JoSTrans
The Journal of Specialised Translation

www.jostrans.org · ISSN: 1740-367X

Jiménez-Crespo, M. (2017). Combining corpus and experimental studies: insights into the reception of translated medical texts. *The Journal of Specialised Translation, 28*, 2-22.
https://doi.org/10.26034/cm.jostrans.2017.231

Combining corpus and experimental studies: insights into the reception of translated medical texts[1]

Miguel A. Jiménez-Crespo, Rutgers University

ABSTRACT

This experimental study combines corpus-based translation studies (CBTS) and Cognitive TS (CTS) in order to study the reception of translated medical websites from English into Spanish in the USA. It builds upon a previous corpus study (Jiménez-Crespo and Tercedor forthcoming) based on the 40-million word TWCoMS comparable corpus (Jiménez-Crespo 2014) that identified significant differences at the lexical and register levels between translated and non-translated medical websites in the US addressed at laymen. This study concluded that Spanish translations contained significantly less Latin Greek terms (LG) and doublets, a LG term accompanied by a reformulation or explicitation, than similar non-translated medical websites produced in Mexico and Spain. This resulted in what is known as known as "register mismatch" (Pilegaard 1997), when the register of the source and target texts are inadequately different.

The question that the present experimental study addresses is whether these translated texts that seem to display lower levels of specialisation and higher percentage of explicitations of medical terms are in fact more understandable and preferred by the target users of the translations in the United States, Spanish speakers. The results of the experimental study demonstrates that despite the preliminary results of the corpus study indicating that translations might be more understandable, subjects preferred by a large margin segments with reformulations of LG terms found in non-translated texts. This brings up the necessity of combining corpus and cognitive empirical studies, both aimed at production and reception, in order to expand the reaches of both subdisciplines following the recent programmatic agenda of leading scholars in CTS (i.e. Alves and Vale 2011; Muñoz 2014; Halverson 2016) already started in previous projects by the author of the paper (Jiménez-Crespo 2013b, 2016).

KEYWORDS

Corpus-based TS, cognitive TS, medical translation, reception studies, web corpora.

1. Introduction

After over two decades of the emergence and consolidation of Cognitive Translatology (Muñoz 2014), empirical research into translation processes has become commonplace in Translation Studies. Among one of the new possible avenues for research, scholars have called for combining the process approach of Cognitive Translatology/CT (Muñoz 2010, 2014) with the product approach of Corpus-Based Translation Studies/CBTS (Laviosa 2014) in order to delve into specific issues that large corpora can provide to design and contrast cognitive processing studies (Halverson 2010, 2016; Jiménez-Crespo 2016; Muñoz 2014; Alves and Vale 2011). Most of the studies and calls for research, nevertheless, take a retroactive approach process, attempting to identify traces of cognitive processes in the products. They follow the steps to attempt to identify the causality or explanation of why translated products display specific features meant to be the result of cognitive processing. Thus, CT

scholars attempt to introduce CBTS to gain insights into cognitive processing of translations, their main goal of this subdiscipline.

The approach taken in this paper is different since it departs from CBTS. It attempts to identify the effects and reception of these distinct features found in translated texts, rather than trace back how they came to appear during the translation process (Halverson 2016). In order to do so, this study continues the triangulation approach taken by Jiménez-Crespo (2016) to combine corpus and cognitive studies. The empirical study addresses the question of whether medical translated texts from English into Spanish, having shown to display lower levels of terminological specialisation and higher percentage of explicitations of Latin-Greek medical terms (Jiménez-Crespo and Tercedor forthcoming), are in fact more understandable and preferred by the target users of the translations in the United States, Spanish speakers. This approach argues that new insights can be gained by starting with studies into specific features in corpus studies that later are used to generate hypothesis and develop testing instruments in cognitive studies. In this approach corpora can also be used to triangulate the results after the empirical study has been conducted.

2. The significance of medical texts addressed at laymen for the combination of corpus and cognitive studies

The translation of general medical information represents a case of expert to non-expert communication where register and lexical usage has to be adjusted to the knowledge and expectations of the target end users (Montalt and Gonzalez-Davis 2007). Nevertheless, research has shown that in some language combinations lexical and register differences between translated and non-translated often make the former more difficult to understand for the average laymen (Askehave and Zethsen 2003; Raynor 2007; Jensen and Zethsen 2012; Zethsen 2013). The reasons behind the differences in register in original and translated texts are that Latin was not similarly incorporated to the same extent in all European Languages (Zethsen 2004: 132). While Spanish and French medical terminology are eminently Latin and Greek in origin, other Northern European languages possess a double-layer medical terminology in which many scientific words possess popular or lower register counterparts (Pilegaard 1997). Thus, "seemingly identical words may indeed be false friends in an interlinguistic context" and they might also present issue related to the "connotative differences, e.g. at the level of formality" (Zethsen 2004: 131-132). For example, English has doublets such as 'clotting' and 'coagulation,' or 'scar' and 'cicatrisation,' with different implications for register and lexical variation in their use, while in Spanish the terms based on Greek and Latin are the only ones that exist, 'coagulación' and 'cicatriz.' These two words are used both high and low registers indistinctively. Therefore, what "in

Latin languages might sound too low a register is perfectly acceptable as scientific terminology in English" (Montalt and Gonzalez-Davis 2007: 242). Thus in English, one of the most common ways in which synonymy occurs in medical and scientific domains is the existence of the technical term with its low register equivalents, such as 'cephalalgia' and 'headache.' It is often understood that these cases of synonymy are "a source of translation problems because languages are not symmetrical in their use: for example, what in Spanish is considered to be low register may be perfectly acceptable in English in the same text genre" (Montalt 2011: 80).

