

Auxiliary verb constructions in Portuguese and Spanish: a comparative study and its applications as second languages

Construcciones verbales auxiliares en portugués y español: un estudio comparativo y sus aplicaciones como segundas lenguas

JORGE BAPTISTA

Faculdade de Ciências Humanas e Sociais
Universidade do Algarve
Faro, Portugal

RAFAEL CRISMÁN-PÉREZ

Facultad de Filosofía y Letras
Universidad de Cádiz
Cádiz, España
rafael.crisman@uca.es

Abstract

This paper presents a comparative study between auxiliary verb constructions in (European) Portuguese and in (Peninsular) Spanish. We compared 98 Portuguese auxiliary verb constructions and 93 Spanish ones. We used natural language processing applications. The results showed a few matches and mismatches. The lexical mismatches observed can contribute to consider auxiliary verb constructions according to their aspectual and modal nature in order to analyze the way each language consider those factors. This can contribute to improve learning processes of Portuguese and Spanish both as a second language according to mixed methodologies.

Keywords: auxiliary verb constructions, Portuguese language, Spanish language, phraseology, corpus linguistics

Resumen

Este artículo presenta un estudio comparativo entre construcciones de verbos auxiliares en portugués (europeo) y en español (peninsular). Comparamos 98 construcciones de verbos auxiliares en portugués y 93 en español. Para realizar este estudio comparativo utilizamos aplicaciones de procesamiento del lenguaje natural. Los resultados mostraron algunas coincidencias y desajustes. Los desajustes léxicos observados pueden contribuir a considerar las construcciones de verbos auxiliares de acuerdo con su naturaleza aspectual y modal para analizar la manera en que cada lengua considera esos factores. Esto puede contribuir a la mejora de los procesos de aprendizaje de portugués y español cada uno como segunda lengua según metodologías mixtas de enseñanza.

Palabras clave: construcciones verbales auxiliares, lengua portuguesa, lengua española, fraseología, corpus lingüístico

Introduction

From a systematic survey of auxiliary verb constructions in Portuguese (PT) (Academia das Ciências de Lisboa, 2001; Baptista, 2013; Baptista & Mamede, 2016), we identified and encoded about 100 auxiliary verb constructions with a description of their structure and their main semantic values (modal, aspectual). The structure of these constructions can be represented by *Aux*, an optional preposition (*Prep*) and a verb *V*, that is, the auxiliary verb constructions which consists of the auxiliary verb (*Aux*), a *connector* -most often a preposition (*Prep*), although it can be optional- and the main verb (*V*).

Auxiliary verb constructions have been heavily researched in the grammar of both languages, namely because of the following: (Alarcos, 1994; Gómez, 1998; Hassler, 2016):

- (a) They reflect the distributional constraints that the main (*full* or *lexical*) verb (here, *ler/leer* 'read') imposes on the lexi-

cal fulfillment of its argument slots; thus, any noun entering the verb *ler/leer* 'read' argument slots (subject and direct complement) will continue to form acceptable utterances:

- (1) PT: *O estudante / deputado / Sindicato está lendo o livro/ relatório/ jornal* ES: *El estudiante / diputado / Sindicato está leyendo el libro/informe/periódico* 'The student/MP/Sindicato is reading the book/report/newspaper'.
- (b) They express several grammatical values (aspectual, modal or temporal), with which they modify the basic (lexical) meaning of the main verb; in the example above, the *estar+gerund* 'be+gerund' construction, besides the present tense and the indicative modality conveyed in the auxiliary inflection, also conveys a durative-imperfective aspectual value.
- (c) They can combine with each other recursively to form longer verbal chains of auxiliary verb constructions, e.g.:

(2) PT: *O Pedro pode estar lendo o livro*
 ES: *Pedro puede estar leyendo el libro* ‘Pedro can be reading the book’.

In this case, the auxiliary verb construction *poder+infinitive* operates on the result of the previous auxiliary *estar+gerund*. This composes a hierarchical structure of auxiliaries (hence forward called *verbal chain*). That means, it is a verbal construction which can combine more than a verbal possibility beyond the concept of verbal periphrasis and auxiliary verb construction. Baptista, Mamede & Gomes (2010) report that in the CETEMPúblico corpus (Santos & Rocha, 2001), the largest verbal chain was composed by a string consisting of five auxiliary verb constructions.

(d) They can be construed as forming a discontinuous morpheme, allowing for different types of insertions (for example, clitic pronouns and adverbs); e.g.:

(3) PT: *O Pedro está sempre lendo o livro*

ES: *Pedro está siempre leyendo el libro* ‘Pedro is always reading the book’;

(e) They occur pervasively in texts; Baptista (2012; 2013) reports that $\frac{1}{3}$ of all the verb instances in a 250000 words fragment of the PAROLE corpus (REF), containing texts from a large varieties of textual types, correspond to auxiliary verb constructions (Bacelar do Nascimento, Bettencourt, Marrafa, Ribeiro, Veloso & Wittmann, 1998).

