the meaning and sounds natural, saving the time to look up a dictionary or to search in parallel texts. On the other hand, we need to recognize those terms whose use may be well standardized in a particular specialist jargon and that will need an exhaustive terminology search.

3.4. Strategies to Overcome Terminology Uncertainty

In the translation profession, a job often needs to be delivered to the client but some terms may have not been satisfactorily resolved. This results in different scenarios and requires different strategies:

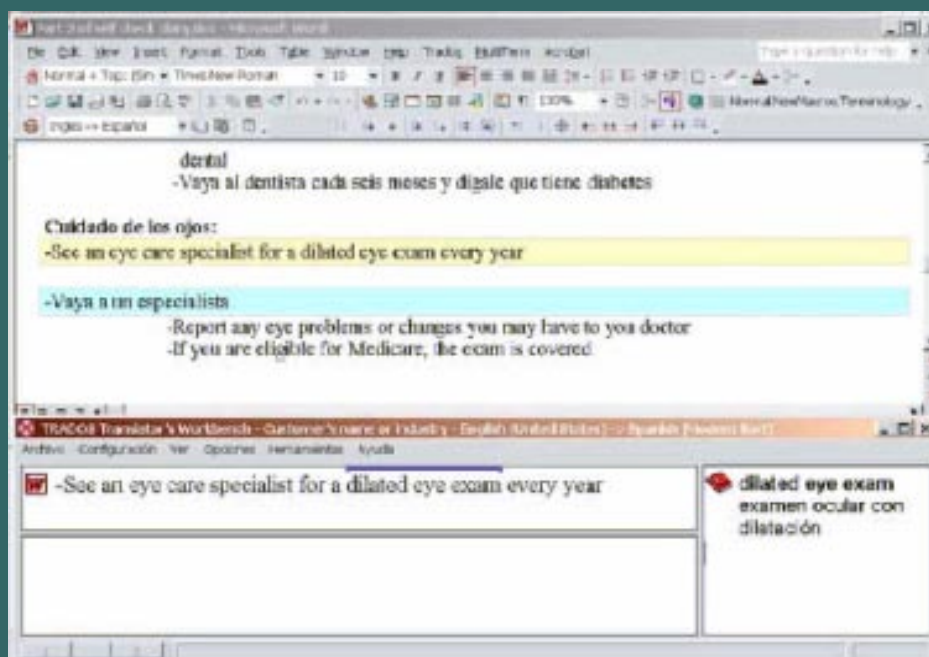
- Scenario 1: The term is well understood but no validated solution (through a primary source) has been found. Strategy: Paraphrase or use a natural term, assuring that it will be understood by the target reader.

- Scenario 2: The term is not understood and no validated solution has been found. We should make all possible efforts to avoid such circumstances, but any experienced translator will sooner or later face this compromising situation. Strategy: 1) Literal translation; Unfortunately, the market is swamped with dubious translations which tend to be literal, avoiding the responsibility and effort of understanding the SL text and standardizing the use of those calques or loan translations; apart from this, it is also the tendency among specialists to use highly technical terms created in an SL (of course, this is specially true for translations from English, due to the use of this language as *lingua franca*). So the

chances of being understood by the target reader are high, especially if translating for a specialist audience. Depending on the type of text it may be acceptable to put the TL term in brackets, so we broaden our chances of being understood, as there may be cases when the target reader knows the terminology in the TL (for example, this will be the case if translating a software magazine from English into Spanish); 2) Undertranslation; This solution may be rendered when the term was not entirely understood but because of the context can be classified as a term belonging to a certain category. For example, if we know that the specific term is a tool, there will be occasions where such an undertranslation will work as the context will provide the target reader with the representation or meaning of the specialized term we are avoiding to mistranslate.

