# A Computational Approach for Processing Diatopic Verbal Fixed Expressions

Arturo Velasco García
María J. Somodevilla García
Ivo H. Pineda Torres
Darnes Vilariño Ayala
Gloria Y. Torrealba Melendez
Concepción Pérez de Celis Herrero
Universidad de Puebla (México)

## ABSTRACT

One of the main problems in the task of a translator are the *fixed expressions* or *phraseological units* (*UFs*), is not possible to establish a meaning from its constituents, in addition to a grammatical structure which can move away from the language rules. Other problems associated with *fixed expressions* are their low representation in thesaurus and dictionaries, also *fixed expressions* have problems with indexing, search and management of their variants. In this paper an alternative method of processing and translate of *verbal locutions* and *verbal syntagms* into a diatopic system of the Spanish language is presented. The proposed approach incorporates computational techniques and tools such as a relational model, expansion and variation of the expressions using key-words and regular expressions to show of examples of use of these expressions in an electronic corpus.

## KEYWORDS

Corpus Linguistics, Diatopic Verbal Fixed Expressions, Relational Databases, Regular Expressions.

## I. INTRODUCTION

The translation process is the reproduction in a terminal language of the message from an original language through its closest and more natural equivalent, trying to preserve meaning and sense Nida, E. and Ch. R. Taber (1986). The objective of translation is that both texts can communicate the same message, at the same time considering aspects like textual gender, context, stylistic conventions, etc. García Y. (1984). For translators, the support of a dictionary and their experience are fundamental tools for the translation, but without doubts, there is an important part of a lexical system that presents difficulties at the moment of translation, this is the case for *fixed expressions*.

Differences in the treatment and not significant percentages of occurrences of these expressions in dictionaries, together with technical problems such as indexation and search of variants and synonym expressions, gave the motivation to develop a method to study how to translate *locutions* and *verbal syntagms*.

A prototype is presented as a versatile computational method in order to help a translator instead of using printed dictionaries.

Some of the problems of phraseological dictionaries are the lack of strict policies at the time of storage and indexing, search problem in absence of clear mechanisms of classification and low frequency appearance of variants and synonym expressions. In addition, examples of use in enriched dictionaries are at the expense of the linguistic competence of lexicographers.

Considering those problems mentioned above, it's proposed to develop a Diatopic Verbal Expressions Digital Dictionary (*DIVEDD*) for Spanish Language (diatopic subsystems of Spain and México) in order to enable the process of translation of verbal expressions in both subsystems. This prototype uses three methods for expansion of verbal expressions (verbal conjugations, regular expressions and keywords), generating synonym and variants expressions, showing through a Corpus examples of real use.

Based on the fact that a digital translator dictionary of diatopic verbal expressions in the Spanish language doesn't exist, gave the reason to come up with an automatic way to translate diatopic verbal expressions.

The second section describes related work with translation of *fixed expressions*; third section presents preliminaries to the work and the delimitation of the problem; in the fourth section the architecture and development of the *DIVEDD is* described; and finally, results, conclusions and ongoing work are presented.

## II. RELATED WORK

The group of *fixed expressions* constitutes an important part of the lexical system, where monolingual and bilingual dictionaries only capture certain number of units, and monolingual and bilingual dictionaries only form certain number of units, often reduced, to an alphabetical process of selection and random description Mogorrón H. (2004). In México there are not recent works of compilation of expressions, some of them are: the *Diccionario breve de mexicanismos*, Gómez de Silva G. (2001), the *Diccionario ejemplificado de mexicanismos*, Steel B. (2000), and the *Diccionario del español usual en México*, Lara L. (2003). The lack of strict rules at the time of integrate these dictionaries brought the introduction of different subsets of *fixed expressions*.

