

Why teach dialectalisms in the foreign language classroom?

Antonio Manuel Ávila-Muñoz

University of Malaga, Spain

amavila@uma.es

Abstract

Dialectal vocabulary can be an interesting pedagogic tool in foreign language teaching. Accepting the need for a “common” language model, we emphasize student contact with dialectal vocabulary. First, the morphological characteristics of such vocabulary will bring students closer to the basic morphological processes of the target language. In addition, such contact prompts the formation of solid, morphological consciousness. Second, the semantic-cognitive properties of this vocabulary contribute to deepening the knowledge and accurate interpretation of both close cultural models and the mental representations of native speaker communities. Our model, which is based on concentric crowns represents a fresh proposal for selecting vocabulary adapted to each proficiency level in the TL. It integrates elements ranging from those most frequently used in any interaction to those considered peculiar to geographic areas, thus including lexical items used in frequent communicative situations.

Keywords: lexical selection; lexical frequency; lexical availability; dialectisms; selection model and teaching foreign language vocabularies

1. Introduction

Choosing the most appropriate linguistic variety to teach foreign languages has given rise to lively debate and considerable controversy among all the actors implicated in the process. It appears logical to have a language model that, while

recognizing and respecting the diatopic, diastratic and diaphasic varieties expressed, offers students the possibility of communicating anywhere in the world where the target language (TL) is used. In the case of languages widely spoken both geographically and socially, like Spanish and English, the correct choice is not an easy task once it becomes clear that this choice may be more of a “*desideratum*” than a reality (Moreno Fernández, 2000).

In this paper, we argue that foreign language students should have contact with what is considered dialectal vocabulary. Acknowledging the need for the language model to be as “common” as possible, our proposal is based on the belief that the morphological characteristics of TL vocabulary in every geographic variety would bring the student closer, in an entertaining way, to basic morphological processes and foster the development of a solid morphological consciousness. By the same token, the semantic-cognitive properties of this type of vocabulary could contribute to greater knowledge and correct interpretation of close cultural models and mental representations made during the learning process. Although dialectal vocabulary has been frequently used in foreign language classrooms, when such vocabulary is coherently and systematically integrated, it can become a powerful and entertaining pedagogical tool.

We use as a point of reference two previous studies concerned with: (1) determining the degree of compatibility of the vocabulary regarding specific groups of experiences that all learners have (Ávila-Muñoz & Sánchez Sáez, 2014) and (2) the morphological and semantic-cognitive nature of dialectal vocabulary (Ávila-Muñoz, 2017). On the basis of these we propose a vocabulary selection model represented in the form of concentric crowns or halo. Overall, the proposal shows the benefits and usefulness of combining common vocabulary and dialectal vocabulary in foreign language classes.

The most frequent vocabulary, independent of situational factors, is placed in the center of the crown and can be updated in any communicative exchange. This core group comprises grammatical words (essentially conjunctions, articles, prepositions, pronouns, adverbs). Around this permanently active central crown, different crowns composed of vocabulary are updated depending on the situational and personal thematic circumstances (Carter, 1987). These more specific lexical items can be selected for each level with the help of appropriate procedures.

2. Linguistic varieties in the foreign language classroom

In the field of foreign language teaching the most accepted and generalized trend has favored the creation of a teaching model using the standard variety at almost all levels. In practice, teachers rarely allow themselves the opportunity

to present their own social or native dialectal varieties and if they choose to do so, it generally happens only at advanced levels. However, the decision to choose the standard variety invites a series of conflicts that are not discussed in depth here, except mentioning something that often goes unnoticed, that is, the coexistence of different forms of prestige present in linguistic use. Among these, the standard variety only represents standard prestige and is not always the most functional in the communication process when compared to regional, covert vernacular prestige types (Villena Ponsoda, 2001, pp. 88-98). It must be remembered that the standardization process of a linguistic variety entails the selection – for non-linguistic reasons – of a variety existing at a given moment. The mere selection of only one variety implies discrimination of the non-selected varieties and, through regular processes (normalization, diffusion), will convert the chosen variety into a standard from then on. In the modern history of Europe, the general socio-linguistic situation has been marked precisely by the coexistence of, on the one hand, traditional dialects and, on the other, standard national varieties, as shown in studies on dialectology. Both models have been guarded zealously by individual speakers and speech communities. It is widely accepted that the dialectal variety has been considered of low prestige while the standard variety is the high prestige code and that rigid separation exists between the two structures in terms of use and mastery. This situation has been widely encouraged in the world of foreign language teaching, and has possibly been inherited from the tradition dominating the teaching of mother tongues. Therefore, it is easy to understand why the use of varieties considered standard has been encouraged and why, more or less intentionally, dialectal varieties have typically been stigmatized. However, if certain communicative skills are to be encouraged in the foreign language classroom, the extended consideration of the concept of linguistic variety must prevail over simpler and reductionist but at the same time traditional proposals.