3.5. Terminology Databases

Once we have a bilingual glossary, we need to think out a way to retrieving terms when we need them. Taking into account that literary translation may account for 2-3% of the translation market, I believe that translators should own any of the current commercial translation memory programs which will improve performance in most translation projects. Translation memories are an efficient tool to gain consistency and save time. They are a word databank of all previous translations which can then be reused or where already translated terms can be looked up. But the important point regarding terminology work is that TMs come integrated with terminology databases. This integration permits a feature in the TMs that automatically recognizes terms stored in our database. So while we are translating, we will have a window where the terms already found in the terminology database will be displayed. Below I have inserted a screen shot showing how this is presented in Trados WorkBench:



On the bottom left corner of this screenshot appears the terminology recognition window, which will pop up when the terms are detected in a new open segment (blue box) and existing in the open database file.

This feature is a step further in terminology management. The traditional method implies at best an automatic search using the Find function to locate a term in a glossary produced with a word processor or a spreadsheet. This way of looking up terminology in a glossary is time-consuming compared to the automatic recognition feature we find in current translation memory applications, not only because we need to copy and paste the term or type it to do the search in the glossary's file, but because sometimes we may expect a term to be included in the glossary only to realise this was not the case once we have invested the time in searching for it.

Besides this, the terminology recognition feature in TM applications carries out fuzzy searches. This means that if a word in the glossary or in the source text is misspelt, the program will still look up for terms very close to the misspelt words, assuring that they will be recognised or plural nouns will get detected if the entry in the glossary was the singular form.

Normally, we will be able to easily import our glossary from a word processor or datasheet into the terminology database, though we will need to convert the glossary file to a .txt file and perform some changes. Note that the terminology databases integrated with TM applications are sophisticated products. This means that we can create a

much greater number of fields (including graphics) than our practical approach will suggest.

4. A MODEL OF A TERMINOLOGY SEARCH

Depending at what stage we decide to create the glossary, we may have a list of terms we need to understand and to translate. We will assume that the decision was to create the glossary before translation, when so may proceed as follows:

1. Context:

Once we have isolated the term in a glossary, we may need to go back to that term in its context. To do so, we can simply do a search on the electronic document we are dealing with. If the project is made up of multiple files, it is advisable to find a way to review the context with a unique search. For example, if we have a batch of text files, we could create a file with all the content. As we proceed to render a translation for the different entries in the glossary, it is important to be able to access the context of the terms in the list to determine in which sense the term is used in the SL text.

2. Understanding the SL term:

It seems obvious that to render the best possible solution and to avoid errors in meaning, the understanding of the SL term is compulsory. There are basically three approaches to this: a) general dictionaries, specialized dictionaries and monolingual glossaries, b) SL written primary source, and c) specialist consultation. Note that it will be possible to take a definition

of a term applicable to our context, because we can access this context to determine which definition applies.

3. Finding a translation:

Here we find again three possible ways to provide a solution: a) bilingual general dictionaries, bilingual specialized dictionaries and bilingual glossaries, b) TL written primary source, and c) specialist consultation. Note that when finding a solution in a secondary source, we will need to validate that solution in a primary source. Many times we may find ourselves at a dead end; we are not able to find in the resources available a reliable solution, sometimes just because we are dealing with terms which have not been rendered into the TL. Please see section 3.4. *Strategies to overcome terminology uncertainty.*

4. Entering information in the glossary:

We will need to enter in the glossary all the information we have been finding. I would suggest creating optional fields in the glossary such as *context, definition and non-verbal elements*. Ideally, at least these fields should be inserted in a glossary, but the reality is that for many professionals this is too time-consuming to pay off. Nevertheless, it will be sensible to create these fields for terms that may be deemed especially complicated, relevant or repetitive. Sometimes we may find that the same word is used as a technical term with different meanings, and possibly different translations, so we need a definition or a context to

make a distinction. On the other hand, if the glossary created will be used by a team, it will be reasonable to add the maximum information possible. These four basic steps are many times interrelated and, depending on individual circumstances, can be performed completely or partially and in a different order. This will always be justified as long as an optimal solution is achieved in the shortest possible period.

