

Suffixal and Prefixal Morpholexical Relationships of the Spanish

Octavio Santana, José Pérez, Francisco Carreras, Gustavo Rodríguez

Departamento de Informática y Sistemas, Universidad de Las Palmas de Gran Canaria,
35017 Las Palmas de Gran Canaria, Spain
{OSantana, JPerez, Fcarreras, GRodriguez}@dis.ulpgc.es
<http://www.gedlc.ulpgc.es>

Abstract. This work is about derivational suffixes, endings and prefixes of the Spanish language, which are useful for the establishment of about 70000 suffixal and 11000 prefixal extended morpholexical relationships deduced from a corpus of 134109 canonical forms. A computational tool is developed capable of solving and answering to any morphological aspect of a Spanish word. The tool encompasses everything related with derivation, prefixation and other nearby aspects. It allows the recognition, the generation and the manipulation of morpholexical relationships of any word and of its related words, includes the recovery of all its lexicogenetical information until arriving at a primitive, the management and the control of the affixes in the treatment of its relationships, as well as the regularity in the established relationship.

1 Introduction

This work aims at obtaining a set of extended morpholexical relationships between Spanish words useful for automatic applications in natural language processing –in a synchronous study with automation on mind, formal or theoretical aspects may do not coincide with those strictly linguistic. There are Spanish words maintaining a strong functional and semantic relationship –the same appearing at the derivational or prefixal level–, that can not be taken as derivation or prefixation, although there is a formal relationship through other stages in the evolution of the languages, so it is indeed considered necessary to include them –*agua* with *acuoso*, *vejiga* with *vesical*, *conejo* with *cunicular*. This concept must be restricted to avoid arriving to the concept of related idea –which exceeds the objectives of this work, *blanco* with *albura*, *sólido* with *endurecer*, *niño* with *pueril*–, therefore a historic-etymological meeting criterion is applied. It is obvious that for the speaker *acuario*, *portuario* and *campanario* are all places equally related with *agua*, *puerto* and *campana* –it must also be so for the automatic data processing–; in order to solve the linguistic boundaries preventing to treat relationships beyond the strict derivation or prefixation, it is necessary to be located at a different level from the morphological level; thus the concept of morpholexical relationship is extended.

2 Lexicon

The corpus handled in this work has been created from: the *Diccionario de la Lengua Española* (DRAE), the *Diccionario General de la Lengua Española* (VOX), the *Diccionario de Uso del Español* (María Moliner), the *Gran Diccionario de la Lengua Española* (LAROUSSE), the *Diccionario de Uso del Español Actual* (Clave SM), the *Diccionario de Sinónimos y Antónimos* (Espasa Calpe), the *Diccionario Ideológico de la Lengua Española* (Julio Casares) and the *Diccionario de Voces de Uso Actual* (Manuel Alvar Ezquerro).

A canonical form is defined as any word with its own identity susceptible of enduring derivational or prefixal processes to form other words. Such a word could be formed from another by similar processes. In the reference corpus a canonical form is any entry word of consulted sources having own meaning –those entries that are appreciative forms of others and do not add any substantial meaning variation are discarded.

The universe of words analyzed in this work is composed of 136795 canonical forms, Figure 1.

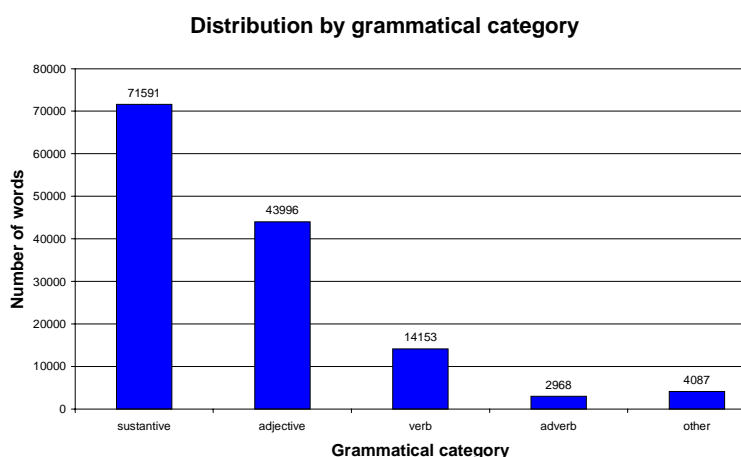


Fig. 1. Canonical forms universe considered by grammatical category

3 Suffixal Relationships

Spanish derivational processes consist fundamentally of suffixal modifications and usually, but not always, it implies a change of the derivative grammatical category with respect to its primitive canonical form.

