

# STUDY ON TERMINOLOGICAL RESEARCH FOR THE TRANSLATION OF SPECIALIZED TEXTS

BLANCA HERNÁNDEZ PARDO<sup>1</sup>

Fecha de recepción: octubre de 2016

Fecha de aceptación y versión definitiva: noviembre de 2016

*RESUMEN:* A menudo nos fijamos en aquellas soluciones que dan los alumnos que comienzan a traducir textos especializados cuando, en realidad, no saben cuál es el equivalente en la lengua materna por no estar especializados en el campo sobre el que habla el texto. Gracias a este estudio se pretende observar las pautas, los métodos y los pasos que siguen los alumnos para encontrar aquel equivalente que creen correcto para la traducción.

*PALABRAS CLAVE:* traducción especializada, investigación terminológica, análisis terminológico, traducción económica.

*ABSTRACT:* It is often the case that our attention is drawn to the solutions provided by our students when they start to translate specialized texts, even though they do not necessarily know what the correct equivalent in the target language is because of their lack of knowledge of the specialized field. This study will analyze the standards, methods and steps that Specialized Translation students use in order to find the equivalent that they consider is the most appropriate for the translation in question.

*KEY WORDS:* specialized translation, terminology research, terminology analysis, economic translation.

## 1. INTRODUCTION

The present article describes the results obtained from a quantitative study carried out in 2012 as part of an academic Degree Dissertation on lexical connections between English and Spanish and developed by specialized Translation students from the *Universidad Pontificia Comillas* [Comillas Pontifical University] during the academic year 2011-2012.

The research project is mainly focused on three disciplines: Translation studies, Terminology and Documentation, with the principal goal being to identify the main problems faced by Translation and Interpreting students

---

<sup>1</sup> Profesora y doctoranda Universidad Pontificia Comillas. Correo electrónico: bhpardo@comillas.edu.

during the process of searching for specialized lexicon. This would allow us to study a systematic protocol that could help both students and professionals search for, and select, certain terminology within specialized contexts.

It is important to highlight that this study is based on a previous article published by Prof. Pilar Úcar and Prof. M<sup>a</sup> Luisa Romana (2011), who proposed in this regard the following flowchart that could be taken as a procedural model to search for and select specialized terminology.

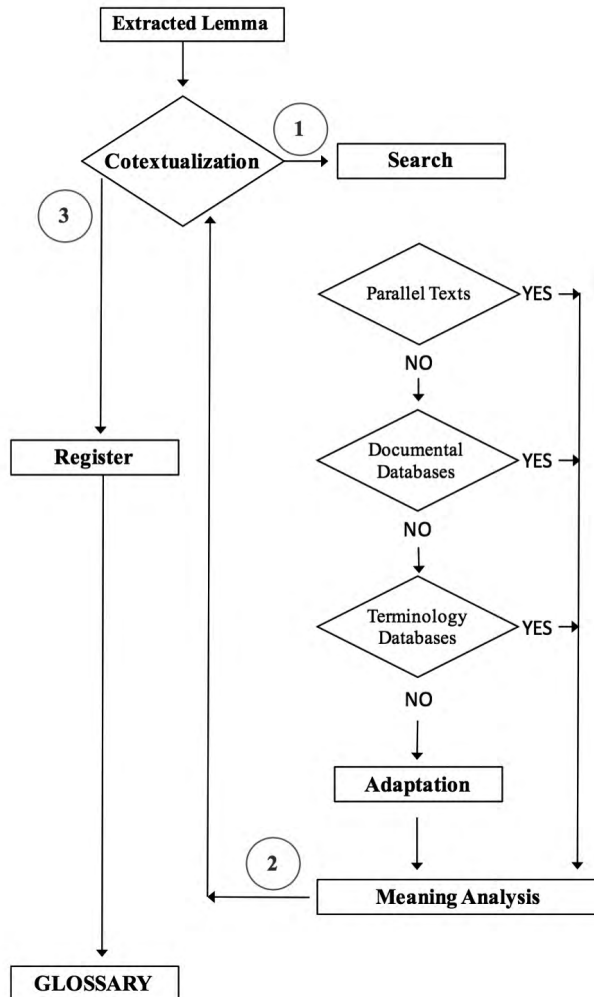


Fig. 1. Flowchart (Romana & Úcar, 2011)

With this proposal, both students and professionals can follow a series of systematized steps in order to find the most adequate terminology for the specialized field the translator is working on, i.e. in a given context.

