Meta

Journal des traducteurs
Translators' Journal

ΜΕΤΑ

Finding Translations. On the Use of Bibliographical Databases in Translation History

Sandra Poupaud, Anthony Pym et Ester Torres Simón

Citer cet article

Poupaud, S., Pym, A. & Torres Simón, E. (2009). Finding Translations. On the Use of Bibliographical Databases in Translation History. *Meta*, *54*(2), 264–278. https://doi.org/10.7202/037680ar

Résumé de l'article

Dans toute étude de traductions, il faut tout d'abord déterminer l'objet qui doit être pris en compte comme une « traduction », puis comment le trouver, en général en ayant recours à des bases de données bibliographiques. Partant de la conception maximaliste selon laquelle les traductions sont potentiellement omniprésentes, nous proposons que les différents processus de distribution imposent une série de filtres sélectifs grâce auxquels certaines traductions sont plus facilement identifiées et accessibles que d'autres. L'étude des traductions doit tenir compte de ces filtres préalables et être à même de les expliquer, parfois de les surmonter. Les processus de recherche imposent ensuite, nécessairement, leurs propres filtres sélectifs, qui peuvent réduire ou augmenter le nombre et le genre des traductions issues des filtres préalables. Nous présentons ici trois projets de recherche dans lesquels les rôles respectifs des filtres préalables et des filtres de recherche sont très différents. Pour des hypothèses très spécifiques et à grande échelle, l'*Index Translationum* peut être relativement efficace. Pour des projets plus détaillés tels que l'étude des flux de traductions entre l'espagnol et le français pendant une période spécifique, une base de données destinée à l'usage des professionnels de l'édition offre des avantages significatifs. Enfin, pour une étude, marquée par la rareté des textes, portant sur les traductions du coréen en anglais après la guerre de Corée, il est nécessaire de combiner plusieurs bases de données, la plus efficace se révélant être celle d'Amazon.

Finding Translations. On the Use of Bibliographical Databases in Translation History

SANDRA POUPAUD
Universitat Rovira i Virgili, Tarragona, Spain
spoupaud@arrakis.es

ANTHONY PYM
Universitat Rovira i Virgili, Tarragona, Spain
anthony.pym@urv.cat

ESTER TORRES SIMÓN
Universitat Rovira i Virgili, Tarragona, Spain
ester.torres@urv.cat

RÉSUMÉ

Dans toute étude de traductions, il faut tout d'abord déterminer l'objet qui doit être pris en compte comme une «traduction», puis comment le trouver, en général en ayant recours à des bases de données bibliographiques. Partant de la conception maximaliste selon laquelle les traductions sont potentiellement omniprésentes, nous proposons que les différents processus de distribution imposent une série de filtres sélectifs grâce auxquels certaines traductions sont plus facilement identifiées et accessibles que d'autres. L'étude des traductions doit tenir compte de ces filtres préalables et être à même de les expliquer, parfois de les surmonter. Les processus de recherche imposent ensuite, nécessairement, leurs propres filtres sélectifs, qui peuvent réduire ou augmenter le nombre et le genre des traductions issues des filtres préalables. Nous présentons ici trois projets de recherche dans lesquels les rôles respectifs des filtres préalables et des filtres de recherche sont très différents. Pour des hypothèses très spécifiques et à grande échelle, l'*Index Translationum* peut être relativement efficace. Pour des projets plus détaillés tels que l'étude des flux de traductions entre l'espagnol et le français pendant une période spécifique, une base de données destinée à l'usage des professionnels de l'édition offre des avantages significatifs. Enfin, pour une étude, marquée par la rareté des textes, portant sur les traductions du coréen en anglais après la guerre de Corée, il est nécessaire de combiner plusieurs bases de données, la plus efficace se révélant être celle d'Amazon.

ABSTRACT

In any study of translations one must first decide what is to be counted as a "translation" and how such things are to be found, usually through recourse to bibliographical databases. We propose that, starting from the maximalist view that translations are potentially everywhere, various distribution processes impose a series of selective filters thanks to which some translations are more easily identified and accessible than others. The study of translation must be aware of these prior filters, and must know how to account for them, and sometimes how to overcome them. Research processes then necessarily impose their own selective filters, which may reduce or extend the number and kinds of translations given by prior filters. We present three research projects where the play of prior and research filters is very different. For one-off large-scale relational hypotheses, the *Index Translationum* is found to be relatively cost-efficient. For more detailed objects such as translation flows from Spanish into French in a specific period, a book-industry database offers significant advantages. And for a study marked by a paucity of texts, as

is the case of translation from Korean into English following the Korean War, a combination of databases is necessary, the most useful turning out to be Amazon.

MOTS-CLÉS/KEYWORDS

histoire de la traduction, définitions de la traduction, méthodologie de recherche, bibliométrie
translation history, definitions of translation, research methodology, bibliometrics

1.   Introduction

When undertaking virtually any study of translations, you have to find, identify and select the translations you want to talk about. The task may seem banal and even straightforward. However, what you find, and especially where you find it, can condition not just your object of study but also, very significantly, the kinds of results you come up with.