There are some works related with translation of expressions in the Spanish language such as *Recopilación de proverbios*, proverbs which were translated into four languages (English, French, German and Italian) by Casado, M. L., Agueda, S., Agueda, B. and Corral, J. (1998). In *Spagnolo-Italiano: Espressioni idiomatiche e proverbi*, there are a summary of idiomatic expressions, proverbs and Spanish and Italian pragmatic sentences. In Zamora M. (1997); were found *877 refranes españoles*, sayings with their correspondence Catalan, Galician, Basque, English and French by Sevilla M. and Cantera J. (1998). Finally, *Divergencias en la traducción de expresiones idiomáticas y refranes* by Sevilla M. (1999), provides a more systematic methodology for the translation of expressions between French and Spanish (Spain). This model of bibliographical record considers different uses, the level of the speaker's registration, antonyms, synonyms, source of the expression and examples among other data. This work considers *Divergencias en la traducción de expresiones idiomáticas y refranes* by Sevilla M. (1999) as a starting point by taking the benefits of a corpus showing use of actual situations.

## III. PRELIMINARIES

Due to the lack of agreement between linguists to establish limits of research of the phraseology and terminology used in this area, we decided to follow Alberto Zuluaga's work. Thus, the phraseology is defined as a branch of linguistics where the object of study is the speech repeated units, commonly known as *UFs* or *fixed expressions*. Zuluaga A. (1980) establishes the presence of two key requirements that *UFs* must have: fixation and idiomaticity. The fixation is a property that has the expressions of being reproduced in the speech like previously defined combinations, i.e. they present certain order in their syntactic structure, Zuluaga A. (1975). On the other hand, the idiomaticity is the property of the *UFs* whose sense cannot be deduced from the sum of the meanings of their constituent elements, Zuluaga A. (1980).

Other important characteristics that have the *UFs* mentioned by Santamaría I. (1998) related to Corpas Pastor are: high frequency of report of their constituent elements, absence of grammatical rules in the expressions and translation problems.

Zuluaga (1980) carries out a classification of the *UFs* based on the actions of the expressions in the speech (See Fig. 1). In the first group, Zuluaga establishes the *locutions* like a stable combination of two or more terms that work as an element in sentences to level of lexeme or syntagm. Inside this classification (*locutions*) he separates those that are in use as grammatical instruments and the expressions that possess semantic sense (*lexical units*). The subset of the *UFs* object of study in this work are the *locutions* and *verbal syntagms* whom belong to the units with lexical sense. The *verbal locution* is equivalent to lexemes, e.g.: *pasar a mejor vida* (to die) or *echar una mano* (to help) and the *verbal syntagm* are equivalent to syntagms e.g.: *pagar los platos rotos* (to suffer the consequences of something).

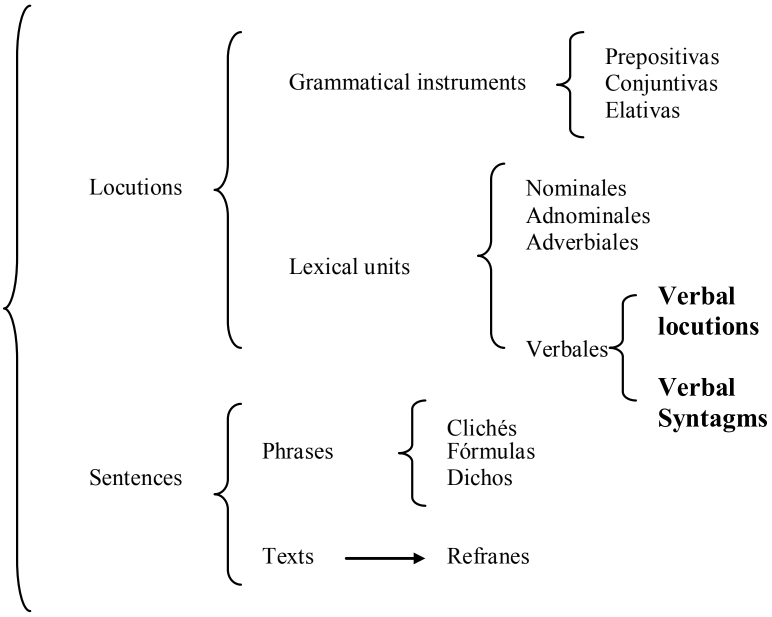

Fig. 1. *UFs* Classifications according to Zuluaga A. (1980).

