12-21-2006

Doing Translation History in EEBO and ECCO

Kristine J. Anderson
Purdue University, kja@purdue.edu

Follow this and additional works at: http://docs.lib.purdue.edu/lib_research

http://docs.lib.purdue.edu/lib_research/74

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Burgeoning interest in the history of translation in English is evident in the recent publication of such works as the ongoing 5-volume *Oxford History of Literary Translation into English* (Oxford University Press, 2005-), Fitzroy Dearborn’s *Encyclopedia of Literary Translation into English* (2000) and the *Routledge Encyclopedia of Translation Studies* (1998). But are such works conceived and produced with adequate rigor? While acknowledging his predecessors’ formidable contributions to the field, Anthony Pym nevertheless criticizes some of them for leaping too precipitously to conclusions not adequately rooted in evidence. Susan Bassnett’s claim in an article¹, for example, that translators became inferiorized as non-original writers during the rise of the printing press is based on only two examples. A wider examination of medieval translation would readily locate such cases several centuries prior to the printing press, Pym asserts. Other prominent translation theorists and historians also fall under his lash, from the 18th century when Amable Jourdain hypothesized a whole school of translators from one example, up to the present day when Pym’s suspicions fall heavily on the claims of prominent systems theorists like Gideon Toury that descriptive translation theories are scientific. Pym also criticizes the method underlying the compilation of the
abovemenioned encyclopedic works and guides whereby an editor draws up lists of the major foreign authors and genres and surveys the translations of each. “Obviously, in the absence of any actual cataloging of past translations, the guides can only show us around the apparently intuitive knowledge of who the most important authors are”(39).

To address what Pym perceives as flawed research practices that can carry us no farther than our own preconceptions, he has written *Method in Translation History*, basing it upon what he has learned from his own trials and errors. Meanwhile, massive corpora of English texts including translations have come on line in collections like *Early English Books Online* (EEBO) and *Eighteenth Century Collections Online* (ECCO), opening up the possibility of expanding research in the early modern history of translation into English beyond the realm of literary translation where the main body of existing scholarship has focused so far. Pym offers a research method that can be usefully applied to these databases as well as to the English Short Title Catalog (ESTC), which catalogs everything in EEBO and ECCO and more. After briefly introducing Pym’s method, I shall use it to formulate a search strategy in these three databases addressing a traditional problem in translation history. The demonstration will illustrate both capabilities and pitfalls in the search engines offered by these two databases while charting how an investigation based on Pym’s method might go.

Pym carefully lays the groundwork for his method by first offering a definition of translation history as “a set of discourses predicating the changes that have occurred or have actively been prevented in the field of translation. Its field includes actions and agents leading to translations . . . , the effects of translations . . . , theories about translation, and a long etcetera of causally related phenomena.”(7). The three stages of
translation history according to Pym are “Translation Archaeology,” “Historical Criticism,” and “Explanation.”

The stage of research I shall be concerned with in this article is “translation archaeology,” but Pym emphasizes the interrelatedness of all three stages. Translation archaeology, as he defines it, "is a set of discourses concerned with answering all or part of the complex question 'who translated what, how, where, when, for whom and with what effect. It can include anything from the compiling of catalogs to the carrying out of biographical research on translators. It . . . often involves complex detective work, great self-sacrifice and very real service to other areas of translation history"(5). Translation archaeology is the foundation of the other parts of the process but also, to some extent, dependent on them. Although “Historical Criticism” is not dealt with extensively in this work, he defines it as “the set of discourses that assess the way translations help or hinder progress,” but “rather than decide whether a translation is progressive for us here and now, properly historical criticism must determine the value of a past translator’s work in relation to the effects achieved in the past“(5). “Explanation” is “the part of Translation History that tries to say why archaeological artefacts occurred when and where they did and how they were related to change”(6). Thus an archaeological question may have an “explanation” as a hypothesis.