These issues are often referred to in the publications that focus on the English to Spanish translation of medical texts. In the case of medical patient guides Campos Andres (2013: 53) indicates from a prescriptive perspective that when translating this genre into Spanish it is not always necessary to use both terms that might appear in source English texts, the Latin and Greek (LG) one and the popular one to refer to the same concept: the Spanish patient or user has a higher chance of being familiar and acquainted with terms of LG origin since they are more common in general language. The same can be said for cases of reformulation or 'determinologisation.' It is one of the most frequent strategies at the lexical level to lower the register and adapt textual genres to non-expert readership. This appears in medical texts both in intralingual translation such as the case of research articles summarised for laymen in the Annals of Internal Medicine (Muñoz-Miquel 2012: 200-2002), or also in translated texts for general audiences (Tercedor and Lopez Rodriguez 2012). This process involves using general language to communicate the meaning of a specialised term (Meyer and Mackintosh 2000), helping to close the gap between specialised knowledge and lay audiences. Montalt and Shuttleworth (2012: 16) refer to determinologisation as

> a process of recontextualisation and reformulation of specialised terms aiming at making the concepts they designate relevant to and understandable by a lay audience. This process is motivated by specific cognitive, social and communicative needs, and takes place as part of a broader process of recontextualisation and reformulation of discourse. (...)

This process involves a large number of potential strategies that are covered under this hypernym such as, explanation, definition, reformulation, exemplification, illustration, analogy, comparison and substitution by a more popular term (Campos Andres 2013; Montalt-Resurrecció and González Davies 2007: 252-253). According to Montalt-Resurrecció and González Davies (ibid), this process can involve a number of strategies that include scientific terms, such as the LG terms under scrutiny in this study, to be:

1. Retained followed explanation, such as "poliuria, aumento de la cantidad de orina" [Polyuria, increase in the volume of urine]

2. Retained in parenthesis after the explanations, such as "aumento en la cantidad de orina (poliuria) [increase in the volume of urine (poliuria)]
3. Retained after a popular term, such as "mal aliento o halitosis" [bad breath or halitosis]
4. Avoided and replaced by explanations or popular terms "patients can experience an increase in the volume of urine".

All these mechanisms can help increase the readability and efficiency of translated medical texts for laymen, but also relate to one of the main general tendencies of translation, "explicitation" (Baker 1993). In previous studies that will be described in the next section (Jiménez-Crespo 2014; Jiménez-Crespo and Tercedor forthcoming), it was argued that if reformulation or determinologisation represents a natural mechanism in intralingual translation in medical genres, the translation process can potentially increase or decrease the frequency and nature of the explicitation strategies present in translated texts addressed at laymen. This mechanism in texts addressed at laymen in both intralinguistic or intergeneric translation (Ezpeleta 2012) might result in a tendency of translated texts to exhibit higher levels of explicitation than non-translated texts.

3. Preliminary corpus-based study (Jiménez-Crespo 2014; Jiménez-Crespo and Tercedor forthcoming)

This paper follows the study by Jiménez-Crespo and Tercedor (forthcoming) relating to terminological variation, lexical features and explicitation in medical texts. The study used the comparable Translational Web Corpus of Medical Spanish (TWCoMS) that includes medical websites translated into Spanish in the US, alongside similar medical websites originally produced in Mexico and Spain (Jiménez-Crespo 2014). It was motivated by previous work in the English to Danish combination (Askehave and Zethsen 2003; Zethsen 2005; Raynor 2007; Jensen and Zethsen 2012; Zethsen 2013) that identified that translated medical texts were on average less lay-friendly and usable than non-translated texts in part due to a direct translation of Latin and Greek (LG) terms into Danish (Zethsen 2004). It is claimed that in languages with doublets, that is, with a LG term and a lay term, such as 'coagulation' and 'clotting' or 'cicatrisation' and 'scaring,' the direct translation of the LG term into certain languages can increase the register level, thus rendering text harder to understand.

In this study by Jiménez-Crespo and Tercedor (forthcoming) in the opposite issue was at stake and in fact, the opposite effect was identified. Translation products showed in principle lower register and specialisation levels than original Spanish non-translated ones. These translated medical texts into

Spanish contained on average 3.2 more times LG terms than comparable non-translated texts. Similarly, another strategy to make texts easier to understand for laymen, the use of determinologisation strategies (Montalt and Shuttleworth 2012: 16), was also consistently higher in translated texts than in non-translated ones.

In terms of percentages, the results of the study showed that on average, LG terms in translated text were accompanied by one of the above mentioned determinologisation strategies 40.59% of the times, while in non-translated texts the frequency went down to 21.23% (Figure 1). The overall results of the study therefore suggest that translated texts into Spanish from English that display a much lower frequency of LG terms (3.2 times less), combined with a much higher rate of determinologisation-explicitation strategies, might in fact logically make them easier to understand for the average Spanish-speaking laymen in the US. This is precisely the starting point for this experimental study. In principle, it could be assumed that the literal translation (Chesterman 2011; Halverson 2015) and the direct transfer of English-language usage of LG terms and their corresponding explicitations-determinologisation would result in texts that might be easier to understand since the register would be lower for target texts in Spanish if compared to naturally produced ones. Also, translation products might show traces of translation-inherent explicitation (Klaudy 1988), that is, by virtue of being translations text might contain even more explicitations due to the nature of the mediating translation process.

If this corpus-based study suggests that in principle, texts with these features (both LG terms and their corresponding determinologisation-explicitation strategies) might be easier to understand, the aim of this study is to address the question of whether an experimental study would confirm that subjects prefer in fact the translated texts to the non-translated texts. This combination of corpus-based studies with experimental studies thus attempts to bridge the gap between this two subdisciplines or paradigms within TS, tending bridges beyond the emerging question of whether corpora can shed in fact light into cognitive process or not.