On other occasions, the convenience of using the so-called “neutral” variety of the TL has been proposed for foreign language classes. In the field of translation, dubbing, subtitling and publicity this has been the variety most widely used – although not without controversy – with the aim of suppressing any differentiating features, thus providing a model that is recognized and accepted by all the speakers in any given place (García Izquierdo, 2006; Iparraguirre, 2014). However, recent studies have concluded that the alleged neutral model can never be exempt or free of localisms or regionalisms and, furthermore, it is rejected by many native speakers. Of late, even translators and publicists have been critical of the use of this artificial, standard variety of language (García Izquierdo, 2006; Iparraguirre, 2014). Clearly, this variety could be considered a useful type of super-standard in certain situations (i.e., bureaucratic,

academic, journalistic) but not in others. In the field of foreign language teaching, we should consider this tree-pruning variety of language a useful tool that is offered to students for other purposes (e.g., avoidance of confusion).

Unquestionably there is a need for a unified language model serving as a basis for structuring foreign language teaching. However, we feel that it is crucial to establish this model on language in use not as a mere theoretical phenomenon but as a reflection of its real-life use. Previous research allows selection of a variety based on real parameters, thus avoiding artificial situations and maintaining certain elements of specific linguistic varieties that could be used in class.

3. The common variety of languages

In recent decades, sociolinguists have directed much attention to the investigation of the dynamic processes of dialectal contact and the convergence/divergence between varieties. The empirical and theoretical work of European social dialectologists (Trudgill, Auer, Kerswill, among others) has revitalized this field and allowed us to understand the patterns of speech recurring across Europe. While dialects are forsaken, generalized access to an apparent common variety implies a considerable heterogenization of the norm and the homogenization of the dialectal panorama. Concurrently, the opposite destabilizing effect is taking place in speech communities with complex situations (Braunmüller, Höder, & Kühn, 2014).

These social changes favor the transformation of traditional varieties into koinés and/or regional varieties of the standard language. *Stricto sensu*, koineization results from the mixing of dialects (i.e., horizontal convergence). This convergence can culminate in a levelling of features that allow the dialects to approximate the standard language. Auer (2005) proposed a possible typology that was applied to the study of peninsular Spanish, qualified by Villena Ponsoda as *español común* or common, everyday Spanish (Villena Ponsoda, 2008). This variety stems from different varieties of peninsular Spanish, particularly from those spoken in southern and central areas, and from transitional dialects (Murcia and Extremadura). In other words, the common language variety reflects the convergence towards a linguistic model devoid of regional features (Villena Ponsoda & Ávila-Muñoz, 2014) worthy of consideration as a model for teaching Spanish to foreigners.

4. Objectives and hypothesis underlying the proposed model

If foreign language teaching aims to offer students a widely accepted language variety, logically the common variety presented in the previous section is a suitable starting point. Nonetheless, dialectal varieties can become powerful pedagogic tools since they contain specific elements that can be employed in instruction. Vocabulary

is one of the most obvious levels. At the lexical level, it is easy to confirm: (1) progressive substitution of forms considered regional, provincial or local, (2) gradual levelling and simplification of dialectal vocabulary towards the standard form that (3) facilitates the appearance of a common variety where the variables marked diatopically are progressively substituted by others belonging to general vocabulary, despite (4) resistance to the disappearance of specific lexical forms in certain traditional fields of experience (Penny, 2004, pp. 77-79), due to the coexistence of distinct linguistic varieties.

Our work proposes establishing a model of lexical selection for foreign language teaching combining common or general vocabulary with dialectal vocabulary. We analyze the present situation of the Spanish spoken in Malaga and use the results to construct a teaching model of Spanish adapted for students learning the language in this city. Should the results prove to be favorable, the proposal could serve as a basis for creating a general model for foreign language teaching elsewhere. It would offer students the possibility of mastering vocabulary capable of being adapted to basic communicative situations regardless of the place where communication takes place. Additionally, they will also be able to use the dialectal characteristics to understand, respect and evaluate specific sociocultural situations. This is because dialectal vocabulary is considered the key to specific psycho-cognitive processes related to the cultural reality of native speech communities.

Bearing this general aim in mind, the following specific objectives of the research can be formulated:

1. To demonstrate that the use of dialectal vocabulary in foreign language classes is really useful.
2. To promote the use and conservation of the non-tangible cultural heritage that this type of vocabulary represents and study its use and functionality as a teaching tool.

The hypotheses of our research can be summarized as follows:

1. *Organization levels.* The model proposed for teaching vocabulary can be represented as a stratified organization on different levels or crowns. All levels can be quantified objectively and adequately, and integrated into foreign language teaching methods.
2. *Functionality hypothesis of the vocabulary crowns.* Each of these levels has a specific function in the teaching of foreign languages.
3. *Entertainment hypothesis.* Working with the specific vocabulary of certain speech communities will motivate teachers and learners alike because dialectal vocabulary is an entertaining teaching tool.

4. *Morphological consciousness hypothesis*. Defined by Carlisle (1995) as a specific aspect of linguistic consciousness, morphological consciousness is the conscious perception by individuals of the morphological structures of words and consequently of their skills in using them (Duncan, Casalis, & Colé, 2009). Consciousness of this kind depends on students receiving instruction that leads them to deliberate and conscious manipulation of the linguistic units (García & Gonzalez, 2006).