## IV. PROTOTYPE ARCHITECTURE

The architecture proposed of the *DIVEDD* is organized in three modules: the database, that contains the essential characteristics of the verbal expression; the corpus, that contains in this first stage of digital texts and transcribed oral language; and the module of expansion of the expressions that is complemented with a list of stop-words and a database storing verbal conjugations. In the Fig. 2 the architecture of the *DIVEDD* is shown.

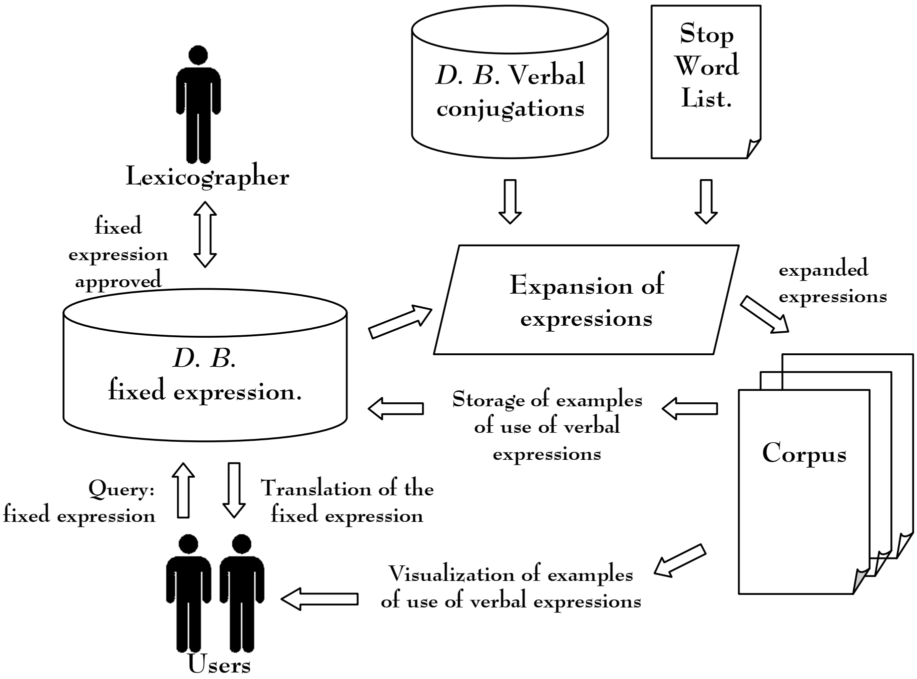
D. B. Verbal conjugations
Stop Word List.
Lexicographer
fixed expression approved
Expansion of expressions
expanded expressions
D. B. fixed expression.
Storage of examples of use of verbal expressions
Corpus
Query: fixed expression
Translation of the fixed expression
Visualization of examples of use of verbal expressions
Users

Fig. 2. Architecture of the *DIVEDD*.

## IV.1. The database

This module is based on a relational model that provides mechanisms that guarantee to avoid duplicity of records and inconsistency problems, it also guarantees the referential integrity and favors improvements of processing of the expressions. The Fig. 3 shows the Entity-Relationship Model corresponding to the *DIVEDD* ´s database, including the entities among that hold the verbal conjugations.