5. USING INTERNET-BASED RESOURCES

This section will show a set of practical resources we can find on the Internet. Following the spirit of this paper, this section does not aim to be a comprehensive display of web-based resources and possible strategies; instead, it looks up to be an informative and stimulating view of the kind of resources we can find on the Web.

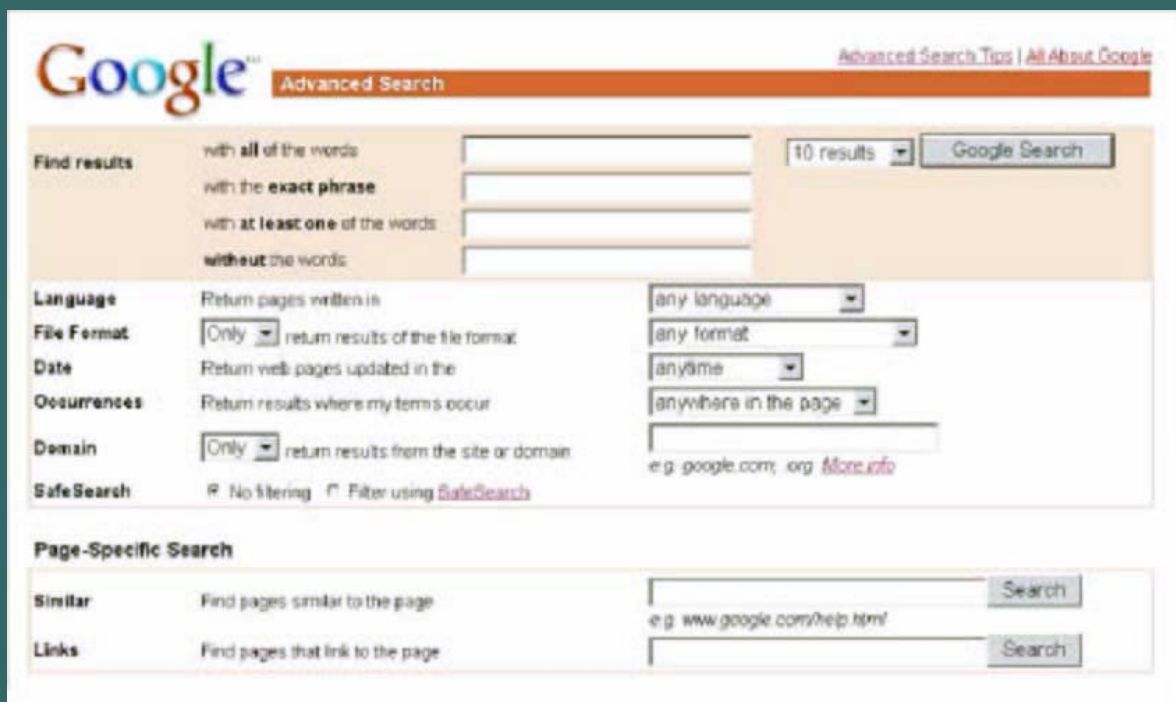
On the Internet, we can find a great number of all types of resources described in section 2. The Internet connects us with such a variety of resources that the translator needs to be cautious not to get lost in this myriad of data. Saving time delivering quality is the goal, so we will need to concentrate on those resources which yield the best results. Most experienced translators will finally select those resources which they have found more helpful.

A very practical point to bear in mind when working with the Internet is the type of connection we have. Note that a dial-up connection can be 5 to 10 times slower than a cable or ADSL connection, which correlates with the search rate per hour at

which we will be able to work with these different types of connection.