Studies on the acquisition and development of morphologic consciousness have usually been carried out in learning the mother tongue, with children being extensively used as participants (Carlisle, 1995; García & González, 2006; González Sánchez, Rodríguez, & Gázquez, 2010). However, we believe that this concept can be used in designing foreign language teaching programs when it is adapted to the special characteristics of the students. In this regard, the morphologic characteristics of dialectal vocabulary could aid building and developing this type of consciousness among foreign language students.

5. Characteristics of dialectal vocabulary: *CONVERLEX Project*

In the case of Spanish lexical variation, researchers are used to working on research projects of international scope (e.g., *Variación Léxica (VARILEX)*, *Proyecto Sociolingüístico del Español de España y América (PRESEEA)*, *Convergencia Léxica (CONVERLEX)*, among others). These macro projects can be used as a framework and provide support for in-depth studies of the different ways in which Spanish is used, thus helping make comparisons between different groups. These macrostudies will assist us in the formulation of a model of real shared language that could also be used to reach specific instructional goals. By the same token, this general research offers relevant information on the specific variants that may have a significant value in particular spheres of linguistic application.

To illustrate how these specific data differ from what we might considerer standard Spanish, we are going to use insights into lexical divergence observed in the city of Malaga. The insights come from the *Proyecto de Investigación sobre la Convergencia Léxica en la ciudad de Málaga (CONVERLEX Research Project on Lexical Convergence on the City of Malaga)*. The aim of this project was to: (1) obtain real examples of active and passive vocabulary from a sample of Malaga city speakers, pre-stratified by sex, age and educational level, and (2) relate the samples to data from other geolinguistic and lexicographical sources. With this information we hoped to discover what proportion of the traditional lexical dialect persists and how it varies according to different strata (Villena Ponsoda, Ávila-Muñoz & von Essen, 2017).

We used a uniform sample of speakers, born in, or residing in the city from an early age, to meet the quota of the planned pre-stratification (N = 54, 27 males and 27 females) (Table 1). Age groups were established using the methods commonly used in research on Spanish and the system used in the pan Hispanic study on Spanish in Spain and America (Moreno Fernandez, 1996; Vida Castro, 2004). In the case of educational organization the speakers were grouped according to whether they possessed at least a secondary education certificate (0), completed high school studies or equivalent (1) or graduated from university (2).

Table 1 Sample of speakers for the study on lexical convergence in Malaga

Women					Men				
Education	0	1	2	Total	Education	0	1	2	Total
Age					Age				
20-34	3	3	3	9	20-34	3	3	3	9
35-55	3	3	3	9	35-55	3	3	3	9
> 55	3	3	3	9	> 55	3	3	3	9
Total	9	9	9	27	Total	9	9	9	27

Note. Education: 0 – Compulsory, 1 – Secondary, 2 – University

In addition to the pre-stratification variables mentioned above (i.e., age, sex, education), we used questionnaires to obtain information on occupation, income and place of residence. Data were also collected on the speakers' complementary education (languages, courses, etc.) as well as on their parents' status and background. The questionnaire provided information on the social network of the subjects and about their opinions about their own linguistic spheres.

Below we present a summary of the morphological and semantic-cognitive results of the study (Ávila-Muñoz, 2017). On the whole, there are four morphological and semantic processes responsible for creating the vocabulary that can be considered dialectal (i.e., the lexical items not found in general dictionaries or marked as *local* or *regional* in them). These are as follows:

1. Very occasionally shortening *cani[jo]* and a phonetic variation of the standard Spanish base (v.g *arenca* for *arenque*).
2. More frequently we found cases of basic composition of words. In the analyzed corpus, the process consisted of the combination of a noun and adjective *aguaviva* (*agua+viva*), *cañaduz* (*caña+dulce*) or a verb and a noun *buscabocas* (*buscar+boca*).
3. The process which is the most widely used for forming words in the dialectal vocabulary is suffixation with some cases of prefixation. These were not affixes characteristic of or exclusive to the area under study nor did we observe affixes used in the composition of technicalisms or abstruse ink-horn terms. On the other hand, we discovered general elements present

in Spanish that sample subjects used for creating local or regional vocabulary. In almost all the cases observed, the lexical base was a noun with occasional use of an adjective or a verb in the morphological process. The overall result in almost all the cases was a noun, as in *victoriano* (typical local anchovy), *lenguetón* (disparagingly of a person who cannot keep secrets), *majarón* (an augmentative disparaging affix for a deranged person), *merdellón* (a disparaging term for a vulgar, common person), *cachorreño* (a bitter orange typical of the region), or *chorra*<(d)>*era* (slide or sledge). In general, the suffixes have diminutive, affective or pejorative meanings and the semantic field usually composed food as in *caña*<(d)>*illa* (murex), *cañaduz* (sugar cane), *calentito* (churro), *gazpachuelo* (fish soup), *lasca* (slice), *salmorejo* (soup), *tejeringo* (churro), *borrachuelo* (typical Christmas honey pastry), or <(d)>*amasquillo* (apricot). Some items also referred to local or regional traditions: *nasa* (pound net), *limoneti* (posh), *piarda* (truancy), or *guiso* (hopscotch). Clearly, simple forms are maintained for concepts which are characteristic for the area under investigation.

