



Universidad de Valladolid

FACULTAD de FILOSOFÍA Y LETRAS
DEPARTAMENTO de FILOLOGÍA INGLESA
Grado en Estudios Ingleses

TRABAJO DE FIN DE GRADO

Some Useful Guidelines for the Translation of a
Medical Text

Juan José Serrano Notivoli

Tutor: Beatriz Méndez Cendón

Curso Académico: 2014-2015

ABSTRACT

The aim of this undergraduate dissertation is to provide possible future translation students with some useful guidelines in order to perform a good translation of a medical text. In order to achieve our purpose, different steps and aspects that have to be taken into account within medical translation are explained all along this undergraduate dissertation. The compilation and design of a corpus according to our purposes is also explained since corpora is one of the most important tools used by translator. Therefore a parallel corpus has been built to illustrate how to exploit a corpus in order to obtain translation equivalents for our translation. The use of further translation tools is also explained in order to show how translation problems can be solved. Our conclusion is that a good translation is based on many aspects. It does not only consist on substituting words from the source text for their consequent equivalents. That is the reason why we are going to explain the different steps that have to be followed in specialized translation: the analysis of the source text, terminology and the compilation of a corpus have to be followed.

Keywords: medical translation, electronic corpus, translation equivalents, translation problems.

RESUMEN: El propósito de este proyecto es proporcionar unas pautas a futuros estudiantes de traducción para ayudarles a realizar una buena traducción de un texto médico. Para llevar a cabo este propósito, vamos a explicar los diferentes pasos y aspectos que tienen que tenerse en cuenta para traducir un texto médico. Puesto que el uso del corpus es vital dentro de la traducción, vamos a explicar como diseñar y compilar un corpus de acuerdo con nuestro propósito. También hemos creado un corpus paralelo para explicar como explotar un corpus para obtener equivalentes de traducción con textos originales. Más adelante, llevamos a cabo a explicación de uso de las diferentes herramientas de traducción usadas para ilustrar como solucionar esos problemas de traducción. Nuestra conclusión es que una buena traducción se basa en una buena organización.

Palabras Clave: traducción médica, corpus electrónico, equivalentes de traducción, problemas de traducción.

Index:

1. Introduction.....	7
2. Language for Specific Purposes.....	9
2.1. LSP Characteristics.....	10
2.2. Scientific-Biomedical Language.....	11
2.2.1. Stylistic characteristics of English Medical Language.....	11
3. Phases within Specialized translation.....	13
3.1. Source Text Analysis.....	13
3.1.1. Tone and Audience.....	13
3.1.2. Register.....	14
3.1.2.1 Communicative setting.....	15
3.1.3. Text types.....	16
3.1.4. Genre.....	17
4. Terminology.....	19
4.1. Types of terms.....	20
5. Corpus design and compilation.....	22
5.1. Classifications of corpora.....	22
5.1.1. Types of corpora in translation according to the number of languages belonging to the texts.....	23
5.1.2 Corpora according to the level of specialization of the texts.....	24
5.2. Comparison parallel, comparable and monolingual corpora for the purpose of translation: Advantages and disadvantages in translation studies.....	25
5.3. Considerations to be taken into account when designing a specialized corpus.....	26
5.3.1. Size.....	26

5.3.2. Subject.....	27
5.3.3. Setting, audience, technicality and authorship.....	28
5.4. My own corpus.....	28
6. Translation Tools: usage and advantages.....	31
6.1. Corpus exploitation tools.....	31
6.1.1. AntConc.....	32
6.1.1.1. Concordance Tools.....	32
6.1.1.2. Frequency list.....	34
6.1.1.3. Collocations.....	37
6.1.1.4. Clusters.....	38
6.1.2. Bilingual concordance tools: ParaConc and MultiConcord.....	40
6.1.3. Wordfast Pro.....	40
6.2. Further translation tools: online dictionaries and glossaries.....	41
7. Solutions to terminological and phraseological problems in our text translation.....	43
7.1. Analysis of the text object of this undergraduate dissertation.....	43
7.2. Searches for terminological and phraseological equivalents.....	43
8. My translation Proposal.....	57
9. Conclusion.....	59
10. References.....	61

List of images

Image 1: Basic concordance for <i>diabetes</i>	34
Image 2: Full context of <i>diabetes</i>	34
Image 3: Clusters of <i>blood</i>	39
Image 4: Clusters of ‘blood’ with a minimum size of 4.....	39
Image 5: Spanish article with problematic words and phrases highlighted.....	44
Image 6: Results for the search <i>inyecciones diarias de insulina</i> in Linguee.....	46
Image 7: Results for <i>insulin shots</i>	47
Image 8: Results for <i>insulin injections</i>	47
Image 9: Number of results for the search <i>daily insulin injections</i>	48
Image 10: Number of results for the search <i>daily insulin shots</i>	48
Image 11: Results for the search of <i>células productoras de insulina</i> in WordFast Pro.....	49
Image 12: Search for <i>cells that produce insulin</i> in the Proz glossary.....	49
Image 13: Results for the search of <i>insulin producing cells</i> in Linguee.....	50
Image 14: Search of <i>Langerhans islets in Emedicinehealth</i>	51
Image 15: Search for <i>Islets of Langerhans</i> in MedlinePlus dictionary.....	51
Image 16: Translation of <i>someterse a un trasplante de células pancreáticas</i>	52
Image 17: Search for <i>someterse a</i> in WordFast Pro.....	53
Image 18: Google search for <i>pancreatic cell transplant</i>	53
Image 19: Google search for <i>transplant of pancreatic cells</i>	53
Image 20: Concordance for <i>develop diabetes</i>	54
Image 21: Search for <i>desarrollar diabetes</i>	54

List of tables

Table 1: Basic frequency list.....	36
Table 2: Frequency list once a stop list has been uploaded.....	36
Table 3: Frequency list once a lemma list has been uploaded.....	37
Table 4: Collocations of <i>Diabetes</i>	38

1. Introduction

The term “translation” can be understood both as a process and as a product.

As a process, “translation is the transference of information between different languages in which terminology plays an important role, since the translator uses frequently terminological units (terms).” (Giovanni Adamo: 2003, M.T. Cabré: 2003)

On the other hand, translation as a product is the text produced by the translator.

The translation of a specialized text differs from the translation of a general language text. There are some aspects such as terminology that have a crucial role in specialized texts. Specialized translators are mediators between the text and the audience; their goal is to transfer the meaning of the original specialized text into the new language so that the reader could read it and understand it. The translator, in order to perform a good translation, must be aware of the changes produced in language and have a knowledge of the terminology belonging to the particular field in which he/she is working in.

To sum up, specialized translation is a complex task in which several steps have to be followed and several aspects, such as terminology and grammar, have to be taken into account in order to produce an appropriate target text.

The object of this project is to explain what translation and LSP are and also state and explain the different steps and aspects that are comprised within specialized translation. Since there are several aspects that have to be taken into account within translation, we have divided our project into two main sections: a theoretical part (chapters 2, 3, 4, 5 and 6) and a practical part- in which the theory is put into practice (chapters 7 and 8).

In this undergraduate dissertation we will be translating a medical text so in the next chapter we are going to explain what is meant by LSP. The first half of this chapter deals with the general characteristics of LSPs while the second half focuses on the stylistic characteristics of medical texts.

In the third chapter, we state the different phases comprised within specialized translation. We focus on the phase of ST analysis since the different characteristics of the ST will determine the style, level of specialization, and so on, of the translation.

The fourth chapter deals with terminology, which is, in my opinion, the most important element of specialized translation. The main aim of this chapter is to point the role of terminology within specialized translation.

These former three chapters deal with the theory about the main aspects of translation. In chapters four, five and six, we are going to state and explain the different tools that can be used in specialized translation.

The fifth chapter explains what a corpus is along with the different types of corpora its uses and the discussion about which type of corpus is the one which best fits our interests, i.e. a translation of a specialized text.

The sixth and final theoretical chapter explains the uses of the different translation programs we used in the translation of the medical text. Their advantages and disadvantages are also stated in order to explain the choices about using one program or other.

In the seventh chapter we are going to solve the different terminological and phraseological problems found in our translation. The different steps followed to solve the translation problems are going to be explained and illustrated.

Finally, in the eighth chapter we present our translation proposal. This chapter contains the final translation obtained by following the different steps explained all along this undergraduate dissertation.

This undergraduate dissertation ends with a conclusion in which we explain that the translation process does not only consist on translating literally, but that there are too many previous steps and aspects that have to be taken in to account before starting to translate.

2. Language for Specific Purposes

Specialized communication differs from general communication in two aspects: in the type of oral and written texts and in the use of specific terminology. (M.T. Cabré, 2000: 105)

As general language (GL), specialized language develops due to society. As Gustavo Mendiluce (2004:31) noted in his thesis “the economical, technological, industrial and scientific development creates new realities that need to be communicated, so society itself satisfies the communicative needs of the different fields by creating these languages.” This is the reason why specialized language exists, to communicate the knowledge produced by the development of these fields. This continuous development results in some specialized knowledge no longer being considered “specialized” as it starts being considered as general knowledge, due to the apparition of new knowledge.

The appearance of new knowledge and the constant conversions of specialized knowledge into general knowledge because of the appearance of this new knowledge makes us see that “there is not an absolute borderline between GL and LSP” Méndez Cendón (2012-2013). There is a constant flow of knowledge and information among general language and language for specific purposes (LSP). Méndez Cendón (2012-2013) also noted that “GL and LSP are considered subsystems of the global language and that they are interrelated because they share a syntax and grammatical structures and they share many words (a lexicon).”

On the other hand, Cabrera (2004: 32) quoted a definition of Cabré in which she defined specialized languages as “the subsets of general language characterized by the subject matter, the users and the communicative settings.”

These characteristics are going to be explained in the following chapters due to the importance they have in the process of translation of a specialized text, more specifically a medical text, which is the aim of this undergraduate dissertation.

2.1. LSP characteristics

Once the definition and the areas involved within LSP are known, we consider important to state the main characteristics of LSP texts. M. T. Cabré and Josefa Gómez (2006: 21) quoted Picht and Draskau in order to state the set of common characteristics of LSPs:

- Monofunctional character
- Restricted number of users
- Voluntary acquisition
- Autonomy in the sense that the variation of LSPs does not affect GL

M. T. Cabré and Josefa Gómez clarified that these characteristics refer to the communicative setting. Both authors also provided a set of common characteristics of LSPs from the semantic point of view (M. T. Cabré & Josefa Gómez, 2006: 22):

- The users want to be informed about the consequent subject field they are interested in
- The sender(s) tend to be specialists in a specific subject field
- The communicative settings tend to be formal and of professional nature
- LSPs have a primary function: informative

Apart from the common characteristics of LSP texts, it is also important to know the different structures that tend to appear in LSP texts (Méndez Cendón, 2012-2013): “Specific terminology, prefixes and suffixes, premodifiers (complex noun phrases), compounds, nominalizations and passives.”

2.2. Scientific-biomedical language

Since the main goal of this undergraduate dissertation is to achieve a good translation of a medical text, we consider crucial to state the main goals of scientific-biomedical language. Gustavo Mendiluce (2004: 70) mentioned in his thesis seven different qualities of scientific- biomedical language that have to be taken into account:

1. Precision: according to the majority of authors this is the most important feature of scientific-biomedical language. In order to be precise, terminology plays a very important role; it is crucial to use adequate terminology in order to avoid ambiguity since it could lead to confusion among its users.
2. Clarity: it is quite related to precision. A text in order to be clear has to have a good organization of ideas. Also special attention has to be paid to the use of polysemic words, the abuse of premodifiers, pronouns and deictics.
3. Fluency: there has to be a logical order among the different ideas. And also, in order to be fluent, a text must have a good connection among these ideas.
4. Concision: this quality is complementary to that of clarity. “There are several aspects that can put in danger the concision of a text”(Gustavo Mendiluce, 2004:71):
5. Redundancy: avoid repeating words that are not necessary
6. Emptiness: avoid words or phrases that do not add new meaning to the text.”
7. Neutrality: “there should not be affective boundaries attached to the messages of scientific texts.” (Gustavo Mendiluce, 2004:72)

2.2.1. Stylistic characteristics of English Medical Language

In the previous subsection we have stated the different features that scientific language must have. Consequently, the characteristics of medical language are going to be related to them. The problem is that instead of stating the main features of medical language, there is a tendency to establish the most common mistakes in medical language.

