

Preparing an Online Dictionary of Business Communication: From Idea to Design*

Sven Tarp, *Department of Afrikaans and Dutch, University of Stellenbosch,
Stellenbosch, South Africa and Centre for Lexicography,
Aarhus University, Aarhus, Denmark (st@bcom.au.dk)*

Abstract: This article discusses the plan for an online English–Spanish dictionary of business communication based on the function theory. After a short overview of lexicographical traditions within this field, it focuses on the philosophy behind the overall dictionary concept, the methods applied to develop this concept, some of the requirements for the corresponding database, as well as the foreseen Internet user interface which may lead the user to ten completely different articles for each word entered in the search field.

Keywords: ONLINE DICTIONARY, BUSINESS COMMUNICATION, BUSINESS DICTIONARY, SPECIALIZED LEXICOGRAPHY, FUNCTION THEORY

Opsomming: Die voorbereiding van 'n aanlyn sakekommunikasiewoordeboek: Van idee tot ontwerp. In hierdie artikel word die plan bespreek vir 'n Engels–Spaans aanlyn sakekommunikasiewoordeboek gebaseer op die funksieteorie. Ná 'n kort oorsig oor leksikografiese tradisies binne hierdie veld word daar gefokus op die filosofie agter die algehele woordeboekkonsep, die metodes wat toegepas is om hierdie konsep te ontwikkel, sommige van die vereistes vir die ooreenstemmende databasis, sowel as die voorsienbare Internet-gebruikerskoppelvlak wat die gebruiker kan lei na tien verskillende artikels vir elke woord wat in die soekveld ingetik word.

Sleutelwoorde: AANLYN WOORDEBOEK, SAKEKOMMUNIKASIE, SAKEWOORDEBOEK, VAKLESIKOGRAFIE, FUNKSIETEOORIE

1. Introduction

In this article, we will discuss the plan for an online dictionary designed to assist English and Spanish business communication, a dictionary which will be produced in collaboration with experts and scholars from various countries. We will focus on the philosophy behind the overall dictionary concept, the methods applied to develop this concept, some of the requirements for the corresponding database, as well as the envisaged user interface which may lead the user to 10 completely different articles for each word (string) entered in the search field.

* This article was presented as a paper at the Twentieth Annual International Conference of the African Association for Lexicography (AFRILEX), which was hosted by the University of KwaZulu-Natal, Durban, South Africa, 6–8 July 2015.

The dictionary project is based on the functional principle combined with the option to individualize, and is inspired by the experience obtained from a similar project including four printed bilingual dictionaries between Danish, on the one hand, and English and Spanish. Before explaining the project in further details, we will give a brief overview of the history of business dictionaries involving English and Spanish.

2. Lexicographical overview

In the European context, there is a several hundred years' old tradition of compiling dictionaries related to the broad field of business and economics, cf. Besomi (2011). Among the early dictionaries are those produced by Savary des Bruslons (1723), Postlethwayt (1751–55), Rolt (1756) and Mortimer (1766) in the 18th century. These dictionaries were generally designed to transmit knowledge about trade and commerce and explain the many terms related to these activities, cf. Tarp and Bothma (2013: 224). This purpose was also the main idea behind Virio's *Collección alfabética* of British duties, tariffs, law extracts, mandates, regulations and rulings, published in Madrid in 1792, cf. Astigarraga and Zabalza (2007: 36-38).

Although trade and commerce between the European countries — and overseas — developed relatively fast in the 18th century (a fact reflected in the above mentioned dictionaries), it was not until the 19th century that the first dictionaries designed to assist interlingual communication saw the light. Strangely enough, from their very appearance these communicative dictionaries seem to have developed along a separate line "with no or only little inter-connection" to the increasing number of cognitive dictionaries designed to transmit knowledge about business and economics, cf. Tarp (2015a: 188).

Among the first known dictionaries designed to assist business communication between the two languages in question is Veitelle's trilingual English–Spanish–French *Mercantile Dictionary* from 1864, which, according to the subtitle, includes "a Complete Vocabulary of the Technicalities of Commercial Correspondence, Names of Articles of Trade, and Marine Terms". An extract from this dictionary where the three languages are treated in separate columns can be seen in Figure 1.

Abandon.	abandonner; délaissér.	abandonar; dejar.
Abandonment.	délaissement, m.	abandono, m.; dejacion, f.
Abatement.	réduction; remise, f.	rebaja; reduccion, f.
Abide the consequences.	subir les conséquences.	sufrir las consecuencias.
Aboard.	à bord.	á bordo.
About.	environ.	cerca de.
to go about (<i>naut.</i>).	virer de bord.	virar de bordo.
Abstract.	extrait; relevé, m.	extracto; resumen, m.
— of account sales.	relevé de compte de vente.	extracto de cuenta de venta.
Absynth.	absinthe, f.	agenjo, m.
Abundant (rice is).	les riz sont abondants.	abunda el arroz.

<p>Accept. Acceptance. general —. partial —. absolute —. — in blank.</p>	<p>accepter. acceptation, f. acceptation totale. acceptation partielle. acceptation pure et simple acceptation à découvert; en blanc.</p>	<p>aceptar. aceptacion, f. aceptacion total. aceptacion parcial. aceptacion pura y simple. aceptacion á descubierto; en blanco.</p>
--	---	---

Figure 1: Extract from the trilingual Mercantile Dictionary (1864)

Another early plurilingual business dictionary to be mentioned here is Graham and Oliver's *The Foreign Traders' Dictionary of Terms and Phrases in English, German, French, and Spanish*, published by MacMillan in 1906. As can be seen in Figure 2, the four languages treated in this dictionary are also presented in separate columns which cover two opposite pages here.