Search engines and directories will help us to locate those web pages which may contain the information we are interested in. Directories (such as Yahoo!) classify web sites in categories; while this may be an efficient way to locate information, search engines are a quicker and more comprehensive tool for this purpose. There is a wide range of search engines we can use: Google, Yahoo, AltaVista, Lycos, MSN, Excite, HotBot, LookSmart, AOL, WebCrawler, InfoSeek and many more. Many of them will be just a change of interface but will render the same searches, for example Yahoo! searches are based on Google algorithms; MSN and AOL are powered by Inktomi (the number two Internet traffic generator), HotBot gives us the possibility to use Inktomi, Google or Teoma to perform a search. As said before, one can get lost easily in front of this range of options. I began searching with Yahoo!, changed to Altavista and finally stuck to Google. Google has gained a lot of the

search engine business, despite having appeared in the industry long after Yahoo! or Altavista. As a matter of fact, now Yahoo! is powered by Google, and Altavista has lost a lot of surfers in favour of Google. The point is that Google's algorithms seems to deliver more pertinent searches, apart from having one of the largest website databases and allowing for sophisticated advanced searches. Below there is a view of the different options to do advanced searches using Google. You get to this window by going to www.google.com and clicking on the link called "Advanced Search." Please read through the screen to get familiar with these options.



The screenshot shows the Google Advanced Search interface. At the top left is the Google logo, and at the top right are links for "Advanced Search Tips" and "All About Google". The main search area includes a "Find results" section with four radio button options: "with all of the words", "with the exact phrase", "with at least one of the words", and "without the words". Each option has a corresponding text input field. To the right of these fields is a "10 results" dropdown menu and a "Google Search" button. Below this are several filter sections: "Language" (Return pages written in), "File Format" (Only return results of the file format), "Date" (Return web pages updated in the), "Occurrences" (Return results where my terms occur), and "Domain" (Only return results from the site or domain). Each filter section has a dropdown menu and a text input field. The "SafeSearch" section has two radio buttons: "No filtering" (selected) and "Filter using SafeSearch". At the bottom, there is a "Page-Specific Search" section with two rows: "Similar" (Find pages similar to the page) and "Links" (Find pages that link to the page). Each row has a text input field and a "Search" button. The "Similar" input field contains the example "e.g. www.google.com/help.html".

Google

While all options may be practical during a search, here it is a summary of those more generally used. Note that to perform quicker queries Google gives the possibility to type in the main search box some search option codes, so we avoid having to go the Advanced Search page. The use of these codes is strongly recommended to further minimize search time.

Whenever we do a search with Google entering more than one word, Google will find pages which have at least all words typed, not considering for the search very common words such as "the", "of", "is", "are" and the like.

Code: +

This code will be useful when we want a very common word to be included in the search. For example, when an acronym is spelt as a very common word (i.e., IS may stand for Intensity Stereo or AS for Associate in Science).

The following screenshot illustrate a search with 4 words, "as", "in", "associate" and "science". As we have introduced the code + with "as" and "in", Google will consider these words while searching, and as we have introduced "associate" and "science", the chances are that we get pages were "AS" stands for "Associate in Science".



The screenshot shows a Google search interface with the query "as in associate science" entered in the search box. The search results are displayed below the search bar, showing several links to educational programs. The first result is "Associate of Science Degree (AS) General Studies" from the University of Phoenix. The second result is "Associate in Science (AS) Programs" from the University of Phoenix. The third result is "CCSC AS Degree in Respiratory Health" from the University of Phoenix. The fourth result is "AS in Computer Science" from the University of Phoenix. The fifth result is "Associate in Science (AS) Degree Requirements" from the University of Phoenix.

Code: -

When using this code, we will get pages where the - coded words will not appear. This may be helpful when we get a high number of pages and want to further fine tune the search. If we look at the pages delivered in the above screen shot, we will see that we obtained searches with "of", which actually introduces noise to our search. With the code -, we can get rid of

this very common word in our search and get a less noisy response. As we will see now in the following screenshot, the first four searches show "Associate in Science", when before we only got two pages with this word pattern.