This article will review the way translations have been found in three quite different research projects. The first is a one-off study of the general relations between translations and the size of target markets, designed to test a specific simple hypothesis; the second is an ongoing research project on translations from Spanish in France in the period 1980-2000; and the third looks at translations from Korean into US English in the period following the Korean War. The practical problems encountered in these projects are extremely diverse, as indeed have been the databases we have worked from. We will be giving consideration to the UNESCO *Index Translationum*, to a professional book-industry database (Electre), and to an online bookseller (Amazon). Our aim is to sketch out the advantages and limitations of each database. The principles and problems, however, are more general, and concern many other kinds of bibliographies and catalogues.

2.   Three principles

In principle (the first principle), translations are everywhere, at all times. It just depends on where you look, and on what you call a translation. More technically, one must assume the ubiquity of the object and then consider the way representations necessarily filter that object into presence. The filtering happens first by the semiotics of metadata and the technologies of distribution (the world calls some texts "translations," then makes some of them easier to find than others), and second by the definitions put to work in each particular research project (some kinds of translations are the ones the researcher wants to find). This actually gives us three principles, which may require brief elaboration.

2.1. Principle of ubiquity: Translations are everywhere

Unless stated otherwise, translations may be the result of any communication, intralingual or interlingual, involving meaning transformation. This is the kind of broad definition currently in vogue in a certain postmodern sociology (see Renn 2006; Akrich *et al.* 2006) and indeed in the vogue of "cultural translation" in a perennially reinvented Cultural Studies. As such, translation may be seen as occurring wherever different languages or indeed different discourses are in contact; translations are

constantly spoken, and of course thought. Translation is thus basic to any situation for which homogeneity cannot be assumed. Indeed, translation would very much be a constitutive feature of the dreams of a publican and his wife happily asleep upstairs in the pub in Chapelizod, Ireland. But we would not know about those dreams, or those particular translations of the self and the world, had they not been published in *Finnegans Wake*.

The only problem with this first principle is that much of the world spends some time and more energy stating otherwise.

2.2. Principle of the prior filter: If you find a translation, someone wanted you to find it

Only some translative activity turns into texts (spoken or written stretches of discourse attributed with macrostructural unity), and only some texts attain significant distribution. They are recorded, reproduced, written down, or otherwise put into a form that can be moved away from the situation of production. Some are then accorded "epitexts" (Genette's term for publicity material, interviews, etc. that promote a text) and a string of "metatexts" (all the texts about the text, including entries in library catalogues, bibliographical listings, and a place in the kinds of catalogues that interest us here). For some, those epitexts and metatexts are also translations, but there is no need to complicate definitions at this stage. Let us bundle together all those technologies, efforts and semiotic representations and call the sum the "prior filter," meaning that these are the selection processes that happened prior to our own intervention in history. Most texts, of course, are filtered *out* by these processes, either because they are considered to be of not enough value to set off extensive reproductive machinery, or because there are active interests in keeping them away from any kind of distribution.

For example, several members of our research group have spent some years trying to record the work of interpreters and cultural mediators in hospitals. We are trying to give the translations a technology (audio or video recording) that can turn them into an object of knowledge. Yet the resistance is insidious: the hospital administrators politely refuse, for a hundred reasons that have nothing to do with the fact that they do not want us to find out why they should spend money on interpreters, and the interpreters themselves refuse, since they are underpaid, thus untrained, and not particularly proud of what they do anyway. And no one got as far as asking the patients. There is thus a kind of institutional consensus that these particular translations should not be looked at beyond their immediate communicative situation. This would be active negative filtering, which one may or may not want to challenge.

There are other kinds of negative filtering. For example, many historical translations into French sought to avoid censorship by having the place of publication printed as "Amsterdam" or "Brussels" or "Strasbourg," or indeed anywhere outside the national territory of France. That way they did not have to send a copy to the government authorities (the *dépôt légal*). The printer could thus not be prosecuted. This meant, however, that the title was then not listed in the Bibliothèque de France, and is thus not listed in any database derived from that inventory. That includes, unfortunately, the *Bibliographie des traductions françaises (1810-1840)* (Van Bragt

1995), which remains a magnificent piece of work awaiting some good research hypotheses (see the review by Pym 1997). The printers avoided the censor, but they thereby avoided all the state mechanisms for metadata and distribution. In the age of nations, one still needs a passport in order to travel.

On the other hand, some translations are relatively easy to find. Not only are they printed or otherwise inscribed, but they are promoted in the media, spawn myriad epitexts, enter all the catalogues and bibliographies, and may be otherwise propelled into the world through the instruments of public or supra-national policy. Those instruments include national and international prizes, collections of "representative works," and state-funded translation programs, most commonly to promote a national literature (or, under many Communist regimes, the literatures of brother countries). All those things require effort and expense. No one works or spends money if they have no interests at stake (altruism is a personal interest). It follows that when you find a translation, a whole series of people have usually had an interest in you being able to find it. The illusion of immediate or fortuitous presence ("Here it is! I have found it!") masks the historical drama of all the forces that have worked either for or against your discovery. Rejoice in prowess for a moment, then reflect on why the task was so difficult.