Trade, commerce, business barter exchange, interchange traffic monopoly	Der Handel, das Geschäft der Tauschhandel der Austausch, Tauschhandel der Handelsverkehr, der Handel das Monopol	Le commerce, les affaires (f.pl.) le commerce d'échange, le troc l'échange (m.), le troc le trafic le monopole	El comercio, los negocios la permuta, el trueque el trueque, el cambio el tráfico el monopolio
Home trade foreign trade import trade export trade carrying [shipping] trade transit [transshipment] trade coasting trade	Der inländische Handel, Binnen- der ausländische Handel [handel] der Einfuhrhandel der Ausfuhrhandel das Rhederei-Geschäft der Transitverkehr der Küstenhandel	Le commerce (à l') intérieur le commerce (à l') étranger le commerce d'importation le commerce d'exportation les affaires (f.) maritimes le commerce de transit le cabotage	El comercio interior [del país] el comercio exterior el comercio de importación el comercio de exportación los negocios marítimos el comercio de tránsito el cabotaje
Wholesale trade retail trade	Der Großhandel, das Engrosgeschäft der Kleinhandel, das Detailgeschäft	Le commerce de [en] gros le commerce de [en] détail	El comercio al por mayor el comercio al por menor

Figure 2: Extract from the quadrilingual *The Foreign Traders' Dictionary* (1906)

Figures 1 and 2 clearly indicate that the lexicographical content and architecture of the first interlingual business dictionaries were rather simple compared to modern standards. Since then various business dictionaries have been published covering both English and Spanish, most of them now bilingual. Noteworthy in this respect is López and Watt's bilingual *Oxford Business Spanish Dictionary*, published by Oxford University Press in 2002 and now only available online. Figure 3 shows an article from the Spanish–English part of the printed version of this dictionary:

cerrar 1 *vt* (tienda, negocio) close;
(definitivamente) close down; (salida) seal off;
~ **un acuerdo** close a deal; ~ **una posición** close
a position; ~ **un trato** make a deal; ~ **una venta**
complete *o* close a sale
2 *vi* close down

Figure 3: Article from the *Oxford Business Spanish Dictionary*

This article reflects the lexicographical development that has taken place since the dictionaries compiled by Veitelle (1864) and Graham and Oliver (1904). Especially useful is the provision of meaning discrimination allowing the users to select the needed equivalent among the ones proposed. However, based on the experience from students' use of a similar Spanish–Danish business dictionary, the syntactic data (*vt* and *vi*), even if written in full form, is probably too abstract to understand for many users without background in linguistics. A more user-friendly solution would therefore have been that these data were exemplified and explained in greater details as the words' syntactic properties constitute one of the main challenges in the redaction of foreign-language texts.

Today the *Oxford Business Spanish Dictionary* is only available for subscribers on the Internet. Here it competes with a growing number of free-access business dictionaries, many of which reflect the same poor article architecture employed in the first printed dictionaries more than hundred years ago. One such example is Andy D. Miles' *Business English to Spanish Glossary* which, according to the webpage, contains 10,000 lemmata. An extract from this online dictionary can be seen in Figure 4.

bill	(1) s. billete (2) s. factura, cuenta (3) s. pagaré, letra, libranza (4) s. anteproyecto de ley (5) v. facturar, cobrar.
bill book	libreta de facturas, talonario de facturas, facturero (Arg.).
bill broker	corredor de letras, casa de descuento de letras.
bill diary	registro de letras, libro diario de letras.
bill for collection	(1) factura a cobrar (2) letra a cobrar, pagaré a cobrar.
bill market	mercado de letras.
bill of debit	pagaré.

Figure 4: Extract from *A-D. Miles' Business English to Spanish Glossary*

What strikes the eye here is that the dictionary does not offer meaning discrimination and thus creates a new and completely unnecessary problem for the user. Neither does it provide data on gender, inflection, syntactic properties or collocations, all of them data that are absolutely necessary to assist an English first-language speaker translating or producing a correct business text in Spanish. This dictionary is emblematic and typical for a large number of specialized dictionaries that are now freely available on the Internet and compiled by people who may or may not be specialists within their respective fields, but who don't seem to have the lexicographical background needed to produce dictionaries of a quality required by the users in the 21st century, cf. Caruso (2011).

To finish the picture, it should be mentioned that there also are a number of monolingual dictionaries designed to assist reception and production of business texts. One example of this is the *Cambridge Business English Dictionary* which is available both in a printed version and online. According to the advertising, this dictionary "has over 35,000 words, phrases, and meanings, and includes business-specific vocabulary" as well as "help with how to use English naturally in business situations like meetings, conference calls, and emails". However, this and other monolingual business dictionaries will not be discussed here.

3. Motivation

Various facts indicate the need for a new online dictionary for business and other types of professional communication between English and Spanish:

- In a still more internationalized world there is a growing demand for a high-quality dictionary within this field.
- English and Spanish are among the three most spoken languages in the world and used as official languages in more countries than any other language.
- Only part of the required lexicographical information can be retrieved from general dictionaries in the two languages.
- The many specialized dictionaries designed to provide knowledge about disciplines relevant to this field do, as a rule, not cater for their users' communicative needs.
- Most of the existing online dictionaries designed to assist English and Spanish business communication are of very low lexicographical quality.
- None of the existing online dictionaries have completely adapted to the Internet and made full use of the new technologies and techniques available to deliver a more personalized lexicographical product.