Frequently, there are pairs of primitive forms in Spanish coming from a same mother tongue in which they went through a derivational process. The current relationship between the members of such pairs is considered as suffixal alteration, since in the current state of the Spanish language the existing relationship between them presents a strong parallelism with the derivational processes between Spanish forms, in its morphological aspects as well as in the semantic and grammatical aspects: for the verb *multiplicar* and the adjective *multiplicable* –both Spanish primitive words derived directly from the Latin– it is commendable to consider *multiplicable* as a deverbal adjectivation of *multiplicar*, Figure 2.

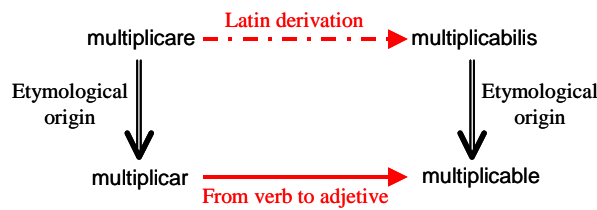


Fig. 2. Derivative process in mother language

There are many words in Spanish which come from a same mother tongue, where they suffered a derivational process from a common element that was never consolidated in the current Spanish; thus, they are etymologically related and are considered derivatives from a non-existent Spanish form, since they show a high similarity in the morphological, semantic and grammatical aspects with analogous Spanish derivatives from an existing form: between the forms *concupiscente* and *concupiscible* –both Spanish primitive words derived directly from the Latin– results feasible to consider an analogous relationships to those between any pair of Spanish deverbal adjectivations formed from a common primitive with the same endings, like *dirigente* and *dirigible* from *dirigir*, Figure 3.

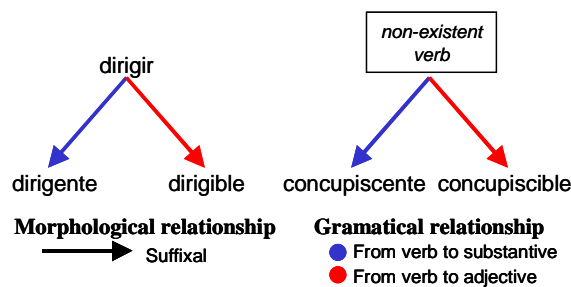


Fig. 3. Derivational process through a non-existent form

Other important aspect about the formation of Spanish words is to decide what words must be considered as primitives and what as derivatives –temporary line of

appearance of the words must be maintained. The main difficulty appears when there is ambiguity between two morphologically related primitives about their diachronic formation; in such case, and for synchronous treatment of the Spanish purpose, the verb, if any, is designed as primitive; when there is not a verb involved, the primitive is selected using the Spanish word formation rules supposing that words are morphologically related in equality terms: *lamento* is considered a deverbal substantivation of *lamentar* knowingly of the fact that both are primitive.

It is necessary to emphasize that there are Spanish words with a close functional and semantic relationship that cannot be directly established through a process classified within the derivational morphology. So, the concept of extended morphological relationship is incorporated and encompasses, in addition to the relationships produced by derivation, other characterized by having a meeting point in their etymological record and by incorporating an ending with the adequate semantic and functional contribution: *audición* is considered suffixally related with *oír*, although *audición* is primitive and possesses a different root from *oír*.

The extended morphological relationship includes to relate words through an ending graphically coinciding with a suffix if having the semantics and functionality corresponding to that suffix: that is the relationship between *miel* and *melaza* ('sediments of honey'), whose ending coincides with the suffix *-azo/a*, in spite of the different formative process; as in *seise* of *seis*, through the ending *-e*, being aware of the fact that it is regressive of the plural *seises* –in definitive, it is a noun closely related to the adjective *seis*.

The extended relationships also includes those cases in which the only impediment to establish it are some characters of the ending: between *diez* and *décimo* –by analogy to the relationship between *mil* and *milésimo*, among others, through the numerical ending *-ésimo*.

Although both the inflection and the appreciative derivation constitute suffixal morphological processes of interest for the development of automatic natural language processors, neither of these aspects will be dealt here, since they have been settled in FLANOM¹.

A suffix is “a phonic sequence added to the end of a word stem, prior to the desinences, if any, that lacks self existence outside the word system, it can not be joined to other morpheme to form a derivative, it is interchangeable with other suffixes and it can be added to other stems” [2]. Basing on the previous definition, the set of suffixes considered in this work to establish relationships can be enumerated –a sequence is considered as suffix if appears in three or more different words. Thus, endings like *-arra*, *-aste*, *-ello*, *-ingo*, *-uz* fulfil the described suffix definition although their suffixal condition in the grammatical sense is questionable.