Throughout the study the viability and utility of the aforementioned flowchart were tested, together with the possibility of ranking the groups of documentation sources proposed by Romana and Úcar (2011).

## 2. WORD AND TERM ARE POLES APART

The starting problem demonstrated by Translation students and which has permeated the whole study was the way in which they carried out the process of search and selection of terminology when translating a text. It was seen that the students did not know or understand the necessary terminological process to translate properly. At this point, the first question arose: How should one train a student to search for and choose the most adequate terminology?

Until that moment, no other protocols that offered systematized steps for the terminological research were known. This proposal seemed to be very practical and useful, if it could be confirmed that it would work for translators.

Every good translator knows that a series of tools are needed to translate a document and, throughout the learning of such an art, he or she realizes that some of these tools, like dictionaries, can help. Nevertheless, the translator also understands that it is not possible to translate with this support alone.

The reader of this article may have heard such expressions as, «Translating is very easy with a good dictionary in your hands...», «Don't you think in the future machines will be in charge of translating? After all, Google Translate is getting better all the time...», or «I have also done some translation projects; if you count on a good bilingual dictionary, it's dead easy!»

According to Cabré (1999a), terms must be studied *in vivo*, i.e. it is necessary to analyze how they act and interact with other words and appear in their natural environment: the specialized documents.

It is easy to take a dictionary and start translating. But the problem starts from the moment the translator does not have enough information to establish the most adequate and accurate term within a specific field and context.

*Word* and *term* are different concepts, which can be appreciated when differentiating *general language* and *specialized language*. The main difference

between both is the type of texts in which they appear. *Terminological units*, which belong to the specialized language and, because of this, will be found in specialized texts, present specific features (whereas in the text, which will be specialized, it is the receptor, who will be an expert in the area of expertise, or the proper use of the term, which will be limited to the field to which it belongs).

Due to this, Cabré (1999a) defines the concept of *term* as a word associated to a considerable semantic, syntactic and pragmatic information; thus, the term will have a certain value depending on the context and use.

In this way, the answer is «Yes, the translator must need to take into account more than an equivalent from a dictionary in order to make the right choice». We need to define the context in which the term appears so that we can see how it implies both in the source language and the target language, and so that we can find the most adequate equivalent in the target language.

### 3. DOCUMENTATION SOURCES

Romana and Úcar (2011, p. 314) rank the sources used in the documentation process regarding the degree of their relationship to the subject: «[...] it is] an ordinal coding system which assigns relevance according to the translator's initial intuitive instinct as regards reliability, from the point of view of a translator to each resource». Nevertheless, in present research, the procedure undertaken was not the same. In this experiment, the documentation sources were not hierarchized, as they are nominal variables. This constituted a methodological obstacle and should be corrected in the event of continuing this study in the future.

Regarding the conclusions from Romana and Úcar (2011, p. 321):

[...] The most important factor for a correct search appears to be reference, given the higher correlation between acceptability and reference and this first correlation may be interpreted as a possible indicator of reliability for each source type. In descending order, the most successful search lies in a) parallel texts (where the reference is an exact equivalence); b) documentary evidence (where the reference has a general equivalence); and c) databases (where all possible references are given).

## 4. ORGANIZATION OF THE EXPERIMENT

In relation to the procedure of the study described in this article, the steps mentioned by Romana and Úcar were followed in order to conduct a successful experiment offering relevant outcomes regarding documentation sources used by a translator when working on a specialized text: *cotextualization*, search, meaning analysis and register verification<sup>2</sup>.

### 4.1. COTEXTUALIZATION

In this phase, a representative series of *cotexts* from the original text where the key term appeared were collected. It is worth pointing out that when using texts from European institutions relevant for the translation circle, it is possible to obtain its official translation in Spanish.

### 4.2. SEARCH

A search within the group of documentation sources of the terms from the original text was carried out.

1. Bilingual texts
2. Documentation databases
3. Terminology and lexicographic databases.

### 4.3. MEANING ANALYSIS

It was verified whether the obtained equivalence had the same meaning as the term found in the translated texts from the same institution that published the original documents.