With this in mind, the author of this article was contacted by a publishing house with a view to preparing a project for an online English–Spanish business dictionary. The project is inspired by the experience obtained from a similar project including four printed bilingual dictionaries between Danish, on the one hand, and English and Spanish, on the other, as well as the experience from more than 25 years of teaching foreign-language business communication. However, although inspired by past experience, the new dictionary concept has to be adapted to the Internet and make use of the new technologies and techniques, and therefore needs to be designed from scratch. The main features of the new concept will be presented in the following sections.

4. Subject field and foreseen user group

Business communication is an amorphous field and, as such, very difficult to delimit lexicographically. This is reflected in the fact that existing business dictionaries contain lemma stocks that vary greatly, much more than can be explained by "normal" competition. It seems that each author has his or her own idea of how the field should be embraced. In this respect, the project will be based on the experience accumulated at the Aarhus School of Business where interlingual business communication has been taught as an academic discipline for more than half a century. Here, the field or "discipline" of business communication is traditionally understood as divided into several genres like mailing, sales letters, enquiries, offers, orders, deliveries, payments, complaints, etc. To this can also be added more recent web-based "about us" texts like company history, visions, mission and corporate social responsibility.

All these genres have their own special vocabulary, characteristics and features which may vary from "daily" language in terms of technicality, formality and requirements to politeness. The corresponding texts may include terms related to any product or business matter, e.g. those Veitelle (1864) called "Names of Articles of Trade" (see above). Such terms can obviously not be incorporated into a modern dictionary focussing on business communication as each of them may appear only in very few texts, whereas they, in their totality, can be counted in tens of thousands, if not more. The tradition from past centuries of lemmatizing the names of the products that are traded cannot be maintained in today's industrial consumer society. However, the above-mentioned texts also contain a large number of frequently used terms related to technicalities such as packaging, transport, payment, banking, insurance, sponsoring, contracts, legal matters, advertising, etc. These terms have to be treated in a dictionary of business communication.

As such, the planned dictionary will cover a specific vocabulary somewhere in the grey zone between general and specialized language as it has been defined by Fuertes-Olivera and Tarp (2014: 7). On the one hand, it will contain frequently used specialized terms belonging to various relevant disciplines, and on the other hand, it will include general-language words and expressions frequently employed in business communication, sometimes with a specific usage.

The main purpose of the dictionary is to assist business people, secretaries, civil servants, translators, students and other possible users who engage in business and other types of professional foreign-language communication in Spanish and English. This group of envisaged users, which is basically composed of first-language speakers of English and Spanish, is very heterogeneous and may have different lexicographical needs that should be catered for in the dictionary design.

5. Basic philosophy and main functions

Taking into account the heterogeneous user group described in the previous

section, the dictionary project is based on the functional principle combined with the option to individualize, i.e. a lexicographical Model T Ford taking the first modest steps towards a future Rolls Royce according to the classification proposed by Fuertes-Olivera and Tarp (2014: 16).

Today, the main relevant user situation is the production of business texts directly in the foreign language or based upon an outline in the source language (mostly done by secretaries). Direct translation of this type of text is mainly restricted to the education system, but as students are also part of the foreseen user group, the support to both foreign-language (L2) text production and L1–L2 translation make up the two main functions of the planned dictionary. Lexicographically, these two functions require very similar treatment. However, as the corresponding data, with only a few additions, may also serve L2 text-reception, L2–L1 translation and L1 text-production, these functions will also be covered by the project, although considered of secondary importance. Taking into account that the user group is composed of first-language speakers of both English and Spanish, it adds up to a totality of 10 lexicographical functions expressed in a combination of monolingual and bilingual lexicographical solutions.

The basic philosophy underpinning the project is that the traditional bilingual dictionary which contrasts two different languages is inconvenient. This kind of dictionary usually provides too little data to fully assist the users in foreign-language text production. And if it does furnish the needed data, the inevitable result would be that many articles would be filled with too much data, thus creating a new problem for its users, namely information stress due to data overload.

With this in mind, the planned project will basically be composed of two monolingual dictionaries with the possibility of bilingual access from L1 to L2 and explication of L2 (by means of equivalents and notes in L1). It can therefore be used not only to assist L2 production and L1–L2 translation, but also L2–L1 translation, L2 reception and even L1 production, with the two former functions being the prioritized ones.

It is foreseen that the dictionary will contain explications of "difficult" technical terms like *letter of credit*, *incoterms*, and *bill of lading* as well as short definitions in case of polysemy but generally it will not be designed to serve L1 text reception. The reason for this decision is that most first-language users will generally know the meaning of the majority of words belonging to the "grey zone" between general and specialized language. However, if this proves to be false or only partly correct, additional definitions can always be included in a later phase of the compilation process.

6. How to satisfy foreign-language text production

As already indicated, the needs related to L1–L2 translation and L2-text production require almost the same lexicographical response, although there may

be some minor differences as the latter, in contrast to the former, is not source-language dependent and, hence, is subjected to more linguistic variation, cf. Tarp (2004). In the following, the two user situations will be treated as one, exemplified with the problems and needs related to L2 production. In this respect, a person writing a business text in a foreign language may experience three types of problems eventually giving rise to a dictionary consultation:

1. The person does not know the word to be used in L2.
2. The person knows a L2 word but is not sure whether it can be used with the specific meaning and in the specific context.
3. The person knows the word but has doubts about orthography, gender, inflection, syntactic properties, collocations, synonyms, antonyms, fixed expressions, etc.