An original word can be any Spanish word with self identity that admits suffixes addition to obtain another related word. Deverbal regressions with addition to the

¹ FLANOM: *Flexionador y Lematizador Automático de Formas Nominales*. Developed by the Grupo de Estructuras de Datos y Lingüística Computacional of the Las Palmas de Gran Canaria University, available on Internet.

endings *-a*, *-e*, *-o* and zero-suffix or empty suffix, as well as plurals and appreciative suffixes with consolidated meaning are also considered in this work. Suffixal elements –suffixes with a strong semantic load on the original word or those possessing self existence– are left for another work, since they are part of the research of composition.

Even though it is true that most of the relationships between words coincide with a formal derivation –mainly the regulars–, the spelling coincidence with a concrete suffix can lead to of relationships detached from such linguistic concept, at least by means of that suffix, but covering the proposed objective. For example, the words *psiquiatra* and *psiquiátrico* are related to *psiquiatría* through the endings *-a* and *-ítrico*, in spite of being aware of the existence of *-iatría* and *-iatra*, the relationship is intended to be established between previously mentioned words and not with a possible root *psiqui-*.

A related word can also be original to another by adding a new suffix: *quej-a • quej-ica • quejic-oso*. The arrow must not to be interpreted as derivation in a strict sense, but as an extended morpholexical relationship between two words: **original** and **related**.

Changes affecting the root such as erudite words relationships or foreign languages influences, among other, are considered irregularities, although they are not pure linguistic derivation, since it notably enriches computer applications. The possible suffix combinations which lead to a relationship are also considered as irregularities, as long as the previously explained continuity criterion cannot be established due to the nonconsolidation or nonexistence of the intermediate form in the corpus. Appearances of interfixes and infixes are considered irregularities.

Now the alphabetical list of studied suffixes is presented together with their corresponding number of extended morpholexical relationship: *-a* (1419), *-áceo/a* (190), *-acho/a* (35), *-aco/a* (27), *-'ada*, *-'ade* (9), *-'ago/a* (11), *-'aino/a* (16), *-'aje* (315), *-'ajo/a* (65), *-al* (1574), *-ales* (6), *-allo/a*, *-alle* (36), *-amen* (34), *-án/a* (38), *-áneo/a* (18), *-ano/a*, *-iano/a* (732), *-año/a* (36), *-anza*, *-enza* (127), verbalizer *-ar* (2635), *-ar* (407), *-ario/a* (532), *-astro/a*, *-astre* (27), *-atario/a* (36), *-ate* (33), *-átíl* (16), *-ato/a* (294), *-avo/a* (29), *-az* (22), *-azgo/a*, *-adgo/a* (100), *-azo/a* (562), *-azón* (75), *-bilidad* (263), *-able*, *-ible* (1088), *-abundo*, *-ebundo*, *-ibundo*, (13), *-ación*, *-ición* (2637), *-'culo/a*, *-áculo/a*, *-ículo/a* (41), *-dad*, *-edad*, *-idad* (1095), *-adero/a*, *-edero/a*, *-idero/a*, (650), *-adizo/a*, *-edizo/a*, *-idizo/a* (118), *-ado/a*, *-ido/a* (16320), *-ador/a*, *-edor/a*, *-idor/a*, (3034), *-edumbre*, *-idumbre* (19), *-adura*, *-edura*, *-idura* (789), *-aduría*, *-eduría*, *-iduría* (42), *-e* (542), *-ear* (1786), *-eser* (102), *-eco/a* (32), *-edo/a* (62), *-ego/a*, *-iego/a* (68), *-ejo/a* (86), *-el* (57), *-elo/a* (63), *-én* (17), *-enco/a*, *-engo/a* (30), *-eno/a* (84), *-eño/a* (381), *-ense*, *-iense* (417), *-enta* (16), *-ento/a*, *-iento/a*, *-ulento/a*, *-olento/a* (125), *-eo/a* (152), *-'eo/a* (128), *-er*, *-ier* (29), *-ería* (881), *-erio* (24), *-erío* (55), *-erizo/a* (23), *-erno/a* (12), *-ero/a* (3151), *-és/a* (207), *-esa* (29), *-ésimo/a* (36), *-ete*, *-eto/a* (394), *-euta* (14), *-ez/a* (374), *-ezno/a* (13), *-grama* (51), *-í* (49), *-íaco/a*, *-iaco/a* (72), *-icio/a*, *-icie* (78), *-ico/a* (35), *-'ico/a* (1690), *-'ide* (18), *-'ido/a* (156), *-ificar* (135), *-'igo/a* (18), *-iguar* (4), *-ijo/a* (59), *-'il* (16), *-il* (155), *-illo/a* (797), *-imonio/a* (6), *-ín/a* (241), *-ina* (174), *-íneo/a* (26), *-ing* (15), *-ingo/a* (13), *-ino/a* (526), *-iño/a* (16), *-'io/a* (499), *-ío/a* (629), *-ión* (279), *-ir* (25), *-is* (9),