### 4.4. REGISTER VERIFICATION

Once a translation option (or several options) was obtained, it was necessary to carry out the «Register Verification» (RV): verification for the real

---

<sup>2</sup> Due to a lack of space, it is not possible to fully develop the procedural steps; a more detailed description can be consulted in Romana and Úcar (2011).

use of the expression or term within the specialized field. To that end, a new search of the Spanish units found was carried out, and the results were ordered according to the following criteria:

1. *Specialized sources.*
2. *Use frequency among them.*

The rv process was done with a general search engine on Internet (Google), introducing the found term with Boolean operators (quotation marks) and writing the necessary data for the register verification.

## 5. EXPERIMENT DESCRIPTION

In order to study the aforementioned ideas and suggestions, an experiment with students in their third year of the Translation and Interpreting degree program was carried out. A calendar was designed with a series of sessions in which students had to collect certain information while doing terminology searches from specialized macroeconomics texts. As such, students were noting down different data about their searches on a previously elaborated file card, as if they had found the equivalent in Spanish of the original term in English, the type and name of sources they had used to find such correspondent term, information about the rv (quantity and reliability) and comments of interest regarding the searches.

The information provided would later be useful to decide what the most reliable group of sources are when translating, and the issues that a translator should pay the most attention to in order to avoid drafting a poor translation.

Regarding the group of documentation sources on which the study would be based, the sources were composed of several webpages considered useful and they were included in one of the three type sources previously discussed.

A file card that an individual student worked on, as well as the information that was collected, is shown below, together with a brief explanation of some of the key fields (for a better understanding of the study conclusions):

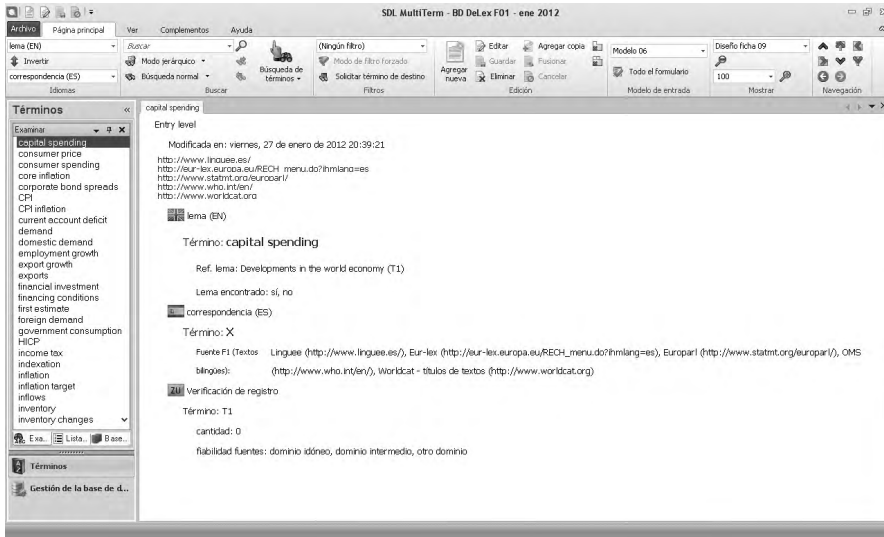


Fig. 2. Multiterm capture.

## 5.1. FIELDS OF THE TERMINOLOGY SEARCHES

- 1) Term reference: a picklist (multiple choice) field where students had to select one option from a list of the different texts with which the different terminology searches would be undertaken.
- 2) Comments: for the students to write a note, a warning or an incident.
- 3) Author: the name of the student.
- 4) Source: a picklist field where students chose the sources for which they could find the equivalent in Spanish of the original term in English (among the group of sources previously assigned; for example, if the student had to look for the term in the bilingual texts group, inside that picklist he would find EUR-lex, Europarl, Worldcat, etc.).
- 5) Found target term: a picklist field where students could only choose between «yes» or «no» depending on whether they had found a satisfactory (from their point of view) equivalent; there was a possibility that students would not be able to find the equivalent before the deadline.
- 6) Register verification: there are two subfields:
  - a) Quantity: the number of entries found in the registry. To verify whether the found term is frequently used in the economics field,

a test of results is obtained with Google: students collect the number of entries (on the economics area) that Google shows when searching for the proposed term in Spanish with Boolean operators. This procedure is useful in a situation where, for example, after doing a search with the different documentation sources, two lexical equivalents in Spanish for the same entry are found. To check which one is more widely used, this procedure can be carried out.