Any first-language speaker of L1 may experience problems related to one of the three scenarios described, but the person with the highest L2 proficiency level is most likely to have problems of type 3 whereas learners at beginners' level most often will face problems of type 1, etc. The three types of problem require three different lexicographical solutions:

Problem 1: An **L1-L2 solution** offering L2 equivalents and meaning discrimination.

Problem 2: An **L2 solution** with meaning explications which could be even more advantageous to the user if they were given by means of L1 equivalents, i.e. a **L2-L1 solution**.

Problem 3: An **L2 solution** including data on orthography, gender, inflection, syntactic properties, collocations, synonyms, antonyms, fixed expressions, etc.

The user with problems of type 3 may be completely satisfied with the proposed solution. The users with problems of type 1 and 2, however, may experience additional needs once they have chosen or confirmed the word to be used, as they may have doubts about gender, inflection, syntactic properties, collocations, synonyms, antonyms, fixed expressions, etc. Such needs can best be solved by:

- A **L2-L1 solution** including data on gender, inflection, syntactic properties, collocations, synonyms, antonyms, fixed expressions, etc.

The reason why an L2-L1 solution with this content would be the best additional solution here is that L1 data, apart from meaning explication, will make it easier for users with low and intermediate L2-proficiency level to navigate in the displayed article and choose the relevant collocations or opt for the most convenient syntactic combinations, among others. In addition, although a user with high proficiency level may be perfectly satisfied with the L2 solution pre-

sented above, an alternative L2-L1 solution as the one proposed as additional to the other user categories would not be bad either (except in case of data overload). For the convenience of this user type, a special button to hide (and unfold) the L1 data could easily be included in the user interface. The article appearing when hiding the L1 data would then be the same as the one a first-language speaker of L2 would get when looking for assistance in text production in this language.

Consequently, the overall concept will, in accordance with the underlying philosophy, consist of two monolingual dictionaries in English and Spanish, each of them with a bilingual dimension expressed in the possibility of access from L1 to L2 and explication of some L2 items in L1. In this way, the project will with elegance and flexibility serve both Spanish and English first-language users of various proficiency levels, not only when they need help in connection with L1-L2 translation and L2 production, but also in relation to L2 reception, L2-L1 translation and L1 production.

7. Lexicographical database

Once the situations to be covered by the planned business dictionary had been established, the next step was to determine the possible types of need which the foreseen user group may experience in these situations. There are various methods to perform this task, cf. Fuertes-Olivera and Tarp (2014: 53-57). In this project, the method chosen was the functional one, i.e. deduction based on the lexicographer's own experience and knowledge of the subject field and the envisaged users' problems when working with foreign-language business texts, cf. Tarp (2013: 148-149). Deduction should not be confused with introspection as it requires a combination of lexicographical expertise and real subject-field knowledge based upon academic studies and research. In this particular case it includes knowledge obtained from teaching business communication and marking thousands of tasks and exercises in searching information as well as subsequent discussions with the students.

The result of the deduction process was a detailed list of user needs related to the different situations covered by the dictionary. Taking this list as the point of departure, it was relatively easy to establish which data categories and which relations between them would be required to cover the detected needs, a precondition for preparing the lexicographical database to be employed in the project.

Each dictionary project has its own personality and it is therefore recommendable to design a lexicographical database from scratch for each new project. The pre-fabricated dictionary writing systems available on the market are, as a rule, too inflexible to be used in projects like the one described here. It is especially important that the corresponding user interface employed by the lexicographers to compile the dictionary is correctly understood, i.e. as a means of production. This requires, among other things, 1) that it guarantees high produc-

tivity and high quality (allowing the preparation of as many correct data as possible in the shortest possible time) and 2) that it, in order to achieve these goals, is as comfortable and easy as possible for the lexicographer to work with, cf. Tarp (2015b).

Apart from this, the overall requirement from the database is, on the one hand, that it (and the dictionary as such, understood as the totality of data that may be displayed in the totality of hypothetical consultations) should include *as much data as possible* relevant to the topic as well as to the detected user needs; and on the other hand, that the individual articles to be displayed on the screen should contain *as little data as possible* in order to avoid information overload, i.e. just as much data as required to satisfy the user's needs in each case, no more no less.

In order to achieve this, the classical method of analysis and synthesis was used. By means of this method, the data categories to be covered were first separated into their smallest relevant parts (i.e. those data items which eventually may be displayed separately on the screen) and then combined in order to establish their relevant mutual relations. All this had then to be communicated to the programmer of the lexicographical database (and of the lexicographers' interface). As this communication takes place between specialists from two different disciplines with different terminologies, it entails a lot of possible misunderstandings and may take unnecessarily long if a common language is not found. In this respect, the main terminology used was *data fields* and their mutual *relations*, i.e. the structures to be represented in the lexicographical database, cf. Bergenholtz and Nielsen (2013) and Tarp (2015b). No specific terms for the corresponding structures were ever used.