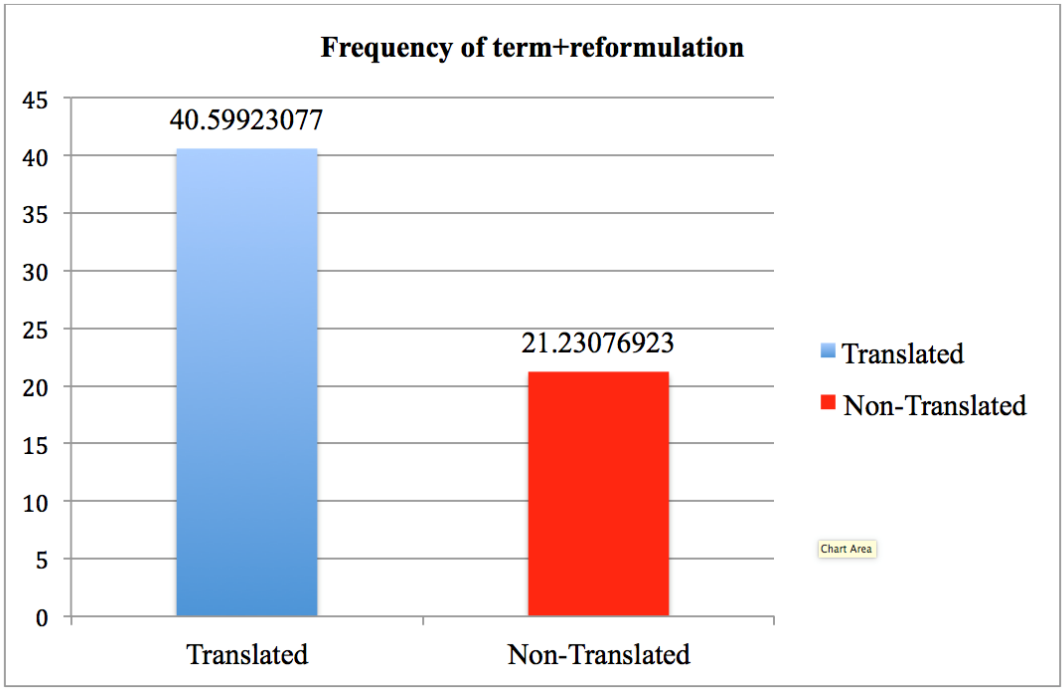

Figure 1. Frequency of Latin and Greek medical terms accompanied by a reformulation-explicitation in translated medical websites (Jiménez-Crespo and Tercedor forthcoming).

The result of the study therefore interrelate the role of lexical variation in medical terminology and the notion of familiarity and specialisation level (Alarcón, Lopez-Rodriguez and Tercedor 2016), with the translation of medical texts and the adaptation to the target audience (Askehave and Nielsen 2012).

4. The intersection of corpus-based translation studies and cognitive translatology

The integration of process and product oriented studies encouraged by some scholars in the cognitive field, such as Muñoz (2014) or Halverson (2010, 2016) is becoming a reality since, as Halverson (2015: 311) indicates this "can hold the promise for a general theory of translation that is general in scope." Over the years, a number of publications have delved into the possibility, or not, that corpus data can provide insights into the cognitive processing of translation. While some propose that with carefully built theoretical constructs offline corpus-data can provide insights into cognitive processing (Olohan 2004; Alves and Vale 2011; Halverson 2015), others argue against that claim, asserting that any hypothetical connection to cognitive processes has to be verified using experimental studies (Halverson 2003, 2010). In this sense, while some studies have attempted to shed light into the specifics of cognition in corpus studies (i.e. Alves and Vale 2011), few studies have actually

combined from corpus and experimental studies such as Jiménez-Crespo (2013b, 2016) attempting to follow Halverson when she indicates that the "objective of empirical Translation Studies is to describe the choices that translators make, and ultimately to understand some of the causes of these choices" (2015: 320). Also, Halverson claims that the difference between process and product-oriented approaches up until recently has been that while product based scholars have primarily been asking what is done and why, process scholars have ben asking the how (2015: 320-321). The present study, nevertheless, extends the reach of this connection to delve not into which specifics cognitive processing details resulted in the observed differences in the corpus study, but rather take a reception approach to observe whether the preliminary conclusions of the corpus study, i.e. that translated texts due to interference and literal translation form the source English texts are in principle easier to understand, do hold true in the light of end users. The hypothesis of the study therefore departs from the conclusions of the Jiménez-Crespo (2014) and Jiménez-Crespo and Tercedor (forthcoming) corpus studies.

5. Hypothesis

Following the previous literature review and in the light of the previous comparable corpus-based study on the lexical features of translated and non-translated medical websites, the hypothesis for this experimental study is that:

> Medical texts translated from English into Spanish that display lower register and lexical specialisation levels than their non-translated counterparts will be easier to understand and preferred by end users, Spanish speakers living in the USA.

6. Methodology

In order to test the hypothesis the triangulation method presented by Jiménez-Crespo (2016) was adopted to develop the study. This triangulation model is intended to combine corpus and experimental research and attempts to interrelate corpus and cognitive research in the discipline. The overall approach entail as a first step to conduct corpus-based studies using corpora and following CBTS principles. Subsequently, the results serve as the foundation to develop experimental studies, both in production and in reception. The actual results and data from the corpus study are then used to develop the testing instruments. Once the experimental studies are conducted, the results can later be triangulated with the overall corpus if necessary, thus providing a triangular model that can extend the reach of both CBTS and Cognitive TS.

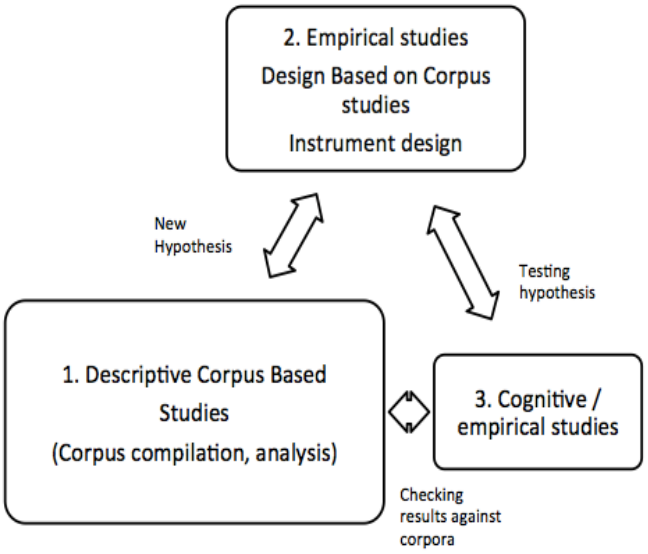

Figure 2. Triangulation methodology to combine corpus-based studies with cognitive/empirical studies (Jiménez-Crespo 2016: 266).