Prior filters are many and various, and there is no sense in drawing up a checklist here. Note, however, that they are to be treated as more than sources of potential error. For example, one quite often comes across pseudotranslations (translations that refer to a source text that never existed) and pseudo-originals (translations presented as non-translations), and one can try to correct the attributions that gave those texts their false genealogies. However, the attributions themselves, the filters that gave or did not give translational status, must be appreciated as the factors that actually create the status of pseudotranslations and pseudo-originals. If the filters make mistakes, they are at least creative in the process, and not always innocent. The same could be said for the correction of the mistakes.

There is rarely any question of simply correcting a prior filter anyway. In most cases, what the researcher does is compare one prior filter with another, spot the omissions and contradictions, then decide the issue one way or another (thus intervening with the researcher's own filter). For instance, if all the library catalogues identify a story as an original, yet the researcher has located a text from which the story would appear to have been translated, they might decide to identify the story as a pseudo-original, thus altering classification by the prior filter. Has the researcher thus introduced truth? Not at all. They have simply decided to prefer, for whatever archeological reasons, the metadata on the text that they now want to call the "original." They have preferred one prior filter to the other.

The fact that researchers can identify pseudotranslations and pseudo-originals indicates the partial and heterogeneous nature of prior filters. There should be no question of assuming just one filter for each culture (does French culture stop at the limits of the Bibliothèque Nationale?), just as there should be no question of trying to add up the results of all the prior filters so as to get some kind of list of "all the translations." That is impossible, since translations are everywhere (our first principle). The object of knowledge will always extend beyond what we grasp. If a researcher thinks they have all the translated books, they can go fishing in the periodicals, and

when they think they have all of *them*, they can try to catch the oral translations from the air of the past.

Researchers do better to spend time defining and applying their own filters.

2.3. Principle of the research filter: One cannot study all translations

The encounter with prior filters only happens once the researcher is actually doing research. That is, they are off in search of something, usually a particular kind of translation. There are two reasons why the researcher only wants a particular kind of translation. First, since translations are potentially everywhere (our first principle), the fact that a researcher is looking for them means, necessarily, that the researcher does not want them all (no one has to look for something that is truly everywhere – one might as well try to grasp the air). Second, even if one does want them all, there is no way to do research on them all (for want of extra lifetimes, to say nothing of the logical fact that the research will itself be a series of translations, thus constantly extending the object). So the researcher only wants some translations. They are thus going to develop and apply their own filter, a specific "research filter." But where should the criteria come from? What is the researcher going to leave out?

Here we take the position that the term "translation" needs to be defined explicitly in each particular research project that involves looking for the beasts, and the way the term is thus operationalized must depend on the particular research project in question (the hypotheses to be tested, the resources available for the testing, and the communicative purpose of the research). The research filter is then nothing more and nothing less than the way the term "translation" has been operationalized.

When doing this, there is no need to confront undue existential dilemmas and the like, perhaps of the kind described by Theo Hermans:

> The question only becomes acute when we try to speak about "translation" generally, as a universal given and therefore supposedly present in all cultures; or when we wish to understand what another culture means by whatever term they use to denote an activity or a product that appears to translate as "translation" – whereby we naturally translate that other term according to our concept of translation, and into our concept of translation; and in domesticating it, we inevitably reduce it. (Hermans 1997: 19; see the discussion of this passage in Halverson 2008)

There is no guarantee that our filter will necessarily "reduce" the other terms (our own research tends to add conceptual extensions), nor is there any surety that a culture (including our own) has just the one concept that could be "domesticated" (as if our own culture were a nice home with just one comfortable consensus). If you think in those terms, you are likely to overlook the active interventions of prior filters (as if each culture had passive concepts, just waiting for us to come along and domesticate them). You are also unlikely to be explicit about your own research filter, which in many cases can (and should) challenge untheorized assumptions about what a translation is or should be. The operationalization of research filters can be a long way from the aporias of cross-cultural encounters. The metaphysical wranglings with definitional responsibility are in any case the luxury of first-world scholars, who too frequently claim abstract "vigilance" (after Derrida) rather than explicit operational decisions.

The one thing a researcher cannot do is fail to operationalize the terms, as if there were some pristine openness to the world of foreign concepts. In more techni-

cal terms, research cannot assume a natural ontology of translation, at least not in any form beyond start-up fictions like our first principle above. One cannot say "I went out into the world looking for translations; I found some; and here are the defining features of what I found: A, B and C." No, that is probably not what happened. It seems more likely that the researcher started with an intuitive idea of what they wanted to find, which led to only a very small part of the available evidence, then the researcher asked some specific questions about those things, and the evidence replied: A, B and C. At a latter stage, the A, B and C (or D, E and F) become the research filter, the researcher's own explicit operative conceptualization of translation. That is an epistemology, not an ontology. If one asks different questions, the world will reply in a different way, and one will be using a different research filter.