As with any other research project, the translation historian begins with a question to which she or he formulates a hypothetical answer. The question should be posed after careful consideration and be important to a wider audience than the researcher alone, although the researcher’s own interest is certainly a much needed motivator. The hypothesis formulated as potential answer is “a statement of what is likely to be
substantiated or falsified”(23). Since research is a project of discovery, a lot of jockeying back and forth may go on as successive hypotheses are proved wrong or the investigation opens up new vistas requiring the researcher to refine the question and produce new hypotheses. Some research questions can be answered easily, from the researcher’s own previously acquired knowledge. Other questions, though, concern a quantitative order requiring the grasp of “more world than the one that gave rise to them”(38), as Pym puts it, going on to define two basic approaches to such problems: reductive, involving the making of small lists from larger ones, and incremental, broadening one’s horizons by working outward from a small area. For the present, I shall focus upon the reductive, list making approach, since it has much in common with the traditional concerns of librarians with enumerative bibliographies as well as the most applicability to databases such as EEBO and ECCO.

Thus the first step in the process is to pose a research question and the second to hypothesize its answer. Since the historical researcher most often finds answers from written texts and there is never time to read everything, our third step, making a list of items that will form our object of study should insure, insofar as it is possible, that we do not waste time on irrelevancies. Pym dwells on the importance of lists as a basic tool for translation archaeology, for they can provide a context for testing hypotheses based on only one case. There are different kinds of lists, ranging from what Pym calls catalogs, purporting to include everything, and corpora, smaller lists distilled from catalogs with the researcher’s particular purposes in mind. The problem with lists, Pym points out, is that they are always based on other lists, and even the large catalogs that smaller corpora are drawn from are often found to conform to a preconceived ideology that may have
eliminated items that would have interested the researcher. To prove his point, Pym takes four translation bibliographies to task, finding their principles of organization particularly unhelpful, concluding that “All that really matters is that indexes allow the information to be reprocessed by the user, who can then extract corpora….Rather than follow fixed period or genre divisions, one should be able to select fields on the basis of actual dates and key-words in titles” (45). Pym goes on to provide a “list of “user’s desiderata,” as follows:

1. The bibliography or index “should be a database, no more, or less” (47).

2. Coverage should be as complete as possible without criteria of quality or quantity.

3. “The existence of possible lacunae should be indicated clearly, along with the exact procedures used to compile the catalog” (47-48).

Both EEBO and ECCO are corpora drawn from a catalog: the ESTC, which now contains MARC records for works printed in England, Scotland, and Ireland and of English books printed abroad from 1475-1800. How well do these three lists fit Pym’s essential criteria? In the first place, they are all databases that can be searched by keyword limited to field. The ESTC comes the closest to a comprehensive source; although it does not include everything, its scope is public knowledge. For example, the compilers freely admitted to excluding periodicals from the original STC. Many translations were published in journals and magazines, but our translation historian will have to go elsewhere to access them. This is acceptable to Pym, as long as the scholar is aware of the gaps and the necessity to look elsewhere. Since both EEBO and ECCO include bibliographical data from the ESTC, much of what Pym calls “paratexts” after
Gerard Genette – “all the textual material that introduces the text proper” including covers, author, title, blurb, table of contents, etc. (62) are also searchable. Some of this paratextual material falls into the category that we librarians today might call “metadata” and can serve to identify the text as a translation if this information cannot be found within the text itself. In the case of ESTC, for example, the information that a particular work is a translation can often be found in the notes field and nowhere else.

As a comprehensive catalog for the period it covers, the ESTC comes close to satisfying Pym’s desiderata. We now, however, have the ability to access not only the catalog, but the texts listed within it in EEBO and ECCO. The three databases differ greatly in the content offered, however. ESTC consists of MARC records only. In EEBO, we can choose to limit our search to MARC records for the full texts in the database, including those texts still available only in microform. This is helpful because the MARC records include the notes which are often the only place where a text is identified as a translation, and give the microfilm citation as well. The records in ECCO are abbreviated, however, and do not offer the option to limit to the MARC records at all. Since the record contains such paratextual material as notes, which sometimes is the only place a work is identified as a translation, it is important to be able to search this. As long as one has access to the ESTC, now free from the British Library, however, ECCO can probably get along without it.