We investigated and found these common mistakes within medical language in both English and Spanish. However, since our aim is to translate a medical text from Spanish to English we are only going to state the problems found in English:

1. Orthographic level: wrong use of punctuation marks.
2. Lexico-semantic level
 - Abuse of `medicalese`. By doing that, specialists of other areas are excluded from the use of medical language.
 - Excessive level of formality
 - Pleonasms
 - Metaphors
 - Abuse of acronyms
3. Discursive level
 - Lack of fluency

(Gustavo Mendiluce, 2004: 74-5)

As we see, instead of establishing the basis of a medical text, authors have dedicated time to mention the common mistakes of medical language.

3. Phases within Specialized translation

Once we know the notions of translation and specialized language, it is important to know the stages through which a translator has to go through. In order to illustrate this section, we are going to name the different phases in the translation process found in *Language-translators-help*¹. They distinguish five phases:

- 1) Preparation: the translator prepares reference materials: glossaries, dictionaries, translation memories that could help him/her with the translation.
- 2) Translation: the translator starts to translate the given text.
- 3) Editing: in this step, the translator compares the ST and the TT in order to check the style and accuracy of the translation. Also, the translator can check the translation by sending it to a second translator or to a specialist in the subject matter in question.
- 4) Formatting: in this phase, the format of the translated text is modified according to the client requirements. The pages, headers and footers are checked in order that there are not mistakes.
- 5) Proof-reading: in this phase, a final check is performed in order to avoid any misspellings or grammar mistakes. The format of the text is also checked.

3.1. Source Text Analysis

Although the analysis of the ST does not appear in the different phases we have provided above, it is one of the most important steps. It does not only involve finding the vocabulary that we do not understand but also identifying the genre, register, tone, audience of text in question.

3.1.1. Tone and Audience

The tone is the attitude of the author towards the reader. If the author is only interested in transmitting information objectively we deal with impersonal tone. On the other hand

¹ See more information at <http://www.language-translation-help.com/professional-translation-services.html> (accessed date 27/11/2014)

if the author provides information from his/her point of view or gives personal advice to the readership, the tone of the text is personal.

Audience also deals with the public. Although it might seem quite similar to the tone it is not the same; audience is concerned with the kind of audience the text is addressed to. If it is directed to a particular kind of audience it is "marked", while if it is addressed to a general audience it is "unmarked".

3.1.2. Register

Most people identify register “with the level of formality of the text(s)” (Isabel García, 2000: 17). I think that this is the most common association concerning register. However, Hasan (1985: 41) explain that “registers are the varieties of language.” Within these varieties we can find also the dialects. The difference is that registers are concerned with the *use of language*, while dialect is concerned with the language *user*.

In order to classify medical texts, we need to know that “in the process of reception and production of scientific texts we find three register levels” (Halliday & Hasan, 1985:12).

1. Field: “what is happening, the nature of the social action that is taking place” (Halliday & Hasan, 1985:12). It can be technical/specialized or common/every day.
2. Mode: Hallyday & Hasan (1985:13) quoted Sinclair & Ball: “The feature ‘mode’ refers to the way in which a text is originally produced.” Oral/written.
3. Tenor: participants in the communication. Chelo Vargas (2005:306)² mentions Cabré to clarify that “only those participants who have a specific knowledge in a professional field acquired through learning can take part in the production-reception process of a specialised communication.” However, the addressee of specialized language can range from experts to non-experts.

² See more information at http://www.academia.edu/971583/A_pragmatic_model_of_text_classification_for_the_compilation_of_special-purpose_corpora (accessed date 8/11/2014)

As Isabel García (2000: 181) concluded: “the three variables that conform the register are related since the choice of a certain field requires the use of a certain channel of transmission and the choice of an adequate level of formality.”

3.1.2.1. Communicative setting

Once we know the three register levels, we are going to state the three different communicative setting that can be found in LSP since the concept of ‘communicative setting’ is related to the concept of ‘register’. We consider important to provide and explain the classification of communicative settings within LSP since communicative settings take into account the field, mode and tenor; there is a relation among the communicative setting and these three register levels. Jennifer Pearson (1998: 52-54) provided the following classification:

- Expert-to-expert communicative setting: both author and reader must have the same level knowledge or similar. The use of specific terminology is crucial in this kind of communication. The meaning of terms and phrases used in the texts is understood by all the members of this communicative setting. Terms (simple or complex) do not have to be explained to any of the members of the communicative setting since they are supposed to be familiarized with the terms.
- Expert-to-initiate/semi expert communicative setting: author and reader do not share the same level of specialization. As in expert-to-expert communication, the members of this communicative setting share some common terminology but many times some unknown terminology has to be explained to the semi experts in order to convey the message correctly. The quantity of terms that appear within this communicative setting is considerably lower than in the former one.
- Expert/ relative expert to the uninitiated/non-expert communicative setting: this setting takes place among experts in the field and mostly mature people who are probably not related to the field but may have a certain interest in it. As a consequence, the level of term density has to be much lower than in the two former settings, so that the reader could read and understand the text without any problems. And in the case of terms that cannot be excluded, the expert can either explain the term to the reader or leave it

without an explanation (if the reader is supposed to know its meaning). However, in most cases general words are used instead of specialized terms to describe concepts since it is very difficult to assume the level of knowledge of the field that the reader has.

3.1.3. Text types

Isabel García (2000: 227) quoted R. de Beaugrande and W. Dressler who defined text type as “a set of heuristics for producing, predicting and processing textual occurrences and hence acts as a prominent determiner of efficiency, effectiveness and appropriateness.”

Since several classifications of text types were found, the classification of E. Werlich and the classification of Hatim and Mason are going to be stated. Isabel García (2000: 228) mentioned Werlich’s classification of text types in which five text types are encountered: “narrative, descriptive, expository, argumentative and instructive.”

On the other hand, Hatim and Mason provided a tripartite classification of text types. Isabel García mentions this classification in his book *Análisis textual aplicado a la traducción* (2000: 236-242):

1. Argumentative texts → “the textual focus of these texts is to evaluate relations between different concepts” (Hatim&Mason, 1990). The evaluation of the relation among concepts is achieved through the use of arguments and counterarguments. Five different types of argumentation are mentioned by Bassols and Torrent (1996:57): simple, multiple, multiple coordinate, multiple subordinate and counterarguments.
2. Expository texts: “their contextual focus can be the analysis of some given concepts or the synthesis of concepts through their constituent elements” (Hatim&Mason, 1990). Within this text type, Hatim and Mason distinguish among conceptual exposition and narrative texts; as we have seen in

Werlich's classification, both of them were not regarded as subtypes of expository texts but as independent text types: conceptual and narrative texts.

3. Instructive texts: "their focus is the formation of future behaviors" (Hatim&Mason, 1990). "They are used to instruct, to influence the receiver/reader. They are directions for use" (Méndez Cendón, 2012-2013).

3.1.4. Genre

Muñoz Torres quoted Hurtado (1997:65), who defined genres as "textual groups which share conventional forms and situations of use." Hurtado also noted that "genres are quite important in translation since texts are grouped according to a set of functions, structural patterns and situations of use." Méndez Cendón (2012-2013) defined LSP genre as "a group of texts belonging to a specific expert community that share an overall structural organization, a communicative purpose, a rhetorical purpose, linguistic and pragmatic characteristics and a type of audience."

However, Muñoz Torres (2012) noted that "the concept of genre within Translation Studies is not clearly defined" but that the proposal for register "has a preference for the existence of common characteristics of a determined group of texts."

"Genre is situated in direct relation with register" (Halliday&Hasan, 1985), although Isabel García (2004: 207) clarifies that "they must not be confused." There is a continuous struggle in order to locate genres since some authors comprise them within the register and others consider both register and genre as "dos sistemes semiòtics separats" (X. Luna, 1994-95: 265).

Now, we are going to name some medical genres that Muñoz Torres distinguished by analysing the medical texts of the corpus that he had compiled. They obtained a total number of 35 medical genres within their corpus: "research article", "medical case reports", "clinical trial", "medical dictionary", "medical results", "clinical protocol", "clinical cases", "clinical recommendations"...

Since we have already stated the different text types, we are also going to provide a Gamero classification of genres by text type (adapted by Muñoz Torres (2012)):

“Argumentative genre with secondary expository focus”, “expositive genre with argumentative secondary focus”, “expositive genres”, “expositive genres with secondary instructive focus”, “instructive genres”, “instructive genres with secondary expositive focus” and “multifunctional genres”.

4. Terminology

For M.T. Cabré (1993: 82) and also J. Pearson (1998:10), who quoted Sager, ‘terminology’ refers to three different related things:

1. The set of practices and methods used for the collection, description and presentation of terms;
2. A discipline, i.e. the set of premises, arguments and conclusions required for explaining the relationships between concepts and terms which are fundamental for a coherent activity under 1;
3. A vocabulary of a special subject field.

According to M. T. Cabré (1993: 21) “terminology as a discipline is not a recent area, but it has experienced a huge reflexion about all of its bases, methods and principles. It has also achieved the recognition it deserves.” The huge growth of economy, the exchange of cultures and the tendency of globalization have provoked terminology into gaining crucial importance. Consequently, terminology is responsible for the naming of concepts.

Terminology plays a very important role in specialized translation since adequate terminology is required in order to correctly express the meaning in LSP (language for specific purposes) texts. As we have stated before, a scientific text has to be concise, precise and natural. Terminology plays a crucial role in the achievement of these aspects.

M. T. Cabré (1993: 167) stated that “terminology is only used within specialized language” but also that “there is a continuous transfer of terminology among general and specialized language” and that we have to distinguish between “belonging and being used in a specialized field” (M. T. Cabré, 1993: 168). A term can belong to a particular field, but it does not mean that it cannot be used in other subfields.

If we consider terminology as the vocabulary of a special subject field, the concepts of ‘term’ and ‘concept’ are quite important. M. T. Cabré (1993: 172) uses the concept of ‘term’ to refer to both the linguistic sign and the concept that it represents. Therefore, a

term can be defined as a linguistic sign which represents a specialized concept within a specific field of knowledge.

Regarding the concept, Wüster suggested that it is the starting point of terminology. According to the Merriam-Webster³ dictionary a concept is “an abstract or generic idea generalized from particular instances to which a denomination (term) is assigned.”

4.1. Types of terms

Terms can be classified according to their form, function, meaning and origin. However, we are going to deal with the types of terms regarding their form. M. T. Cabré (1993: 176) provided two main descriptions concerning:

1. The number of morphemes:
 - Simple term: ‘acid’
 - Complex Term: ‘acidification’
2. The types of morphemes:
 - Simple: ‘ulcer’
 - Derivative: ‘ulcerous’
 - Compound: ‘paperweight’

Terms can be also classified according to their components; we can distinguish among simple terms and compound terms. Since there are several patterns of formation of compound nouns, we are going to state the most common patterns in medical language (Françoise Salager, 1985: 6-11):

- N1+N2→
 - ✓ ‘Wall thickness’
 - ✓ ‘Laboratory test’

³ <http://www.merriam-webster.com/dictionary/concept> (accessed date 05/10/2014)

○ N1+N2+N3→

- ✓ 'Serum salicylate level'
- ✓ 'Phenobarbital plasma concentration'

○ N1 +(N2) -ing+N3→

- ✓ 'Drug metabolizing enzyme'
- ✓ 'Arrhythmia producing concentration'

○ N1+V-ed +N2→

- ✓ 'Virus induced necrosis'
- ✓ 'Iron-induced anaemia'

○ Adv+ adjective→

- ✓ 'Exceptionally valuable'
- ✓ 'Easily identified'

○ ADJ-ING+NOUN→

- ✓ 'Paralyzing enzymes'
- ✓ 'Developing cancer'

5. Corpus design and compilation

In this chapter the definition and the different types of corpus will be explained, along with which kind of corpus best suits our needs.

Sinclair (1991: 171) defined a corpus as “a collection of naturally occurring language texts, chosen to characterize a state or variety of a language”. In 2005 he defined corpus as “a collection of pieces of language texts in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research.”

This is one of the several existing definitions of “corpus” and also one of the best ones. Jitka Hrusková (2008: 8) uses corpus to refer to the “an ordered set of texts in electronic form that must have certain characteristics”.

Corpora have become crucial in order to analyse real language and describe its structure and its use for different applications such as translation, lexicography, grammar, sociolinguistics and language learning/teaching.

It is also important to know that corpus linguistics is used to study language, to analyse “real word” texts in order to find linguistic features which are characteristic of the genre. Corpora have become crucial in order to analyse real language and describe its structure and its use for different applications (translation, lexicography, grammar, sociolinguistics, language learning/teaching...).