Research filters can be quite formalized and abstract, for use at very general levels (Pym 2007). Some examples would be the three postulates by which Toury sets out to identify "assumed translations" (1995: 33-35), or the "complete interpretative resemblance" that Gutt (1991: 186) sees as the assumption created by "direct translations," or even the maxims of quantity and of first-person displacement proposed by Pym (2004). In most projects, however, the filter can be of a more local, pragmatic kind, often as a rule of thumb. It may prove efficient, at least in an initial survey, to accept as a translation everything classified as such in one particular prior filter. Then the researcher works on it, asking questions and getting answers as they go, comparing the results of different filters, and thus developing their own filter.

How this can be done is best illustrated by example. Here we present three.

3. What the *Index Translationum* is good for (Anthony Pym)

It has become mildly fashionable to use the UNESCO *Index Translationum* in its online version, and then complain about its qualities as a database. Its first and prime quality must nevertheless be its online availability. In seconds (or perhaps minutes, depending on the connection speed) one can have data on translations to and from a lot of languages over a fair number of years (since 1979). That is a very big advantage. For earlier periods, one can also consult the paper-based versions in the UNESCO statistical yearbooks, which give numbers on translations and a lot more besides.

The basic disadvantage of the database is that, like the entire United Nations system, it is only as good as its member states. And the data-gathering capacities of those states are highly variable. This can be seen by plotting the numbers year-on-year, which often gives wild fluctuations that can only be attributed to inconsistencies in the census techniques. Beyond that, there is no universal agreement on the basic categories: each contributor can define "book" as it likes, and the meaning of the category "translation" would similarly seem unregulated. Under those circumstances, or at least without checks through other filters, great care should be taken whenever attempting to compare data from one country with those of another.

The way around these problems is fairly straightforward. First, to ride out the fluctuations, pick a period of at least three consecutive years and work on the means (in the example below, we have used a period of five years, selected so as to come prior to the withdrawal of the United States from UNESCO in 1984). Second, within each country (and preferably each year), privilege proportional data. For example, if one only looks at the proportion of books published to books translated (i.e., the

percentage of translations), it does not matter too much how that country defines what a book is, or how enthusiastic it is about collecting data. Presumably the definition and the enthusiasm will be roughly the same for both the numbers presented. Third, only use the database as a rough guide to large-scale quantitative relationships, where aspects like different cultural concepts of "translation" are not likely to be of major consequence. Examples of this kind of use can be found in Heilbron (1999), which tests the validity of classifying languages in terms of their central or peripheral status within a world system, or Pym and Chrupala (2005), which tests the relation between translation percentages and the relative size of the publishing system concerned. Figure 1 shows one of the graphs from the latter, where the aim was to test the very specific hypothesis that the bigger the publication space of the target language, the lower the percentage of translations in that language (so the low percentage of translations in the United States could be due to the high number of books published, rather than a direct consequence of cultural hegemony). In fact, the purpose of the exercise was to question the common assumption that the relatively low percentages of translations into English are a direct indicator of hegemony (see Venuti 1995). Note that the nature of this hypothesis requires us to accept as a "book" anything that the various countries choose to call a "book," and the same for "translation" as well. The resulting pattern, however, is so clear that little further investigation is needed on those points of relativist definition.

The *Index Translationum* is a convenient low-effort first step, suitable to large-scale testing of one-off hypotheses, where little detail is needed in the background and the sheer numbers of titles will outweigh any need for accuracy. Of course, if the resulting pattern were not convincing, or some more complex hypothesis is at stake, then better filters are required. As is mostly the case.

FIGURE 1

Percentage of translations by percentage of books published in language.
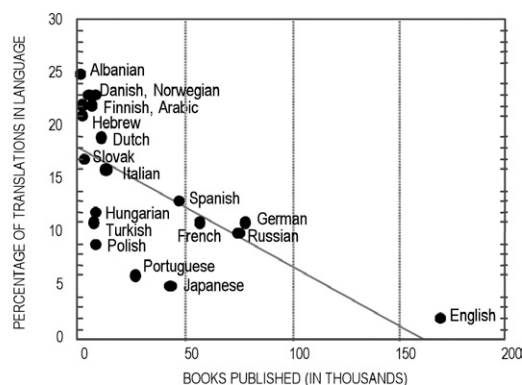**UNESCO data for 1979-1983.**

Figure 1 leaves little doubt that the more books are published in a language, the lower the percentage of translations in that language (this is crudely indicated by the diagonal line, which should actually be an algebraic curve – see Pym and Chrupala 2005). In this case, that is all we were looking for. Of course, a lot remains to be

explained. For example, why are the translation percentages for German, French and Italian not in the mid to high 20s, as is more usually the case? Why does Albanian, of all languages, get top mark for percentage of translations? Why is the percentage for Japanese so low? For all those questions, indeed for any kind of in-depth study, one inevitably has to turn to a more detailed and more reliable set of filters.

4. Filters in the Electre database (Sandra Poupaud)

My research deals with translations of literary works from Spanish in France, and more specifically with the role of the various agents involved in these translation practices. Since the focus is on translators rather than translations as such, different conceptualizations of the term "translation" tend not to be problematic. The purpose is to draw a general map of the literature translated from Spanish, not to compile an exhaustive list of all translations. However, it was through the translations that the translators and publishers had to be located, at least in terms of a general background study. When trying to find a suitable database for this, I faced two main problems, which are fairly common when working with databases. One problem is linked to the data, their availability and reliability. The other problem has more to do with the design of the interface, the search options and the organization of the data.