The real advantage to EEBO and ECCO, of course, is not the ability to search MARC records but to access entire full texts. EEBO offers photographic images of nearly everything listed in the ESTC from 1475-1700 plus the Thomason Tracts, 1641-1700. The images are not searchable by keyword, however, unless they are part of the
TCP (Text Creation Project) subset that has been keyboarded in and tagged according to TEI (Text Encoding Initiative) standards. Other tools are currently being developed to use with TCP texts, like the “Virtual Modernization Tool” developed by Professor Martin Mueller at Northwestern, which enables users to find instances of particular keywords in all their variant spellings. Since we want our corpus to be drawn from as large a universe as possible however, we do not want to limit ourselves to TCP texts at this stage, however.

All the texts in ECCO have been OCR’ed, on the other hand, and are therefore fully keyword searchable. Searches can be limited to full text, keyword, Author, Title, Front Matter, Main Text, Back-of-Book indexes, Publisher, or Place of Publication. The advanced search screen only offers three Boolean operators: AND, OR, and NOT, but a glance at Search Tips shows that Proximity operators are also available: “W” which will “find documents containing the specified words in the specified order within the number of words you indicate” and “N” which “locates documents containing the works you specify within the number of words you specify, but the words can be in any order.” In the main it handles variant spellings with the “fuzzy search” option which applies algorithms guessing at what you possibly meant, but a version of the Virtual Modernization Tool to work with ECCO TCP texts is also being prepared.

In all three databases -- ESTC, EEBO, and ECCO the researcher makes lists by formulating a search strategy, choosing keywords and limiting them to the fields which seem most promising. Pym goes into a lot of detail about the process of forming “working definitions” which means “developing and applying some basic concepts concerning the object to be studied” (55) and involves, among other procedures, deciding
what is and what is not a translation, or member of a specific genre. Working definitions may evolve during the course of actual empirical work and not become stable until near the end of the project. It is nevertheless a crucial step in extracting a much smaller corpus from the catalog in question and will be used to determine keywords and to formulate a search strategy.

Both EEBO and ECCO provide a gold mine for the translation historian. The years covered by EEBO, 1475-1699 witnessed the rise of vernacular English to the status of a national language and the effects of the invention of the printing press, both on the standardization of English and the availability of literary works in critical editions from other vernacular languages. William Caxton himself, an early printer in England, translated many of the works he printed. Other writers took advantage of the proliferation of printed critical editions to hone their skills by translating the Greek and Roman Classics, culminating in versions of the likes of Homer and Ovid from the pens of now canonical figures like John Dryden and Alexander Pope in the 18th Century, the period covered by ECCO. Yet, according to Stuart Gillespie in *The History of Literary Translation in English* (3:135-140), there were even more translations from the French than from any other language during this period, since it often also served as an intermediary language for Oriental and Middle Eastern works. Next in quantitative importance as sources for English translation are Italian and Spanish, though full-length works of translation from these number in the 100’s rather than the thousands.

These two databases clearly offer enormous potential to the translation historian working with the languages mentioned above. For my first attempt at applying Pym’s method, however, I chose to work with a language that results in a much smaller list –
Danish. It is a traditional archeological translation history problem, asking simply, “what was translated from a particular language, in this case Danish, into another particular language --English-- during a particular period – in this case the early years of printing, from 1475-1800, the years covered by the ESTC, EEBO and ECCO?

Gillespie does not even include Danish in the volume of the *Oxford History* covering 1660-1790, reserving its brief discussion for the following volume, since the British fascination with the far North is really largely a 19th century phenomenon. Nevertheless some works were translated from Danish, as the work of Elias Bredsdorff shows.