-ismo (1365), -ista (1325), -ístico/a (122), -ita (66), -ita (83), -ito/a (258), -'ito/a (18), -itud (57), -ivo/a, -ativo/a, -itivo/a, (692), -izar (554), -izo/a (95), -ma, -ema (20), -ambre, -imbre, -umbre (23), -mente (2432), -amento, -imento, -amiento, -imiento (1876), -ancia, -encia (532), -ander/a, -endero/a (28), -ando/a, -endo/a, -iendo/a, -ondo/a, -iondo/a (87), -ante, -ente, -iente (1617), -o (2740), -ol (17), -olo/a, -ol (73), -ón/a (1334), -ongo/a (15), -or/a (286), -'ora (4), -orio/a (89), -oso/a (1286), -ote/a, -oto (125), -arro/a, -orro/a, -arrio/a, -orrio/a (55), -s (46), -asco/a, -esco/a, -isco/a, -usco/a (213), -tano/a, -itano/a, -etano/a (44), -aticio/a, -iticio/a (15), -'tico/a, -ático/a, -ético/a, -ítico/a, -ótico/a (382), -atorio/a, -itorio/a (298), -triz (28), -ucho/a (23), -uco/a (17), -udo/a (238), -uelo/a (134), -ujo/a (17), -'ulo/a (45), -uno/a (60), -ura (202), -uro (22), -uto/a (33) and other lower frequency endings.

4 Prefixal Relationships

Prefixation does not produce grammatical category change; normally, it shades, amends, modifies, in short, it guides the meaning of the word. In addition to the prefixation through traditional elements, there is the certain composition through prefixal elements, among other forms of composition –these elements will not be taken into account in this work due to the strong semantic load they provide.

A prefixal morpholexical relationship between a pair of primitive forms is established when, formally, one could be obtained by prefixation from the other and it shows semantic relationship inherent in the prefix –*culpar* and *inculpar* came directly from the Latin *culpare* and *inculpare* but one can be obtained from the other by adding the prefix *in-*.

Independently from the existing suffixal morpholexical relationships, a prefixal relationship between two original words is projected to the corresponding derived pairs –the prefixal relationship between *amortizar* and *desamortizar* is applied to *amortización* with *desamortización* and to *amortizable* with *desamortizable*, Figure 4.

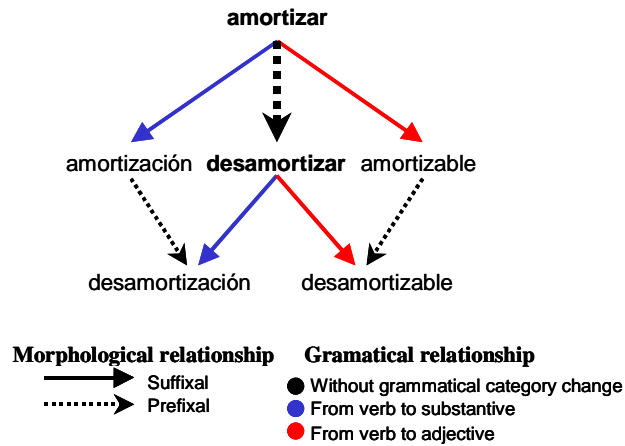


Fig. 4. Prefixal morpholexical relationship projection

Apart from the prefixal morpholexical relationships, the suffixal morpholexical relationships of a word are projected to their corresponding prefixed forms, since their mutual morphological, semantic and grammatical relationships are equivalent –*sobrecalentamiento* and *sobrecalentar* are prefixal forms between which the same suffixal morpholexical relationship existing between *calentamiento* and *calentar* is established, Figure 5.

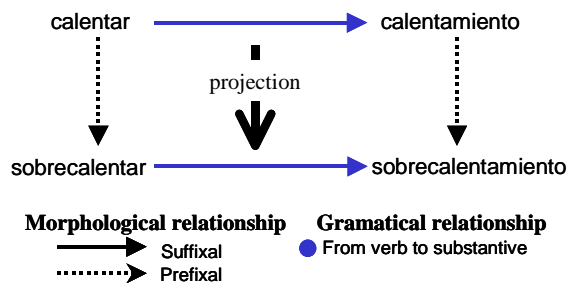


Fig. 5. Suffixal morpholexical relationship projection

The same considerations as in the suffixal alterations are applied to extend the concept of prefixal morpholexical relationship: *coepiscopo* is considered to be prefixally related with *obispo*, though *coepiscopo* is primitive and, from a synchronic perspective, it possesses a different base from that of *obispo*.