In this particular case, the communication was both oral and written using a diagram as the one reproduced below:

1. English lemma
 - 1.1. Homonymy number
 - 1.2. Polysemy number
 - 1.3. UK/US
 - 1.4. Full form
 - 1.4.1. Full form note in English
 - 1.4.2. Full form note in Spanish
 - 1.5. Abbreviation
 - 1.5.1. Abbreviation note in English
 - 1.5.2. Abbreviation note in Spanish
 - 1.6. English definition
 - 1.6.1. Link to English Internet page
 - 1.6.2. English text to link

- 1.7. Spanish definition
 - 1.7.1. Link to Spanish Internet page
 - 1.7.2. Spanish text to link
- 1.8. Part of speech in English
- 1.9. Part of speech in Spanish
- 1.10. Inflection
- 1.11. Syntactic mini-rule(s)
 - 1.11.1. Example sentence in English
 - 1.11.2. Example sentence translated into Spanish
 - 1.11.3. Mini-rule note in English
 - 1.11.4. Mini-rule note in Spanish
- 1.12. Synonym(s)
 - 1.12.1. UK/US
 - 1.12.2. Synonym note in English
 - 1.12.3. Synonym note in Spanish
- 1.13. Antonym(s)
 - 1.13.1. UK/US
 - 1.13.2. Antonym note in English
 - 1.13.3. Antonym note in Spanish
- 1.14. Spanish equivalent(s)
 - 1.14.1. Equivalent note in English
 - 1.14.2. Equivalent note in Spanish
- 1.15. Collocation(s)
 - 1.15.1. UK/US
 - 1.15.2. Spanish translation
 - 1.15.3. Collocation note in English
 - 1.15.4. Collocation note in Spanish
- 1.16. Fixed expression(s)
 - 1.16.1. UK/US
 - 1.16.2. Spanish translation
 - 1.16.3. Fixed expression note in English
 - 1.16.4. Fixed expression note in Spanish
- 1.17. Memo field
- 1.18. Reserve field

- 1.19. Reserve field
- 1.20. Reserve field
- 1.21. Reserve field

Diagram 1: Data fields in the English part of Business Dictionary

The diagram presented above was initially given to the programmer in charge of designing the lexicographical database. As can be seen, to support the English part of this dictionary the database requires a total of 50 fields of which 45 are defined data fields, one is a memo field (to be used for notes and internal communication between the lexicographers), and four are maintained as reserve fields (if something new shows up during the compilation process). The Spanish part of the dictionary contains a similar number of fields.

It should, however, be noted that the relations reflected in this diagram represent only a fraction of the many relations relevant to the project. In order to design the database needed to have the final product exactly as required, all relevant relations between the data types had to be described and communicated to the programmer. A diagram reflecting all the relevant data and their mutual relations would be extremely complex and the risk of misunderstanding or missing some of these relations would be considerable. The written diagram was therefore combined with oral explications in a complex interdisciplinary communication.

8. Dictionary portal and search options

The dictionary portal to be accessed via the Internet is, at this point, foreseen to include a traditional search field and eleven search options which can be seen in Figure 5.

A user of the Business Dictionary who has accessed the web-portal shown in Figure 5 will first have to enter a word (or string) in the search field and then chose one of the eleven search options (buttons). A click on the first of these buttons (the one next to the search field) initiates a general search in the database, whereas a click on any of the other ten buttons starts specific searches in a pre-selected number of database fields, thereby filtering the data to be displayed. Although these ten specialized buttons, by first glance, may seem to represent each of the ten functions covered by the dictionary, this is actually not the case as it will be argued in the following.

The first thing to be noted in the dictionary portal shown in Figure 5 is that users can carry out specific searches based upon their first language. This possibility provides a first important adaptation to each of the different characteristics of the heterogeneous user group described above. After deciding the preferred language (and without performing any action), the users can choose

between five search options in accordance with the activity giving rise to the consultation.

Business Dictionary

– Diccionario de Comunicación Empresarial

Enter a Word...	SEARCH
-----------------	--------

Haga clic en su problema para ser mejor atendido
Mi lengua materna es español y necesito ayuda para

- escribir un texto en inglés
- traducir un texto del español al inglés
- comprender un texto en inglés
- traducir un texto del inglés al español
- escribir un texto en español

Click on your problem in order to get better assistance
My first language is English and I need assistance

- to write a Spanish text
- to translate from English into Spanish
- to understand a Spanish text
- to translate from Spanish into English
- to write an English text

Figure 5: Planned dictionary web-portal

If English first-language speakers, for instance, enter a Spanish word in the search field and click on the first option, i.e. *to write a Spanish text*, then a Spanish article with English equivalents and notes will appear on the screen (with the aforementioned possibility of hiding the English items changing it into purely monolingual). By contrast, if the same users write an English word in the search field and activate the second option, i.e. *to translate from English into Spanish*, they will be directed to an "in-between" article with no other data than Spanish equivalents, meaning discrimination, and part of speech. Here, they can decide to stop the consultation if an equivalent is all they need or, if they also require additional information, they can click on the preferred equivalent and continue to a Spanish article identical to the aforementioned one in order to get data on gender, inflection, syntactic properties, collocations, synonyms, antonyms and fixed expressions (see the corresponding argumentation in Section 6).

However, a little "trick" will be incorporated in the search buttons. If the users, instead of a Spanish word, enter an English word and activate the first option — or a Spanish word (instead of an English one) and click on the second

option — then the immediate search results would be reversed. The idea behind this little "trick" is that the dictionary should be as easy as possible to operate without reading any lexicographical instructions. Hence, if English first-language speakers were experiencing a problem when writing a Spanish text, and if this problem consisted in that they did not know the Spanish word to use in a specific context, then they may — just may — enter a known English word in the search field and click on "to write a Spanish text" because this activity is actually the one they are engaged in. A similar thing could happen if they were translating an English text into Spanish. In order to reduce the number of frustrated or superfluous consultations, the system (dictionary) simply takes such situations into consideration and performs "intelligent searches" based on the language to which the search word belongs. The guiding principle is that the users, as a rule, should not perform more than two actions (entering a search string and clicking on a specific option) before getting quick and easy access to the data required in each case — and one more action (a click) if they need additional data.