4. The semantic processes of lexicalization create dialectal vocabulary from a general meaning in Spanish that, because of the contextual and experiential conditions of the speech community studied, acquires specific subtleties. Often the generalization or limitation of sense processes were fixed by metonymic or metaphoric cognitive processes (*búzano* from *buzo* = diver; *corralón* from *corral* = large courtyard; *gurripato* = rookie airman; *avenate* from *a+vena+ate* = make one's blood boil; *calentito* for *churro*).

As can be seen from the above, the use of dialectal lexical forms was helped by very basic morphological processes deeply rooted in common Spanish that did not go beyond simple combinations of morphological or lexical units. The forms that manifested the most resistance to lexical convergence were those closely related to local or regional life and representing basic morphological composition, always akin to other non-technical or unspecialized forms of general Spanish. The pedagogical implications of the differences between dialect and standard language are considered in the following section.

6. Discussion: Lexical selection in foreign language classrooms

6.1. General considerations

The vocabulary of any language consists of a constantly active nucleus of frequent lexical items that do not depend on specific communicative situations. The group of non-thematic elements is essential to maintaining communicative

activity. Very specific methodological procedures allow us to identify with statistical precision the vocabulary speakers use most often to construct their messages, irrespective of the subject of discourse. To measure this lexical inventory, we take the total number of times a word appears in the corpus under consideration (absolute frequency), we weight its distribution in the sphere of interest (dispersion) and we find the real use of the element (frequency x dispersion). Such information can serve different purposes although it is most frequently used in planning vocabulary teaching of mother tongues and foreign languages (Davis, 2005; Lonsdale & Le Bras, 2009). The most frequent (and stable) vocabulary of the most popular languages studied is usually composed of non-thematic lexical items used frequently by speakers irrespective of the subject of conversation, such as conjunctions, prepositions, adverbs, articles, pronouns, verbs and some common adjectives and nouns (Figure 1).

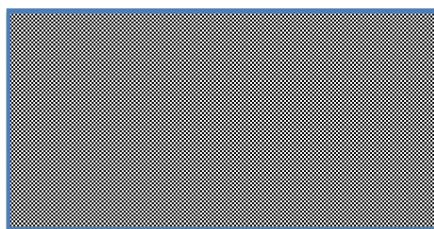


Figure 1 Representation of the most active lexical nucleus of a language

Vocabulary associated with specific communicative contexts, which is essential in foreign language teaching, is distributed around this language nucleus. Examining the lists of frequent lexical items, we can observe that in some cases very common words are missing or they attain very low frequency rates and are not considered as belonging to the basic repertoire (Savard & Richards, 1970). To compensate for these irregularities, the core is complemented by lexical speech inventories that are used in specific communicative situations. The concept of situational frequency is employed for creating lists of lexical frequency centered on frequent lexical items. In fact, lexical availability justifies the existence of some lexical items in accordance with the maxim that commonly used words in a language are closely related to the presence or absence of subjects of conversation. Gathering available, but not necessarily frequent, vocabulary is done by developing associative tests around stimuli or centers of interest. Specific, related vocabulary arises around these thematic nuclei, and forms the inventory the speaker really uses if, at any given time, a conversation takes a certain course. The total number of tokens included in these lists of available lexical items is limited either because of the speakers' response time or because the lists are closed after a certain number of items has been collected. It is assumed that the most available

words come first to mind and thus such words are put at the top of the list. Clearly, the available lexicon is part of the essential mental lexicon but is not in general used in everyday conversation unless a specific topic is brought up (Figure 2).

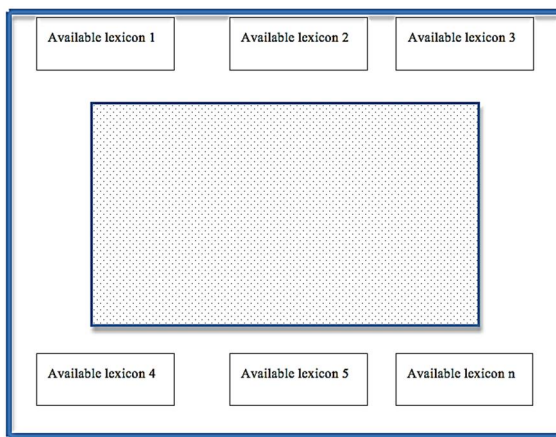


Figure 2 Representation of the second concentric lexical crown: Available lexicon

In the outer part of the lexicon, very specific elements appear depending on the precise characteristics of the communicative conditions (e.g., unusual) or of the speakers themselves (e.g., dialectisms, localisms, jargon) (Figure 3). When it comes to grammar, it is acquired by assimilating frequent core lexical items across linguistic components from outside the lexical sphere (i.e., syntax, morphology, grammar).