Code: " "

This is probably one of the most helpful search options we have in Google. As a matter of fact, we find this option in most search engines. It is highly helpful when validating expressions or searching for specialized web contents. For example, in the educative example provided above, we could just enter a search using quotation

marks such as "associate in science (AS)" as this screenshot shows:



Google [Advanced Search](#) [Preferences](#) [Language Tools](#) [Search Tips](#)

Associate in science (as)

Web Images Groups Directory News

Searched the web for "Associate in science (as)" Results 1 - 12 of about 2,010 Search took 0.92 seconds

Under County Community College - Programs
Associate in Science (AS) Programs. Associate in Science Degrees are available in the following fields: Advertising Design and Graphics Technology ...
www.undercountycc.edu/asprograms/as.asp - 18k - [Cached](#) - [Similar pages](#)

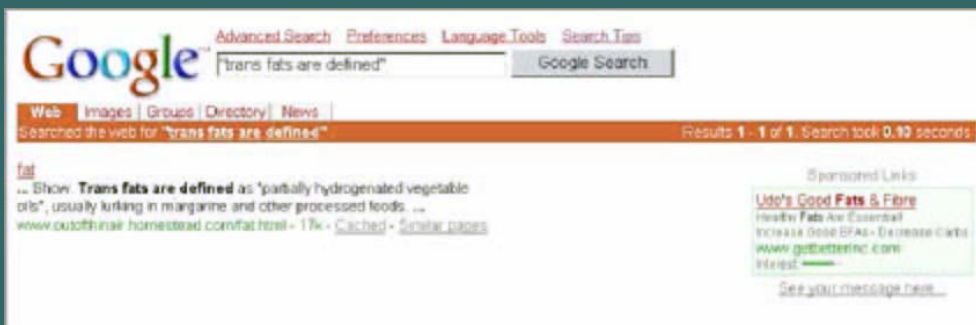
Associate in Science (AS) Degree Requirements
 ... **Associate in Science (AS) Degree Requirements**. Credits in General Education Core Curriculum: 58 Additional Mathematics/Science credits: 15 Credits in addition ...
www.ccd.edu/academic/academic/Cert_Dept/degree/req.htm - 18k - [Cached](#) - [Similar pages](#)

Alameda Community College - Associate in Science (AS)
 Alameda Community College, Advanced Search | Site Map Popular ACC Sites ...
www.alameda.cc.edu/degrees/associate-science.htm - 24k - 3 Jul 2003 - [Cached](#) - [Similar pages](#)

North Hennepin Community College - Associate in Science (AS)
Associate in Science (AS) Degrees. General Information | Catalog | Associate Degree Programs / **Associate in Science (AS) Degree** ...
www.nhcc.edu/academic/associate-in-science-degree.htm - 18k - [Cached](#) - [Similar pages](#)

Associate in Science Degree
 Associate in Science Degree. The Associate in Science degree program is for students who plan to transfer to senior colleges and ...
www.scholarship.usd.edu/general_information/associate.htm - 11k - 3 Jul 2003 - [Cached](#) - [Similar pages](#)

The use of quotation marks gives us the possibility to search for definitions, which we may not have found any dictionary, by locating them in specialized texts or finding contexts which will help us to infer a definition. The following example illustrates this point:



Google [Advanced Search](#) [Preferences](#) [Language Tools](#) [Search Tips](#)


trans fats are defined

Web Images Groups Directory News

Searched the web for "trans fats are defined" Results 1 - 1 of 1 Search took 0.92 seconds

fat
 ... Show **Trans fats are defined** as "partially hydrogenated vegetable oil", usually lurking in margarine and other processed foods ...
www.outthineat.com/fat.html - 17k - [Cached](#) - [Similar pages](#)

Sponsored Links

Udo's Good Fats & Fibre
 Healthy Fats Are Essential
 Increase Good Cholesterol - Decrease Cholesterol
www.getbetterinc.com
 Interest 

[See your message here...](#)

Code: `site:domain`

In Google we can make a search within a single domain. This may be helpful when we have no glossary from the client but we need to be consistent with the clients' terminology, with Google and this function we can easily check if a term is used in the clients' website. For example, we could check if *motherboard* is translated in Intel's website as *placa madre* or *placa base*, two different possible solutions. We could the use of this terminology by entering in Google's search box: `site: intel.com "placa base"`; and in a second search, `site:intel.com "placa madre"`.