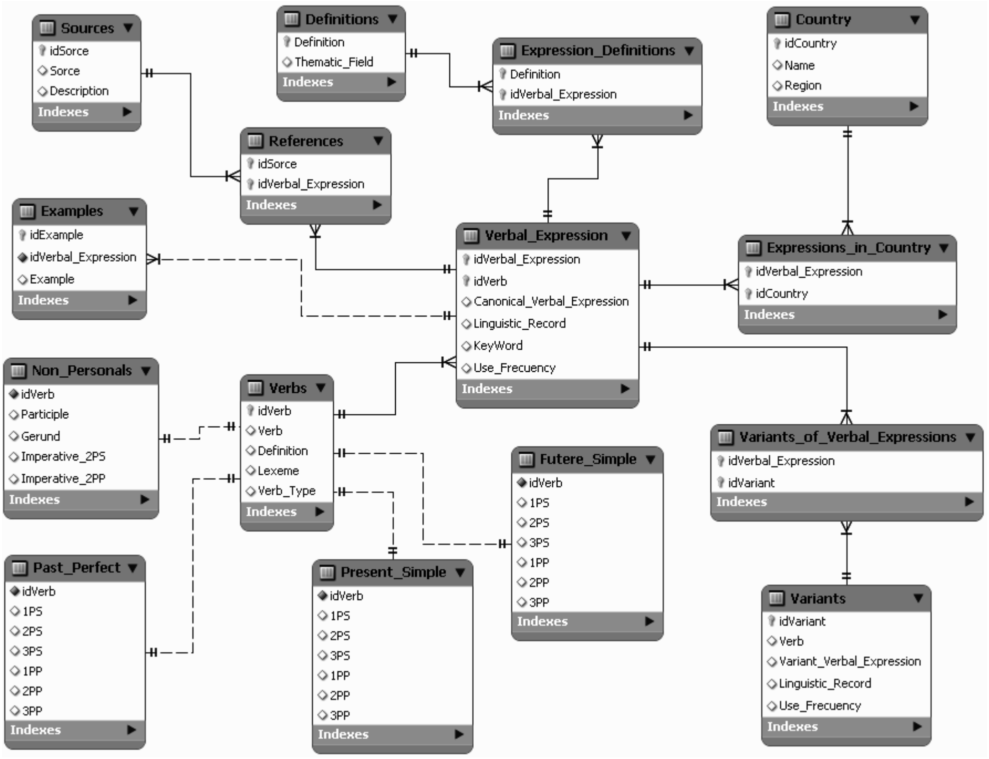

Fig. 3. Entity-Relationship model of the *DIVEDD* ´s database.

Table 1 shows the most important attributes of verbal expressions. This table does not correspond to any table generated from the E–R model in Fig. 3, only it is intended to be a descriptive character.

| ATTRIBUTE | DESCRPTION |
|---|---|
| *Verb* | Main verb used in the expression. |
| *Canonical_ Verbal_ Expression* | *Locution* or *verbal syntagm* in its canonical form. |
| *Definition* | Definition of the verbal expression. Field used to make the translation among the diatopic verbal expressions. |
| *Source* | Resource where the expression was extracted. |
| *Use_ Frecuency* | The number of frequencies of appearance of the expression in the corpus. |
| *Linguistic_ Record* | Level of registration of the expression. |
| *Country* | Country of origin of the expression. |
| *Region* | area or region of use of the expression. |
| *Thematic_ Field* | Thematic field of the expression |
| *Key-Word* | Alexical component unit of the expression. Useful to distinguish among synonym expressions. |
| *Variant_ Verbal_ Expression* | Field of the table *Variant* that stores the variants of the canonical verbal expressions. |
| *Example* | Field of the table *Examples* that stores the examples provided by the lexicographer and extracted of the corpus. |

Table 1. More important attributes of the verbal expressions.

### *IV.1.1. Handling of variants and synonyms in the database*

Zuluaga A. (1980) mentions the difference between variants and variations. The variants are presented because of a complete degree of fixation does not always exist, i.e. these units can vary or omit some of their constituents, also, Zuluaga points out that they should belong to the same functional language and do not to present semantics variation. The variations, can present semantics variation when changing some lexical element in diatopic, diaphasic or diastratic systems.

In this prototype a slight change to the definitions of Zuluaga was made: *Variants:* those that vary or omit any of its closed lexical elements without having semantic change. *Synonyms*: Those that have changed in their non closed lexical element i.e. key-word or those which do not contain any element in common, but they do not have a semantic change.

Thus, in the first group of Zuluaga´s classification, variants and synonyms were distinguished. In the second group, Zuluaga considers a semantic change, which will have to be reflected in a change in *Definition* of the table *Definitions*, so these expressions are different.