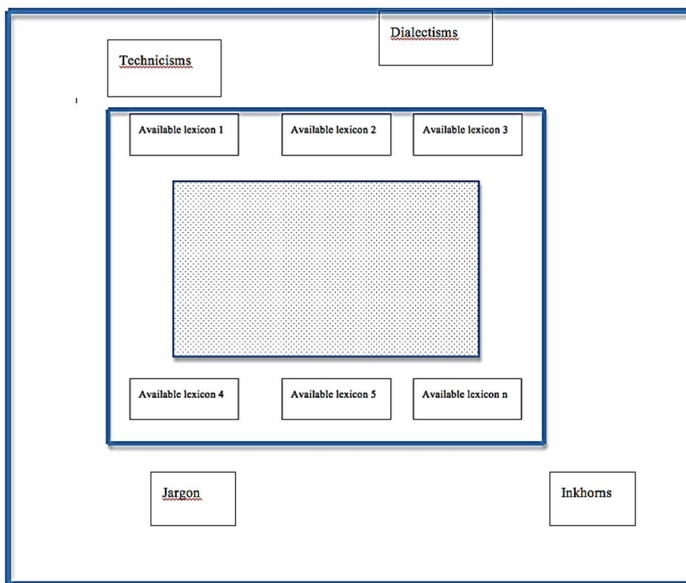


Figure 3 Representation of the third concentric lexical crown: Dialectal lexicon

6.2. A specific example

By analyzing the data of the available lexicon of Spanish spoken in Malaga, we proposed a model of cognitive lexical selection to teach Spanish according to the so called *Fuzzy sets* model (Ávila-Muñoz, 2016). This approach allowed us to work with common vocabulary adapted to the level of the students learning Spanish as a foreign language (ELE *español como lengua extranjera*). Having established this basic vocabulary with the help of objective tools (i.e., *fuzzy expected value*, FEV), we proposed creating a syllabus for ELE courses based on the lexical items employed by speakers in certain fields of experience because this lexicon is moderately common to all speech communities (Ávila-Muñoz, 2016).

Using the *Common European Framework of Reference* (CEFR, Council of Europe, 2001), we compared the proposal of specific notions present in the *PCIC* or Cervantes Institute Syllabus Plan (2006) with the centers of interest traditionally used in studies of lexical availability. Using the tools mentioned in the previous section, we were able to observe how communal prototypes are created and how terms can become more or less compatible with the collective prototype associated with specific fundamental notions. This allowed us to adjust the index of lexical availability and convert it into a parameter capable of differentiating between various levels of compatibility in the available lexicon of the relevant center. The frequency of several items at each level increased as the degree of compatibility decreased. To establish the compatibility threshold, we used a variant of the *fuzzy expected value* to determine the compatibility value of the *Fuzzy set*: *Weighted fuzzy expected value*. In effect, we were able to establish, on the one hand, a characterization limit of the suitability value of each item in regard its center of interest, and, on the other hand, we managed to establish parameters to identify elements that were very or not really characteristic of each center. Thanks to the FEV model, we ascertained the characteristic value for belonging to a *Fuzzy set* with respect to a specific measure, which in our case was the relative size of the groups of items that exceeded the compatibility level (cut-off group marks). The FEV allowed us to determine the value for belonging in a specific center, establishing a balance between the number of items that exceeded it and the value itself (see Ávila-Muñoz & Sánchez Sáez, 2014, for a detailed explanation of these procedures).

Since the lists of available lexicons are usually gathered from native speakers, these parameters allow us to identify the words that the subjects find the most available. Therefore, it is relatively unproblematic to apply the compatibility index to the selected vocabulary field, which is essential to the design of vocabulary learning programs for foreign languages. Obtaining the compatibility index for each lexical item in its associated semantic field also allows objective

selection of appropriate words at each level. Consequently, we established a road map which enables adequate selection of lexical items adjusted to the main ELE levels. The resulting list of lexical items contains the common (or most compatible) elements for each center of interest.

In all the analyzed cases, words with a higher compatibility index belonged to the basic user level (A1-A2; novice – lower Intermediate), which corroborated the assumption that core vocabulary sets the limits for the basic shared vocabulary of a language. Nevertheless, we did observe differences between the PCIC proposal and the words contained in the compatibility lists. Several of these indicated that some very compatible elements did not appear even at the basic levels proposed by the PCIC, while others referred to possibly inappropriate inclusion of words which, according to the compatibility calculation, should be contained at other levels.

Table 2 shows the prototypical lexicon of the *food* center of interest, ordered according to the six levels of compatibility after applying the FEV model to the lexicon results obtained from a group of 72 native speakers of Spanish in Malaga, stratified according to age, sex and educational level. The third column shows the level according to the CEFR while the next presents the equivalent level of the American Council on the Teaching of Foreign Languages (ACTFL). The last column provides the specific theme assigned by the PCIC.

Table 2 Level of reference of core lexicon according to the *PCIC*. *Center of Interest 5. Food*

<i>Word</i>	<i>Compatibility</i>	<i>CEFR Level</i>	<i>ACTFL</i>	<i>Specific PCIC Notion</i>
LEVEL 1				
<i>pan</i>	0.958877	A1	<i>Novice</i>	5.3. <i>Foods</i>
LEVEL 2				
<i>leche</i>	0.803341	A1	<i>Novice</i>	5.2. <i>Drinks</i>
<i>tomate</i>	0.803224	A2	<i>Lower intermediate</i>	5.3. <i>Foods</i>
<i>carne</i>	0.799878	A1	<i>Novice</i>	5.3. <i>Foods</i>
LEVEL 3				
<i>lenteja</i>	0.692349	B1	<i>Intermediate mid</i>	5.3. <i>Foods</i>
<i>patata</i>	0.656898	A2	<i>Lower intermediate</i>	5.3. <i>Foods</i>
<i>pescado</i>	0.634333	A1	<i>Novice</i>	5.3. <i>Foods</i>
<i>arroz</i>	0.610909	A2	<i>Lower intermediate</i>	5.3. <i>Foods</i>
<i>garbanzo</i>	0.606576	B1	<i>Intermediate mid</i>	5.3. <i>Foods</i>
<i>huevo</i>	0.599870	A1	<i>Novice</i>	5.3. <i>Foods</i>
<i>lechuga</i>	0.588998	A2	<i>Lower intermediate</i>	5.3. <i>Foods</i>
LEVEL 4				
<i>pimiento</i>	0.554343			5.3. <i>Foods</i>
<i>verdura</i>	0.499876	A1	<i>Novice</i>	5.3. <i>Foods</i>
<i>cebolla</i>	0.434353	A2	<i>Lower intermediate</i>	5.3. <i>Foods</i>