- b) Source reliability: a picklist in which the answer is given, depending on the reliability of the source where the Spanish equivalence was found. Options for this picklist are the following:
  - I. Exact domain.
  - II. Intermediate domain.
  - I. Other domains.

Afterwards, all this information was organized in databases and codified. As can be seen, many of the aforementioned fields offer non-quantitative but nominal information (for example, *source reliability*). In order to clearly study all of this data, it is necessary for each piece of information to have a numerical value. Thus, when talking about quantities, the variable is numerical, but when it contains a concept, the variable is nominal (for example, the searched term or the correspondent found term). The third possibility is the ordinal variables, which represent a hierarchy; that would be the case of the source type. All these values are presented in the table below:

VARIABLE	TYPE OF VARIABLE
Term [EN]	Nominal
Original text	Nominal
Author	Nominal
Source type	Ordinal
Found term (y/n)	Ordinal
Correspondence [ES]	Nominal
Quantity	Scalar

## 5.2. OTHER SEARCHING DATA

In addition to the data provided by the students, in each file card the following information was added:

- 1) *Searched translation*: term used in the professional translation from the OT, the official published text. Nominal variable.



- 2) *Outcome accuracy* (regarding the rv): distance between the proposed equivalent by the student and the version found in the rv. Ordered as follows: 0 = completely different, 1 = partially different, 2 = exactly same. Ordinal variable.
- 3) *Translation mark*: mark given by the teacher. Scalar variable.

## 6. ANALYSIS OF THE RESULTS

After all the information was collected, organized and codified, the analysis and study of the experiment, together with the corresponding relevant conclusions, was initiated.

A correlation is defined as the measure of the dependency existing among random variables. The following table of correlations was the one obtained from the information collected:

		<b>Trans. mark</b>	<b>Source type</b>	<b>Found term</b>	<b>Search accuracy</b>	<b>Quantity RV</b>	<b>Reliability RV</b>
<b>Trans. mark</b>	Correlation (Pearson)	1	<b>-0.492(**)</b>	0.046	-0.085	<b>-0.236(**)</b>	0.021
	Mean. (bilateral)		0	0.553	0,267	0.005	0.795
	N	172	172	172	172	142	158
<b>Source type</b>	Correlation	-0.492(**)	1	<b>-0.228(**)</b>	0.066	<b>0.142(*)</b>	-0.118
	Mean.	0		0	0.312	0.045	0.081
	N	172	234	234	234	200	219
<b>Found term</b>	Correlation	0.046	-0.228(**)	1	<b>0.147(*)</b>	0.021	<b>0.373(**)</b>
	Mean.	0.553	0		0.025	0.771	0
	N	172	234	234	234	200	219
<b>Search accuracy</b>	Correlation	-0.085	0.066	0.147(*)	1	-0.014	<b>0.272(**)</b>
	Mean.	0,267	0,312	0.025		0.849	0
	N	172	234	234	234	200	219
<b>Quantity RV</b>	Correlation	-0.236(**)	0.142(*)	0.021	-0.014	1	<b>-0.180(*)</b>
	Mean.	0,005	0,045	0.771	0.849		0.012
	N	142	200	200	200	200	193
<b>Reliability RV</b>	Correlation	0.021	-0.118	0.373(**)	0.272(**)	-0.180(*)	1
	Mean.	0,795	0,081	0	0	0.012	
	N	158	219	219	219	193	219

As a first comment regarding these data, it is observed that the type of source is highly related to the mark given for the translation ( $corr = -0.492 / p = 0.000$ ), i.e. in general terms, the more specialized sources, the higher the quality the target text will be. The negative sign confirms that the quality of the target term is indeed related to the source type and, moreover, to the order of the documentation sources previously foreseen in the hypothesis (regarding the reliability of use for the specialized translation): the texts whose terminology was searched for within bilingual texts have obtained a better mark than those searched for within documentation databases, and these last have a better mark than the ones searched within the terminology databases. On the other hand, the quality of the text is also related to the quantity obtained in the *rv* ( $corr = -0.236 / p = 0.005$ ), which proves the utility of this method.