The "duality" hidden in the two first search options clearly indicates that these options — although the accompanying text may insinuate it — do not correspond directly to specific functions, but to possible steps required to fully serve the two main and prioritized functions covered by the dictionary, i.e. L2 text production and L1-L2 translation. As to the three other search options in the portal presented in Figure 5, they do correspond directly to the functions with the same name. Clicking on them will lead the users to three different articles. Finally, if they, instead of the specialized search options, choose the general search button placed on the right of the search field, they will get a "traditional" article with unfiltered data, i.e. all the data attached to the word in question, including various homonyms. However, whether or not this general option will be maintained will depend on the users' reaction once the dictionary is available for access via the Internet.

9. Two examples of master articles

Apart from the web-portal, it is foreseen that the completed dictionary project will include eleven other user interfaces, ten for each of the search options discussed in the previous section and one to be used in the case of homonymy. Each of these eleven interfaces will represent a master article containing fields for all the data categories to be presented in each case. In this section we will briefly discuss the two master articles directly related to the prioritized functions of the dictionary.

The following master article is the one planned to serve an English user consulting the dictionary in order to get assistance when having problems in the production of Spanish business texts:

Spanish lemma homonymy number *part of speech in English*

<Inflection>

polysemy number

DEFINITION

- English definition
hyperlink to external source
English text to hyperlink

ABBREVIATED FORM

- abbreviation
(abbreviation note in English)

FULL FORM

- full form
(full form note in English)

EQUIVALENTS

- (equivalent note in English) **English equivalent(s)** (UK/US)

SYNTACTIC PROPERTIES

- **mini-rule(s)**
example sentence in Spanish
example sentence translated into English
(mini-rule note in English)

COLLOCATIONS

- **Spanish collocation(s)**
English translation
(collocation note in English)

FIXED EXPRESSIONS

- **Spanish fixed expression(s)**
English translation
(fixed expression note in English)

SYNONYMS

- Spanish synonym(s)

ANTONYMS

- Spanish antonym(s)

Example 1: Master article for English user writing a Spanish text

This master article includes all the data categories that an English first-language speaker may need when having problems related to text production directly in Spanish. The items underlined represent hyperlinks 1) referring the users to an external source, 2) directing them to another article in the dictionary, or 3) activating a pop-up window (which is a more convenient way of showing, for instance, the large number of inflected forms of Spanish verbs). The plural forms of some items in Example 1 indicate that there could be more than one item of this data category included in the displayed article. Apart

from this, one will note that there is an extensive use of metatexts written with small caps (DEFINITION, ABBREVIATED FORM, etc.) which serve as article-internal indicators or headlines separating sections with data of the same category.

The data to be lemmatized will include words, word combinations, terms, abbreviations, acronyms, names of relevant institutions, fixed expressions, and routine formulae. As indicated above, it is envisaged that a number of "difficult terms" will be defined. In some of these cases the definitions will be accompanied by links to relevant external sources. For instance, the users consulting one of the *incoterms* will, apart from a definition, be offered a link to the corresponding webpage of the *International Chamber of Commerce* which is the organization in charge of regulating these terms, and where they can get updated information. Users consulting the term *letter of credit* will be referred to a website of a financial institution where they, apart from additional information, can download a letter-of-credit form to be filled in, and so on.

As it is presented in example 1, the master article prepared for an English first-language speaker writing Spanish business texts may appear relatively overcrowded with data categories. In this respect, it is important to remember that it is a master article, and that the concrete articles displayed on the screen in all cases will include less data categories. As already explained above, definitions are only foreseen in relatively few articles (apart from cases of polysemy). In addition, external links, abbreviations, full forms and fixed expressions will also be irrelevant in most cases. The data on abbreviation and full form will be mutually exclusive, etc. Hence, if the English users referred to in Example 1 choose to hide the Spanish items, the most typical data categories to appear on the screen in this case would be the ones shown in Example 2:

Spanish lemma *part of speech in English*
<Inflection>
polysemy number definition

SYNTACTIC PROPERTIES

- **mini-rule(s)**
example sentence in Spanish
(mini-rule note in Spanish)

COLLOCATIONS

- **Spanish collocation(s)**
(collocation note in Spanish)

SYNONYMS

- Spanish synonym(s)

ANTONYMS

- Spanish antonym(s)

Example 2: Typical article for English user writing a Spanish text

If the English users instead write an English word in the search field and click

on the button *to translate from English into Spanish*, an article based upon the following master article will be displayed:

English search word(s)
(meaning discrimination) **Spanish equivalent** PART OF SPEECH
etc. ...

Example 3: Master article for an English user translating an English text into Spanish

As discussed in Section 6, an article of this type may solve the problem for the English first-language users who only need a Spanish equivalent and nothing more. For other users who will need additional data, the article will serve as an "in-between" page, from where a simple click on the preferred equivalent will lead them to an article based on the master article shown in Example 1.

Finally, it should be noted that none of the master articles shown in the above three examples reflects the graphical design of the articles eventually to be presented to the users. No decision has yet been taken in this respect. With a well-prepared database, the use of the available technologies allows almost any design — in this case the one considered most appropriate, taking into account the characteristics of the target users — and it also permits its modification or complete replacement with another design if the first one does not live up to the expectations. This is one of the many advantages brought to lexicography by the introduction of new technologies and media.