Table 2 shows an example of synonym expressions, therefore, the expressions: *ir al bote (in canned)*, *ir al tambo(up the river)*, *ir a la sombra (in the shadow)* and *ir tras las rejas(behind bars)* are synonyms, because their key-words changes but they have the same definition. The same situation applies to the expressions *hacer la barba*, *hacer la pelota* and *hacer la rosca*, but in addition, *hacer la barba* is the translation into Spanish (México) of the expressions of Spain *hacer la pelota* and *hacer la rosca*.

Synonyms

| *Canonical_ Verbal_ Expression* | *Definition* | *Key-word* | *Thematic_ Field* | *Linguistic_ Record* | *Country* |
|---|---|---|---|---|---|
| Ir al bote | Meter a alguien en la cárcel | Bote | Behavior | Informal | México |
| Ir al tambo | Meter a alguien en la cárcel | Tambo | Behavior | Informal | México |
| Ir a la sombra | Meter a alguien en la cárcel | Sombra | Behavior | Informal | México |
| Ir tras las rejas | Meter a alguien en la cárcel | Rejas | Behavior | Informal | México |
| Hacer la barba | Lisonjear a alguien | Barba | Behavior | Informal | México |
| Hacer la pelota | Lisonjear a alguien | Pelota | Behavior | Informal | Spain |
| Hacer la rosca | Lisonjear a alguien | Rosca | Behavior | Informal | Spain |

Table 2. Handling synonym expressions in the *DIVEDD*.

Table 3 shows the variants through regular expressions. A regular expression is a set of pattern matching rules encoded in a string according to certain syntax rules Stubblebine T. (2007). Thus, it is possible to describe or represent a set of strings without need to enumerate all of its elements.

| *Canonical_ Verbal_ Expression* | *Definition* | *Key-word* | *Variant_Verbal_Expression* |
|---|---|---|---|
| Ir al bote | Meter a alguien en la cárcel | Bote | [ir,llevar,meter] al {bote}<br>[refundir] en el {bote} |
| Ir al tambo | Meter a alguien en la cárcel | Tambo | [ir,llevar,meter] al {tambo}<br>[refundir] en el {tambo} |
| Ir a la sombra | Meter a alguien en la cárcel | Sombra | [ir,llevar,meter] a la {sombra}<br>[refundir] en la {sombra} |
| Ir tras las rejas | Meter a alguien en la cárcel | Rejas | [ir,llevar,meter,refundir] tras las {rejas} |

Table 3. Variants expressions in the database of the *DIVEDD*.

## IV.2. Corpus Processing for DIVEDD

The corpus of the *DIVEDD* is a textual, conformed by written language and transcribed oral language, Procházková P. (2006), based on the recommenda-tions of Sinclair J. (1996, 2004). The subcorpus of written language is built from digital Mexican newspapers and Mexican literature. On the other hand, the subcorpus of transcribed oral language is from the Sociolinguistic Corpus of the México City (*CSCM*), developed by Colegio of México.

The procedure carried out for preprocessing of both subcorpuses follows. Each step is identified with the prefix C (Corpus):

C1. Standardization of texts. Texts are transformed in ASCII files of plane text (. txt).

C2. Removal of symbols or non Alphabetical signs: 56 symbols were eliminated initially that the *Real Academia Española* presents in the appendix 4 (Lists of symbols or signs non alphabetical).

C3. Standardization of spaces and new lines: If the text presents more than one space or tabulator between words these are reduced to just one space. Also, spaces between lines or paragraphs are eliminated.

C4. Text indexation with verbs used in the document: The indexation allows to speed up the process of searching of examples of use of the expressions.

C5. Removal of labels such as Phonic-acoustic labels and other comments: This process only applies for the transcribed oral texts.