For these research purposes, one of the main drawbacks of the *Index Translationum* (besides its notorious unreliability) is very literally a problem with the search filter: no bibliographic search according to the country of the source text can be carried out. This is a marginal difficulty when dealing with languages such as German or Italian, but it becomes truly problematic when considering international languages such as English or Spanish. This project requires identification of works coming from Cuba, Mexico, Spain, and so on. But if one uses the *Index*, the only way to isolate the country of origin of the translated title is, unfortunately, to do so manually. As a further complication, the Literature category of the *Index* (category 8 of the UDC) also includes Children's Literature, which had to be excluded from this project as it corresponds to different agents and publishing strategies. In a first pilot study carried out over a three-year period, I used data from the online *Index Translationum*, excluded Children's Literature manually, and did not look closely into the countries of origin of the translated titles. This was highly time-consuming and not really satisfying.

There are some alternatives. The online catalogue of the Bibliothèque Nationale de France (BNF), while very useful for an isolated search, does not allow one to obtain bibliographic data on translations over a period of time. In fact, the source language is not even a search criterion, while the search options for the country are only those of the country of publication of the translated book. The CD-ROM version of the BNF does allow searches according to the source language, but once more the country of origin is not mentioned.

Another option was offered by the Centre de sociologie européenne in Paris, which was conducting a research project on translations in France under the direction of Gisèle Sapiro (Sapiro 2008). The Centre had reached an agreement with the professional database for the French book industry, Electre, by which Electre graciously provided them with the data on translations published in France between 1984 and 2002.

Electre is a professional database created by the Cercle de la Librairie, the French booksellers' professional association (information available in French only at www.electre.com). It is the bibliographic reference tool used by booksellers when they check for the availability of a given title. Originally paper-printed, the database has had an electronic version since 1984 and became available online in 1997. It is available in Canada in the *Memento* database, which was created in 2005 through a partnership between Electre and the Banque de titres de langue française (BTLF). Similar databases exist in other countries, often in connection with ISBN agencies which provide a convenient means to locate them.

The Electre database contains over 900,000 books published in French, including 12,000 forthcoming titles and data on unavailable books published since 1984. Updated daily thanks to the information provided by publishers, it contains detailed bibliographic notices. It also contains information about the availability and selling price of the title. All this information is provided by the French distributors.

The Electre website extols the virtues of their database as follows: it takes its data from the source (publishers); it is exhaustive, structured and respects given norms; it has two thematic indices (Dewey and Rameau) and a powerful search engine. To ensure the coherence of data and a normalized access, it uses authority records and the notices are written using the French bibliographic norm Afnor Z 44-073.

Of course, one of the problems with the Electre database is that it is not readily available to the general public. The solution in this case has been through cooperation with the Centre de sociologie européenne, as they needed someone to work on the translations from Spanish. They provided the data, and in exchange I wrote an article for their research project, analyzing this specific set of translations (Poupaud 2008). It should be pointed out that I was not given direct access to the Electre database but received the data in Excel files that had been pre-processed by Anaïs Bokobza and the Centre de Sociologie Européenne. The data supplied by Electre to the CSE were in DB format, which the Centre de Sociologie européenne then proceeded to reorganize and transfer to Excel files.

At first sight, one of the main benefits of this database is that it contains all the information needed for the project, and even more.

The Excel files concerning Spanish were divided into two main files: one for the new titles and one for the paperbacks. The file contains 37 fields with, among others: author, title of the book, sex of the author, genre (novel, poetry, theater, etc.), country of origin, year of publication, publisher, collection (if any), existence of a paperback edition, name and sex of the translator(s), previous translation(s) (if any), price, thematic index. A number of fields had not been filled in, such as the print run (which publishers generally refuse to communicate), the date of publication in the source language, and the literary prizes or publication subsidies received by a given title. Other fields, such as the presence of a paratext and the name of its author, were only partially filled in.

Having the data in an Excel file allows for highly flexible analysis and classification options at the macro and micro levels. Using pivot tables in Excel, it was possible to study the breakdown of translations according to the country of origin, the plots over time of translations coming from a given country (see Figure 2 for Spain) and to isolate various phenomena specific to that country (the rise of Cuban literature, for instance).