6.1. Corpus-based study

The corpus study used the Translational Web Corpus of Medical Spanish, TWCoMS (Jiménez-Crespo 2014), a corpus project was conceived as a tool to study medical translation in the United States. This corpus contains approximately 40 million words from medical information websites addressed at general audiences in the EEUU and a comparable section of websites for Mexico and Spain. The corpus includes exclusively web genres, those that emerged exclusively for the web (Santini 2005: 2; Jiménez-Crespo 2013a: 66-100). The TWCoMS comprises two subcorpora in Spanish. One of them includes medical websites or portals in the United States translated into Spanish (32.330.52 tokens) and another subcorpus of similar websites produced originally in Spain or Mexico (8.701.867 tokens). It is important to highlight that the corpus contains full websites, that is, the entire structure of the website is downloaded and analyzed. The compilation process was carried out using the Httrack.

The TWCoMS comprises two interrelated subcorpora in Spanish. One of them includes medical websites or portals in the United States translated into Spanish and another subcorpus of similar websites produced originally in Spain, Mexico or for a general Latin American audience.

1. The translational subcorpus contains 32,330,052 tokens. It comprises four distinct subsections of texts that could be considered cases of

"intrasocial translations", that is translations addressed at members of the same society speaking another language, such as the case of Spanish speakers in the US. The four sections are (1) US government websites (i.e. Center for Disease Control, Womenhealth); (2) websites with medical information from the different Departments of Health at the state level; (3) general medical websites from national organisations (i.e. MedlinePlus); and (4) medical history forms available on the Internet, a category that can be used for contrastive purposes (Gonzalez-Darriba 2013).

2. The non-translational subcorpus contains medical websites originally produced in Spanish such as Mapfresalud (Spain), Universomedico (Mexico) and Geosalud (Latin America). The comparable section contains 8,701,867 tokens.

The main results of the corpus study were described previously in section 3. In order to test the hypothesis, the experimental study focuses on reformulations or explicitations previously identified in both sections of the corpus that have been listed in order of frequency of use. These reformulations are selected since they, in fact, have the objective of making texts 'easier to understand' taking expert knowledge closer to the average layman. The objective is to identify whether subjects perceive as more understandable and usable translated reformulations rather than their non-translated counterparts (see Table 1). In doing so, it will help shed light into whether, even when translated texts in the corpus study show a lower level of specialisation and register, and thus in principle could be perceived as easier to understand, they are in fact preferred by potential end users, Spanish speakers in the US, over those found in original and non-translated texts. As an example, the different reformulation or determinologisation renderings for the term 'dyspnea' in Spanish original texts were the following in frequency of use, including their approximate literal translation into English in order to grasp the potential variation:

1. Dificultad para respirar [Distress in breathing]
2. Dificultad respiratoria [Respiratory distress]
3. Falta de aire [Lack of breath]
4. Dificultad al respirar [Distress when breathing]
5. Falta de aire [Shortness of breath]
6. Dificultad en la respiración [Respiratory distress]
7. Sensación de falta de aire [Feeling of shortness of breath]

The first three options were the most frequently used in the original Spanish corpus, clearly indicating higher levels of lexicalisation. The interference from

the source English texts is also clearly shown in the translational subcorpus, where the most frequently used reformulations are:

1. Respiración entrecortada [Shortness of breath]
2. Dificultad para respirar [Respiratory distress]

The first reformulation represents a clear case of literal rendering of a source text lexicalised lexical unit, 'shortness of breath,' that does not appear in the non-translated text.

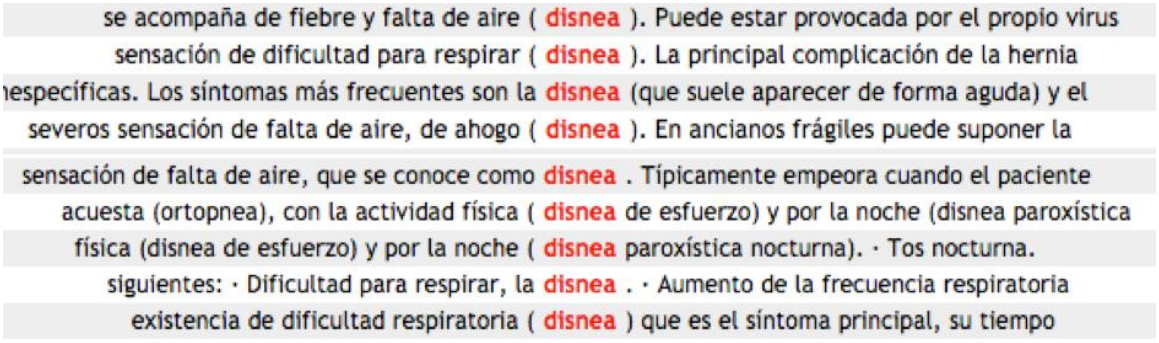

Figure 3. Snapshot of concordance lines in the original corpus for the concept 'dyspnea' with its corresponding reformulations-explicitations using Sketchengine.

Similarly, Table 1 shows the reformulations found in both corpora for the concept 'hypoglycemia' in their order of frequency.