10. Some techniques to be employed

It is impossible to treat all features and technicalities related to the project in the framework of this contribution. Here we will limit ourselves to a brief discussion of a few techniques that may be used either to make the displayed articles more readable or to further reduce the amount of data displayed in the first instance in order to avoid information overload and adapt the dictionary to smaller devices like tablets and smartphones. Most of the techniques will be known from already published online dictionaries though they may still not be sufficiently generalized.

One of the many challenges in modern online lexicography is to handle the contradiction between the almost unlimited freedom from space restraints in the database and the undeniable space restraints at the level of the screenshots, i.e. a contradiction very similar to the one between storage space and presentation space in the terminology proposed by Lew 2012). This contradic-

tion is also present in the planned dictionary project which applies techniques that, on the one hand, are more space-consuming, and on the other, reduce the total amount of data on the screen, i.e. only the data displayed on the screen at a given moment but of course not all the data needed by the users in each consultation.

Among the space-consuming techniques are 1) less compact article formats with different data placed in separate lines, 2) abolition of metalexicographical abbreviations and their replacement with full forms, and 3) use of metatexts to introduce sections with specific data categories. These techniques are employed in order to improve readability and are therefore justified in spite of being more space-consuming.

Among the data-reducing techniques are 1) data filtering based on identification of the user's first language and the type of activity where the need occurs, 2) use of hidden data to be unfolded when required, and 3) use of pop-up windows to present additional data. These techniques are applied in order not to burden the users with too many data which, if presented simultaneously on the screen, may increase the consultation time as well as the risk that the users retrieve incorrect information or eventually abort the consultations.

The challenge is to balance the two types of technique for the benefit of the users. This implies that all lexicographical data required by the latter in the first instance should be included in the screenshot without the need to scroll down or sweep the screen to the right or left in order to find out whether or not the required data can be found there. Of course, as data underload can never be the alternative, this may frequently imply that the users have to perform an extra action in order to access the data needed to get the right information, but it should be absolutely clear from the data presented in the first screenshot that the needed data are included in the dictionary and can be accessed by clicking or moving the mouse over a specific item, or even by scrolling down or sweeping the screen (in this case with the knowledge that the data are there). This principle is especially relevant in dictionaries for communicative purposes which are usually consulted by people experiencing problems while doing something else and therefore demand a quick answer in order to continue with their main activity, e.g. foreign-language text production.

Data-reducing techniques are important in all types of online dictionary used on all types of screen but they are especially relevant when the dictionaries are accessed from devices with small screens. In such devices there are two ways to avoid scrolling down or sweeping the screen, namely to make the letter types smaller or to reduce the amount of displayed data. The first of these solutions cannot be recommended as "it is quite convenient to see a lengthy entry on a computer screen, but it is not convenient to see it in a smartphone screen" (Kwary 2015: 203-204). The Business Dictionary therefore opts for the second solution which is achieved by a flexible combination of data-reducing and space-consuming techniques, i.e. hidden data combined with metatexts:

Spanish lemma

INFLECTION
DEFINITION
ABBREVIATED FORM
FULL FORM
EQUIVALENTS
SYNTACTIC PROPERTIES
COLLOCATIONS
FIXED EXPRESSIONS
SYNONYMS
ANTONYMS

Example 4: Possible first-instance article to appear on smartphone screen

This master article foreseen to be used in smartphones is based on the one presented in Example 1. The metatexts will only be shown when they actually indicate the presence of data that are hidden; in case of zero data they will disappear. The idea is to inform the users where there may be relevant additional data which can be unfolded "on demand" by simply touching the screen over the respective metatexts. This master article still has to be refined and tested, as its convenience or inconvenience in the last instance depend on the users' reaction. But as indicated in the previous section, it can easily be modified or even replaced by another if this appears to be relevant. A special smartphone option for people experiencing problems in oral communication and only offering equivalents is also a possibility that is considered.

Apart from the mentioned space-consuming and data-reducing techniques, a few other techniques with other purposes also deserve to be highlighted here. These include improved search algorithms which, among other things, will allow the users to access words and word combinations that are not lemmatized. It is not envisaged, for instance, that all equivalents (used to explain L2), synonyms and antonyms are treated as lemmata, neither are the collocations, but the users will nevertheless be able to search for them and, thus, get access to useful data. Finally, the use of log files and interactive options is also foreseen in order to get feedback from the users with a view to continuously improving the dictionary both in terms of available lexicographical data and functionality.

11. Conclusions

This paper aims at sharing some reflections and discussing some steps in the process from idea to design of an online dictionary of Spanish and English

business communication. A meticulous, theory-based approach in this preparatory phase is crucial for guaranteeing a successful lexicographical product when it eventually is placed at the disposal of users on the Internet. Major mistakes and shortcomings will necessarily affect the final product as any major reprogramming of the underlying database will delay the project and challenge the project budget once the compilation process has started, and therefore only can be recommended in exceptional cases.

However, a well-considered dictionary design, or concept, is not in itself a guarantee of success. Just as important are meticulous efforts in the remaining phases of the overall dictionary project, cf. Tarp (2014). These phases include choosing the empirical basis and determining the different methods which the lexicographers should employ to select and prepare the data to be incorporated in the database as well as the compilation process itself. The selection of lemmata, equivalents, collocations, etc., will be based on a combination of existing business dictionaries, glossaries published by various relevant institutions, indexes in textbooks, and corpora — an empirical basis recommended by Bergenholtz and Tarp (1995) as the most appropriate to be used in specialized lexicography. By contrast, the definitions will be written based on expert knowledge and sometimes also on the material published by normative bodies such as the International Chamber of Commerce. All this will hopefully be treated in a later contribution if everything goes as planned.