Bredsdorff can be regarded as an archaeological pioneer in Pym’s sense, having produced a comprehensive bibliography of *Danish Literature in English Translation* in 1950. Using this work to locate translations from the early modern period, however, encounters some of the same problems Pym encountered in the four bibliographies he criticizes. Bredsdorff’s goal was simply to find all literary translations from Danish to English up to the time he had to stop writing and submit the book to a publisher. The first question we must ask is “how complete is it”? For the word “literary” implies that a certain ideology is being applied. Bredsdorff, in fact is forthright about his definition of Danish literature in his preface, broadening the concept beyond fiction, poetry and drama, to include Danish theologians, historians, critics and philosophers as well as “a number of prominent scientists, explorers and scholars before 1800” (7-8). Another characteristic of his definition is that it covers Danish writers, not just the Danish language. A greater problem, however, is Bredsdorff’s arrangement into three sections:

I. Books (and some articles) about Danish literature
II. Anthologies of Danish literature

III. English translations from Danish authors, and books and essays on individual Danish authors.

Since it is not a database, limiting the translations listed to those done during a particular time period has to be done painstakingly, examining each entry for 1) author’s dates and 2) date of translation, listing only those translations done before 1799.

Like Gillespie, Bredsdorff found very few translations from Danish to English that were made between 1475 and 1800, offering as a possible explanation the fact that since Latin was the “lingua franca” of the time, an author who wanted to attain a wider audience for his works would simply write it in Latin. This was the certainly the case for the most famous Danish work of the 18th century, Ludwig Holberg’s Niels Klim, and indeed, many of Bredsdorff’s Danish authors were translated from Latin, not Danish.

Since the present exercise is partly a test of how well these databases can be made to do, a hypothetical answer to my question might therefore be, “all the translations from Danish into English done before 1800 that Bredsdorff found, plus whatever may have been located since around 1950, or that Bredsdorff may have excluded as not being literary” although he did admit to a very loose interpretation of “literary.”

To begin, I have chosen to search only the MARC records in the ESTC under “translat*” truncated to cover all instances of “translation,” “translator” or “translated.” and “Danish” in the records. In this case my definition of translation includes anything that is identified as such somewhere in the bibliographic record. (Of course if it does not identify itself as such, it will not come up in the search). Other possible terms to use could include “imitation” or “adapt*,” or other words indicating it was not a word-for-
word translation. Another term frequently used for translations into English was “Englished” or “made English.” Also, during the time period covered by EEBO, Danish was sometimes spelled “Dence.” The latter spelling, however, had a number of other meanings as well, as can be seen from a brief search on the Virtual Modernization Tool.

For the purposes of making a list, however, I have determined that the information identifying a particular book-length work as a translation from Danish into English is going to be either in the title or elsewhere in the record if anywhere. Thus if I want to search only “paratexts” to determine the number of translations from Danish to English during this time period, the ESTC itself is the ideal place to begin. I did two searches here, one searching for my terms only in title fields, the other simply as keyword. The first search brought up 24 entries, the second 50, a number which included other editions of the same work. This search found 9 items from the period covered by EEBO, where the same search in EEBO itself found only 7.

The 41 remaining entries are from the 18th century, or the period covered by ECCO. A Boolean title search in ECCO under the same terms, translat* and Danish, brings up only 16. This discrepancy can be explained by several factors, including those already mentioned plus the fact that ECCO purposely excludes texts found in other databases such as the Evans Early American Imprints collection.

Going back to our original list of 50 works found in ESTC and comparing it with those in Bredsdorff actually published in English between 1475 and 1799, we find that there are items on each list that are not on the other. ESTC includes the following not listed by Bredsdorff:
Horrebow, Niels. The natural history of Iceland . . . Translated from the Danish original of Mr. N. Horrebow. And illustrated with a new general map of the island. London: printed for A. Linde, D. Wilson, T. Durham, G. Keith, P. Davey [and 4 others]: 1758.

Løvenørn, Poul. Sailing directions for the Kattegat to be used with the new improved chart containing descriptions of the lights, soundings, grounds and shoals, . . . published from the Royal Danish marine archives… by P. de Løvenørn. . . translated from the Danish original by Frederick Schneider. . . Copenhagen: printed for the Archives by J.F. Schultz; 1800.

Schlegel, J.F.W. An examination of the sentence in the case of the Swedish Convoy, pronounced in the High Court of Admiralty of England, on the eleventh of June, 1799; together with a previous historical sketch of the European, and in particular the English system of capture. Translated from the Danish, as written by Professor Schlegel, London: printed by W. Wilson, 1800.