5.1. Classifications of corpora

Authors like J.Sinclair (1996) proposed a classification taking into account the following different aspects:

- Size of the texts conforming the corpus
- Linguistic variety and level of specialization of the texts
- Purpose of the texts

- Number of languages

L.Bowker and J. Pearson also proposed their own classification of corpora. They distinguished among (2002: 11-13):

- General reference corpora vs. special purpose corpora
- Monolingual vs. multilingual corpora
- Open vs. closed corpora
- Learner corpora

5.1.1. Types of corpora in translation according to the number of languages belonging to the texts

As we are dealing with a translation from Spanish to English, the type of corpus to be chosen has to depend on the number of languages involved. There are two types of corpora in translation concerning languages:

- Monolingual corpora: A collection of texts in the same language. Its main uses in translation would be obtaining phraseology and terminology, as well as the study of language and its characteristics in a particular field.
- Bilingual corpora: A collection of texts in which two languages are involved.
- Multilingual corpora: A collection of texts in more than two languages. There are two types:
 - Comparable corpora, which is a collection of “similar texts” in different languages. The criterion is compiling texts with a similar size, date of publication, level of specialization, genre and text type.
 - A Parallel corpus consists of a collection of original texts together with their translations into one or more language(s).

The main difference stated by the Eagles- Expert Advisory Group on Language Engineering Standards Guidelines (1996)⁴ is that a comparable corpus “is used to compare different languages or varieties in similar circumstances of communication but avoiding the inevitable distortion introduced by the translations of a parallel corpus.”

5.1.2. Corpora according to the level of specialization of the texts

In general terms, we can find two types of corpora according to the level of specialization of the texts:

- General language corpora: According to Sinclair (1991:17) a general language (GL) corpus “is a collection of material which is broadly homogeneous, but which is gathered from a variety of sources.” Bowker and Pearson (2002: 11-12) also defined a GL corpus as “one that can be taken as representative of a given language as a whole and can therefore be used to make general observations about that particular language.”
- Specialized corpora “focuses on a particular aspect of a language. Because of its specialized nature, such a corpus cannot be used to make observations about language in general” (Bowker&Pearson, 2002: 12)

So, if we have understood Bowker and Pearson, the difference among GL corpora and specialized corpora is that the former one deal with pieces of ordinary language. The main example of specialized corpora is any corpora belonging to a specialized subject field, such as medicine or engineering. Specialized corpora are different from GL corpora in the sense that the first contains “a high proportion of unusual features” (Sinclair, 1994a:7). Bowker and Pearson (2002: 14) also state that “general corpora and specialized corpora can be used in a comparative fashion to identify those features of a specialized language that differ from general language.”

⁴ See more available information at <http://web.lettras.up.pt/bhsmaia/belinda/pubs/CL2003%20workshop.doc> (accessed date 09/10/2014)

5.2. Comparison Parallel, comparable and monolingual corpora for the purpose of translation: Advantages and disadvantages in translation studies

Parallel, comparable and monolingual corpora are useful in translation studies. All of them can be used with corpus exploitation tools in order to analyse and obtain valuable information for a translation. First of all we are going to state the main differences among parallel and comparable corpora since both deal with more than one language.

Parallel corpora have the disadvantage that translators might make mistakes and therefore these mistakes can be transmitted into the translated texts. On the other hand, by using a comparable corpus this problem disappears since it comprises original texts. By using comparable corpora, we deal with authentic language, not with translationese. This turns out to be very helpful in order to translate a text since the main goal in the field of translation is the naturalness of texts.

The problem of using a comparable corpus is that since it comprises original texts, translation equivalents cannot be extracted as easily as in parallel corpora. A parallel corpus provides us with those translation equivalents because we are dealing with original texts and their translations. Electronic linguistic tools can be used to extract those equivalents.

Besides, parallel corpora allow us to see how words and phrases have been translated in the past. Another advantage of using parallel corpora is that “a corpus of original texts and translations can be a rich source in the study of translations patterns”, according to Johansson (2003:137).

Finally, since English is the language of science, most scientific and technological texts are written in English. So, the problem with parallel corpora, as some researchers note (Maia, 2003; Varantola, 2003), is that “specialized texts are more often available in English but to a much lesser extent in other languages.” Therefore, there will be more problems by the time of compiling parallel corpora due to the lesser availability of texts that are not written in English.

On the other hand, monolingual corpora have some advantages over parallel and comparable corpora. That is the reason why monolingual corpora are used by the majority of translators.

Time is one of the most determinant advantages of using monolingual corpora. It takes too much less time and it is quite easier to compile a monolingual corpus than a bilingual or parallel corpus. Other important aspect is that by analysing a monolingual corpus we can obtain valuable information for our translation: terms, equivalents, patterns, collocates and so on.

To sum up, the main and most important advantage of using monolingual or comparable corpora is that since you are dealing with original texts you can extract authentic terminology and phraseology by using linguistic tools such as AntConc. On the other hand, by using parallel corpora, “you can enter a keyword or phrase in one language and the bilingual concordancing software will retrieve all of the sentences where this word or phrase appears” (Bowker&Pearson, 2002: 93).

5. 3. Considerations to be taken into account when designing a specialized corpus

As we have been saying through the whole project, our aim is to translate a medical text, so we have to create a corpus according to this purpose. We are going to create and design a specialized corpus. Bowker and Pearson (2002: 46) stated that in order to “design a special purpose corpus, special attention has to be paid to issues such as size, number of texts, medium, text type, authorship, language and publication date.” Bowker&Pearson (2002: 51) explain: “the types of texts that you include in your corpus wil depend on what you wish to study.”

5.3.1. Size

The first and one of the most controversial considerations to design a corpus is its size. There are and there have been many debates about to the number of tokens required to compile a corpus.

According to Bowker and Pearson (2002: 45) “there are no hard and fast rules to determine the ideal size of a corpus.” So, they named three crucial factors in order to decide the size of a corpus (2002: 45-46): “the needs of your project, the availability of data and the amount of time that you have.”

We cannot include all corpora within the same group, not all of them share the same characteristics and more importantly, not all types of language have the same number of texts available.

As Bowker and Pearson said, the representativeness of a corpus will be related to the availability of data. For instance, writing about the songs of a specific singer is not the same as writing about medicine. In the former case, the corpus would be representative if you take all or almost all of the songs that exist; if there were only 10 songs, you might have only 1,000 tokens for example, but if you take the 10 songs, the corpus will be representative of this particular genre.

In the case of medicine, you know that there are thousands and thousands of texts dealing with medical issues, so a 1,000 token corpus will not be representative. You will have to compile many texts concerning medical issues. For example if you were going to make a corpus of general medicine, you would have to compile texts coming from every branch of medicine: cardiology, urology, pediatric and so forth.

Since a corpus has to be designed according to the needs of the compiler, the size of the corpus will also vary. Our corpus will be used for translation purposes since the aim of this undergraduate dissertation is to translate a medical text. The topic of the text is very specific (a disease), so, the size of our corpus will be related to the availability of texts about diabetes. Further details about the size of our corpus will be given in section 5.4.

5.3.2. Subject

The next criterion is the topic of the texts compiled. Bowker and Pearson stated that “although may seem too obvious to mention, the text that you include in your corpus need to be about the specialized subject that you are studying”.

5.3.3. Setting, audience, technicality and authorship

The texts must also share the same setting, audience and level of technicality which is “an attribute based on the degree of specialist/technical knowledge of the author and target readership” (Atkins et al. 1992:8). However, it is unnecessary that they share the same author since then it would be impossible to achieve representativeness unless you are analysing the features of the language used by that specific author. But as J. Pearson states (1998:60) “the author must be an acknowledged individual or institution.”

5.4. My own corpus

Now that the criteria for corpus design and compilation has been stated, we are going to explain in this section the characteristics and purpose of our corpus.

The first aspect to be taken into account is quite clear, the purpose of the corpus. Our aim is to translate a journalist’s text about diabetes published in *Elmundo.com*. So, the purpose of the corpus is to help us with the translation. From this point, we have to design our corpus according to the characteristics of the source text in order to obtain the best results. In order to be clear, we are going to explain each characteristic of our corpus by keeping the order we have followed in this chapter.

For our purpose, the first choice is quite clear. We have to design a specialized corpus in order to obtain adequate information for our translation. We have to compile texts about a specific subject field: diabetes. So, once the corpus is compiled, valuable information can be extracted and analysed in order to satisfy our needs.

Concerning the number of languages, we have to decide between compiling a monolingual or a bilingual corpus. As it is explained in this chapter, both corpora have their advantages and disadvantages. A monolingual corpus can provide us with translation equivalents which are very useful for our translation. While in bilingual corpora the advantages vary depending on if we choose a parallel or comparable corpus.

A very decisive factor at the time of choosing a type is time. Due to the accessibility of original texts, compiling a comparable or monolingual corpus takes far less time than compiling a parallel one since it is very difficult to find texts together with their translations. This is the reason why comparable corpora are more used than parallel corpora in translation. Anyway, some translators (although it takes a lot of time) combine the use of these three types of corpora. So, by doing it, the translator can make use of all the advantages that these types of corpora provide.

We have finally decided to compile a bilingual parallel specialized corpus with original English texts and their translations in Spanish. The reason for compiling this kind of corpus is that we can take advantage of the characteristics of both a monolingual and a bilingual corpus.

On the one hand we have a bilingual parallel corpus with English original texts and their translations in Spanish. By aligning the texts of the parallel corpus we can use WordFast (explained in the translation tools section) in order to see how certain terms or phrases have been translated. This was a reason for compiling a parallel corpus, to obtain already made translations.

On the other hand, we can take only the English original texts; we can separate them from the translations in Spanish in order to have a monolingual corpus. As we are going to explain in the Translation tools section, programs such as AntConc and Wordsmith tools were used to exploit the monolingual corpus in order to extract terminological and phraseological units.

When we reached the matter of size, we face the problem that the number of resources belonging to students is narrower than the resources of a professional translator. So, we basically compiled texts from the following official webpages: emedicinehealth, NIH, Medscape and Niddk. We selected these webpages because both the original English texts and their translations into Spanish were available in these webpages. So, we were able to compile a whole bilingual parallel corpus getting the texts from these pages.

Medscape is a medical database in which you can find texts addressed to specialists, experts and also consumers. *Emedicinehealth* also provides consumers with health information, so we have an expert to non-expert communication. *NIH* (National Institute of Diabetes) is an American governmental institution based on medical research in which we are going to find texts from expert communication. And finally *NIDDK* (National Institution of Diabetes and Digestive and Kidney Diseases) is also based on medical research (performed by experts and also by students). The *NIDDK* researches are directed both to experts and also to the public (patients, families, and friends). Many other medical webpages with a similar genre and register could have been used; however they required annual or monthly payment.

As we see, all the texts compiled from these webpages do not share the same exact genre or register since there are some differences among them. Although most of them deal with expert communication, we can also find articles about findings or developments in medicine. These articles about medicine are not written by experts but by journalists.

Finally we were able to compile a 70,000 token corpus in which we can find both texts from expert-expert communication and texts from expert/semi-expert-layman communication. So, by the time of translating the text we will have both instances about specialized language and journalist medical language.

Overall, despite the differences among some texts, our translation of the medical text from *El Mundo* can be performed successfully through the use of linguistic tools such as dictionaries, glossaries, online translators, online translation memories and also corpus exploitation tools. In chapters 6 and 7 we are going to explain and illustrate the use of each of these translation tools.

6. Translation Tools: usage and advantages

All the former chapters explained the different aspects (what translation is, terminology, corpus etc) and steps (corpus design and compilation) that have to be performed before starting a translation. This chapter deals with the the usage of tools used in this undergraduate dissertation in order to translate the text about ‘diabetes’.

The first aspect to be taken into account is the availability that a student has regarding translation tools. Many of them are not free and therefore they cannot be used unless you pay a huge amount of money. So, we are going to list and explain those tools that are free and available for all beginner translators and translation students.

It is obvious that those tools that are not free would provide with further more options and advantages than those which are free of charge. However, this does not mean that we are not going to be able to translate the same text correctly. There are several tools that are available online and also other ones such as specialized dictionaries or glossaries that are not very expensive.

Since nowadays everybody has access to the Internet, we are able to manage the available tools to write our translation. First we are going to explain the different tools, their functions and how they work. In the next chapter the results and searches for translating the text obtained by using these tools are going to be explained.