Extended morpholexical relationship includes to relate words through a start graphically coinciding with a prefix if having the semantics and functionality corresponding to that prefix: a relationship is established between *eficiente* and *deficiente*,

because the initials of the second word coincide with the prefix *de-* and it verifies with the rest of the criterion with respect to the first one.

Also, other graphic elements providing some type of relationship between words are studied, such is the case of the article of Arabic origin *al-* that permits to establish extended morpholexical relationship between *juba* and *aljuba* as variants one from the other.

The relationship between an original word and a prefixed form is semantic, functional and formal. In the semantic aspect, the prefix incorporates its own specific nuance, but the main semantic load corresponds to the original word. Syntactic and grammatical functions tend to be maintained in the prefixed form, though eventually it might produce a grammatical category change, due more to use than to the prefixational process. The formal differences are adjusted to the general rules of prefixation and to those which are specific to each prefix, and, though irregularities exist, they have a much lower incidence than that of alterations in other positions of the lexical base.

Prefixal element, henceforth prefix, can be defined as an affix attached to the front of the original word. This definition is so wide that produces two problems when delimiting what prefixes must be considered in this work without approaching composition.

Some prefixes have their own functional identity in Spanish –generally prepositions and some adverbs–, they are called vulgar prefixes and some authors consider them compositional elements. However, a prefix coming from one preposition can be detached from this since it does not fulfil the same grammatical function, though sometimes they agree in meaning. The semantic contribution of prepositions is a nominal subordination and that of prefix is a semantic addition and the similarity between them is reduced at present, and in a synchronic study, to a phonetic issue. In this work, all prefixes of this type are treated in a similar way –each them is accompanied with the cardinality of its extended morpholexical relationships–: *a-* (615), *ab-* (29), *ad-* (52), *al-* (58), *ana-* (14), *ante-* (106), *con-* (516), *contra-* (313), *cuasi-* (6), *de-* (287), *des-* (1 815), *e-* (50), *en-* (228), *entre-* (136), *ex-* (85), *extra-* (64), *in-* (1 317), *para-* (61), *por-* (3), *post-* (89), *pro-* (102), *sin-* (28), *so-* (91), *sobre-* (254), *sota-* (14), *sub-* (249), *ultra-* (64).

The second issue under discussion is what prefixes of the erudite prefixes, prefixoids, prefixal elements or compositive elements should be included. It is opted for generously studying those changing the meaning of the term to which they are joined, in an objective or subjective way, and those providing a pronominal or adverbial sense to the base. They do not appear as independent terms, and their semantic value are generic and applicable to any grammatical category. The prefixes of this type are: *abiso-* (2), *acro-* (4), *ambi-* (4), *anfi-* (6), *auto-* (185), *bar-* (7), *bati-* (5), *bi-* (111), *circa-* (0), *circun-* (16), *di-* (87), *diali-* (3), *ecto-* (3), *endo-* (27), *equi-* (14), *eu-* (3), *exo-* (18), *hemi-* (7), *hetero-* (32), *hiper-* (97), *hipo-* (72), *infra-* (21), *iso-* (28), *macro-* (44), *maxi-* (3), *mega-* (24), *meso-* (15), *meta-* (33), *micro-* (116), *mini-* (58), *mono-* (64), *multi-* (54), *omni-* (10), *opisto-* (1), *pan-* (44), *pen-* (4), *pluri-* (11), *plus-* (4), *poli-* (49), *preter-* (5), *proto-* (23), *retro-* (33), *semi-* (159), *super-* (220), *supra-* (21), *tele-* (70), *uni-* (23), *vice-* (42), *yuxta-* (3). Erudite prefixes whose semantic

contribution is strong are discarded: *bio-* (life), *foto-* (light), *metro-* (measure), among other, and apocopes acting as prefixal elements, like *auto-* (of *automóvil*), *tele-* (of *televisión*) among other.

Of course, prefixes that the employed sources define as such are considered too: *anti-* (416), *apo-* (12), *archi-* (44), *cachi-* (10), *cata-* (8), *cis-* (6), *citra-* (1), *dia-* (21), *dis-* (86), *epi-* (36), *es-* (82), *inter-* (161), *intra-* (35), *ob-* (20), *per-* (74), *peri-* (20), *pre-* (243), *re-* (995), *requete-* (3), *res-* (9), *tatara-* (3), *trans-* (258), *za-* (7).