**New literary translations, Spanish into French, from Spain, 1985-2002.
Source: Electre.**
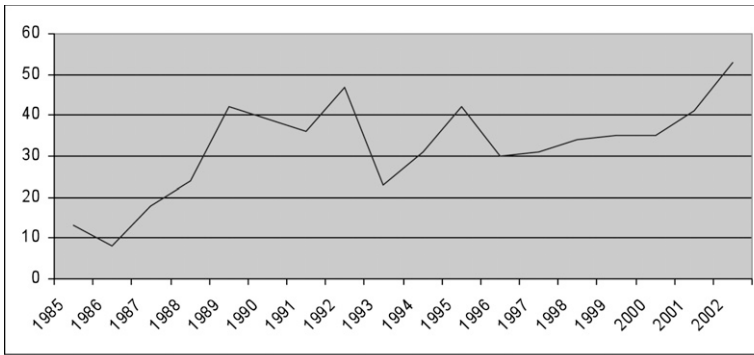
This also helped highlight the impact of certain events on the number of translated titles. For instance Spain's entry into the EU in 1986 triggered a rise in the number of translations, as did the 1995 Salon du Livre in which Spain was the guest of honor. It was also possible to see rapidly who were the most translated authors or genres. This highlighted the interest raised by Manuel Vázquez Montalbán, with 34 new translations over the 1985-2002 period, and the ongoing success of the authors of the Latin-American boom: Vargas Llosa, Fuentes, and Cortázar still being among the most translated authors. This flexibility made it possible to choose a more detailed level of analysis, which is particularly suitable for the ongoing research. It allowed study of the main trends mentioned above but also to follow in a more microscopic manner the actions of the various agents, here the publishers and translators. I was thus able to see if a given publisher had a specific strategy in terms of countries, authors, genres (Gallimard specializes in the authors of the Latin-American boom; Christian Bourgois privileges contemporary Spanish authors, for instance), or if there existed any kind of loyalty between authors, translators, and publishers (very often, there is not).

There were also a few problems that are difficult to avoid when dealing with an institutional filter like this. One problem is linked to the design of the database. Since this is a commercial service, the initial feeding in of the data is not compulsory. It is carried out voluntarily by publishers, hence the absence of those reluctant to use this service, mainly small publishers. This is where the prior filter built into the database leads me to question my own research filter, since a decision to adopt without further questioning the results obtained from Electre is likely to mean leaving out the less commercial titles, which have been filtered out by Electre's definition of a book (print run over 500, regular commercial distribution, no self-publication, among other criteria). Looking at the Electre database for the period between 1985 and 2002, one thus gets the impression that the Spanish poet Antonio Machado was not translated at all, while the *Index* lists three titles for the period. One title was published by the translators themselves, and another by an art gallery, and both had thus been filtered out by Electre. Since the database is mostly used by publishers and booksellers to market and sell books, there is little interest in listing publications that are off the

main distribution circuit. At this point the researcher has to decide if the research filter can accommodate the restrictions imposed by the prior filter, in which case it should be made clear that the term "translation" has been operationalized as "commercial translation," for lack of a better word. If researchers reject this negative filtering and wish to include more marginal translation practices, they have to go beyond the limitations imposed by a commercial database and pursue their quest using less flexible institutional tools such as the *Index Translationum* or the catalogue of the Bibliothèque Nationale.

Another problem with the use of Electre is that in some cases the country of origin had not been mentioned in the original input data and was then given as being Spain by default. This meant I had to go over the whole list of entries for Spain in order to check and correct when Spain was not the proper country of origin of the translated title. The classification by genres also has to be used carefully as it can be slightly fluctuating. Another problem is linked to the treatment of new editions and reprints. Reprints have been excluded when processing the data extracted from the Electre database, while they can give useful information on the success (or lack thereof) of a book. The existence of new editions is mentioned, but their date has not been indicated, so it is in fact impossible to determine how many books are published in total for a given year. Finally, some components have been excluded from the Literature category (notably cartoons), so it makes it difficult to compare the data extracted from Electre with other databases using a different definition of Literature. This problem does not exist for Children's Literature, as it is identified easily in the Electre database; one can choose to include or reject it, according to needs.

In short, researchers have to know precisely how the prior filter has been defined before they carry out any type of comparison. If not, they run the risk of comparing apples and pears.

All in all, the Electre database has proven a rewarding source of data for my research purposes, bearing in mind the limitations induced by its prior filter. The coherent and detailed data it provides, allied with the possibility of analyzing them in Excel, has been a great help. Beyond this specific case study, this type of professional database can be a valuable source of data for researchers, always bearing in mind the potential filters induced by their commercial nature.

5. Using Amazon (Ester Torres Simón)

The original aim of my research was to sketch the image projected by translations from Korean into English in the United States during the Cold War. This would ideally reveal something about the role of translation as image builder.

This project did not concern exchanges between major languages, as do the ones presented in Pym's chart or Poupaud's French-Spanish study. It dealt with cultural exports from a country that was barely recovering from a period of colonization and that was opening to the world for the first time. One of the direct results of this situation was a lack of organized information on translations. This absence of structured databases led me to look at the *Index Translationum* as a source of information on translation flows. As it happened, the *Index*'s distribution by country was helpful in order to isolate "translations in the United States," which thus became operationalized as "translations published in the United States."

The research corpus was built from the translations from Korean published in the United States from 1950 (the beginning of the Korean war) to 1974 (the establishment of diplomatic relationships with China). However, the decision to use the *Index Translationum* proved to be wrong, for the very same reasons that had made me opt for it.