A non-causal relationship was also found between the source type and the fact of having found or not the equivalent term ( $corr = 0.228 / p = 0.000$ ). Thus, it will be more likely to find terminological units in textual databases, like EUR-lex, DTSearch or Translation Memories (TM) *ad hoc*.

Regarding the information on the *rv*, it is important to highlight that the reliability information is related to the accuracy of the search ( $corr = 0.272 / p = 0.000$ ). This reinforces the quality of the analyzed documentation protocol, in this case, in relation to the second criterion of the register verification: the source reliability. Nevertheless, it must be said that it would be surprising not to find any correlation, given that one expects to achieve a more reliable outcome when the source itself is more reliable. The procedure is additionally confirmed due to the relationship found between the two subcomponents of the *rv*, quantity and reliability ( $corr = 0.180 / p = 0.012$ ). However, taking into account the low correlation, this result indicates the high level of specialization of the selected terms. On the other hand, the quantity of the *rv* seems to be related to the source type too, with the smallest correlation ( $corr = 0.142 / p = 0.045$ ). At the same time, this correlation is one of the most informative results, since fewer results are obtained on Internet when the source is more specialized. Again, this can be seen as an indicator of the *technolectal* specialization. Nevertheless, in order to obtain definitive conclusions, it would be necessary to know also the balance or imbalance among the quantity of the three source types. Finally, the correlation obtained between the reliability of the *rv* and the fact of having found an equivalent term ( $corr = 0.373 / p = 0.000$ ) also indicates that, if we found the equivalent term in any source, regardless of the quality of such source, it will be more likely to get a higher *rv*: the procedure of directly looking for a Spanish formulation, common among experienced translators, omitting the

bilingual research, will be much more successful if we try to find the original (English) term first.

## 6. CONCLUSIONS

Taking into consideration the results of the present study for the verification of Romana and Úcar's study (2011), the reinsertion of a new element in the general schema is recommended: A filter, a procedure that could be named «equivalence adaptation» and that, at some point prior to the rv is analyzed, enables to confirm that the proposed translation has all the information contained in the term.

In this way, if we decided to conduct a research on the term «capital gains tax», we should first take into account that inside this terminological unit there is another one, which is «capital gain» and which may lead to a change regarding the main concept. After looking for the term in Google, we can see a result of 428,000 pages (a high quantity), and that the main results come from Governments in countries such as the United Kingdom, Australia and the United States. Thus, we should focus on the concept within the macroeconomics field and look at the meaning of the term in use, i.e. its context. Then, we should move through the first group of sources, the documentation sources (Linguee, EUR-lex...) and several results would be observed for the same original term: for the terminological unit «*impuesto sobre el incremento de capital*», Google provides with 9 results; for «*impuesto sobre ganancias de capital*», there are 12,2000 results; for «*impuesto sobre plusvalías*», we find 59,200 webpages; and for «*impuesto de plusvalía*», the search webpage shows 125,000 results. According to these results, it is possible to state that the last two options will be the most reliable ones when translating «capital gains tax» into Spanish. Moreover, if we analyze the type of results provided by Google with respect to «*impuesto de plusvalía*», it is seen that the main ones belong to governments, councils, banks and official institution webpages, which confirms the reliability for this translation.

However, it should be noted that, when doing the register verification with Google, a problem was found: this search engine does not always respect the use of Boolean operators to restrict the searches. In order to obtain this desired effect when checking the information found, we would recommend using another search engine called Exalead, a tool which enables the translator to look for terms using filters that restrict even further the possible results. In this way, it is possible to add different filters once the results

have been provided (in the left side of the webpage, there are options to be chosen, such as language, related terms, type of webpage, category or year). Additionally, it is possible to do initial advanced searches (exact words or sentences, optional terms, order of words within a sentence or expression, proximity...), to indicate if we want to obtain published results on Internet up to or from one specific date, etc. This search engine has been operating in Spain since 2007 and has an index of over 8 million webpages.

To sum up, the main and most useful conclusions from the study are the following:

- 1) The translation work will be improved if terms are searched for according to the following source order:
  - a) Bilingual text databases (EUR-lex, DTSearch, translation memories).
  - b) Documentation databases (corpus, monolingual webpages).
  - c) Terminology databases (IATE, glossaries and other bilingual terminology publications).

What is more, from the terminological point of view, it will be more likely to find a successful solution if this searching order is respected.