Why teach dialectalisms in the foreign language classroom?

<i>manzana</i>	0.423343	A2	<i>Lower intermediate</i>	5.3. Foods
<i>naranja</i>	0.423230	A2	<i>Lower intermediate</i>	5.3. Foods
<i>fruta</i>	0.422985	A1	<i>Novice</i>	5.3. Foods
<i>aceite</i>	0.422873	B1	<i>Intermediate mid</i>	5.3. Foods
<i>plátano</i>	0.419893	A2	<i>Lower intermediate</i>	5.3. Foods
<i>queso</i>	0.419211	A2	<i>Lower intermediate</i>	5.3. Foods
<i>mantequilla</i>	0.408971	A2	<i>Lower intermediate</i>	5.3. Foods
<i>pera</i>	0.400021			5.3. Foods
LEVEL 5				
<i>jamón</i>	0.399873	A2	<i>Lower intermediate</i>	5.3. Foods
<i>melón</i>	0.396565			5.3. Foods
<i>filete</i>	0.395443	A2	<i>Lower intermediate</i>	5.3. Foods
<i>sandía</i>	0.389854			5.3. Foods
<i>azúcar</i>	0.287899	A2	<i>Lower intermediate</i>	5.3. Foods
<i>yogur</i>	0.278987	A2	<i>Lower intermediate</i>	5.3. Foods
<i>tortilla</i>	0.277652	A1	<i>Novice</i>	5.3. Foods
<i>agua</i>	0.268987	A1	<i>Novice</i>	5.2. Drinks
<i>pasta</i>	0.265654	A2	<i>Lower intermediate</i>	5.3. Foods
<i>pollo</i>	0.259056	A2	<i>Lower Intermediate</i>	5.3. Foods
<i>zanahoria</i>	0.258789	A2	<i>Lower Intermediate</i>	5.3. Foods
<i>sal</i>	0.249091	A2	<i>Lower Intermediate</i>	5.3. Foods
<i>ensalada</i>	0.248112	A1	<i>Novice</i>	5.3. Foods
<i>legumbre</i>	0.234904	B2	<i>Intermediate high</i>	5.3. Foods
<i>calabaza</i>	0.233211			5.3. Foods
LEVEL 6				
<i>chorizo</i>	0.210091	B1	<i>Intermediate mid</i>	5.3. Foods
<i>chocolate</i>	0.199813	A2	<i>Lower intermediate</i>	5.3. Foods
<i>café</i>	0.198721	A1	<i>Novice</i>	5.3. Drinks
<i>fresa</i>	0.198109	A2	<i>Lower intermediate</i>	5.3. Foods
<i>espagueti</i>	0.187909			5.3. Foods
<i>macarrón</i>	0.186491			5.3. Foods
<i>salchicha</i>	0.186555			5.3. Foods
<i>galleta</i>	0.175621	A2	<i>Lower intermediate</i>	5.3. Foods
<i>salchichón</i>	0.156741			5.3. Foods
<i>cereal</i>	0.155542	A2	<i>Lower intermediate</i>	5.3. Foods
<i>tarta</i>	0.149087	A2	<i>Lower intermediate</i>	5.3. Foods
<i>maíz</i>	0.144386	A2	<i>Lower intermediate</i>	5.3. Foods
<i>mandarina</i>	0.134512			5.3. Foods
<i>atún</i>	0.122017	A2	<i>Lower intermediate</i>	5.3. Foods

When we examine the table, we can immediately observe a close relationship between the vocabulary compatible with the elementary levels proposed by the CEFR and the ACTFL. Only *lenteja*, *garbanzo*, *aceite* and *chorizo* correspond to level B1 or *Intermediate mid* and *legumbre* to level B2 or *Intermediate high*. The words *sandía*, *calabaza*, *espagueti*, *macarrón*, *salchicha*, *salchichón* and *mandarina* are not classified in the PCIC. The lexical items *pimiento*

and *pera*, situated at compatibility level 4, should receive the same pedagogic consideration as others situated at the same cut-off level, representing vegetables or fruit (A1, *Novice*), such as *cebolla*, *manzana*, *naranja*, *plátano*, *queso*, *mantequilla* (A2, *Lower intermediate*) or *aceite* (B1, *Intermediate mid*). The two words should be included in one of the levels where the latter are situated. It is also interesting to compare the compatibility indexes of *pimiento* (0.20800) and *pera* (0.19019) with those of vegetables (0.20438) or fruit (0.19218), included at A1 level or that of *Novice*. In the light of these results, our hypothesis is corroborated, especially when we realize that even the term *pimiento* is more compatible than the two words included in the PCIC. Similarly, *melón* (0.15528), *sandía* (0.13798) or *calabaza* (0.10458), situated at level 5, should be given the same attention as others situated at this level. For example, *ensalada* only reached the value of 0.09245 in terms of compatibility and yet it is considered by PCIC to be at level A1. Logically, we could say the same about the words found at level 6 that do not appear in any of the PCIC levels, that is, *espagueti*, *salchicha*, *macarrón* and *mandarina*.