A related word can be original to another by adding a new prefix: *emitir* • *transmitir* • *retransmitir*. The arrow must not be interpreted as prefixation in a strict sense, but as an extended morpholexical relationship between two words: **original** and **related**.

Changes affecting the root such as erudite words relationships or foreign languages influences, among other, are considered irregularities, although they are not pure linguistic prefixations, since they notably enrich computer applications. They are also treated as irregularities the prefix combinations causing a relationship, when the previously explained continuity criterion cannot be established due to the nonconsolidation or nonexistence of the intermediate form in the corpus.

5 Extended Morpholexical Relationships Organization

The joint formed by an original word and all its morpholexically related words is designated as **family**. Since a word can be related to an original word and at the same time to be original word in relationships linked with other words, kinship relationships between different families are established through this word. All the families related in this way compose a **clan**.

In order to represent the different types of relationships produced by the rules to form Spanish words and by the applied extended criteria, a directed graph has been chosen; nodes identify the Spanish words, edges express existence of extended morpholexical relationships, the direction of each edge determines the relationship between the nodes and the edge labels classify the type of extended morpholexical relationship. Spanish words become grouped into disjoint sets of mutually related elements –connected components of the graph, or clans–, Figure 6.

Graph of extended morpholexical relationships

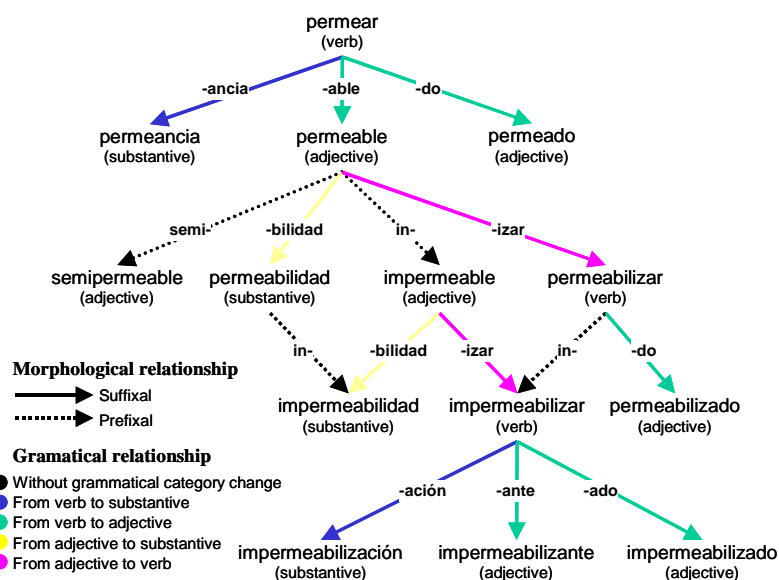


Fig. 6. Extended morpholexical relationships clan of *permeable*

6 Navigation on the Extended Morpholexical Relationship Graph

The graph can be gone through in any direction: from one particular node, every other node in the same connected component of the graph can be reached, knowing at any time the extended morpholexical relationships –crossed edges– until arriving at the destination. Depending on the direction chosen –ascending, horizontal or descending– the words' morphology and the distance between them are categorized. From one word, it is possible to obtain those which have suffered fewer formative processes –ascending–, those which have suffered the same number of alterations –horizontal– and those which have suffered more formative processes –descending.

6.1 Derivation

Derivation is the process by which words are formed, or are related through the extended criterion, by alteration of the structure or meaning of others –original–, generally with a different grammatical category, though they may be obtained from others with the same functionality. By descending the graph one level through suffixal

edges, derivatives of the wished grammatical category are found. Substantive, adjective, verbal and adverbial derivation are envisaged.

6.2 Direct Ancestry

Direct ancestry permits to obtain the original word with which a concrete word has been related –the inverse to derivation or prefixation. Ancestors are found ascending the graph one level. Thus, the direct ancestry of the verb *tutear* is the personal pronoun *tú*, that of verb *preconcebir* is the verb *concebir* and in the clan of *permeear*, the direct ancestry of the substantive *permeabilidad* would be the adjective *permeable*. If the direct ancestry process is applied twice, the original from the original from the current node is obtained; so, the direct ancestor in two levels of the verb *permeabilizar* is *permeear*; on the other hand, the verb *tutear* has not this option because the pronoun *tú* is the root of the graph.