	Reformulations for term: Hypoglycemia	
	Translational corpus	Non-Translational corpus
Reformulations	nivel bajo de azúcar en la sangre azúcar bajo en la sangre* bajo azúcar en la sangre* azúcar baja en la sangre concentraciones bajas de azúcar en la sangre* glucemia baja * concentración anormalmente baja de azúcar en la sangre* niveles bajos de azúcar* los bajos niveles de azúcar en la sangre* (baja del azúcar sanguíneo)* bajo nivel de azúcar en la sangre (baja del azúcar en la sangre)* baja del nivel de azúcar en la sangre* la disminución de los niveles de azúcar en la sangre* (disminución del azúcar en la sangre)* niveles bajos de azúcar* disminución abrupta del nivel de azúcar en la sangre*	nivel bajo de glucosa en sangre azúcar en la sangre demasiado baja disminución de los niveles de glucosa en la sangre baja en el azúcar bajada de glucosa bajo nivel de azúcar en la sangre un nivel bajo de azúcar en la sangre cuando la concentración de glucosa sanguínea es inferior a 50 mg/dL cuando baja la glucosa en la sangre cuando los niveles de azúcar en la sangre están demasiado bajos descenso de los niveles sanguíneos de azúcar descenso del nivel de azúcar en sangre disminución excesiva del nivel de glucosa en sangre valores de azúcar muy bajos valores de glucosa en la sangre muy bajos valores muy bajos de azúcar en la sangre bajada de azúcar en sangre bajada de los niveles de glucosa descenso excesivo de glucosa en sangre disminución de los niveles de glucemia en sangre los niveles de azúcar en sangre bajos azúcar baja en la sangre glucosa baja en sangre disminución de glucosa en sangre

Table 1. Contrastive analysis of reformulations of the term 'hypoglycaemia.' Starred reformulations do not appear in the non-translated texts.[2]

It is for example observed that over 80% of use of reformulations in the translational subcorpus do not appear in naturally produced texts. This implies that end users might not have been exposed to that reformulation in the past. This does not mean that they might not be understandable, quite the opposite. They can be adequately be understood by native speakers. This does bear implications for lexical variation and the role of interference in translation processes.

Obviously, not all corpora might be suited for a deep exploration of cognitive issues related to translation, mostly due to the lack of information about production, authoring, translation process, editing, etc. Halverson, for example, indicates that the "limitations of corpus data is a serious weakness with regards to cognitively oriented theoretical approaches, and must be remedied in coming generations of corpora if cognitive frameworks are to be properly tested (2015:316). Nevertheless, the approach taken in this paper is not necessarily geared towards a retroactive identification of production issues but rather the relationship between corpus-based approaches and empirical studies on reception. As previously mentioned, this is the opposite of the previous approaches by Alves and Vale (2011) or the programmatic proposal by Halverson (2016). The findings of the corpus based study are tested not in order to delve into potential causality relationships in production, but rather to test whether the suggested results and distinctive features observed in corpus studies do hold true in reception studies with users. In this sense, the empirical and cognitive part of the study would be closer to the reception studies in audiovisual text such as Kruger, Fox and Doherty (2016).

6.2. Testing Instruments of the empirical reception study

In order to test the hypotheses, two separate instruments were developed using the results of the corpus study. The first instrument presented a simple decision task between two choices of reformulations for a LG term, while the second decision presented a variable list of existing reformulations and subjects had to select two out of the entire group. The instructions for both of them indicated that they should choose the reformulation that subjects thought would be best to explain the medical concept to other Spanish speakers in the United States. This allowed to focus the study in reception issues and to avoid any distortion due to potential dialectal differences between Spanish speakers in the US. By extending the reception to other Spanish speakers in the US subjects needed to take into account not only their personal preferences but also the general "standard" or "international" Spanish variety in the US.

The first instrument contained a decision task that presented two possible choices to select from the most frequent reformulations for each LG term in each section of the corpus. If both sections of the corpus had the same most frequent reformulation, then the next most frequent reformulation was selected. All the reformulations only appeared on one subcorpus or the other, but not in both. For example, for the concept "dyspnea" subjects could select either "respiración entrecortada" (trans corpus) or "dificultad para respirar" (original corpus). This means that "respiración entrecortada" did not appear in the original corpus, while "dificultad para respirar" was not identified in the translational corpus. Subjects had therefore to select either a translational or

non-translational rendering. Subjects were presented with the LG term (disnea) and they had to select one of the two choices.

Disnea	Respiración entrecortada	□
	Dificultad para respirar	□
Dismenorrea	Periodos dolorosos	□
	Dolores menstruales	□

Figure 4. Example from the first instrument with two selection items, 'dyspnea' and 'dysmenorrhea.' Subjects were instructed to select only one reformulation.

The second testing condition included a selection between a large variable number of reformulations for each LG term. Some were present only in the original corpus and the others only in the translational corpus. If one refo'mulation appeared with high frequency in both subcorpora, it was included in the instrument and coded as 'both.' The number of options varied according to the number of possibilities identified in each subcorpus. For example, the range of reformulations per term varied from eight for 'hysterectomy' to only three in 'polydipsia.' In this last case, subjects were only presented with 'sed excesiva' [excessive thirst], 'sed intensa' [intense thirst] and 'exceso de sed' [excess in thirst]. The first one appeared in both subcorpora, while 'sed intensa' was the most frequent reformulation found in non-translated texts, and 'exceso de sed' was found only in the translational corpus.

Histerectomía	Extirpación del útero □ Extirpar el útero □ Cirugía para extirpar el útero □ Operación para extraer el útero □ Extirpación quirúrgica del útero □ Cirugía en la que se extrae el útero □ Cirugía para extraer la matriz □
Polidipsia	Sed excesiva □ Sed intensa □ Exceso de sed □

Figure 5. Example from instrument two with multiple-choice options. Examples for 'hysterectomy' and 'polydipsia.' Subjects were instructed to select two reformulations in each.

6.3. Subjects

Twenty-five subjects took part in the experimental study. The subjects were all Spanish speakers living the State of New Jersey. The average age was

25.22 and the average number of years living in the USA was 16.12. From the cohort, 44% of them received all education in the USA, meaning that they could be considered full heritage speakers of the language, while 56% came to live in the US in different point in their lives. This means that the study closely represents the two groups of Spanish bilinguals that live in the USA, heritage speakers and native speakers of Spanish that moved to the USA from different countries. Most subjects were female, 80% of them, while 20% of them were male. From those born abroad, seven countries were represented, Mexico, Spain, Ecuador, Peru, Argentina, Panama, Puerto Rico. All students were college students in the last year of their studies, graduate students or had received already their BA or graduate degree.

