Designing Pedagogical Support Materials in The Area of Lexis: A Technical and Sub-Technical Terminological Corpus

Alicia I. Pérez de Pereyra y N.Beatriz Aguilar de Espinosa
Universidad Nacional de Córdoba (Argentina)

The purpose of this article is to describe the steps that are being followed to elaborate a bilingual (English-Spanish/Spanish-English), technical and sub-technical terminological corpus in the field of the agricultural sciences and to delineate the next phases required for its completion.

I. Introduction

One of the most challenging projects undertaken by the Language Laboratory team at the Facultad de Ciencias Agropecuarias (Córdoba National University) is the compilation of a corpus of agricultural terms.

In the majority of South American countries, most university graduates find it very difficult to keep abreast of the profession unless they are able to read updated bibliographical material which, as is well-known, is mostly written in English. The lack of knowledge of the language sometimes becomes an unsurmountable barrier; thus, they are faced with the need to develop reading comprehension skills to a considerable degree.

As university teachers we have been asked many times to organize courses for this purpose. In our country, a course of these characteristics must be subject-specific. The reasons for this are, concurrently, motivation and lack of time. Our students, both graduates and undergraduates, are generally pressed for time due to academic and work demands; so they are only motivated by content that is closely related to their field. Students are often more stimulated to improve their language proficiency if they see the connection between language and their field of work.

II. Rationale

A course in reading comprehension aims at developing in the learner the capacity to construct meaning from a text. As Michael Lewis (1993:33)

points out "more of the meaning [of the text] is carried by lexis than grammatical structure. Focus on communication necessarily implies increased emphasis on lexis, and decreased emphasis on structure"; although, traditionally, grammar has been considered of greater importance than vocabulary.

This is a highly controversial topic; however, we think that the mastery of a large lexical set within each specialism greatly facilitates the access to the required specific literature. Krashen (1983:155) also reflects on the importance of the lexical system in the comprehension of the message; although he refers to "utterances", his consideration applies equally well to written discourse :

> Vocabulary is basic to communication.....Indeed, if our students know
> the morphology and syntax of an utterance addressed to them but do
> not know the meanings of the key lexical items, they will be unable to
> participate in the communication.

And this is exactly what has become evident to us throughout our extensive teaching experience. Both during the period of material selection for the development of our regular ESP courses and during the classes, it was observed that there was a lexical corpus of technical and sub-technical terms which imposed a severe limitation to the fluency and precision in the students' reading ability.This prompted us into the creation of the said corpus which we will make available to the readers in the form of a bilingual dictionary of agricultural terms.

III. Methodology
We follow Sager (1990:3) in his definition of technical vocabulary as " a collection of terms which has a certain coherence by the fact that the terms belong to a single subject area" and as regards sub-technical vocabulary, we find Trimble's (1989:128-129) definition to be the most satisfactory one for our purposes. Trimble traces the origin of the term to Dr. Ronayne Cowan who defines it as "context-independent words which occur with high frequency across disciplines" to which Trimble adds another dimension:

" ...those 'common' words that occur with special meanings in specific scientific and technical fields". He considers that " Together, the two sets of words make up the English sub-technical vocabulary".

These technical and sub-technical terms are lexical items (content words) as opposed to grammatical items (function words) (Lado,1964:115; Halliday 1989:61). In Halliday's discussion (1989:61) about lexical density in written language he points out that " there are twice as many lexical words as there are grammatical words. .....This is a characteristic difference between spoken and written language. Written language displays a much higher ratio of lexical items to total running words". This highlights the importance of a thorough understanding of the meaning of content words, in our case, the vocabulary of the agricultural sciences.

Faced with the problem of selecting appropriate materials for our courses, we searched for adequate sources of reference such as specific bilingual dictionaries, lexicons, glossaries and we found there are only a few available and those do not fully meet our students' needs. Apart from the many general bilingual dictionaries, to our knowledge, there only exist in our milieu: an Agronomic Glossary compiled in the University of Paraná (Argentina), a Glossary of Seed Terms used in the International Center of Tropical Agriculture — Cali, Colombia and a Glossary of Soil and Water Conservation from the Centro Regional A.I.D., Mexico.

Although we find the attempts made to fill this gap praiseworthy, we have noticed most of them have a limited coverage of specific terms and/or fields, not to mention regionalisms which are often completely foreign to the Argentinian variety of Spanish.

How have we approached this work?

In the first place, we deemed appropriate to make a needs analysis of the prospective reader. On this basis we found that the most commonly consulted bibliographic material were mainly journals, specialized magazines, study books, technical texts, scientific papers and abstracts.

A study of the style, medium and purpose of the sources was then undertaken; the style is formal, the medium is written and the purpose is scientific communication.