Finally, our work concluded with a proposal to compile extensive corpora, including the largest possible numbers of lexical items from the Hispanic sphere. In this way, we would create a sufficiently representative database of common Spanish, enriched with lexical items peculiar to the different geographical areas where the data were collected and containing every one of the words obtained in the different centers of interest. The names of the centers of interest, the indexes of lexical availability and compatibility and the pertinent CERM and ACTFL levels would then be associated with this list of lexical units. Consequently, teachers of Spanish as a foreign language would have the opportunity to select, at any given moment, the words best suited to the needs of their students. Even more importantly, the model we propose favors the kind of instruction that is adjusted to the real needs of students.

To sum up, the proposal for vocabulary selection presented above is based on the study of real shared lexical items that students will be able to use in specific communicative situations. Lexical availability studies offer us the possibility of teaching a wider range of lexical material in a quick and easy manner. After gathering relevant data and applying procedures for determining compatibility presented in this section, lexical selection adapted to the official levels of foreign language teaching can be undertaken. Given the nature of such sampling, a greater percentage of lexical items is shared by members of the general linguistic community, although occasionally the lists of availability contain specific geographical terms (Arnal, 2008; Hernandez Cabrera & Samper, 2003; Valencia, 2005). To complete the corpora of available lexicon with others that are more specific, we need to create complementary strategies for gathering such more specific items. Thus, the *CONVERLEX project* could be an excellent research supplement.

7. Final proposal

Our lexical selection proposal is based on the organization of vocabulary as structured in concentric crowns, levels or halos. The first crown would be formed by a nucleus of active, very frequent lexical items, independent of the situation and the subject. The following crown would comprise subject-dependent vocabulary (-differential +available), surrounded by an outer ring including specific vocabulary (+differential -available). Thus, the core of the most frequent, non-thematic vocabulary, that is vocabulary unaffected by situational or personal factors, lies at the center. The second concentric crown around this active core or nucleus of common vocabulary is made up of common vocabulary shared by the linguistic community. This vocabulary is linked to certain frequent situations of language use. To select and adapt this vocabulary to different teaching levels we converted the indicator of lexical availability into a quantitative parameter capable of measuring the compatibility of each lexical items with each of the fields of experience or notional cores most widely employed in curriculum programs. Once the common vocabulary, adjusted to each instructional level, has been selected, gradual introduction of local vocabulary can be useful. Words belonging to the latter crown refer to firmly rooted habits and traditions and will thus ease cultural integration in an enjoyable manner.

To conclude, the first nucleus supports general structure dealing with the essential non-grammatical items in any communicative situation (frequent vocabulary, such as conjunctions, prepositions, articles, adverbs, etc.). The second crown includes shared vocabulary, irrespective of the varieties that appear in the TL but related to the thematic situation (i.e., available lexicon). The outside ring would include the most specific vocabulary, including that typical of various dialects. Such vocabulary will allow fostering morphological consciousness of foreign language students and help them to learn and respect the cultural and semantic-cognitive customs of the area or the speech communities where they are learning the TL.

In order to demonstrate the effectiveness of our theoretical proposal, we show finally an example of lexical selection based on the concentric crowns. Table 2 showed the more available vocabulary mentioned in the notional field of *Food* in the lists in question. If teachers had this material at their disposal and added it to the material in Table 3 that includes common dialectal vocabulary (*CONVERLEX Project*), they would have full lexical material for working with the *Food* notional field in their Spanish language classes.

Table 3 Dialectal vocabulary corresponding to Food center of interest (CONVERLEX Project)

Meaning	Dialectal form
murex	<i>búsano</i>
sugar cane	<i>cañaduz</i>
fried dough fritter	<i>calentito</i>
sea shell	<i>curruco</i>
gazpachuelo (fish soup with mayonnaise)	<i>gazpachuelo</i>
slice	<i>lasca</i>
salmorejo (cold thick tomato soup)	<i>salmorejo</i>
fried dough fritter in rings	<i>tejeringo</i>
sword fish	<i>agu<(j)>apala<(d)>a</i>
herring	<i>arenca</i>
small fresh anchovy	<i>vitoriano</i>
honey covered pastry filled with spaghetti squash jam	<i>borrachuelo</i>
apricot	<i><(d)>amasquillo</i>
Seville orange	<i>cachorreña</i>

Lastly, our methodological proposal places emphasis on the use of extensive and well-designed linguistic corpora that contain real instances of language use. In the case of teaching Spanish teaching as a foreign language, the use of such databases as CONVERLEX, PRESEEA or the VARILEX corpus could be of vital importance. The development of such corpora would offer access to data necessary to adapt the model proposed here to any specific speech community and any teaching unit.