The code `site: domain` is very useful when dealing with different variants of a language. For example, if we are translating for an international audience, what happens often when translating into Spanish or into English, we may need to know if a term is used in a locale or had different uses in other variants. If we investigate the word *computer* in Spanish, which should be translated as *computadora* for Latin America and as *ordenador* for Spain, we will see that *computadora* appears in 11.500 pages in ".es" domains ("es" is the regional domain for Spain) versus 69.300 pages in ".mx" domain pages ("mx" is the regional domain for Mexico), on the other hand, *ordenador* appears in 137.000 ".es" pages versus 4.110 ".mx pages".

This type of constrain may be very useful to found more relevant and reliable resources. By using the code `site: domain` we may filter a lot of web sites, which are themselves translations and that should be considered less reliable than web sites produced by native-speaking specialists. Regional domains are more expensive, so we can expect more quality in regional web

pages. Another example, if we are looking for documentation in English, is that we can search just ".edu" web sites, which will belong to a post-secondary accredited educational institution in the US.

Code: `-site:domain`

This option is just the reversed option of the previous one. There may be cases in which it is useful to be able to avoid pages of a specified domain.

Code: `OR`

Using the code `OR` we will obtain pages which contains one of the two words or expressions searched. We can use `OR` to simultaneously search in various domains. For example, we may decide to validate our IT technology by using the Spanish websites of three major IT vendors, such a Microsoft, Intel and Hewlett Packard, on the basis that their multinational presence and power will have standardized the use of some IT terms. The next screenshot illustrates this:



Glossaries and dictionaries

On the Internet we find available a great deal of on-line specialized dictionaries, especially if our SL or TL is English. Here we find high quality dictionaries on-line such as Dorland's and Stedman's medical dictionaries or What is? and CCI Computer computing dictionaries.

A wonderful resource to find definitions in English monolingual dictionaries is Onelook, a portal that acts as a gateway to 933 on-line dictionaries. We can expect this kind of portals in other languages, and if not, it is very likely that this kind of solution will be popular in some years. At this point, there is not any resource comparable to Onelook for the Spanish language.

Multilingual or bilingual dictionaries are not so comprehensive as a rule. We find a practical example of this in Spanish. Two good resources are Eurodicautom, the official dictionary of the EU, and LOGOS, a multilingual dictionary made available on the Web by the translation company of the same name. Onelook also offers the

possibility to rendered translations for term searches, though it is not very efficient. None of this source may be considered as reliable as the high-quality published bilingual dictionaries. Unfortunately, we may found very poor translations in the mentioned on-line dictionaries. In any case, as we said in section 2, secondary sources should be always validated through a PS, more if we are dealing with SSs found on the Internet.

An impressive tool to locate specialized dictionaries in a great number of language combinations is Lexicool, which has indexed over 3000 Internet dictionaries. The following screenshot illustrates the type of search results we can obtain with Lexicool:

lexicool.com (Database of physical and multilingual dictionaries)

Search Add Software Comments Links About Us

Language(s): English-Spanish and Subject: Medical-Health

Title	Entries (approx.)
1 EURODICALTOM - European Terminology Data Base (MULTI)	5500000
2 DeCS - Health Sciences Descriptors (EN<->ES<->PT)	24941
3 Diccionario Terminológico de Oftalmología (ES<->EN)	6537
4 DOIA - Dermatology Online Atlas (DE-EN-ES-FR-PT)	-
5 PeDOIA - Pediatric Dermatology Atlas (DE-EN-ES-FR-PT)	-
6 Cercoterm (CA-EN-ES-FR-IT)	-
7 Glosario de las Drogas (DE-EN-ES-FR)	-
8 LEXICO - Lexique Multilingue de Termes de Pathologie Professionnelle (FR<->DE<->EN<->ES)	408
9 Dictionary of Demographic and Reproductive Health Terminology (EN-ES-FR)	-

Click HERE to add this search to your Favorites!