### IV.3. Module for expansion of expressions

This module controls creates all the possible variants and synonyms from a verbal expression. As an entry it receives the verbal expressions resulting as a of a user´s query to the database. The steps for this process are identified by the letter E (Expansion), to identify it from the subcorpus process in 4.2. The procedure is the following:

E1. Find the canonical synonym expressions of each expression generated by the query and to store them temporarily. The synonym expressions are extracted when the comparission in the *Definition* field of the table *Verbal_Expression* for both expressions is true.

$$VE = \{VE_1, VE_2, \ldots\ldots, VE_N\} \qquad (1)$$

E2. For each canonical expression stored (E1), find each one of their possible variants through of the field *Variant_Verbal_Expression*, of the table *Variants*.

$$VVE_i = VE_i + \{variant\} \qquad (2)$$

The system offers the option of eliminating the stop-words of verbal expressions. This process takes an input the group of expressions obtained in (E2), by eliminating in this way articles, prepositions, adverbs, etc. a stop-word list of 310 words, which was generated from 4 lists, was taken as a reference.

E3. For each expression (*VVEi*) generated in (E2), to substitute the verb of the expression in their canonical form, for each one of their stored conjugations. It is necessary mention that the number of conjugations stored for each verb is twenty-two. Four of them correspond to non personal forms (participle, gerund and second person of the singular and plural in the imperative form); the eighteen conjugations remaining belong together to the six conjugations of the indicative perfect past, present indicative and indicative simple future (to see Fig. 3).

E4. Each expression generated in step (E3) is taken and then and look for them in the corpus. In this process it is important to consider some characteristics of syntax in the text that could alter the expression.

The process of expansion of synonyms (step E1) is responsible for searching the entire database occurrences of expressions with the same definition. This method of linking synonym (to see Table 3) presents a drawback: the lexicographer must enter syntactically the same definition of a synonym expression previously stored in database. Therefore, we decided to validate an expression before is stored. Table 4 shows queries (*Qi*) performed to validate matches between expressions.

| | *Canonical_Verbal_Expression* | *Definition* | *Thematic_Field* | *Key-word* | Similarity | Message / Action |
|---|---|---|---|---|---|---|
| *Q1* | ✓ | ✓ | ✓ | ✓ | ★★★★★ | Expression already stored. Not save. |
| *Q2* | ✓ | ✓ | | | ★★★★☆ | Error, *Thematic_Field* or *Key-word*. Not save or edit. |
| *Q3* | | ✓ | | | ★★★☆☆ | To enter a synonym. Save. |
| | ✓ | | | | | Expression already stored. To save diatopic expression or with semantic differences. |

| | | | | | | |
|---|---|---|---|---|---|---|
| ***Q4*** | | ½ (nsw) | ( ✓ OR ✓ ) | | ★★☆☆☆ | To save an expression lightly different in syntax in the field *Definition* to another expression already stored. Save or edit. |
| | ½ (nsw) | | ( ✓ OR ✓ ) | | | To save an expression lightly different in syntax in the field *Canonical_Verbal_ Expression* to another expression already stored. Save or edit. |
| ***Q5*** | | ½ (sw) | | | ★☆☆☆☆ | To save an expression with a minimum coincidence in the field *Definition* with another expression already stored. |
| | ½ (sw) | | | | | To save an expression with a minimum coincidence in the field *Canonical_Verbal_Expression* with another expression already stored. |

Table 4. Similarity level among verbal expressions.

The symbol ✓, means exact match among the attributes of the expression to store with regard to those stored in the database. The number of shady ★ in Similarity (Table 4) represent the degree of similarity.

*Q1* shows those expressions stored in the database that match the following fields: (*Canonical / Variant*)*_Verbal_Expression*, *Definition*, *Thematic_Field* and *Key-word*. For example the following code *MySQL* corresponds to *Q1*:

```
"SELECT expresionf1.IdEF, EFCanonica, expresionf1.Definicion, CampTem, Clave
        FROM expresionf1, expresionf2
        WHERE expresionf1.Definicion = expresionf2.Definicion AND
                        expresionf1.Definicion = '$definicion' AND
                        expresionf1.EFCanonica = '$efcanonica' AND
                        Clave = '$clave' AND  CampTem = '$camptem'
                        ORDER BY Verbo";
```

*Q2* shows those verbal expressions stored in the database that match the following fields: (*Canonical* / *Variant*)*_Verbal_Expression* and *Definition*.