References

- Arnal, M. L. (2008). Los dialectalismos en el léxico disponible de los estudiantes aragoneses. In M. L. Arnal (Ed.), *Estudios sobre disponibilidad léxica en los jóvenes aragoneses* (pp. 17-49). Zaragoza: Institución Fernando el Católico.
- Auer, P. (2005). Europe's sociolinguistic unity: A typology of European dialecto/standard constellations. In N. Delbecque, J. Van der Auwera, & D. Geeraerts (Eds.), *Perspectives on variation: Sociolinguistic, historical, comparative* (pp. 7-42). Berlin – New York: Mouton de Gruyter.
- Ávila-Muñoz, A. M. (2016). El léxico disponible y la enseñanza del español. Propuesta de selección léxica basada en la teoría de los conjuntos difusos. *Journal of Spanish Language Teaching*, 3, 31-43.
- Ávila-Muñoz, A. M. (2017). Entre limonetas y malaguitas: Estudio del léxico divergente de proximidad. Nuevos datos del Proyecto CONVERLEX-Málaga. *Hispania*, 100(4), 538-553.
- Ávila-Muñoz, A. M., & Sánchez Sáez, J. M. (2014). Fuzzy sets and prototype theory: Representational model of cognitive community structures based on lexical availability trials. *Review of Cognitive Linguistics*, 12(1), 133-159.
- Braunmüller, K., Höder, S., & Köhl, K. (2014). *Stability and divergence in language contact. Factors and Mechanisms*. Amsterdam – Philadelphia: John Benjamins.
- Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189-209). Hillsdale, NJ: Lawrence Erlbaum.
- Carter, R. A. (1987). *Vocabulary: Applied linguistic perspectives*. London – New York: Routledge.
- Davis, M. A. (2005). *A frequency dictionary of Spanish: Core vocabulary for learners*. London: Routledge.
- Duncan, L. G., Casalis, S. & Colé, P. (2009). Early metalinguistic awareness of derivational morphology: Observations from a comparison of English and French. *Applied Psycholinguistics*, 30, 405-440.
- García, J. N., & González, L. (2006). Diferencias en la conciencia morfológica, la escritura y el lenguaje en función del desarrollo y el nivel educativo del niño. *Psicothema*, 18(2), 171-179.
- García Izquierdo, I. (2006). El español neutro y la traducción de los lenguajes de especialidad. *SENDEBAR*, 17, 149-167.
- González Sánchez, L., Rodríguez, C., & Gázquez, J. (2010). Aproximación al concepto de conciencia morfológica: Concepto, desarrollo y relación con la lectura y la escritura. *Revista de Formación del Profesorado e Investigación Educativa*, 2010, 135-146.

- Hernández Cabrera, C., & Samper, J. A. (2003). Los dialectalismos en el léxico disponible de Gran Canaria: Análisis de un centro de interés. In F. Moreno Fernández et al. (Eds.) *Lengua, variación y contexto: Homenaje a Humberto López Morales* (pp. 339-353). Madrid: Arco Libros.
- Iparraguirre, C. (2014). Hacia una definición de español neutro. *Síntesis*, 5, 232-252.
- Londsale, D., & Le Bras, Y. (2009). *A frequency dictionary of French: Core vocabulary for learners*. London: Routledge.
- Moreno Fernández, F. (1996). Metodología del Proyecto para el estudio del español de España y de América (PRESEEA). *Lingüística*, 8, 257-287.
- Moreno Fernández, F. (2000). *¿Qué español enseñar?* Madrid: Arco/Libros.
- Penny, R. (2004). *Variación y cambio en español*. Madrid: Gredos.
- Savard, J. G., & Richards, J. (1970). *Les indices d'utilité du vocabulaire fondamental français*. Québec: Presses de l'Université Laval.
- Valencia, A. (2005). Dialectalismos en el léxico disponible chileno. In *Memorias del XIV Congreso Internacional de la Alfal*, Monterrey (México), edition in CD, tomo I: 230-239.
- Vida Castro, M. (2004). *Estudio sociofonológico del español hablado en la ciudad de Málaga. Condicionamientos sobre la variación de /-s/ en la distensión silábica*. Alicante: Universidad.
- Villena Ponsoda, J. A. (2001). *La continuidad del cambio lingüístico: Tendencias innovadoras y conservadoras de la fonología del español a la luz de la investigación sociolingüística urbana*. Granada: University of Málaga.
- Villena Ponsoda, J. A. (2008). La formación del español común en Andalucía: Un caso de escisión prestigiosa. In P. Martín Butragueño & E. Herrera (Eds.), *Fonología instrumental: Patrones fónicos y variación* (pp. 211-256). México: El Colegio de México.
- Villena Ponsoda, J. A., & Ávila-Muñoz, A. (2014). *Estudios sobre el español de Málaga*. Málaga: Sarriá.
- Villena Ponsoda, J. A., Ávila-Muñoz, A. M. & von Essen, M. C. (2017). Efecto de la estratificación, la red social y las variables de pequeña escala en la variación léxica. Proyecto de investigación sobre la convergencia del léxico dialectal. *Léxico Español Actual*, V, 209-239.