6.4. Testing conditions

The empirical study was approved by the Human Subject Protection Program at Rutgers University and received its corresponding IRB. Subjects completed a brief personal questionnaire with personal information, signed an approved consent form and then were provided with instructions for the first instrument. After the completion of the first instrument, subjects were provided with the multiple selection second instruments. The time for completion of the study was not recorded and subjects were reminded to take as much time as needed. The tests were carried out individually and subjects had no access to any reference materials of any kind.

7. Results

The results of the experimental study will be shown following the progression described in the methodology section. The first instrument requested the subjects to select the reformulation or explicitation that they would use most likely to explain the concept or LG term to another Spanish speaker. Subjects could select only one and therefore needed to select among the most frequent reformulations used both in the translated and in the non-translated subcorpus. The possible responses were coded according to whether subjects selected a reformulation from the translational or non-translational corpus. Figure 6 shows the results of the study:

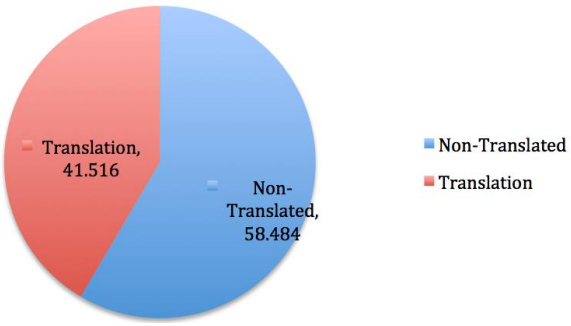

Figure 6. Percentage of reformulations preferred by subjects from the translational and non-translational corpus.

In general, subjects preferred the reformulations found in the original or non-translated corpus. In 58.48 percent of the cases, the responses selected were the most frequent reformulations found in the non-translational corpus, that is, reformulations that have been produced without a translation mediation process, and that do not appear in the translational subcorpus. Meanwhile, 41.51 percent of the times subjects selected the most frequent reformulations found in this latter corpus. It is of interest that the distribution closely matches the percentage of participants that are heritage speakers and received all their education in the USA (44% heritage speakers born in the USA, 56% speakers of Spanish born abroad), pointing at a higher degree of acceptance of literal translations for heritage speakers.

The previous task only allowed participants to select one of the two reformulations presented. After the completion of the experimental task several subjects commented on the difficulty in selecting just one reformulation since both seemed correct to them, and they were reassured of the instructions provided to select the reformulation that they though they would better use to explain this LG term to another Spanish speaker in the US. The next step in order to delve into the different degrees of acceptability involved allowing participants to select two reformulations out of a variable pool of options extracted from both corpora.

Figure 7 shows the results of this decision task. The data were coded according to the three possible options, whether the reformulation appears only in one or the other subcorpus, or whether the reformulation appears in both of them. Subject preferred by a small margin those found in non-translated texts [Non-translated=38.66; Both= 35.73; Translated= 25.6%]. The second choice would be reformulations found in both corpora. These types of reformulations are of interest since given that they are the most frequent in both corpora, they often represent both the most frequently used forms in non-translated texts, but also in translated ones. Overall, they only represented 19.1% of the overall number of reformulations in the entire instrument presented to

subjects, and therefore could be identified a type of reformulation that subjects tend to prefer. The last group is represented by the 25.6% of reformulations only found in the translated texts, often the result of literal translations of the English sourced text.

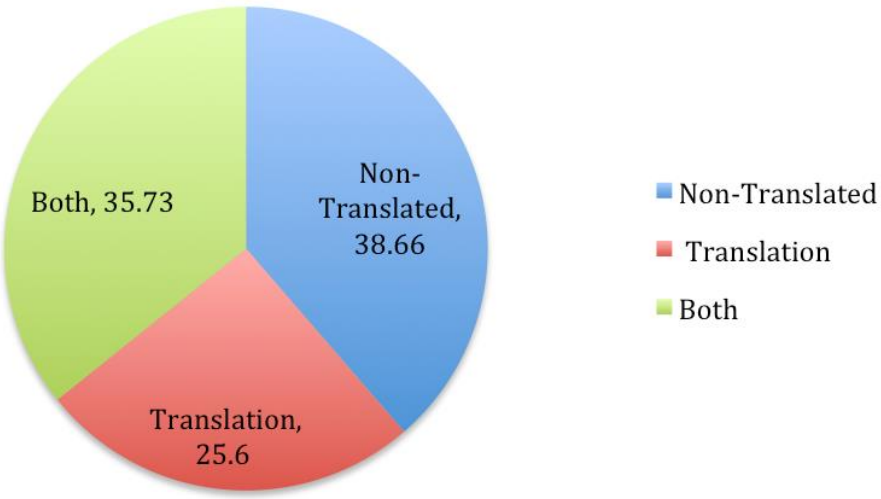

Figure 7. Results of study of preference for reformulations of Latin –Greek terms in translated and non-translated medical websites.

These results suggest that translated and non-translated texts display features at different levels that are shared by both textual populations, but others that appear exclusively in either one or the other. The data obtained in this experiment suggests that these formulations or features that appear exclusively in either textual population in the corpus-based study have different levels of acceptance by the target members of the discourse community at which they are addressed. It is assumed that shared formulations and renderings in both corpora are equally acceptable for speakers of the target language since they appear prominently in the non-translated corpus. However, for those formulations that appear exclusively in one corpus or another, it is non-translated ones that are preferred by subjects to a greater degree [Non-translated=38.66; Translated= 25.6%]. It should again be borne in mind that 44% of the subjects can be considered heritage speakers and although completely bilingual, they have been raised and educated in the United States and therefore can display different levels of acceptability towards literal translations and segments that display traces of interference (Toury 1995). Nevertheless, overall the results from both tests suggest that by large non-translated reformulations-explicitations are preferred. In a context in which corpus-assisted translation is still lagging in the professional world (Gallego-Hernández 2015; Frerot 2016), it is important to insist in the benefits of using corpus-driven methods in professional translation since they could help identify the most recurrent formulations to

help produce more "natural sounding translations" (Bowker 1998; Zanettin 1998; Bowker and Barlow 2008).