*Q3* shows those verbal expressions stored in the database that match the following fields: (*Canonical* / *Variant*)*_Verbal_Expression* or *Definition*.

*Q4* returns those verbal expressions stored in the database that match in the field *Definition* in half or more of the words of the definition, removing stop-words *(½ (nsw))*, and match with *Thematic_Field* or *Key-word*. Also, idem with *(½ (nsw))* in the field (*Canonical* / *Variant*)*_Verbal_Expression*.

*Q5* idem that the query *Q4*, except in this case is half or more than a half of the words *(½ (sw))* (counting the stop-words).

Storage of variants of a canonical verbal expression is done through regular expressions provided by the lexicographer. The syntax used is a subset of a standard notation for regular expressions. Table 5 shows basic syntax.

| Operator | Function |
|---|---|
| ‘[‘ y ‘]’ | Denotes the set of verbs that can be used in the expression. |
| ‘(‘ y ‘)’ | Denotes the set of connector–words between the verb and key-word. |
| ‘{‘ y ‘}’ | Denotes the set of key-words that are used in the expression. |
| ‘,’ | Separator of a set of words (verbs, key-words, connector–words). Can be use ‘,’ instead of ‘\|’. |
| ‘\|’ | Performs the same function as ‘,’. |
| ‘_’ | Joint two or more nonseparable words in an expression |
| ‘ ’ | The blank space denotes the separation between groups. |

Table 5. Operators of the regular expressions.

Considering the operators used in regular expressions specified in Table 5, the canonical verbal expression formed by *hablar más que un loro*, where the key-word is *loro* and the regular expressions are denoted by:

*[hablar,platicar] (más_que_un,como,como_un) {loro,perico,merolico}*

*[hablar,platicar] (más_que_una,como_una) {cotorra}*

The set of variants of *hablar más que un loro* are: *hablar como loro, hablar como un loro, platicar más que un loro, platicar como loro* and *platicar como un loro.*

The set of all its synonyms are: *hablar más que un perico, hablar como perico, hablar como un perico, platicar más que un merolico, platicar como merolico, platicar como un merolico, hablar más que una cotorra, hablar como una cotorra, platicar más que una cotorra* and *platicar como una cotorra*.

As it can be seen the properties of regular expressions help us to match any possible variation of the expressions in the corpus without necessity of having enumerated each one.

## V. RESULTS

The *DIVEDD* database was developed in *MySQL 5.0.18*. All the processing and corpus indexing are implemented with *AWK 3.1.3*. The module of generation of synonyms and variations was developed on *PHP 5.1.2* with the exception of step E4, which is implemented by an *AWK* script, which receives as input a list of all expressions generated by E3. Subsequently E4 creates a file for each expression in a list, where the examples are extracted from the corpus and stored. Finally, the examples of use are displayed, allowing the storage of them in the field *Example* of table *Examples*.

Initial tests of the prototype have been done using a sample of 5 interviews conducted by the Colegio of México to men between 20-34 years with a high educational level and all of them living in México City. The interviews turn out in conversations, which have different themes: friends, family, school, work, interests, etc. The size of the corpus under test is 38,197 words. In manual fashion the verbal expressions were identified in 5 interviews, later to generate synonyms and variations of expressions and the extraction of examples in the corpus were applied. The results are encouraging, because it could extract all the examples of use identified in a manual fashion considering a reasonable computational time.

On the other hand, the tests of translation have only been carried out between verbal expressions of Spain and México. The results were the expected ones and the problem to store other definition for each synonym expression it is validated by the lexicographer with support to the process of validation and compatibility of expressions (Table 4).