Struensee, Johann Friedrich greve. The trial of Count Struensee, late Prime Minister to the King of Denmark, before the Royal Commission of Inquisition, at Copenhagen. Translated from the Danish and German Originals. London: printed for the translator; sold by J. Whitaker, J. Fox; T. Waters; 1776.

Ziegenbalg, Bartholomaeus and Pluntscho, Heinrich. Propagation of the Gospel in the east: being a farther account of the success of the Danish missionaries, sent to the East-Indies, for the conversion of the heathens in Malabar. Extracted from the
letters of the said missionaries, and brought down to the beginning of the year MDCCXIII.

These five items, all from the latter part of the 18th century, reveal a variety of recorded “non-literary” works by Danes apparently thought at the time to be of interest to English monolinguals.

At this point, however, it seems appropriate to explore the full-text capabilities of EEBO and ECCO to ascertain whether our short list can be expanded. The same search, searching “front matter,” in ECCO, one of its options, brings up 233 items. The problem with the latter search, though, is that not only does “front matter” cover quite a bit of territory, our Boolean AND search requires only that these two keywords be in the same book. One of the capabilities in ECCO, the ability to search for keywords within each text, proves helpful here, revealing that the word “translation” is the one that comes up most frequently in each text, and hardly ever anywhere near “Danish” which may only occur once in each book. At this point it becomes crucial that the scholar has read all the “search tips,” for it is clear that a proximity search is called for, and it is not readily apparent from the main search screen that these are available. At this point I do the search again, trying a full text search: “Danish n10 translat*,” specifying that I want the word Danish to be within 10 spaces of “translat*”, but the order does not matter. This reduces the number considerably, to a mere 44, not at all a difficult number to go through. I can reduce this number even more by limiting the search to front matter. This brings the list down to 10. One interesting text is found when repeating this search using Danish and English* to cover those instances when “English” or “made English” may have been the synonyms for “translated”:
Schneider, Christian Frederik. *Danish grammar adapted to the use of Englishmen... by Fredk Schneider...* Copenhagen, [1799]. 332pp

a work which contains a section of parallel English and Danish texts.

Of course “false hits” are bound to come up, necessitating that each title be examined carefully before being included on the list. A look at some of these can be instructive: *Tales, translated from the Persian of Inatulla of Delhi,* which comes up on both ESTC and EEBO turns out to have a variant title, “Bahar-i danish” which undoubtedly means something different in Persian than in English! Another apparent false hit, however, *Odes of Anacreon, translated into English Verse, with notes,* by Thomas Moore, when subjected to full-text searching in ECCO turns out to have a Latin poem by a Danish poet in a footnote with Moore’s reference to “the Danish poets collected by Rostgaard.” This does suggest that a keyword search through ECCO full text might well reward us on another level in leading us to other texts that have not been found through our other methods of searching.

As we saw in the case of *Anacreon* there is always the possibility that a full-text search could bring up the citation of a translated Danish text that we have not found before. At this point, we decide to return to EEBO to see what a full text search limited to the TCP portion (the only portion that can be searched full-text) has to offer. EEBO has now incorporated the Virtual Modernization Tool into its search interface, supplying a checkable box to indicate that you want to search variant spellings. When this box is checked for the search *Danish nr10 translat*®, EEBO brings up only one variation of “danish”: *danysh* (but not *Dence*, found when searching in the Virtual Modernization Tool itself) and nothing for *translat*®. The search produces 18 hits in 112 records. When
the “Context of Matches” link is clicked, showing the lines in which the search terms can be found, only one book, Animadversions on the Pretended Account of Danmark in the Year 1692 holds any promise. Unfortunately for our present purpose, however, the translations listed are only from other languages into Danish.