- 2) The translation work will be improved if target terms pass through a register verification or rv: a monolingual search on Internet, after which the quantity of the results obtained and the reliability of the sources are written down.
- 3) Finally, the more specialized a text is, the fewer the number of results the translator will obtain on Internet for the desired terms, which makes him pay even more attention to the type of source.

In case of failing to solve the reliability of the translation through technology, the translator must resort to a personal source that can help him or her to decide whether it is possible or not to use this translation method when passing from simple to complex units, either because the meaning would modify the idea or simply because it is not used in the specialized field.

In any case, it is confirmed that there is a hierarchy when establishing what type of sources are more reliable within the context of translating texts in the field of macroeconomics. Such progress may help both current professionals when translating texts, and Translation and Interpreting students who are still learning and acquiring the linguistic and conceptual skills in order to become professional translators able to undertake quality documentation searches. Even then, it is necessary to keep on researching within the field of documentation and terminology searches so as to contribute with relevant information for the translation profession and translation studies.

## REFERENCES

- ALCARAZ VARÓ, E. (2000). *El inglés profesional y académico*. Madrid, España: Alianza.
- CABRÉ, M. T. (1993). *La terminología: Teoría, metodología, aplicaciones*. Barcelona, España: Antártida/Empúries.
- (1999a). *La terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos*. Barcelona, España: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada.
- (1999b). Hacia una teoría comunicativa de la terminología: Aspectos metodológicos. *Revista Argentina de Lingüística* 15, 24-38.
- DE BESSÉ, B. (1997). Terminological Definitions. *Basic Aspects of Terminology Management*, 63-74.
- ESTOPÁ, R. (1999). Eficiencia en la extracción automática de terminología. *Perspectives: Studies in Traductology*, 7(2), 277-286.
- LORENTE, M. y BEVILACQUA C. (2000). Los verbos en las aplicaciones terminográficas. *Actas del VII Simposio Iberoamericano de Terminología RITerm 2000*. Lisboa, Portugal: ILTEC.
- MICROSOFT SUPPORT OFFICE (17 de octubre 2016). Crear un gráfico de principio a fin [Publicación oficial de apoyo teórico y práctico]. Recuperado de <https://support.office.com/es-es/article/Crear-un-gr%C3%A1fico-de-principio-a-fin-0baf399e-dd61-4e18-8a73-b3fd5d5680c2>.
- (17 de octubre 2016). Usar las Herramientas para análisis para realizar análisis de datos complejos [Publicación oficial de apoyo teórico y práctico]. Recuperado de <https://support.office.com/es-es/article/Usar-las-Herramientas-para-an%C3%A1lisis-para-realizar-an%C3%A1lisis-de-datos-complejos-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6>.
- ORELLANA, L. (2001). Estadística descriptiva. *Material didáctico*. Buenos Aires: Departamento de Matemática de la Universidad de Buenos Aires.
- PÉREZ, C. (2000). *Explotación de los corpora textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento* (Tesis doctoral inédita). Universidad de Málaga, Málaga, España.
- ROMANA, M. L. (2009). *La sintaxis en la traducción económica (inglés-español)* (Tesis doctoral inédita). Universidad Pontificia de Comillas, Madrid, España.
- & Úcar, P. (2011). «Al final dejé esto: las decisiones léxicas en la traducción», en *Revista de Lingüística y Lenguas Aplicadas*, 6(1), 311-328.
- SAGER, J. C. (1993). *Curso práctico sobre el procesamiento de la terminología*. Madrid, España: Fundación Germán Sánchez Ruipérez.
- & Somers H. (1996). *Terminology, LSP, and translation: studies in language engineering in honour of Juan C. Sager*. Amsterdam, Países Bajos: J. Benjamins Pub. Co.
- SINCLAIR, J. M. (1996). The Empty Lexicon. *International Journal of Corpus Linguistics*, 1(1), 99-119.
- SUBIRATS, R. C. (2001). *Introducción a la sintaxis léxica del español*. Madrid, España: Iberoamericana.
- WÜSTER, E. (1974). The Road to Infoterm: Two reports prepared on behalf of Unesco; inventory of sources of scientific and technical terminology; a plan for establishing

an International Information Centre (Clearinghouse) for Terminology. Pullach, München, Alemania: Verl. Dokumentation.