8. Conclusions

This study was initiated with the goal of implementing Jiménez-Crespo's (2016) triangulation model to interrelate empirical research in CBTS and CT following the calls to include corpus research in CT (Alves and Vale 2011; Muñoz 2014; Halverson 2016). A reception study of reformulations-explicitations of Latin and Greek terms in translated medical websites in the US was conducted following the results of Jiménez-Crespo (2014) and Jiménez-Crespo and Tercedor´s (forthcoming) corpus studies. The results of these previous studies identified that, contrary to the results in previous similar studies in the English to Danish combination (Askehave and Zethsen 2000a; Zethsen 2005; Jensen and Zethsen 2012), translated medical texts in the English to Spanish combination show lower register and specialisation levels in the lexical treatment of specialised terminology. Thus, the limitations of corpus studies came to the fore, if the translation process resulted in texts that could then be in principle 'easier to understand,' would they in fact be more understandable than the non-translated textual population with higher register and higher use of specialised LG terms? The premise of the paper was thus that, in the light of the necessity of lay-friendliness in translated medical texts (Montalt and Gonzalez-Davis 2007), only experimental studies in the context of CT could help extend the confines of CBTS and address this hypothesis.

The experimental study was designed using the data from the corpus study, and the instruments were developed using this same data extracted from the corpus study. The results of the study showed that both in the dual decision task and in the multiple choice decision task subjects that were the target users of the translated texts, bilingual Spanish/English speakers living in the USA, preferred the most frequent reformulations-explicitations found in non-translated texts. It is of interest to point out that the results of the previous corpus study (Jiménez-Crespo and Tercedor forthcoming) could preliminary indicate that given the lower register and the higher percentage of use of explicitation strategies, Spanish medical texts translated from English could be perceived as more readable, usable or comprehensible. After all, using lower register terms and determinologisation are the most frequent strategies at the lexical level to lower the register and adapt textual genres addressed at non-expert readership. Nevertheless, when corpus and cognitive studies are combined, a clearer picture emerges. The use of corpus-based methods resulted in preliminary results that pointed at one direction, while when these same data were used in an experimental task the results pointed at the other. These results therefore would require further study, using full paragraphs or

small texts, as well as extending the study to an experimental task with professional translators. It would also be beneficial to separate heritage speakers from dominant Spanish speakers in the US, and to enlarge the study population. Despite the initial resistance in both Cognitive Linguistics and CT to deny the possibility of corpus studies showing any connection to actual cognitive processing, the need for a solid interrelation of both subdisciplines or paradigms, both at the theoretical level (i.e. Halverson 2015) and at the experimental level (i.e. Halverson 2011; Jiménez-Crespo 2013b, 2016) is more necessary than ever. It is hoped that this paper will contribute to the debate in this direction and in the emergence of more studies combining these two subdisciplines.

Bibliography

- **Alarcón, Esperanza, López-Rodríguez, Clara Inés and Maribel Tercedor** (2016). "Variation dénominative et familiarité en tant que source d'incertitude en traduction médicale." *Meta* 61, 117-144.

- **Askehave, Inger and Zethsen, Karen K.** 2000a. "Medical Texts Made simple – Dream or Reality?" *Hermes, Journal of Linguistics* 23, 63-74.

- **Askehave, Inger and Zethsen, Karen K.** 2000b. *The Patient Package Insert of the Future. Report for the Danish Ministry of Health*. Aarhus, The Aarhus School of Business.

- **Askehave, Inger and Zethsen, Karen.** 2003. "Communication barriers in public discourse: The patient package insert*." Information Design Journal* 4 (1), 23-41.

- **Baker, Mona.** 1993. "Corpus Linguistics and Translation Studies: Implications and Applications." Mona Baker, Gill Francis and Elena Tognini-Bonelli (eds.) (1993*), Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins, 233-250.

- **Bowker, Lynne.** (1998). "Using Specialized Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study." *Meta* 43(4), 631-651.

- **Bowker, Lynne and Barlow, Michael.** (2008). "A Comparative evaluation of Bilingual Concordancers and Translation Memory Systems." Elia Yuste Trigo (ed) (2008), *Topics in Language Resources for Translation and Localization*. Amsterdam-Philadelphia: John Benjamins, 1-22.

- **Campos Andres, Olga.** (2013). "Procedimientos de desterminologización: traducción y redacción de guías para pacientes." *Panacea* 14, 48-52.

- **Chestermann, Andrew**. (2011). "Reflections on the Literal Translation Hypothesis." Cecilia Alvstan, Adelina Held and Elisabeth Tisselius (eds.), *Methods and Strategies of Process Research*. Amsterdam-Philadelphia: John Benjamins, 13-23,

- **Ezpeleta, Pilar.** (2012). "An Example of Genre Shift in the Medicinal Product Information Genre System." *Linguistica Antverpiensia, New Series Themes in Translation Studies* 11, 139-159.

- **Frerot, Cecile.** (2016). "Corpora and Corpus Technology for Translation Purposes in Professional and Academic Environments. Major Achievements and New Perspectives." *Cadernos de Traduçao* 36(1), 37-61.

- **Gallego-Hernández, Daniel.** (2015). "The Use of Corpora as Translation Resources: A Study Based on a Survey of Spanish Professional Translators." *Perspectives* 23(3), 375-391.

- **González-Darriba, Patricia.** 2014. *English to Spanish Translated Medical Forms: A Descriptive Genre-Based Corpus Study*. MA Thesis, Rutgers University. https://rucore.libraries.rutgers.edu/rutgers-lib/42391/PDF/1/ (consulted 11.01.2016).

- **Gutiérrez Rodilla, Bertha.** (2014). "El lenguaje de la medicina en español: cómo hemos llegado hasta aquí y qué futuro nos espera." *Panacea* 15, 86-94.

- **Halverson, Sandra**. (2011). "Schematic Networks in Translation: Bringing together Process and Corpus Data." *Paper presented at Text-Process-Text*, Stockholm, 17-19 November 2011.

- **Halverson, Sandra.** (2016). "Cognitive Translation Studies and the merging of empirical paradigms." *Translation Spaces* 4, 310-340.