The methodology employed for the collection of this lexical corpus was, first, to select seven thematic areas. The criterion for selection was pertinence to the Agricultural field determined:

1. on the basis of the subjects in the curricula of several Schools of Agricultural Sciences from different Argentine universities and
2. on the basis of consultations to researchers and teachers of the above mentioned institutions.

From the linguistic point of view the selection was carried out not only considering the inclusiveness of the fields but also taking care to avoid overlapping of the lexical areas; nevertheless, it has to be borne in mind that in the same way different disciplines overlap, lexical fields intersect.

The second step was to compile a balanced list of English terms representative of each of the thematic areas; to date, approximately two thousand items. Given that lexical items are part of an open set which extends indefinitely and that language is dynamic, we can only consider this a preliminary list. By means of a qualitative analysis only terms with a normal or special(Lado,1964:120) degree of difficulty were selected.

These items were chosen from a textual corpus of paragraphs extracted randomly from the sources mentioned above which at present contains approximately 200,000 words. The linguistic distribution of a word can only be validated by adequate and varied contexts; thus, for each text we selected complete paragraphs up to 500 words each. The paragraph was taken as the selectional unit as the only means to retrieve meaning objectively because "it is necessary to work with real empirical knowledge of the language,(....)because only a linguistic corpus of data would produce the desired objectivity" (Lara,1987:13).

The criterion of organization of the macrostructure (Alvar Ezquerra, M., 1983: 121) of the corpus is the alphabetic sequence. That is to say, the only link each word has with the preceding or the following one is a strong formal resemblance regardless of the specific area to which it belongs.

Polysemes, homographs, synonyms, compound words and derivatives

were weighed to determine their treatment as separate entries, taking into account that "it is important to decide whether the entry term represents the concept or is simply the linguistic form of the term" (Sager 1990, 145). In the case of this work they were incorporated as separate entries considering that entry terms represent concepts.

Considerable progress has been done as regards the selection of Spanish equivalents to the English items. To determine the pertinence of these equivalents an analysis based on two axes —interlinguistic and intralinguistic— was carried out. For the intralinguistic selection we confered with specialists in the different subjects, chosen for their experience as post-graduate students or fellows in English-speaking countries; we consulted relevant bibliography in Spanish; and we resorted to our extensive experience in the field.

A careful selection of a the equivalent ensures its effectiveness. This is feasible because the connotation of the term becomes narrower given the high specificity of the fields considered the choice of the appropriate equivalent was made taking into account the context in which the English lexical item occurred. Wherever no one-to-one equivalent existed, we resorted to "inclusive paraphrase" (M.Alvar Ezquerra, 1983: 124).

How do we plan to continue?

There are several stages to go before we can feel we have completed our task, in the next academic periods we plan to:

- finish off the search for Spanish equivalents to the existing English terms
- add to the existing list of English terms
- carry out a frequency count plus a distributional study to confirm the pertinence of the terms; to this purpose, the best available software will be used and computation specialists will be consulted
- make this corpus available to the public both in printed and machine - readable form

IV. Conclusion

We expect this corpus to be of help to graduates and students alike, not

only in the field of the Agricultural Sciences but also in related fields such as Biology or Chemistry. It can also prove useful to professionals, fellows, researchers and the staff of business concerns and organisations related to agriculture.

We believe that once the complete corpus is available we will be in a better pposition to address the question that, as teachers, ultimately drives our reasearch:

Can a lexical corpus like the one in preparation promote more efficient reading comprehension?

During the coming years, we hope to test our hypothesis empirically.

WORKS CITED

Alvar Ezquerra, M. 1993."Lexicografía" in López Morales, H. (ed.) *Introducción a la Lingüística Actual*. Madrid: Editorial Playor.

Halliday, M.A.K. 1989. *Spoken and Written language*. Hong Kong: Oxford University Press.

Krashen, S. & **T.Terrell.** 1983. *The Natural Approach*. England: Alemany Press and Pergamon Press.

Lado, Robert. 1964. *Language Teaching —A Scientific Approach*. U.S.A.: McGraw-Hill.

Lara, Luis F. 1987. "Methodology in a Non-Spanish Dictionary of the Spanish Language: The Diccionario del Español de México" in Robert F. Ilson (ed.) *A Spectrum of Lexicography*. Amsterdam: John Benjamin's Publishing Co.

Lewis, Michael. 1993. *The Lexical Approach*. England: Language Teaching Publications.

Sager, Juan. 1990. *A Practical Course in Terminology Processing*. Philadelphia: John Benjamin's Publishing Co.

Trimble, Louis. 1989. *English for Science and Technology — A Discourse Approach*. Cambridge: Cambridge University Press.