Since the project concerned an object that was small in terms of cultural distribution, the unreliability of the data became immediately visible. When working on a large cultural distribution, a fluctuation of one, two or ten volumes does not change the general picture. However, for translations from Korean published in the United States, the book-form version of the *Index Translationum* from 1950 to 1974 gave such a low number of titles (14 in 25 years) that it was not even possible to talk about a flow. Of those 14 listed titles, two volumes were repeated in different years (and they were not re-editions), two books were listed as translations from Korean and they were later proven to be translations from Chinese and German, and another title was a Korean language textbook. This reduced the total list to just nine titles.

Assuming that there is a relationship between increasing interest in a culture and translations from that culture, logic dictated that there should be more translations. The interest was certainly there: enormous casualties, injured and POWs during the Korean War, thousands of United States soldiers living in Korea, and many Korean immigrants in the United States. In the period prior to 1974, the Korean War was not yet the Forgotten War.

It became clear that some assumptions had been taken too much for granted. First, the volumes published in the United States were not the only ones to reach the American public. Volumes published in the United Kingdom, Japan or Korea could have reached the United States easily, as many publishing houses had distribution arrangements with local companies. The research filter I required was perhaps not well served by the prior filters built into the *Index Translationum*.

A search for titles translated into English and published in Korea (which appear under Korea with an asterisk) offered only one result to add to the previous nine. Once again, this information in the *Index Translationum* was expected to be defective as it was actually provided by the Library of Korea, the organism that provides information on all books published in Korea. Any country recovering from a Civil War is bound to have several priorities considered more important than cultural organization. Thus, for example, there is no information at all for 1968.

Second, analysis of the Korean section of the Library of Congress revealed differences between its filters and those of the Library of Korea. While the United States considered "volume" a synonym of "book," Korea widened its sense to include "speeches" and "bulletins." This filter was especially interesting for this project, as most of the works on Korea could be expected to be technical and informative and not necessary in book form.

I looked for other databases in order to double check the fluctuations. I located the Korea Literature Translation Database (LTI) established by the Korean Culture and Arts Foundation. The LTI was born in March 2001 with the acquisitions and integration of functions and responsibilities previously held by the Literature Department of the Korean Culture and Arts Foundation and the Korea Translation Foundation. The LTI database is thus more exhaustive with respect to books. It gives no less than 56 results for the timeframe of this project, including volumes translated

in Korea, the Philippines and the United Kingdom. The ten volumes (nine published in the United States and one in Korea) previously located in *Index Translationum* were also included in the LTI. The concept of "translation" was also wider here, as retellings of Korean oral traditions were considered translations, even though the actual texts were written originally in English. Some indirect translations could also be found, notably via Japanese.

Another source of data was the University of Yonsei in Korea, an institution that has always promoted translations and has published many of them. The search engine of the university's main library produced a further five titles. The engine allowed the subject "Korean – translations" and language discriminations. No clear reason could be found to explain why these further five titles only appeared here, other than the limited distribution of the volumes.

I thus eventually had around 60 volumes to work with, many more than the nine that resulted after consulting the *Index Translationum*. The problem was that I could not be completely sure about the distribution of these translations in the United States. If the titles did not reach the United States, they could not tell us much about the American image of Korea.

This is where Amazon.com became the solution. On the one hand, Amazon gives an accurate image of what is available to the American public. It not only provides their own selection of books, but also includes information on volumes held by other bookstores, second-hand bookstores and private sellers. Once the information on a book is entered into the database by any of the possible sellers, it is kept for future reference, even if the book is out-of-print or unavailable. Government documents are also listed, becoming a good reference database for grey literature. There were thus guarantees that the books had been available to the American public at some point in time, even the minor technical reports. This helped widen the corpus.

Amazon's most valuable asset is its advance-search engine. On top of the usual options of author, title and keyword search, it allows for discrimination by publication date ("before-during-after year") and the topic search is precise. The results can be presented according to different criteria: best-selling, publication date, author a-z/z-a, title a-z/z-a, and the total number of results are given with the list. On the left-hand side, the titles are organized under secondary subject headings, which include the entry "other languages." It is also possible to discriminate them by "New – Used – Collectible." For a more specific search, there is Boolean Search for all these fields. Both search possibilities are easy to use and well-explained.

A "Power Search" for literature published from 1950 to 1974 in English under the subject "Korea" gave no fewer than 640 results. Most of the titles were technical works (485, several of them translations), followed by translations of Korean leaders' speeches (82), literary works (71, of which 33 were translations) and war-related books (12, of which five were possible translations).

As Amazon was not designed to be a translator researcher's tool, it presents some limitations. The most important is that "translation" as such does not constitute a field. Often, the translator appears as another author of the book and is included in the author field search, and only some books are marked as translations in the book review section. Further, the flexibility of the subject organization provides wider results on a first search but may lead to erroneous conclusions when consulting the secondary subjects. These secondary subjects are not reciprocally exclusive. For

example, books listed below Korea-Non-fiction may be listed as well under Korean-Military, making the numbers given by the secondary subjects unreliable.

These two drawbacks make the corpus-building process slower, as the results need to be compared for very accurate analysis. On the other hand, the researcher is free to apply their own definition of translation. As mentioned above, some of the results provided by the Language Translation Institute Database were retellings in English of Korean folk tales. Translation databases often do not consider such books translations, but originals. The bookseller's database, however, is more interested in being exhaustive instead of restrictive. The "translation" filter thus becomes the responsibility of the researcher.