6. 1. Corpus exploitation tools

As we have seen in chapter 4, once a corpus has been compiled under the criteria corresponding to our purposes, it can start being searched. We compiled a parallel specialized corpus with original English texts and their translations into Spanish.

Corpus exploitation tools such as *WordFast pro* and *AntConc* were the main tools used in order to achieve helpful information for the translation of the text. *Wordfast pro* needs to be fed with translation memories and that is the reason why we compiled a

parallel corpus. The other reason for compiling this kind of corpus was to have an English monolingual corpus so that by using *AntConc* we could obtain terms together with their context.

6.1.1. AntConc

Both *AntConc* and *Wordsmith Tools* allow us to analyse any corpus. In terms of functions they are quite similar: you can produce a word list, a keyword list and you can see clusters and collocates in concordance lines. However, we find some limitations since we are dealing with demo versions of the programs.

Due to this problem and since both programs are similar, we decided to use *AntConc* and leave aside *Wordsmith Tools*. The reason was quite simple; *Wordsmith tools* had more limitations than *AntConc*. The only problem was that with *Wordsmith tools*, although it provided us with the same functions), it did not provide us with a sufficient number of instances to be representative. It only provides us with thirty or forty examples of each word (its collocates, patterns and keyword list). These “instances” of each word are called “concordance lines.”

This problem was solved with the latter tool (*AntConc*). *AntConc* provided us with all of the possible collocates, and with words lists with all of the words that occurred in the original texts. By using concordance lines we were able to extract information of terms in their contexts and also equivalents. In order to be clear and not confuse the reader, the diverse available functions of *AntConc* are going to be explained and illustrated by using some screenshots.

6.1.1.1. Concordance tools

The most important and widely used tool is the concordancer. This function provides us with concordance lines which let us see the term in context. It is very important to state what the concordancer does: you can look for a word or phrase in the concordancer and

the program will show you on the same screen the number of hits and all the instances in which that word or phrase occurs.

This is a very helpful tool as normally the word or phrase you have looked for in the concordancer appears highlighted in bold in the centre of each concordance line, so it is very easy to locate and consequently analyse it within the context. Image 1 below shows some concordance lines for the search term *diabetes* in our corpus. It is an ordinary search.

However, in order to see the behaviour of a term and its usage, we cannot only base our result on the immediate words that follow the term. Hunston (2002:58-60) noted that “only by looking further than the immediate co-text” you can analyse the usage and functions of an item in the text. Image 2 below shows how by clicking on the word in bold in the concordance lines, immediately the full context of the word occurs.

In order to analyse these concordance lines, we have taken some of the steps that John Sinclair described in his book *Reading Concordances* (2003):

1. Take a look at the words that are on the right and left of the searched term.
2. See if there is any kind of relationship among the repeated words
3. Write a hypothesis
4. Try to confirm that hypothesis
5. Look for unusual aspects in the rest of the concordance lines

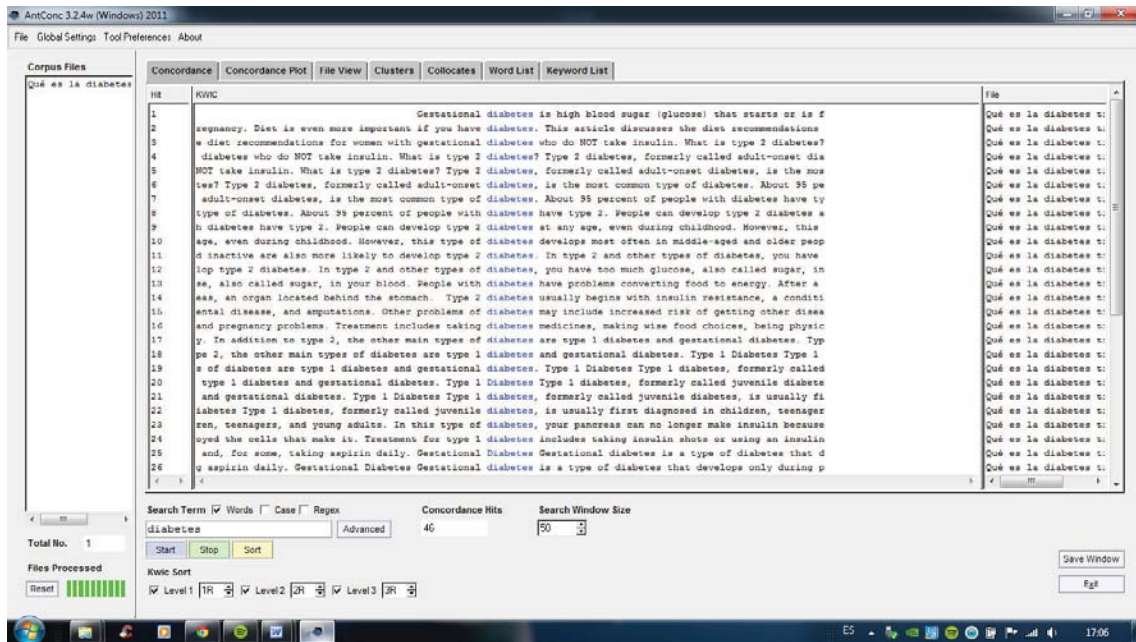


Image 1: Basic concordance for *diabetes*

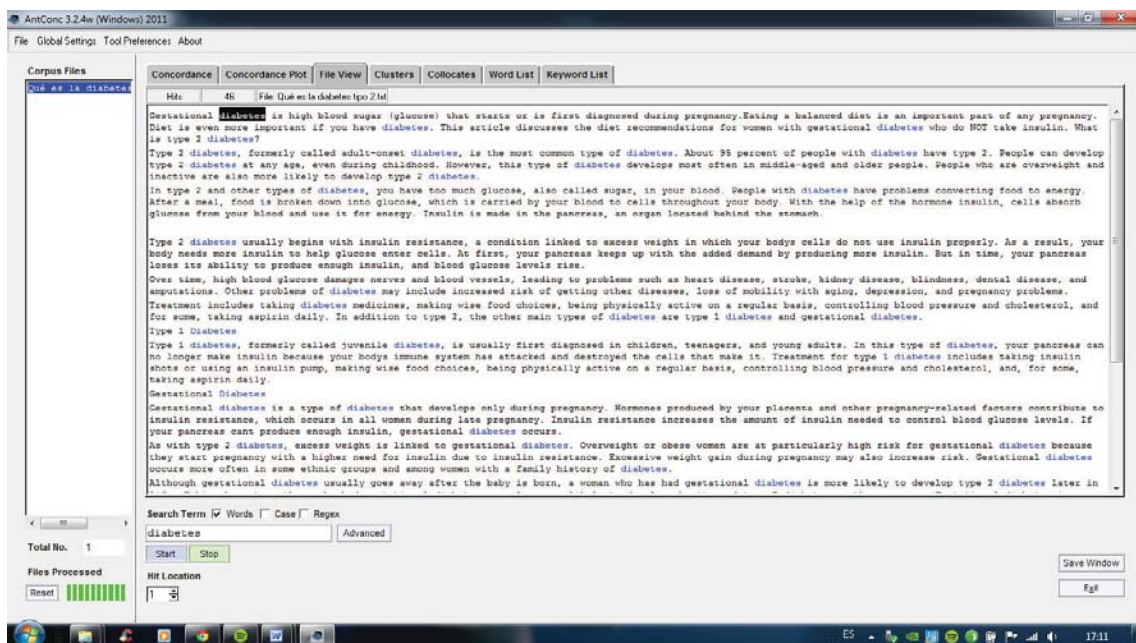


Image 2: Full context of *diabetes*

6.1.1.2. Frequency list

AntConc also has the possibility of making a frequency list with all the words that occur in our corpus. This list provides us with the number of occurrences of each word in the

corpus. Table 1 below shows a frequency list without any sort of modification (from words with the highest number of occurrences to words with the lowest number).

Among the categories of words contained in the corpus there are, of course, lexical words and grammatical words. Grammatical words are words that do not have a lexical meaning but a functional one, that is to say auxiliary and modal verbs, prepositions, conjunctions and determiners.

The lexical category comprises words that have a full meaning: noun, adjectives, adverbs and verbs. It is important to know that grammatical words are the most frequent ones in these lists; however, these words do not provide us with helpful information.

Frequency lists give us the number of occurrences of all words. Therefore a stop list (a list of adverbs, prepositions, auxiliary verbs) must be uploaded to the *AntConc* program so that these grammatical words do not occur in the frequency list. Table 2 shows the results of the frequency list after uploading a stop list.

Another problem occurs with *lexemes* and *lemmas*. The frequency list will also show you the singular and the plural of the same word in separate entries. For example, the word *cell* in its singular form occurs in the frequency list with a frequency of 300 occurrences whereas its plural form *cells* may occur in a different entry with 200 occurrences. The same occurs with tenses in verbs i.e. develop-developing-developed.

So, in order to know the total number of occurrences of a word (singular and plural or the different tenses) a lemma list also has to be uploaded. Table 3 below shows a frequency list in which all lexemes are joined under the same lemma.

Total No. of Word Types: 4325		
Total No. of Word Tokens: 65274		
Rank	Frequency	Word
1	1799	your
2	1719	The
3	1698	To
4	1685	And
5	1294	Of
6	1258	A
7	1232	Diabetes
8	1200	You
9	1171	Or
10	1071	Blood

Table 1 Basic frequency list

Rank	Frequency	Word
1	1232	Diabetes
2	1071	Blood
3	638	Glucose
4	335	Problems
5	301	Type
6	300	Care
7	295	Doctor
8	277	Sugar
9	272	Insulin
10	267	Health

Table 2 Frequency list once a stop list has been uploaded

Rank	Frequency	Word	Lemma word form(s)
1	1232	Diabetes	
2	1071	Blood	
3	638	Glucose	
4	369	Problem	Problem 34 Problems 335
5	338	Type	Type 301 Types 37
6	327	Level	Level 140 Levels 187
7	305	Doctor	Doctor 295 Doctors 10
8	304	Care	Care 300 Cares 1 Caring 3
9	293	Foot	Feet 205 Foot 88
10	279	Sugar	Sugar 277 Sugars 2

Table 3 Frequency list once a lemma list has been uploaded

6.1.1.3. Collocations

Collocates are “actual words in habitual company”. Collocation is a mechanisms to achieve lexical cohesion in which certain words are more likely to occur in combination with other words in certain contexts. A collocate is therefore a word which occurs within the cotext of another word Firth (1957:14).

These “words in habitual company” can be found in the concordance lines. Words that constantly appear near the searched word are the right or left collocates of that word

respectively. Collocates can be classified according to the strength of the collocations (Lewis&Hill, 1998: 2):

- Weak collocation (fast car, have dinner...)
- Medium strength collocation (magnificent house, relatively strong...)
- Strong collocation (budding author, avid reader...)

As Lewis & Hill (1998:2-3) state “most collocations lie in the middle round of the cline: there are very few strong collocations. “ By looking for collocations, you are going to discover specific language patterns. Table 4 shows some collocations of *Diabetes*.

Key term	Diabetes		
Collocations	Verb + key term ...taking diabetes... ...have diabetes... ...developing diabetes... ...getting diabetes... ...avoiding diabetes...	Key term + verb: diabetes has... diabetes may include... diabetes is delayed... diabetes treated... diabetes can lead to...	Adj + key term: Gestational diabetes Juvenile diabetes

Table 4: Collocations of Diabetes

6.1.1.4. Clusters

The “clusters” tool is used in order to generate a list of sets of words that occur next to the search term. Instead of reading each concordance line in order to look for these clusters, this function already provides you with a list. The minimum length of the cluster and also the minimum number of occurrences can be selected in order to choose those which we are looking for. Image 3 below shows the clusters of *blood*. As we can see in the figure, we can obtain multiple complex terms that derive from a simple term.

As we will see in the results, this function allows us to extract terminology and search for equivalents.