Bredsdorff’s list includes Danish authors who wrote in Latin and were translated from that language. Our original question specified English translations from the Danish language, so this information is not relevant, unless, of course, we decide to revise the question. Were we really interested only in translations from the Danish language, or were we interested in the reception of Danish writers in England? At this point we might want to re-evaluate our purpose in asking the question. The apparent need met by the translation of Latin works by Danish authors may be significant. An author search for each one of them listed in Bredsdorff in EEBO and ECCO finds them there, even though they did not come up under our original search. An early translation of a work by Tycho Brahe even comes up in EEBO, although Bredsdorff did not find anything by him translated before the 19th century. Devising a search for other translated items by Danish authors who wrote in Latin that may have been missed by Bredsdorff presents a formidable challenge to our general practice of Boolean keyword searching, however. Authors are not always identified by their nationality, neither in the MARC record nor in the text, and Boolean logic cannot be made to distinguish the direction of the translation. For example, our simple keyword “Danish and translat*” search in ESTC unexpectedly brings up a number of works translated from Danish to English, thereby proposing yet another direction our investigations could take us in.
Concluding Observations.

My application of Pym’s method in the search for Danish to English translations in EEBO, ECCO and ESTC has thus encountered a number of pitfalls but holds a great deal of promise as well. A very simple, common archaeological problem was chosen to test how Pym’s method could be used in designing searches to extract items of a defined nature from such full-text databases of primary works as EEBO and ECCO. The field chosen, translations from Danish to English during the period covered by these two databases predictably yielded a short list suitable to a first scouting foray. The much longer lists that could be derived from an investigation of French to English, or Latin to English would have more statistical validity and could be used, for example, to plot frequency curves, another, later step in Pym’s method, to be applied after the lists are successfully compiled.

Although Pym speaks with great longing for databases capable of being searched by keyword in place of the hard-copy bibliographies he was forced to use for lack of anything better at the time he was working, keyword searching is not a panacea, either. Boolean syntax is not very specific, unable, for example, to distinguish the target language from the source language. And its successful use depends upon a number of other factors as well, such as the depth of the researcher’s historical knowledge, especially of the terminology used during the period being investigated, the level of subject indexing present in the records, and the standards according to which the electronic texts have been tagged. Despite these caveats, however, massive electronic corpora such as EEBO and ECCO present far more possibilities than anything existing prior to their existence.
In summary I offer the following observations gleaned during my experiment:

1. Searching the records in ESTC will be the best initial approach as it is likely to contain newly found documents not yet in EEBO or ECCO and has the most complete paratextual material.

2. A great deal more work would have to be done before the researcher could rest assured that all translations in the database had been found. The initial list is only a crude approximation of the number of translations published. The researcher cannot stop at EEBO and ECCO, but must venture forth into other databases and bibliographies to find texts that are clearly out of scope for EEBO and ECCO. As Pym later admits, however, a complete list is not always necessary: some hypotheses will require only a few contradictions to be dismissed.

3. The search interfaces offered up front by vendors differ greatly from database to database and do not display all possible search options. Some of these are hidden away behind “Search Tips” or “Help” screens. The Chadwyck-Healey interface used by EEBO provides a number of blank spaces in which to write your keywords, but no indication that they will all be connected with Boolean “ands.” You can in fact do Boolean searching in EEBO, but you must select a blank space, whether it is next to “title”, “keyword,” “author” or “subject” and write it all into that one space. The EEBO “advanced search” simply gives you a wider variety of blank spaces to write things in. The ECCO advanced search interface gives you unlabeled blanks with drop-down menus enabling you to select your own field limiter, as well as a drop-down box offering the three major Boolean search terms. Fuzzy search menus are also offered as drop-down boxes, enabling
you to gauge the level of fuzziness to your preference; none, low, medium, or high. Even here, however, you must still go to the Search Tips to learn how to formulate proximity searches.

In conclusion, EEBO and ECCO do offer an excellent resource to researchers of translation history. The production of smaller corpora for study of a particular aspect of the phenomenon is not automatic, though, but one that requires a great deal of work, thought and ingenuity in formulating search strategies as well as a thorough knowledge of older forms of the English language and its terminology.

Works Cited


3 Pym’s definition actually expands on Genette’s – as he points out, “...Genette strangely fails to look at the translator’s name or other translational signs, which are clearly parts of paratexts”(62).