Page 1 of 5 54 matches (maximum)

Copyright © lexicool.com 2008-03. Site last updated: 17 June 2003

Regarding glossaries, we can find multiple specialized glossaries published on the web by public and private organizations. Many times a search with the term and the word "glossary" may return a page with a glossary and the word itself. See example:

Google™ Advanced Search Preferences Language Tools Search Tips

motherboard glossary site: edu Google Search

Web Images Groups Directory News

Searched the web for **motherboard glossary site: edu**. Results: 1 - 10 of about 783. Search took 0.19 seconds.

(PDF) Intro to PC Hardware Glossary of Terms
File Format: PDF/Adobe Acrobat - [View as HTML](#)
Intro to PC Hardware **Glossary** of Terms CD-ROM: Abbreviation of Compact Disc ... **Motherboard**
A **motherboard** is the physical arrangement in a computer that contains ...
[www.wfu.edu/Library/IT/Choosing/pchardware/pchardwareglossary.pdf](#) - [Similar pages](#)

PC Survival Guide - Glossary
... The original IBM PC was configured by means of DIP switches, tuned on the **motherboard**. ... Not all of these are present on every **motherboard**. ...
[www.qvct.com/net/edu/pcguide/glossary.html](#) - 38k - [Cached](#) - [Similar pages](#)

Nebula Glossary
... in a word or phrase to get more information. University of Chicago has a fairly complete technical **glossary**. ... **Motherboard** The main circuitboard inside a computer. ...
[www.washington.edu/nebula/glossary.html](#) - 18k - [Cached](#) - [Similar pages](#)

Basic Concepts Glossary
GLOSSARY ... Back to Top. H. Hardware A device which can be literally touched such as a printer, a monitor, a chip, a cord, a **motherboard**, etc. Back to Top. I ...
[www.nvcc.vccs.edu/home/shot/1st100dw/glossary.htm](#) - 14k - [Cached](#) - [Similar pages](#)

Note that this search was done only in web sites with domain ".edu" in order to increase our chances to obtain a reliable source.

Professional web sites

Nowadays, there are a number of web sites for translators offering freelancers and agencies to appear in their service directory (www.proz.com; www.gotranslators.com; www.aquarius.net and many others). The leading translators' web site is Proz.com, there you will find a large glossary of terms entered by other colleagues, beside you may post a query in real time to all the translators accepting questions. This same concept has been implemented by Aquarius, but at this moment the terminology activity at Proz.com is much higher. Translators are motivated to answer as they will score points to show potential clients their expertise, so answers should be taken with caution and always validated.

6. BIBLIOGRAPHY

Dagan, I. and Church, K. 1994. *Termight: identifying and translating technical terminology*. In Proceedings of Applied Language Processing, pp. 34-40.

El Hadi, M et al. 2001. *The ARC A3 Project: Terminology Acquisition Tools: Evaluation Method and Task*. In ACL-2001 Workshop on Evaluation Methodologies from Language and Dialog Systems, pp. 42-51.

Harris, R. 1997. *Evaluating Internet Research Sources*. In <http://www.virtualsalt.com/evalu8it.htm>

Kübler, N. 2002. *Creating a Term Base to Customise and MT System: Reusability of Resources and Tools from the Translator's Point of View*. Proceedings of the First International Workshop in Language Resources for Translation Work and Research. Paris: ELRA (European Association for Language Resources).

Yuste, E. 2002. *Language Resources and the Language Professional*. In Yuste, E. (Ed.) Proceedings of the First International Workshop in Language Resources for Translation Work and Research. Paris: ELRA (European Association for Language Resources).