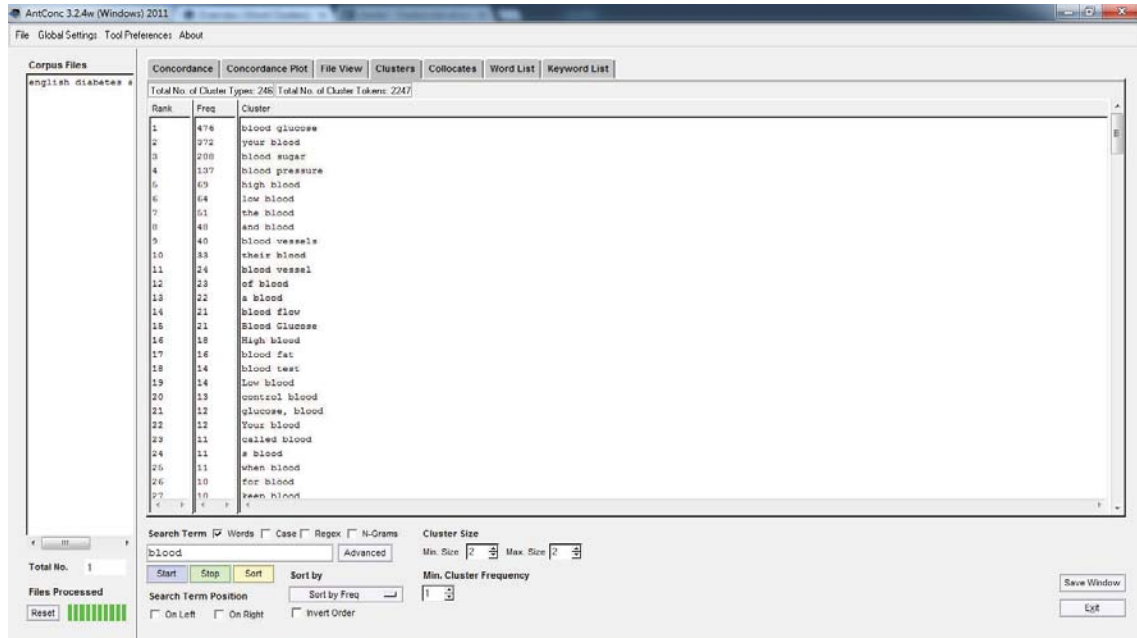


Image 3 Clusters of blood

Figure 3 shows clusters of blood with a minimum size of 2, while Image 4 shows the results improved by implementing the minimum size to 4 or 5.

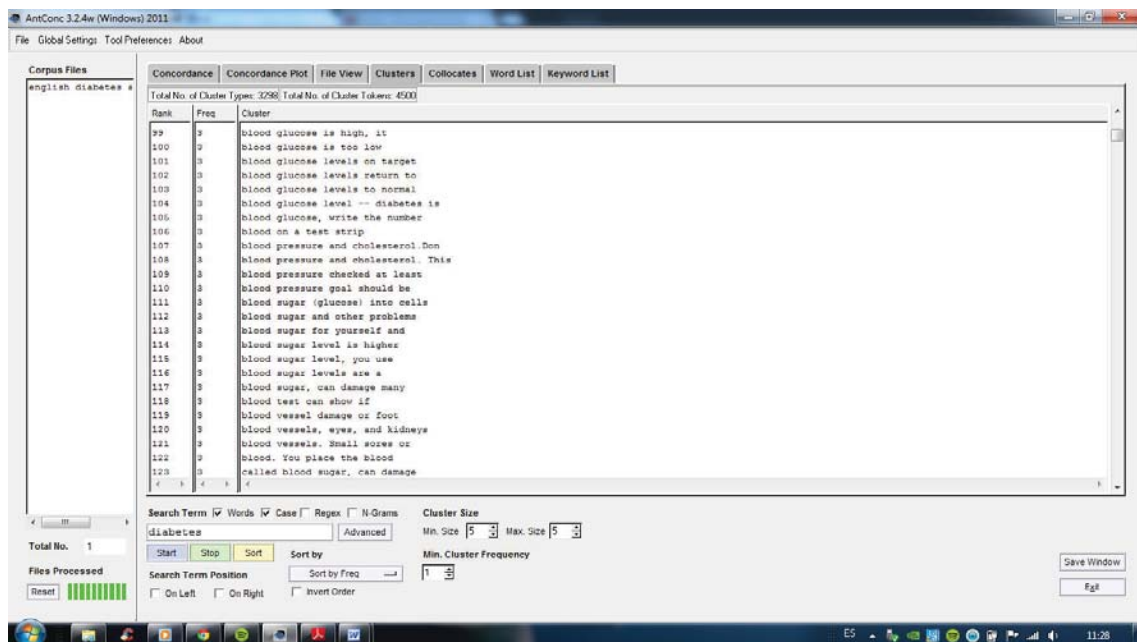


Image 4: Clusters of 'blood' with a minimum size of 4

As we see in the figure above, we can appreciate compound terms derived from ‘blood’: ‘blood glucose level’, ‘blood sugar level’, ‘blood test and ‘blood vessels’. So, the cluster function provides us with a very useful tool in order to extract terminology.

6.1.2. Bilingual concordance tools: ParaConc and MultiConcord

In order to exploit our corpus, it was also possible to use a bilingual concordance tools. These tools need to be fed with a parallel corpus. Many of these programs require the alignment of the texts before starting to analyse them.

ParaConc and MultiConcord are the main bilingual concordance tools and allow us to see how certain terms or phrases in the source language have been translated into the target language(s). The problem with these programs is that due to their level of technology and specialization they are not free. These programs have not been used for the translation of the given text, but it is useful to know that there are also tools in order to see how terms have been translated into other languages without doing it manually.

6.1.3. Wordfast Pro

Wordfast pro, since it is fed by a translation memory (a linguistic database that “learns” from the translator) provides us with terminological units (complex or simple) or phrases that have been already translated.

In order to create a translation memory, translators need to align their translations manually or automatically. Alignment⁵ is the process by which a translation memory is created through the source texts and their translations.

From translation memories we can obtain a “100%” or a partial match. I mean, by using *Wordfast Pro*, when you enter a phrase, the program may provide you with a complete translation of the phrase (100% match). If only a part of the phrase has been translated in the TM, the program will provide you with a partial match, will provide you with the

⁵ In this webpage you can automatically align texts: <http://www.freetm.com/> (accessed date 05/09/2014)

part of the phrase that is already translated. So once you have a good translation memory it is very helpful as it can translate many parts of the text automatically. However, there is the possibility that translators make mistakes and transmit their own style into the translation so you have to be very careful when using this program. Before copying and pasting the translation of the respective phrase or term you have to check if it is correct.

This program is very helpful when you have your own translation memory since the results you have are a product of your own work. But in this case since I have compiled the texts and its translation from the Internet, I have decided to not trust the results and make use of the monolingual corpus. Therefore, the main tool I have used to exploit the corpus is AntConc.

6.2. Further translation tools: online dictionaries and glossaries

In order to complement and improve the accuracy of our terminological decisions, online medical dictionaries and glossaries together with computer-aided translator tools have been used. As we stated in the theoretical part of this undergraduate dissertation, terminology plays an important role in translation. That is the reason why the use of specialized dictionaries and glossaries is absolutely necessary. Since using a dictionary is a very well known task and almost everybody knows how to do it, we are going to avoid the explanation of its usage and advantages. However, the use of a glossary, although it seems similar, is a bit different from the use of a dictionary.

One of the most popular online translator tools is *Linguee*. Since it is easier to understand through examples, we are going to explain the usage of this program in the next chapter. Now we are only going to explain what *Linguee* is. Most people believe that *Linguee* is an online translator, however this is not true. *Linguee* is a combination of a dictionary and an online translation memory. You can search for terms or also for some phrases and if they have been translated by anyone, it will appear on the left of the original text with the phrase, word or term highlighted in yellow and on the right of the

translation the search phrase, word or term also highlighted. Since it is not an automatic translator it will not translate complete phrases or texts.

We could say that *Im translator* is close to being an automatic translator. When you enter a text, phrase or term in this online program, several translations from several translation engines are obtained and compared in order to provide the user with several possible translations. As we will see in the results, some translations are quite good. However, special care must also be taken.

7. Solutions to terminological and phraseological problems in our text translation

As it is explained in the previous sections, we can solve problems dealing with terminology or phraseology by different means. In this chapter we are going to explain the different searches made in order to find the correct equivalents for unknown terms. But before that, we are going to analyse the characteristics of our ST in order to make our translation adequate for them.

7.1. Analysis of the text object of this undergraduate dissertation

Since we have already explained the notions of genre, audience and register, we are only going to provide a simple analysis of the ST with some explanations about why it belongs to a particular genre, register, audience or tone.

The text is unmarked since it is directed to a general audience. The author uses an impersonal tone because he only wants to transmit information objectively; the author only provides information and explains facts about diabetes. Regarding the register, it is technical since we are dealing with a medical text, which belongs to the expert to non-expert communicative setting. The information is transmitted through a text, so the mode is obviously written. Finally it is an expository conceptual text since facts about diabetes are presented and explained without being evaluated by the author.

7.2. Searches for terminological and phraseological equivalents

When we started to translate our text about diabetes we found both simple and complex terms and also phrases that were problematic. Many of the terms were unknown and therefore their translation to the target language was a challenge. We couldn't simply take a glossary or dictionary and select the first available equivalent.

In order to choose the correct equivalent, further research had to be done. Since analysing and explaining the different researches of each unknown term made by several different translations tools is a very tedious task, in this section we are only

providing a full description and illustration of the search procedures of the most complex terms. Among these search procedures we are going to include screenshots of the search of these terms in different programs: *Linguee*, *Im translator*, *Wordfast pro*- or on online medical dictionaries and glossaries etc. This is to show to possible future translators the different available solutions to solve terminological problems.

Firstly we are going to explain how we solved terminological problems of our text. As it is shown in Image 5, words which could be a problem for the translator are highlighted. Once we know the possible translation difficulties, we are going to provide solutions to them by using the several translation programs explained in the previous chapter. Since we are going to illustrate how we solved terminological problems and how we obtained the equivalents of a few source text terms, we are going to explain the different possible ways to do it.

PROBADA EN ANIMALES
Una nueva esperanza para la diabetes tipo 1

Unas magnetocápsulas ayudan a 'vigilar' los trasplantes de células pancreáticas.

CRISTINA G. LUCIO

MADRID.- Decir adiós a las inyecciones diarias de insulina. El sueño de millones de diabéticos podría estar un poco más cerca gracias a una nueva técnica desarrollada por científicos estadounidenses. Sólo se ha probado en animales, pero permite transplantar de forma más efectiva células pancreáticas en un cuerpo que no produce insulina.

Hoy en día, algunas personas con diabetes tipo 1 se someten a un trasplante de células pancreáticas. Concretamente reciben una agrupación de células - denominadas islotes de Langerhans- que permiten regular los niveles de insulina en su organismo. Lamentablemente, en muchas ocasiones, el trasplante se rechaza al poco tiempo sin que se conozca el porqué.

Según los expertos, parte del problema reside en que, una vez introducidas en el organismo, es imposible seguir la pista de estas células o saber si el sistema inmune del paciente las ha destruido.

Proteger y vigilar las células

Precisamente para superar este escollo, investigadores del centro Johns Hopkins de Estados Unidos diseñaron una nueva técnica que consiste en rodear las agrupaciones de células de un material que impide que el organismo las reconozca como un cuerpo extraño y las ataque, pero permite que puedan seguir segregando insulina. A este material le añadieron una sustancia magnética que contiene hierro, llamada **Feridex**, que es visible a través de una resonancia magnética.

Los científicos probaron la técnica en ratones diabéticos y en cerdos y, en ambos casos, las magnetocápsulas, tal como las denominan sus creadores, se mostraron efectivas ya que pudo controlarse su evolución desde el exterior y no fueron eliminadas por el sistema inmune que, en el organismo de un diabético, ataca a las células productoras de insulina.

"La importancia de nuestros hallazgos es que ahora podemos utilizar técnicas de imagen convencionales para vigilar la evolución de células transplantadas, como los islotes pancreáticos," explica a elmundo.es Aravind Arepally, profesor adjunto de radiología y cirugía en el centro estadounidense y uno de los autores de la investigación.

Los resultados de su trabajo se publican en el último número de la revista 'Nature Medicine' (<http://www.nature.com/nm/index.html>).

Experimentos en ratones y cerdos

Los investigadores midieron en primer lugar el efecto de las magnetocápsulas, cada una de las cuales contenía entre 500 y 1.000 células productoras de

Image 5: Spanish article with problematic words and phrases highlighted

As the text was too big to include all pages with the highlighted problematic words, we are only going to include the first page. Also, as stated before, due to the process of explaining all the searches of all terms being a tedious task, only the searches of a few selected terms are going to be explained. The first problematic term chosen was ‘inyecciones diarias de insulina’. So the process we followed to find its equivalent is the following:

Firstly, since we were not very familiar with the term, we used *Linguee* in order to obtain the possible different ready-made translations. As we see in Image 6 below the term “inyecciones diarias de insulina” has been already translated and we obtain the following translations: ‘daily insulin injections’, ‘daily injections of insulin’, ‘daily injecting of insulin’ and ‘daily insulin shots’.