12. References

12.1 Dictionaries

- Combley, Roz.** 2011. *Cambridge Business English Dictionary*. Cambridge: Cambridge University Press.
- Graham, James and George A.S. Oliver.** 1906. *The Foreign Traders' Dictionary of Terms and Phrases in English, German, French, and Spanish*. London/New York: Macmillan.
- López, Sinda and Donald Watt.** 2002. *Oxford Business Spanish Dictionary. Spanish–English, English–Spanish*. Oxford: Oxford University Press.
- Miles, Andy D.** 2012. *Business English to Spanish Glossary*. [<http://www.andymiles.com>]
- Mortimer, Thomas.** 1766. *A New and Complete Dictionary of Trade and Commerce*. London: Printed for the author, and sold by S. Crowder, and J. Coote, and J. Fletcher.
- Postlethwayt, Malachy.** 1751–1755. *The Universal Dictionary of Trade and Commerce*. London: John and Paul Knapton.
- Rolt, Richard.** 1756. *A New Dictionary of Trade and Commerce*. London: T. Osborne and J. Shipton.
- Savary des Bruslons, Jacques.** 1723. *Dictionnaire universel de commerce*. Paris: Chez J. Etienne.
- Tarp, Sven in cooperation with Birger Andersen, Karin Balsgart and Morten Pilegaard.** 1999. *Dansk–engelsk erhvervsordbog*. Aarhus: Systime.
- Tarp, Sven in cooperation with Birger Andersen and Karin Balsgart.** 2006. *Engelsk–dansk. Erhvervsordbog*. Copenhagen: Gyldendal.
- Tarp, Sven in cooperation with Virginia Hvid and Anne Lise Laursen.** 1999. *Dansk–spansk erhvervsordbog*. Aarhus: Systime.

- Tarp, Sven in cooperation with Virginia Hvid and Anne Lise Laursen.** 2003. *Spansk-dansk erhvervsordbog*. Copenhagen: Gads Forlag.
- Veitelle, I. de.** 1864. *Mercantile Dictionary: A Complete Vocabulary of the Technicalities of Commercial Correspondence, Names of Articles of Trade, and Marine Terms, in English, Spanish, and French, with Geographical Names, Business Letters and Tables of the Abbreviations in Common Use in the Three Languages*. New York: D. Appleton & Co.
- Virio, Juan Bautista de.** 1792. *Colección alfabética de los aranceles de la Gran Bretaña, y extractos de las Leyes, Reglamentos, Órdenes y Providencias expedidas en aquel Reyno para el régimen de sus Aduanas, y fomento de su comercio*. Vol. 1-4. Madrid: J. Ibarra.

12.2 Other literature

- Astigarraga, Jesús and Juan Zabalza.** 2007. Los diccionarios de Comercio y Economía en el siglo XVIII español. *Revista de Historia Industrial* 35: 13-46.
- Bergenholtz, Henning and Jesper Skovgård Nielsen.** 2013. What is a Lexicographical Database? *Lexikos* 23: 77-87.
- Bergenholtz, Henning and Sven Tarp (Eds.).** 1995. *Manual of Specialised Lexicography: The Preparation of Specialised Dictionaries*. Amsterdam/Philadelphia: John Benjamins.
- Besomi, Daniele (Ed.).** 2011. *Crises and Cycles in Economic Dictionaries and Encyclopaedias*. London/New York: Routledge.
- Caruso, Valeria.** 2011. Online Specialized Dictionaries: A Critical Survey. Kosem, Iztok and Karmen Kosem (Eds.). 2011. *Electronic Lexicography in the 21st Century: New Applications for New Users*: 66-75. Ljubljana: Institute for Applied Slovene Studies.
- Fuertes-Olivera, Pedro A. and Sven Tarp.** 2014. *Theory and Practice of Specialised Online Dictionaries: Lexicography versus Terminography*. Berlin/Boston: De Gruyter.
- Kwary, Denny A.** 2015. Challenges of Online Learner's Dictionaries to be Used in Smartphones. *Estudios de Lexicografía* 4: 199-206.
- Lew, Robert.** 2012. How Can We Make Electronic Dictionaries More Effective? Granger, Silvia and Magali Paquot (Eds.). 2012. *Electronic Lexicography*: 343-361. Oxford: Oxford University Press.
- Tarp, Sven.** 2004. Reflections on Dictionaries Designed to Assist Users with Text Production in a Foreign Language. *Lexikos* 14: 299-325.
- Tarp, Sven.** 2013. What Should We Demand from an Online Dictionary for Specialized Translation? *Lexicographica* 29: 146-162.
- Tarp, Sven.** 2014. Theory-based Lexicographical Methods in a Functional Perspective: An Overview. *Lexicographica* 30: 58-76.
- Tarp, Sven.** 2015a. On the Disciplinary and Functional Status of Economic Lexicography. *Ibérica* 29: 179-200.
- Tarp, Sven.** 2015b. Structures in the Communication between Lexicographer and Programmer: Database and Interface. *Lexicographica* 31. Forthcoming.
- Tarp, Sven and Theo J.D. Bothma.** 2013. An Alternative Approach to Enlightenment Age Lexicography: The Universal Dictionary of Trade and Commerce. *Lexicographica* 29: 222-284.