6.3 Indirect Ancestry

Indirect ancestry gets the morphologically related words with the direct ancestors and that are found at the same level on the graph. Words of the clan having suffered one alteration less than the current word can be obtained in this way. In the clan of *permeear*, the indirect ancestors of the adjective *impermeable* are the adjective *permeado* and the substantive *permeancia*. As occurs with direct ancestry, it is possible to navigate other levels of morphological relationships: second level indirect ancestors of *permeabilizado* are the substantive *permeancia* and the adjective *permeado*, the same results as with only one level for *permeabilidad*.

6.4 Horizontality

Words morphologically related with the same original word constitute the horizontal direction –they are words with the same number of alterations. They are achieved recovering the direct ancestors and descending one level through all the edges of that node. This option recovers all members of a family from one of them –it does not include the original word. From the adjective *impermeabilizante*, the adjective *impermeabilizado* and the substantive *impermeabilización* are obtained.

6.5 Descendants

Descendants from a word are the other members of the family for which it is the original word. The level two descendants are the descendent family of each one of the members of the family of the original word. In the clan of *permeear*, the descendants from the adjective *permeable* are the substantive *permeabilidad*, the verb *permeabilizar* and the adjectives *impermeable* and *semipermeable*. The level two descendants

from the adjective *permeable* are the adjective *permeabilizado*, the substantive *impermeabilidad* and the verb *impermeabilizar*.

7 Filters

Filters in the extended morpholexical relationships permit selective discrimination of navigation response. Results of different types of navigation are susceptible of being submitted to various diverse nature filters –functional, regularity and by affixes.

7.1 Functional

It consists of accomplishing a selection by grammatical category. Thus, in the clan of *permeable*, the only one substantive descendent from the adjective *permeable* is *impermeabilidad*.

7.2 Regularity

It consists of selecting the regular morpholexical relationships or the irregular morpholexical relationships that maintain the words with respect to the original word. If wanting to explore the irregular horizontal relationships of one word, horizontal navigation and the irregular filter are applied.

7.3 Affixal

The affixal filter makes the selection according to the affixes establishing the morpholexical relationships –discrimination by one or more affixes can be applied.

8 Application

The application interprets and handles with versatility the most relevant aspects of the extended morpholexical relationships. It represents a form of showing the system power, without damage of its integration in other useful natural language processing tools. The interface of *RELACIONES*, Figure 7, facilitates exploration of the extended morpholexical relationships of a canonical form of any Spanish word. The *Entrada* permits the user to introduce any word. As a result of recognition, only canonical forms that possess extended morpholexical relationships are located in the *Forma canónica* combo box –it permits to request its relationships.



Fig. 7. Extended morpholexical relationships interface

A set of check boxes permits navigation by the clan of related words while filtering the results. The related words are shown on the list boxes located on the right of the interface which are organized by grammatical category.

Check boxes grouped under *Regularidad* filter the response according to the regularity: regular relationships, irregular relationships or both.

Dirección and *Profundidad* check boxes establish which words morphologically related with the canonical form are shown. These two check boxes groups are linked since when establishing the search direction on the graph, it is needed to specify its depth level. The result shows the union of all related words for each one of the indicated directions with each one of the depth levels chosen –they are classified in list boxes by grammatical categories.

Three buttons appear under each grammatical category list box –*Todos*, *Puros* and *Mixtos*– which permit to select the words of that list only by the list grammatical function –*Puros*– or those having another grammatical function in addition to the one defines by the list –*Mixtos*–, or well to show all words having at least the grammatical function defined by the list –*Todos*.



Fig. 8. The interface of extended morphological relationships: Suffix

In the interface of *RELACIONES*, Figure 8, appears the *Sufijos* tab sheet –it permits to configure suffixal filter. The considered suffixes are shown classified by the grammatical category that they produce and by its appearance frequency. Each group appears alphabetically ordered to facilitate location. The *Otros* tab sheet collects the verbalizer suffixes, adverbializer suffixes and others not easily classifiable irregular suffixes –they do not appear by appearance frequency.