Among the first six results, three of them translated the multi-word term as ‘daily injections of insulin’ and the three other results were divided among the remaining multi-word terms. Therefore, ‘daily injections of insulin’ may be the correct translation. However, many terms can be translated into several ways. So, further research has to be done in order to check the correct equivalent. It is also likely that all the results are correct; in this case we will choose the one that is more widely used and therefore its frequency of occurrence is higher.



Image 6: Results for the search *inyecciones diarias de insulina* in Linguee

The main tool we used for extracting terms and equivalents was AntConc tools, so we exploited our English monolingual corpus in order to verify which translation is correct. ‘Daily injecting of insulin’ was excluded from the search; we considered that it was an incorrect translation since a verb (instead of a noun) is acting as the nucleus of the multi-word term. So, we performed the search of the three remaining terms.

It was clear that ‘daily’ was the correct equivalent for ‘diaria.’ And since we did not obtain any results in our corpus for any of the terms, ‘daily’ was removed from the multi-word term. Therefore Image 7 and 8 show the results for ‘insulin shots’ and ‘insulin injections’ respectively.

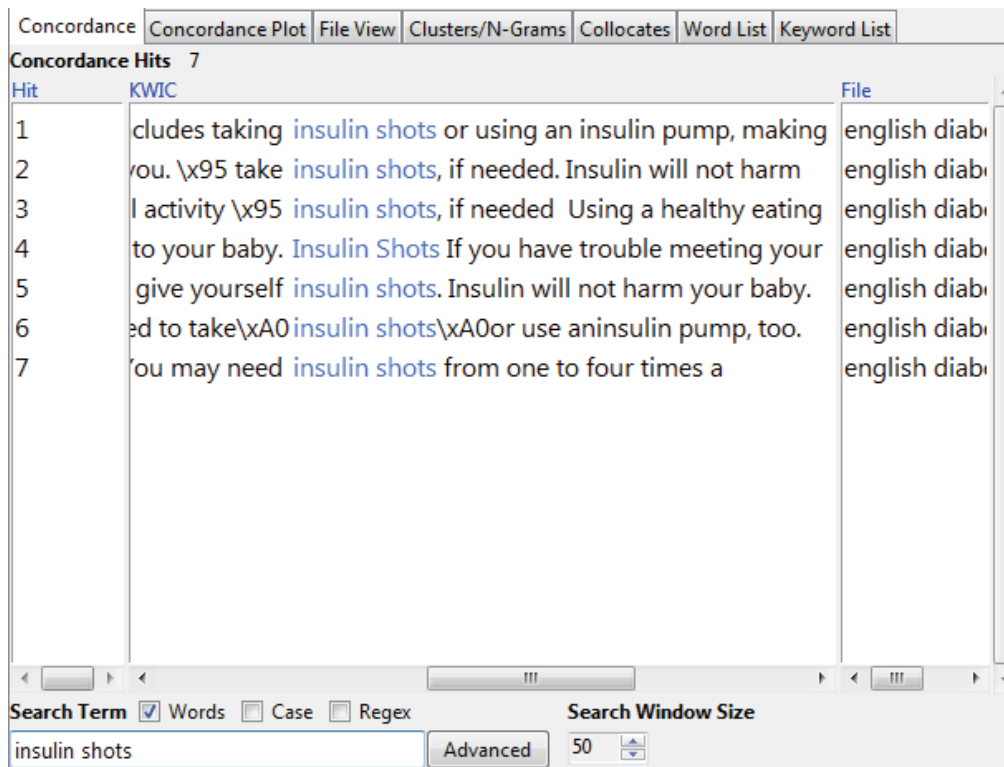


Image 7: results for *insulin shots*

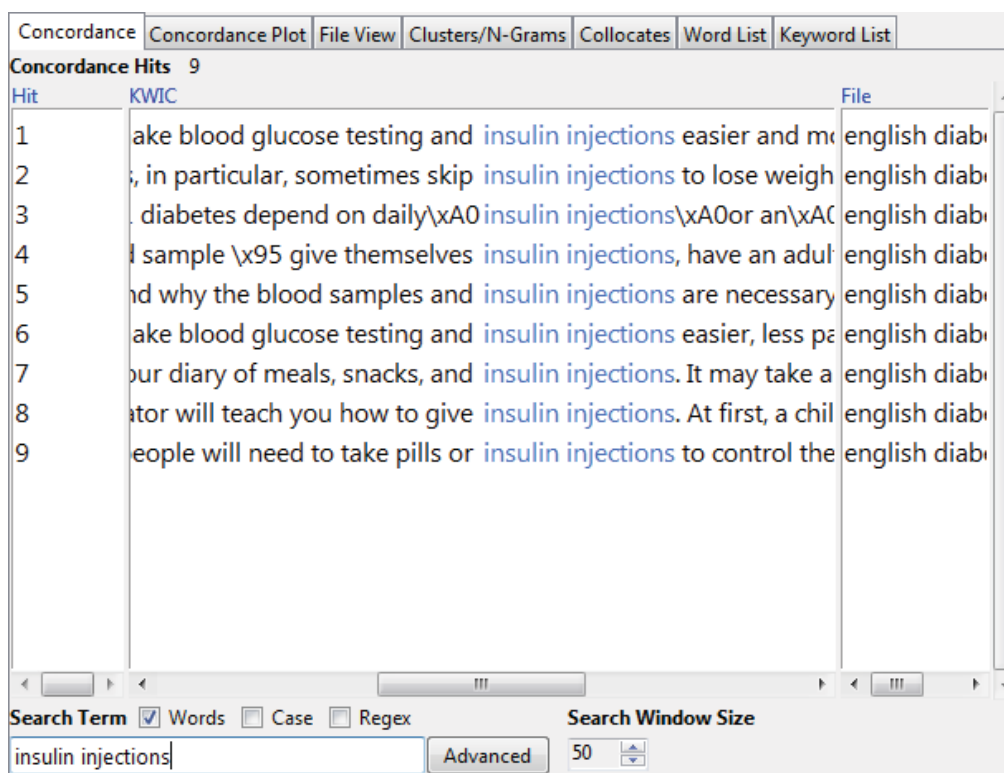


Image 8: Results for *insulin injections*

As ‘injections of insulin’ did not obtain any hits within the corpus we have avoided its image. However, ‘insulin shots’ and ‘insulin injections’ obtained 7 and 9 hits respectively as we can see in Images 7 and 8.

Finally, two options remained: ‘daily insulin shots’ and ‘daily insulin injections’. We had to choose between these two terms. Both of them were correct and we could not make a decision according to our corpus since the number of hits of each one was quite similar. So, in order to undo this equality, we searched them on Google in order to see which one was the most used.

As we can see on the Images 9 and 10, ‘daily insulin injections’ obtained around three million results while ‘daily insulin shots’ obtained a much lesser amount (215,000). ‘Daily insulin injections’ was therefore the term we used in our translation.

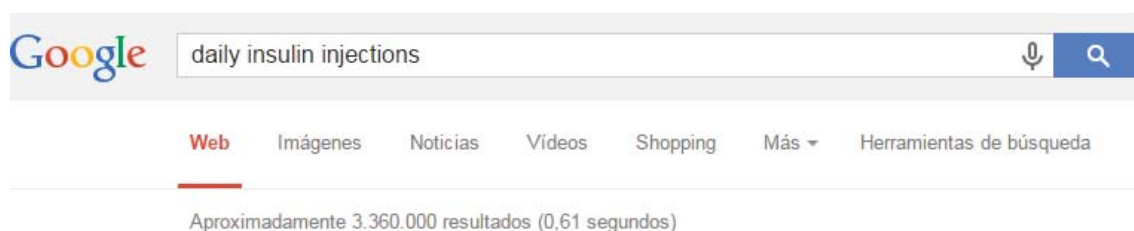


Image 9: Number of results for the search *daily insulin injections*

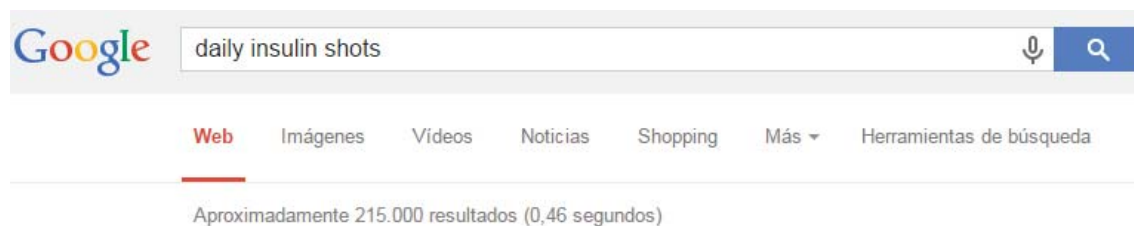


Image 10: Number of results for the search *daily insulin shots*

As we have seen in the process of finding the correct equivalent for ‘inyecciones diarias de insulina’, we cannot pick the first translation we find. It does not matter if you have found it in your corpus, in a specialized dictionary or in a translation engine. Further research has to be done in order to check the “correctness” of the equivalent.

The term ‘células productoras de insulina’ was also problematic. In order to not be repetitive we are going to show a slightly different process. Since in our corpus we compiled similar texts to the one we had to translate, we use our corpus in order to find if the equivalents of the terms are used in any text within the corpus.

On this occasion we are going to add the search of the term in Wordfast pro. We also built a parallel corpus with original texts and their translations in Spanish, so if the term is translated within the corpus, the WordFast Pro tool will provide us with the consequent equivalent. Image 11 below shows that the exact term was not found in the original texts. However, a close term was found: ‘células productoras de insulina del páncreas’. It has been translated as ‘cells in the pancreas that produce insulin’. We only had to delete the translation of ‘del pancreas’ and we got the equivalent for our term.

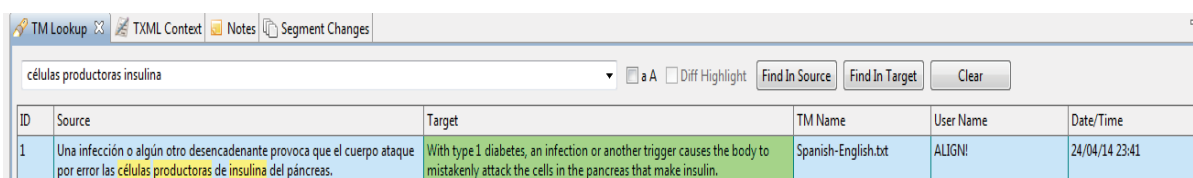


Image 11: Results for the search of *células productoras de insulina* in WordFast Pro

Although the translation seemed correct, we had to check it. We searched ‘cells that produce insulin’ in the Proz glossary and it did not occur in any of the 1,232 results obtained. However, as we can see on Image 12, one possible equivalent of the term is ‘insulin producing cells’.

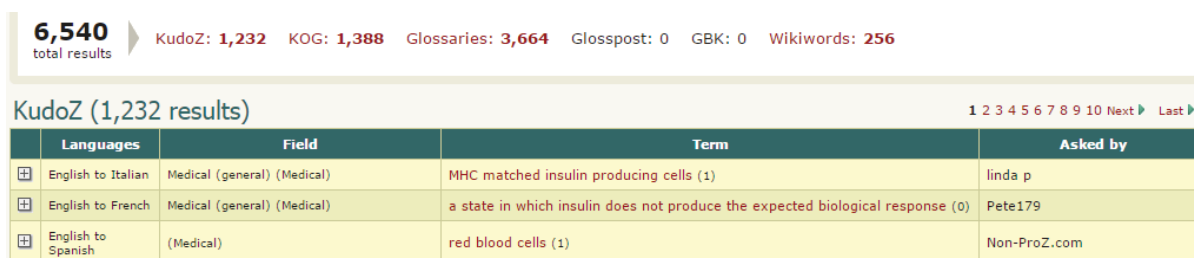


Image 12: Search for *cells that produce insulin* in the Proz glossary

We discarded ‘cells that produce insulin’ for our translation and therefore we proceeded to check if the last term was correct. For that purpose, since we only obtained one

matching ‘insulin-producing cells’ in our corpus, we used Linguee in order to see if anyone had translated the term. Image 13 shows clearly that the term was absolutely correct.