The use of Amazon thus provided a good picture of the imports from Korea, and perhaps of public expectations in the United States. It was now possible to plot a rise in interest in Korea since the Korean War:

FIGURE 3
Titles on Korea available in the United States, 1956-1980
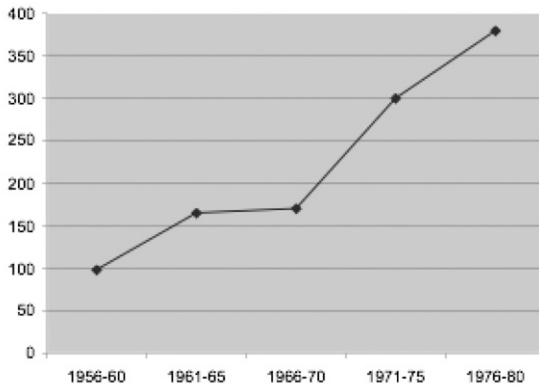
The corpus had risen from the original nine volumes appearing in *Index Translationum* to more than 120 confirmed volumes and 100 more possible ones. Without further study of the possible translations, accurate figures cannot be presented, but a general position can. Translation assumed three tasks at presenting Korea to the United States. First, it was a tool to track the positions of the two Koreas by translating speeches and statements from leaders of South and North. Second, translating local economic, political and technical texts allowed them to be compared with non-Korean studies on the peninsula's development. And third, translations introduced Korean folklore and fairy tales to the American public. This increase in the number of translations is framed by an increase in literature on Korea in general.

As a conclusion, information from bookseller's databases can supply exhaustive data on available titles, giving researchers the chance to apply their own research filter. Research thus becomes less dependent on prior filters.

6. A general conclusion

The three research projects presented in this article are very different, sometimes with approaches quite opposed to the use of bibliographical filters. However, the first

important point is that all three cases work on the same tension: the authorities of prior lists conflict with the priorities of research; the given filters compete with the need for our own. The second main point is that, in all three cases, we have used the lists in order to think beyond them. Research has been more than the repetition of data.

After all, the lists are only of things. History, especially translation history, is about people. To get to the people one has to go beyond the lists. One has to uncover the drama of distribution and concealment, the conflict of human interests weighing for and against the movement of objects across time and space. Those are the struggles that make the passion of history. Bibliographical databases are no more than the traces of such stories.

REFERENCES

Akrich, Madeleine, Callon, Michel and Latour, Bruno (2006): *Sociologie de la traduction. Textes fondateurs*. Paris: Presses de l'École des Mines.

Genette, Gérard (1978): *Seuils*. Paris: Seuil.

Gutt, Ernst-August (1991): *Translation and Relevance. Cognition and Context*. Oxford: Basil Blackwell.

Halverson, Sandra L. (2008): Translations as institutional facts: an ontology for "assumed translation." *In*: Anthony Pym, Miriam Shlesinger and Daniel Simeoni, eds. *Beyond Descriptive Translation Studies. Investigations in homage to Gideon Toury*. Amsterdam and Philadelphia: Benjamins, 343-361.

Heilbron, Johan (1999): Toward a Sociology of Translation: Book Translations as a Cultural World-System. *European Journal of Social Theory*. 2(4):429-444.

Hermans, Theo (1997): Translation as Institution. *In*: Mary Snell-Hornby, Zuzana Jettmarová and Klaus Kaindl, eds. *Translation as Intercultural Communication*, Amsterdam: John Benjamins, 3-20.

Poupaud, Sandra (2008): Du réalisme magique à la récupération de la mémoire historique. La littérature traduite de l'espagnol. *In*: Gisèle Sapiro, ed. *Translatio. Le marché de la traduction en France à l'heure de la mondialisation*. Paris: CNRS Éditions, 231-256.

Pym, Anthony (1997): Review of *Bibliographie des traductions françaises*, Katrin Van Bragt *et al.*, *Target*. 9(2):365-369.

Pym, Anthony (2004): Propositions on Cross-Cultural Communication and Translation, *Target*. 16(1):1-28.

Pym, Anthony (2007): On History in Formal Conceptualizations of Translation. *Across Languages and Cultures*. 8(2):153-166.

Pym, Anthony and Gregorz Chrupala (2005): The Quantitative Analysis of Translation Flows in the Age of an International Language. *In*: Albert Branchadell and Lovell Margaret West, eds. *Less Translated Languages*. Amsterdam: John Benjamins, 27-38.

Renn, Joachim (2006): Indirect access: Complex settings of communication and the translation of governance. *In*: Arturo Parada and Oscar Diaz Fouces, eds, *Sociology of Translation*. Vigo: Servizio de Publicaciones do Universidade de Vigo, 193-210.

Sapiro, Gisèle, ed. (2008): *Translatio. Le marché de la traduction en France à l'heure de la mondialisation.* Paris: CNRS Éditions.

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

Venuti, Lawrence (1995): *The Translator's Invisibility. A History of Translation*. London: Routledge.