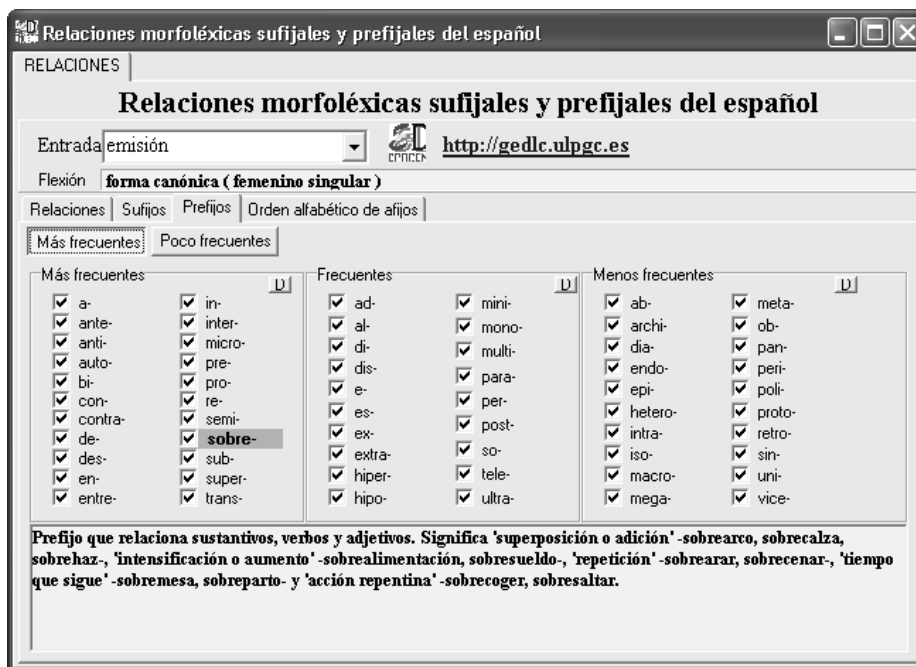


Fig. 9. The interface of extended morpholexical relationships: Prefix

In the interface of *RELACIONES*, Figure 9, appears the *Prefijos* tab sheet –it permits to configure prefixal filter. The considered prefixes are shown classified by its appearance frequency. Each group is alphabetically ordered to facilitate location.

References

1. Alarcos Llorach, E.: *Gramática de la Lengua Española*, Espasa-Calpe, Madrid, Spain (1995)
2. Almela Pérez, R.: *Procedimientos de formación de palabras en español*, Ariel, Barcelona, Spain (1999)
3. Alvar Ezquerro, M.: *La formación de las palabras en español*, Cuadernos de lengua española, Arco/Libros, Madrid, Spain (1993)
4. Bajo Pérez, E.: *La derivación nominal en español*, Arco/Libros, Madrid, Spain (1997)
5. Bosque, I., Demonte, V., Lázaro Carreter, F.: *Gramática descriptiva de la lengua española*, Espasa, Madrid, Spain (1999)
6. Dee, J. H.: *A lexicon of latin derivatives in Italian, Spanish, French and English*, “Vol. I Introduction and Lexicon”, Olms-Weidmann, New York, USA (1997)
7. Dee, J. H.: *A lexicon of latin derivatives in Italian, Spanish, French and English*, “Vol. II Index”, Olms-Weidmann, New York, USA (1997)
8. Faitelson-Weiser, S. 1993. “Sufijación y derivación sufijal: sentido y forma”, *La formación de palabras*, Varela (ed.), Taurus, Madrid, Spain (1993)

9. García-Medall, J.: *Formaciones prefijales en español: morfología derivativa del verbo*, Ph. Degree Thesis, University of Valencia, Valencia, Spain (1991)
10. Lang, Mervyn F.: *Formacion de palabras en español. Morfología derivativa productiva en léxico moderno*, Cátedra, Madrid, Spain (1992)
11. Malkiel, Y.: “El análisis genético de la formación de palabras”, *La formación de palabras*, Soledad Varela (ed.), Taurus, Madrid, Spain (1993)
12. Moreno De Alba, J. G.: *Morfología derivativa nominal en el español de México*, Autonomous National University of Mexico, Mexico (1993)
13. Santana, O., Hernández, Z., Rodríguez, G., Pérez, J., Carreras, F., Bogliani, S.: “Reconocedor automático de formas verbales que trata conjugación y pronombres enclíticos”, *Lingüística Española Actual*, 16-1, Arco/Libros, S.L. (1994) 125–133.
14. Santana, O., Pérez, J., Hernández, Z., Carreras, F., Rodríguez, G.: “FLAVER: Flexionador y lematizador automático de formas verbales”, *Lingüística Española Actual*, 19-2, Arco/Libros, S.L. (1997) 229–282
15. Santana, O., Pérez, J., Carreras, F., Duque, J., Hernández, Z., Rodríguez, G.: “FLANOM: Flexionador y lematizador automático de formas nominales”, *Lingüística Española Actual*. 21-1, Arco/Libros, S.L. (1999) 253–297
16. Soledad Varela (ed.): *La formación de palabras*, Taurus, Madrid, Spain (1993)
17. Varela, S., Martín, J.: “La prefijación”, I. Bosque, V. Demonte (eds.), *Gramática descriptiva de la lengua española*, Espasa-Calpe, Madrid, Spain (1999) 4993–5040