This involves transplanting insulin-producing cells into the body of a person with diabetes. <small>↳ levinechildrenshospital.org</small>	Consiste en el trasplante de células productoras de insulina en el organismo de la persona con diabetes. <small>↳ levinechildrenshospital.org</small>
[...] every year thereafter, and tested each for the presence of three antibodies capable of harming insulin-producing cells . <small>↳ vnaicarenewengland.org</small>	[...] entonces, también analizaron a cada uno en busca de tres anticuerpos capaces de dañar a las células productoras de insulina . <small>↳ vnaicarenewengland.org</small>
[...] indefinitely, which is why people who've had transplants of insulin-producing cells eventually must go back on insulin. <small>↳ womenshealth.gov</small>	[...] razón por la cual las personas que han recibido trasplantes de estas células , con el tiempo deben volver a recibir insulina. <small>↳ womenshealth.gov</small>
This lack of insulin occurs when the bodys immune system attacks and destroys the insulin-producing cells of the pancreas, a process called islet autoimmunity (IA). <small>↳ vnaicarenewengland.org</small>	Esta falta de insulina sucede cuando el sistema inmunológico del cuerpo ataca y destruye a las células productoras de insulina del páncreas , un proceso llamado autoinmunidad de islote (IA). <small>↳ vnaicarenewengland.org</small>

Image 13: Results for the search of *insulin producing cells* in Linguee

We also searched in *MedicalTerminologyDB* and also in the online medical glossaries of the following webpages: *MedlinePlus* and *Emedicinehealth*. However, we did not find any matches on these online databases or dictionaries, maybe because it is a very complex term.

So, in order to show the usefulness of online glossaries and dictionaries we are going to illustrate the search of the equivalent of ‘Islotes de Langerhans’. I had never heard or seen this term before, so I had to guess what its equivalent might be in English. My first and only guess was translating it literally: Langerhans islets. Firstly, we searched the term in the dictionaries of *Emedicinehealth*. Image 14 shows that in one of the results in the medical dictionary the correct term appeared: *Islets of Langerhans*.

The screenshot shows the eMedicineHealth website interface. At the top, there are four featured articles with images: 'Inflammatory Bowel Disease', 'Foods That Help or Harm Sleep', 'Joint-Friendly Exercises', and 'Black Widow vs. Brown Recluse'. Below these is a 'Most Popular Topics' list with 9 items, including 'Psoriasis Pictures' and 'Hep C Medications'. The main content area is titled 'Search Results:' and shows '19 articles on "Langerhans islets" found.' Below this, there is a section for 'Slideshows, Images and Quizzes (found 1)' with a link to 'Picture of Pancreas - Image'. The 'Medical Dictionary (found 18)' section lists several entries: 'Islets of Langerhans', 'Langerhans, islets of', 'Pancreas', and 'Alpha cell, pancreatic', each with a source attribution to MedicineNet.

Image 14: Search of *Langerhans islets* in *Emedicinehealth*

‘Islets of Langerhans’ was used in our translation as these kinds of dictionaries are reliable; they are created by experts of medical associations. Nevertheless, we made a final check in the other medical webpage mentioned above: *Medline Plus*.



Medical Dictionary

One entry found.

Main Entry: **islet of Lang-er-hans**

Pronunciation: \-'lāŋ-er-,hānz, -hān(t)s\

Function: *noun*

: any of the groups of small slightly granular endocrine cells that form anastomosing trabeculae among the tubules and alveoli of the pancreas and secrete insulin and glucagon—called also *islet*

Lang-er-hans \-'lāŋ-er-,hāns\ **Paul (1847–1888)**, German pathologist. Langerhans is notable for his studies of human and animal histology. He was among the first investigators to successfully explore this new area of research using staining techniques and other innovative methods. Most important was his work on the pancreas. He described the cell islands of the pancreas in a paper published in 1869. This paper presented the first careful and detailed description of the histology of the pancreas. The islets of Langerhans were first given that name in 1893 by the French histologist G. E. Laguesse. Langerhans also studied the anatomy of the skin and its innervation.

Image 15: Search for *Islets of Langerhans* in *MedlinePlus* dictionary

Obviously, the equivalent we found in the *Emedicine* dictionary was verified in the *MedlinePlus* dictionary as we can see on the image above.

Finally, since we have only explained how to obtain term equivalents, we are also going to explain how to obtain equivalent phraseology. In order to do that we are going to use our own monolingual corpus together with translation engines in order to see how they work and if they can be reliable. ‘Someterse a un trasplante de células pancreáticas’ is our first search. We entered this phrase within the *Imtranslator* engine and it provided the following translations:

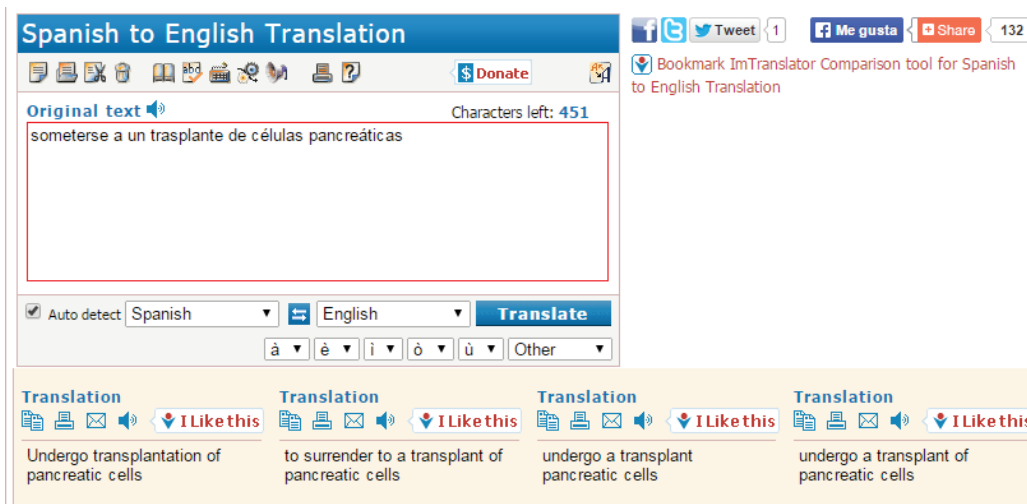


Image 16: Translation for the phrase *someterse a un trasplante de células pancreáticas*. As we see in the image above, *Imtranslator* provides us with four translations. Since all of them are different from each other we are going to see the differences and similarities among them. In three of the four translations the verb ‘someterse’ has been translated as ‘undergo’ while the remaining translation is ‘surrender’. In the case of the noun phrase we see:

1. That the first result differs mainly from the other ones in the translation of ‘trasplante’. It is translated as ‘transplantation’ in the first result and as ‘transplant’ in the respective three remaining results.
2. The translation of ‘células pancreáticas’ is not a problem as in all four translations it has been translated as ‘pancreatic cells’.

- The third translation, due to it being grammatically incorrect (missing ‘of’ → transplant of...), is only taken into account in order to know how the verb has been translated.

If we consider the statistics, ‘undergo’, ‘transplant’ and ‘pancreatic cells’ should be the equivalents in English. So, we may infer that the correct equivalent of the phrase might be: ‘undergo a transplant of pancreatic cells’ or ‘undergo a pancreatic cell transplant’. However, we are going to follow the same process as with terminology; we are going to check the equivalent. It is very difficult that a translation tool provides us with the translation of a whole phrase. This is the reason why we have combined the results obtained from our corpus by using *Wordfast Pro* and the results obtained from the Internet.

Image 17 shows how ‘someterse a’ has been translated in our corpus as ‘undergo’. Images 18 and 19 show that ‘Transplant of pancreatic cells’ has more results than ‘pancreatic cell transplant’.

ID	Source	Target
1	El médico o el personal de enfermería deben hablar sobre el estrés emocional de someterse a un trasplante de médula ósea.	Your doctor or nurse will likely discuss the emotional stress of having a bone marrow transplant.
2	Muchos pacientes de este tipo de cáncer recaen tras someterse a la quimioterapia y necesitan un trasplante autólogo de células madre.	Many patients with this type of cancer relapse after undergoing chemotherapy and require an autologous stem cell transplant.

Image 17: Search for *someterse a* in WordFast Pro

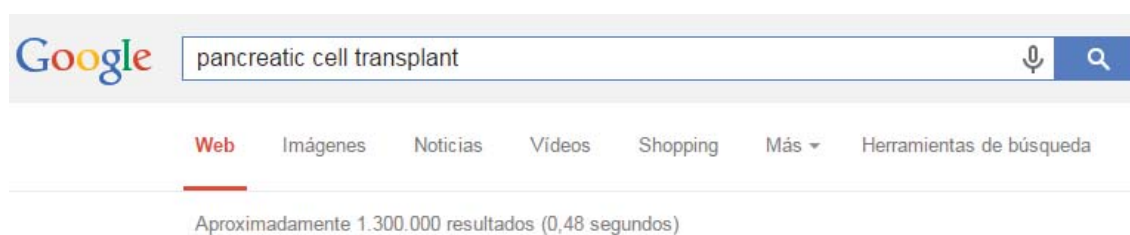


Image 18: Google search for *pancreatic cell transplant*

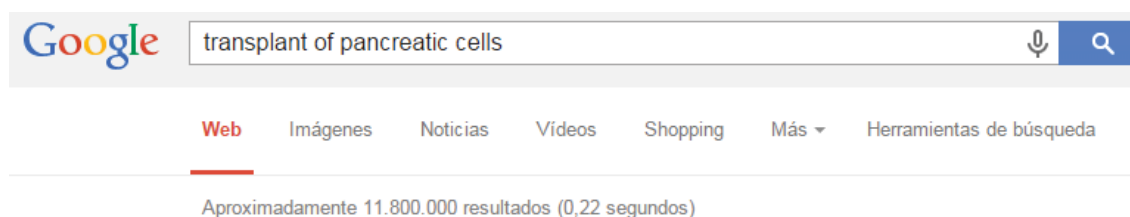


Image 19: Google search for *transplant of pancreatic cells*

So, by comparing all these results we decided to translate the phrase as ‘undergo a transplant of pancreatic cells’.

To conclude the explanation of how to solve translation problems, we are going to translate another phrase ‘desarrollar diabetes’. As in all the previous examples, we are going to use our corpus. On this occasion since we are dealing with a less complex phrase, we obtained the equivalent easier and earlier. We guessed that ‘develop diabetes’ could be the correct translation and we actually validated this guess by using the corpus.

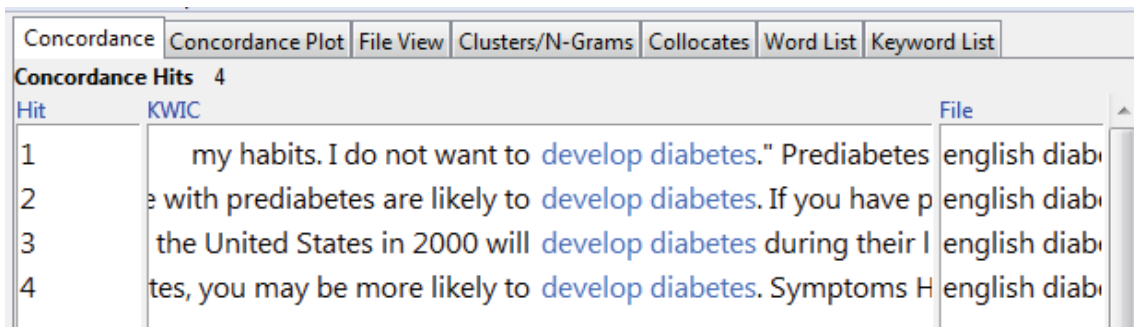


Image 20: Concordance for *develop diabetes*

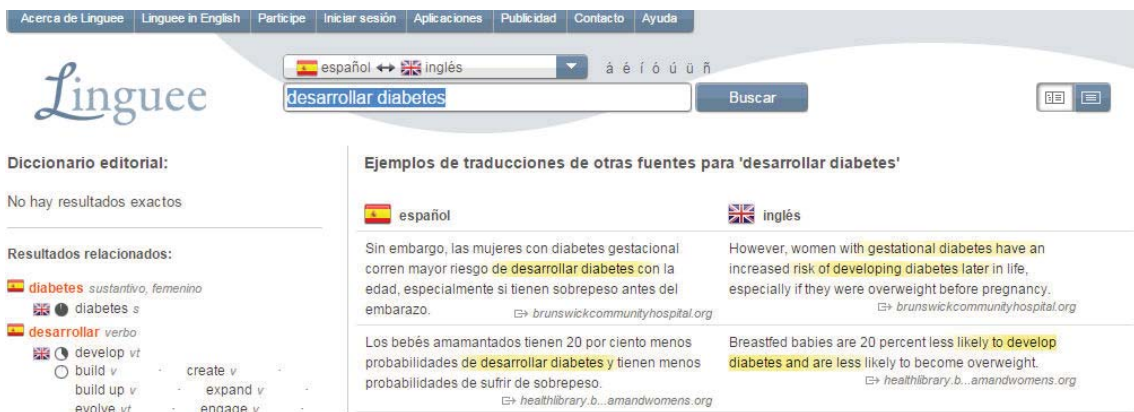


Image 21: Search for *desarrollar diabetes*

As we have seen, translating a phraseological unit is much more complex than translating a term. Translator-aided tools will probably always provide you with the corresponding equivalent of the term you are looking for. However, it is very difficult to

find a whole sentence already translated. These are the reasons why so many steps have to be followed to translate a phraseological unit correctly.