## VI. CONCLUSIONS AND ONGOING WORK

The *DIVEDD* appears to be a system for human translation assisted by computer, providing a definition and basic characteristics of verbal expressions. The *DIVEDD* does not try to be a detailed dictionary but it is as a mechanism reliable of storage of phraseological information enforcing structure and integrity of data, reducing times of search and translation. On the other side, the mechanisms of search expressions by expressions´ attributes and its combinations make the *DIVEDD* a flexible tool.

The absence in the database of verb conjugations in its subjunctive form is due to their low use frequency use in the corpus.

The use of a corpus as a support in the process of understanding of the expressions, sample seems to be significant in providing examples of use without the intervention of a lexicographer's linguistic.

The work of collecting articles from digital newspapers is still in process. In addition, a module for storing verbal expressions, for which there is not translation available, is under construction.

## NOTES

[1] In this work the term verbal expressions or simply expressions refer to both subsets of fixed expressions being studied in this work.
[2] This proposal can be extended to the translation of diatopic verbal expressions in general.
[3] Initially, a small Mexican corpus has only been gathered in which examples of actual use of expressions have been extracted.
[4] Referential integrity ensures that a record of the database is always linked with other valid records. This ensures that data are consistent, avoid unnecessary redundancy, missing data etc.
[5] Changes in the linguistic record of the expressions do not denote a change in semantic, i.e., verbal expressions can have the same meaning but in different spoken register.
[6] http://lef.colmex.mx/Sociolinguistica/CSCM/Corpus.htm
[7] http://buscon.rae.es/dpdI/apendices/apendice4.html
[8] http://ranks.nl/ (71 words); http://dot-seo.com/ (23 words); http://snowball.tartarus.org/ (52 words) and http://www.elwebmaster.com/ (164 words).
[9] Management system of relational database of free licence. It is the *DIVEDD* ´s search engine.

## BIBLIOGRAPHY

Casado, M. L., Agueda, S., Agueda, B. and Corral, J. (1998). Recopilación de proverbios. Alcobendas. ISBN: 9788471436450.

Franco C. (2008): Programa de Introducción a la Lingüística. Facultad de Cs. de la computación Puebla, México.

García Y. (1984): Teoría y práctica de la traducción. 2 vols. Madrid: Gredos.

Gómez de Silva G. (2001): Diccionario breve de Mexicanismos. 1a ed., México, FCE.

Lara L. (2003): Diccionario del español usual en México. 1a ed. ISBN: 9789681207045.

Mogorrón H. (2004). Los diccionarios electrónicos fraseológicos, perspectivas para la lengua y la traducción. Universidad de Alicante.

Mogorrón H. (2008). Análisis contrastivo multilingüe de las expresiones fijas: su traducción. Elaboración de corpus electrónico.

Nida, E. and Ch. R. Taber (1986): La traducción. Teoría y práctica. Madrid: Ediciones Cristiandad.

Procházková P. (2006). Fundamentos de la lingüística de corpus, concepción de los corpus y métodos de investigación con corpus.

Santamaría I. (1998): Tratamiento de las unidades fraseológicas en la lexicografía bilingüe español-catalán**.** ISSN 0212-7636, Nº 12.

Sevilla M. and Cantera J. (1998): 877 refranes españoles con su correspondencia catalana, gallega, vasca, francesa e inglesa. Madrid: EUNSA.

Sevilla M. (1999). Divergencias en la traducción de expresiones idiomáticas y refranes (francés-español).

Sinclair J. (1996): Preliminary Recommendations on Corpus Typology. EAGLES Document EAG-TCWG-CTYP/P.

Sinclair J. (2004): Developing Linguistic Corpora: a Guide to Good Practice Corpus and Text—Basic Principles.

Steel B. (2000): Breve Diccionario Ejemplificado de Mexicanismos.

Stubblebine T. (2007): Regular Expression, Pocket referente. O`really 2nd edition.

Zamora M. (1997): Spagnolo-italiano: espressioni idiomatiche e proverbi, Milano, EGEA.

Zuluaga A. (1980). Introducción al estudio de las expresiones fijas. Frankfurt: Peter Lang.

Zuluaga A. (1975): La fijación fraseológica en: Thesaurus XXX.