- **Jensen, Matilde.** (2013). *Translations of Patient Information Leaflets: Translation Experts or Expert Translators? A Mixed Methods Study of Lay-Friendliness*. PhD Thesis. Aarhus School of Business, Denmark.

- **Jensen, Matilde and Zethsen, Karen**. 2012. "Translation of Patient Information Leaflets: Trained translators and Pharmacists-cum-Translators – a Comparison." *Linguistica Antverpiensia. New Series* 11, 31-50.

- **Jiménez-Crespo, Miguel A.** (2013a). *Translation and Web Localization*. London-New York: Routledge.

- **—** (2013b). "Crowdsourcing, Corpus Use, and the Search for Translation Naturalness: A Comparable Corpus Study of Facebook and Non-Translated Social Networking Sites." *Translation and Interpreting Studies (TIS)* 8, 23-49.

- **—** (2014). "Medical Translation and the Web: Medical genres in the TWCoMS Corpus Project." *Conference of the American Association of Translation and Interpreting Scholars*. New York University, April 1-3, 2014.

- **—** (2016). "Testing Explicitation in Translation: Triangulating Corpus and Experimental Studies". *Across Languages and Cultures* 16, 257-283.

- **Jiménez-Crespo, Miguel A. and Maribel Tercedor** (Forthcoming). "Lexical Variation and Register in Medical Translation: a Comparable Corpus Study of Medical Terminology in US Websites Translated into Spanish." *Translation and Interpreting Studies*.

- **Klaudy, Kinga.** (1998). "Explicitation." Mona Baker (ed.) (1988), Encyclopedia of Translation Studies. London: Routledge, 80-85.

- **Kruger, Jean Louis, Fox, Wendy and Stephen Doherty.** (Forthcoming). "Multimodal Measurement of Cognitive Load during Subtitle Processing: Same-language Subtitles for Foreign Language Viewers." Isabel Lacruz and Ritta Jääskeläinen (eds). *New Directions in Cognitive and Empirical Translations Process Research*. Amsterdam/Philadelphia: John Benjamins.

- **Laviosa, Sara.** (2002). *Corpus-Based Translation Studies. Theory, Findings, Applications*. Amsterdam-New York: Rodopi.

- **Laviosa, Sara.** (2013). "Corpus Linguistics and Translation Studies." Carmen Millan and Francesca Batrina (eds.) (2013). *Routledge Handbook of Translation Studies*. New York-London: Routledge, 228-240.

- **Meyer, Ingrid, and Mackintosh, Kristen**. 2000. "When Terms Move into our Everyday Lives: an Overview of Determinologization." *Terminology* 6, 111-138.

- **Montalt, Vincent.** (2011). "Medical Translation." Carol A. Chapelle (ed.) (2011), *Encyclopedia of Applied Linguistics*. Hoboken: Wiley.

- **Montalt, Vincent and Maria González Davies**. (2007). *Medical Translation Step by Step. Translation Practices Explained.* Manchester: St. Jerome Publishing.

- **Muñoz Martin, Ricardo**. (2010). "Leave no Stone Unturned. On the Development of Cognitive Translatoloy." *TIS Translation and Interpreting Studies* 5 (2), 145–162.

- **Muñoz Martín, Ricardo**. (2014). A Blurred Snapshot of Advances in Translation Process Research. *MonTI Special Issue – Minding Translation* 2014, 49-84.

- **Muñoz-Miquel, Ana.** (2012). "From the Original Article to the Summary for Patients: Reformulation Procedures in Intralingual Translation." *Linguistica Antverpiensia, New Series Themes in Translation Studies* 11, 187-206.

- **Pilegaard, Morten.** (1997). "Translation of Medical Research Articles." Anna Trosborg (ed) (1997), *Text Typology and Translation*. Amsterdam-Philadelphia: John Benjamins, 159-184.

- **Raynor, Theo D.T.** (2007). "The Importance of Medicines Information Leaflets." *Prescriber* 18 (2), 60-62.

- **Tercedor, Maribel**. (2016). "The Way we Call Realities: Terminological Variation in Medical Language." In preparation.

- **Tercedor, Maribel, and López Rodríguez, Clara Inés**. (2012). "Access to Health in an Intercultural Setting: the Role of Corpora and Images in Grasping Term vVriation." *Linguistica Antverpiensia, New Series--Themes in Translation Studies* 11, 247-268.

- **Toury, Gideon**. (1995). *Descriptive Studies and Beyond*. Amsterdam-Philadelphia: John Benjamins.

- **Zanettin, Federico**. (1998) "Bilingual Comparable Corpora and the Training of Translators." *Meta* 43(4), 616-630.

- **Zethsen, Karen K.** (2005). "Latin –Based Terms: True or False Friends?" *Target* 16, 125-142.

Biography

Miguel A. Jiménez-Crespo is an Associate Professor in the Department of Spanish and Portuguese at Rutgers University, where he directs the MA program in Spanish Translation and Interpreting. He is the author of *Translation and Web Localization* (Routledge, 2013). He has published extensively on web localization in peer-reviewed journals such as *Target*, *Perspectives*, *META*, *Translation and Interpreting Studies*, *Linguistica Antverpiensia*, *JoSTrans*, *Localization Focus*, *Journal of Internationalization and Localization* or *Tradumatica*.

He can be reached at miguelji@rci.rutgers.edu

[1] This research was possible thanks to a grant by the Rutgers University Council Grants program. It was also possible thanks to the CombiMed Project (FFI2014-51899-R), funded by the Spanish Ministry of Economy and Competitiveness.
A preliminary version of this paper was presented at the IATIS fifth conference in 2015.

[2] It should be mentioned that in Spanish both 'azúcar baja' and 'azucar bajo' [low blood sugar], with the adjective in the masculine and feminine form are possible in Spanish. They are prominent in the translational corpus and only one instance appears in the non-translational corpus. A search in the dialectal variation site Diatopix showed that the masculine form, 'azúcar bajo' is preferred in all Spanish–speaking countries except in Colombia. In any case, both forms are acceptable for Spanish speakers.