Several terminological and phraseological problems were found when we performed the first reading of our text. As we have explained how to solve the problems with a few terms, we are going to make a list with the remaining terms and their equivalents:

- ‘Diabetes tipo 1’ → ‘type 1 diabetes’
- ‘Magnetocápsulas’ → ‘magnetocapsules’
- ‘Diabéticos’ → ‘diabetics’
- ‘Recibir’ → ‘receive’
- ‘Agrupación de células’ → ‘grouping of cells’
- ‘Niveles de insulina’ → ‘insulin levels’
- ‘Sistema inmune del paciente’ → ‘patient’s immune system’
- ‘Vigilar’ → ‘monitor’
- ‘Segregar’ → ‘secrete’
- ‘Feridex’ → ‘Feridex’
- ‘Resonancia magnética’ → ‘magnetic resonance’
- ‘Modificado genéticamente’ → ‘genetically modified’
- ‘Niveles de glucosa en sangre’ → ‘blood glucose levels’
- ‘Animales trasplantados’ → ‘transplanted animals’
- ‘Hiperglucémico’ → ‘hyperglycemic’
- ‘Ramificaciones sanguíneas’ → ‘blood vessel branches’
- ‘Técnica de imagen por resonancia magnética’ → ‘technique of magnetic resonance imaging’
- ‘Células de islote pancreático’ → ‘pancreatic islet cells or islet cells of the pancreas’
- ‘FDA’ → ‘FDA’
- ‘Agencia estadounidense del medicamento’ → ‘The Food and Drug Administration’
- ‘Práctica clínica’ → ‘clinical practice’

- ‘Visualizar’ → ‘visualize’

Once all the terminological and phraseological problems have been solved we can continue with the translation. The next chapter will contain the text already translated with all the equivalents of the terminological and phraseological problems highlighted in yellow.

8. My Translation Proposal

TESTED ON ANIMALS

A new hope for type 1 diabetes

Some magnetocapsules help to 'monitor' pancreatic cell transplants

CRISTINA G. LUCIO

MADRID. - Say goodbye to daily insulin injections. The dream of millions of diabetics could be a little closer thanks to a new technique developed by U.S. scientists. It has only been tested on animals, but it allows the transplant of pancreatic cells in a body that does not produce insulin in a more effective way. Nowadays, some people with type 1 diabetes undergo a transplant of pancreatic cells. They concretely receive a grouping of cells – named islets of Langerhans – that allow regulating the levels of insulin in their organisms. Unfortunately, in many occasions, the transplant is rejected before long without knowing the cause. Experts say that part of the problem is that once they are introduced into the body, it is impossible to track these cells or to know if the patient's immune system has destroyed them.

Protect and monitor the cells

In order to accurately overcome this obstacle, researchers from the Johns Hopkins centre in the United States designed a new technique that consists of surrounding the clusters of cells with a material that prevents the organism from recognizing them as a foreign body and also from attacking them, However, this technique still allows them to continue secreting insulin. A magnetic substance that contains iron and is visible in a magnetic resonance, called Feridex, was added to this material. Scientists tested the technique in diabetic mice and pigs, and in both cases, the magnetocapsules, as its creators call them, were effective because it was possible to control their evolution from the outside and they were not eliminated by the immune system, which in a diabetic's organism attacks the insulin-producing cells. Aravind Arepally, assistant professor of radiology and surgery in the American centre and one the authors of the research explained to elmundo.es that the importance of their findings is that now they can use conventional imaging techniques to monitor the evolution of transplanted cells, like the

pancreatic islets. The results of his work are published in the latest issue of the journal 'Nature Medicine'.

Experiments in mice and pigs

First of all, the researchers measured the effect of the 'magnetocapsules', each of which contained between 500 and 1,000 insulin-producing cells, in a group of genetically modified mice, so that they developed diabetes. Half of them received a transplant and the rest did not undergo any procedure. One week after the operation the blood glucose levels of the transplanted animals returned to normality and remained constant up to eight weeks later. In contrast, mice that did not receive insulin-producing cells either died or remained hyperglycemic. In a second experiment, the scientists wanted to test the effectiveness of the 'magnetocapsules' in a more similar organism to that of humans, for which 40,000 units were introduced in pig livers. The reason why the pancreas was not chosen is that the liver contains many blood vessel branches and it can rapidly lead the insulin to the rest of the body. Through a technique of magnetic resonance imaging, three weeks after the intervention the researchers found that the capsules were still intact and functioning.

"This finding is directly applicable to the current improvements of the transplant of pancreatic islet cells, particularly because our magnetocapsules are composed of clinically applicable materials", assert researchers in their study. "We hope that soon it could be used on humans. This is a promising fact, because all our methodology is done with products and imaging techniques approved by the FDA [The Food and Drug Administration]. Our aim is to move this to the clinical practice as soon as possible", mentions Arepally. "We will do further studies in pigs in order to check if the islets can continue producing insulin for longer periods of time. Also, we will refine the imaging techniques in order to visualize the encapsulated cells to a greater degree. And finally, we will have to determine the final destination of our magnetocapsules, that is to say, to know if they remain intact or if they break with the passing of time", concludes this expert.

9. Conclusion

The process of translation does not only consist of translating word for word or translating a source text literally. As we have seen, there are many steps involved within translation. In my opinion, the two most important ones are the textual analysis and the search for terminology and phraseology.

In the textual analysis, all the categories (genre, register etc) analysed are related to one another. A. Trosborg (1997: 12) notes that “there is not only a relation among genre and text type but also among these variables and register”. This affirmation has been sustained by numerous authors such as Lemke (1985: 278-279) and Beaugrand e& Dressler (1981). Isabel García (2000: 243) also clarifies that “although these categories are interrelated, we have to know how to distinguish them since a genre can be expressed through different text types and the same text type can be representative of more than one genre”.

In summary, the choice of a particular register will determine the genre, audience and tone. At the same time, the choice of a certain genre will also determine the same parameters, that is to say register, audience and tone.

Within terminology we find that it develops at the same speed as the fields in which it is used because of the constant growth of technology and science. “Due to the unceasing apparition of new forms and new concepts, terminologists are encountered in the trouble of solving the mess of the relations among concepts and forms” (Rondeau, 1983). Rondeau (1983) also confirmed this role of terminologists: “terminology is a tool that must serve to undo the ambiguity in the scientific and technical communication.”

As a matter of fact, terminology and specialized translation are interrelated disciplines. Specialized translation needs the use of terminology. Translators should pay more attention to terminology when translating specialized texts as a high-quality translation requires the use of adequate terminology. Also the terms chosen by the translator must be appropriate to the level of specialization of the texts as we have explained in the section 3.1.2.1.

However, we believe that the process is not reciprocal since terminology does not need

specialized translation. The reason is that translation is the final step while terminology is the first one. Terminology aims to represent specialized language and it conveys specialized meanings in specialized texts since, as we have explained in chapter 4, a term is the linguistic representation of a concept.

Also, searching for the equivalents of terms in the target language never deals with making a literal translation of these terms. The interpretation of each one of the terms (definition and usage in each context) is necessary to look for the equivalents. All of this does not mean that the translator is professional or an expert in the concerned area, but he/she has to assure that the knowledge and information of the original text has been transferred to the target text in a manner that the target reader is able to understand it.

Several tools can be used in order to ease and improve the task of translation: from corpus exploitation tools to lexicographic tools, such as dictionaries and glossaries. This way, corpus linguistics has been considered as one of the most helpful disciplines for translation.

The usage of a corpus (monolingual, bilingual, multilingual, specialized, general, etc.) can provide the translator with a huge amount of information about the language he/she wants to study. Depending on the kind of corpus compiled, the information obtained can range from equivalent terminology and phraseology in the target language to ready-made translations.

To sum up, the process of translation is a complex process in which the translator has to take many aspects into account and also has to have a certain knowledge of the field in which he/she is working.

In my opinion, the most important aspect is that the decisions made by the translator, such as the choice of terms, phrases and so forth, will determine the quality and accuracy of the translation.

10. References

- Atkins et al. (1992) *Corpus design criteria*. Oxford: Oxford University Press.
- Bassols, M., & Torrent, A. (2003). *Modelos textuales :Teoría y práctica* (2a ed.). Barcelona: Octaedro.
- Bowker, L., & Pearson, J. (2002). *Working with specialized language :A Practical Guide to Using Corpora*. London: Routledge.
- Cabré, M. T., & Gomez de Enterría Sánchez, J. (2006). *La Enseñanza de los Lenguajes de Especialidad: La Simulación Global*. Madrid: Gredos.
- Cabré, M. T., & Sager, J. C. (1993). *La terminología: Teoría, metodología, aplicaciones*. Barcelona: Antártida.
- Congost, N. (1994). “Visión pragmática de la traducción científico- técnica aplicada a language médico”. Ed. Universidad de Alicante. *Problemas de la traducción técnica: Los textos médicos en inglés* (pp. 31-59) Alicante: Universidad de Alicante, pp. 31-59.
- Corpas, G. (2001). “Compilación de un corpus ad hoc para la enseñanza de la traducción inversa especializada.” Málaga, Universidad de Málaga, pp. 155-184.
- De Santiago, P. “Introducción a la Traducción Automática.” Class notes 2013-2014. Universidad de Valladolid.
- Firth, J. (1957). *Papers in Linguistics*. Oxford: Oxford University Press.
- Garcia Izquierdo, I. (2000). *Análisis textual aplicado a la traducción*. Valencia: Tirant lo Blanch.
- Halliday, M.A.K. & Hasan, R. (1985) *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Victoria: Deakin University Press.
- Hill, J. (2002). *Revising priorities: from grammatical failure to collocational success*. In Michael Lewis (Ed.) *Teaching collocation: Further developments in the lexical approach*. Hove, England: Language Teaching Publications

McEnery, T., & Wilson, A. (2001). *Corpus linguistics :An introduction* (2nd ed.). Edinburgh: Edinburgh University Press.

Méndez Cendón, B. “Traducción de Lenguajes Especializados II (Inglés/Español: Ciencia y Técnica).” Class notes 2012-2013. Universidad de Valladolid

Mendiluce, G. (2004). Estudio comparado inglés/español del discurso biomédico escrito: La secuenciación informativa, la matización asertiva y la conexión argumentativa en la introducción y la discusión de artículos biomédicos escritos por autores nativos y no-nativos. Ph D Thesis available at <http://bib.cervantesvirtual.com/FichaAutor.html?Ref=7789> (accessed date 15/12/2014) .

Muñoz, C. (2011). “Análisis contrastivo y traductológico de textos médicos (inglés-español). El género caso clínico”. Barcelona: Universidad Autónoma de Barcelona. Ph D thesis available at <http://www.tdx.cat/handle/10803/48658> (accessed date 12/11/2014)

Olohan, M. (2004). *Introducing corpora in translation studies*. London: Routledge.

Pearson, J. (1998). *Terms in context*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Rey, A., Sager, J. C., & de Bessé, B. (1995). *Essays on terminology*. Amsterdam: John Benjamins.

Sager, J. C. (1990). *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: OUP.

Sinclair, J. (2005). “Corpus and text: Basic principles.” In M. Wynne (Ed.), *Developing Linguistic Corpora: A Guide to Good Practice*. Oxford: Oxford Books. Págs. 1-16.

Sinclair, J. (2003) *Reading Concordances*. London: Pearson Education Limited.

Somers, H., & Sager, J. C. (1996). Terminology, LSP, and translation: Studies in Language Engineering in Honour of Juan C. Sager. Amsterdam: John Benjamins.