In the present paper, we compare these constructions with Spanish (ES). We proposed answer three research questions:

1. Which are the differences between both languages according to auxiliary verb constructions and which are the semantic and grammar values of them?
2. How can the empirical study of the linguistic corpus contribute to this question?
3. How the concept of auxiliary verbal construction and its comparison with verbal periphrasis can bring advantages to teaching of Portuguese as a second language and Spanish as a second language?
4. Which possibilities of verbal chaining do exist in both languages beyond those two concepts?

To date, and to the best of our knowledge, there are different grammatical comparative studies between PT and ES (Markic, 2011; Torijano, 2014). However, different compare no systematic comparison of these constructions in both languages has been made by a linguistic corpus. The semantic nuances introduced by the auxiliaries in the utterance overall meaning as well as the subtle differences appearing between the auxiliary verb construction of each language constitute a challenge both for second language learners (Kratohvilova & Jindrova, 2017). Likewise, this occurs to machine translation and other natural language processing applications (Aparecida, 2005; Gelbukh & Kolesnikova, 2013). The present study aims to clarify the translation possibilities in order to avoid mistaken transfers. So then, by the present paper students of PT as a second language and ES as a second language could consult a list of possibilities of auxiliary verb constructions according to their specific values in each language.

Many auxiliary verb constructions do not have any connecting element and the auxiliary verb is directly connected with the main verb (e.g. *poder ler/leer* ‘can read’, *ter lido/ter leído* ‘have read’, *estar lendo/estar leyendo* ‘be reading’). The main *V*, in its stead, must be one of the non-inflected, so-called nominal forms, $V = \{K, G, W\}$, namely, *K*: the invariable, masculine-singular form of the past participle (*ter lido/ter leído* ‘have read’), *G*: the gerund (*estar lendo/estar leyendo* ‘be reading’) and *W*: the non-inflected infinitive (*ter de ler/ter de leer* ‘have to read’). In complex verbal chains, formed recursively by the operation of more than one auxiliary verb, the formal structure of the auxiliary verb constructions of each successive *Vaux* constraints on the choice of the connector and the nominal form of *V*. This is imposed on the previous auxiliary verb constructions with the left most *Aux* carrying the person-number and tense inflections.

Auxiliary verb constructions were classified into three types according to Baptista, Mamede & Gomes (2010) theoretical framework. These researchers consider the main grammatical and semantic value conveyed by the auxiliary verb: VMOD (modality), VTEMP (tense) and VASP (aspect). For each type, several grammatical values are further distinguished:

- (a) Tense values for VTEMP include: *past* and *future*; in this case, although we grouped these constructions in a separate set more closely associated with the concept of *tense*, a *tense-aspect* combination of values may also be considered as relevant linguistic information,

pertinent for this taxonomy, namely the aspectual distinction between *perfective/imperfective*; hence, we encoded these values as *past-perfective*, on the one hand, and *future-imperfective*, on the other hand; e.g.:

- (4) PT: *O Pedro tinha lido o livro*
 ES: *Pedro había leído el libro* ‘Pedro had read the book’ (past-perfective)
 (5) PT: *O Pedro vai ler o livro*
 ES: *Pedro va a leer el libro* ‘Pedro is going to read the book’ (future-imperfective)
 (b) Aspectual values for VASP include: *inchoative* and *terminative* constructions. E.g.:
 (6) PT: *O Pedro começou a ler o livro*
 ES: *Pedro comenzó a leer el libro* ‘Pedro started reading the book’ (inchoative)
 (7) PT: *O Pedro acabou de ler o livro*
 ES: *Pedro acabó de leer el libro* ‘Pedro ended up read the book’ (terminative).

We consider the formal structure of auxiliary verb constructions (Alarcos, 1994; Gómez, 1998) (i.e. the connector and the nominal form of *V*) and identified unequivocally each construction within the set of auxiliary verbs here studied:

- (c) Modality values for VMOD include: *epistemic*, *deontic* and *imperative*; e.g.:
 (8) PT: *O Pedro chegou a ler o livro*
 ES: *Pedro llegó a leer el libro* ‘Pedro got to read the book’ (epistemic)
 (9) PT: *O Pedro tem de ler o livro*
 ES: *Pedro tiene que leer el libro* ‘Pedro has to read the book’ (deontic).

We add the unique case of an imperative modal auxiliary:

(10) PT: *Toca a levantar, meninos*

ES: *¡Toca levantarse, chicos!* ‘Let’s get up, kids!’ (imperative)

One should notice that the same verb can enter multiple auxiliary verb constructions; each one conveyed a different grammatical value. For example, the verb *acabar* ‘finish/end’ appears in two distinct auxiliary verb constructions, only different by the connecting prepositions:

(11) PT: *O Pedro acabou de ler o livro*

ES: *Pedro acabó de leer el libro* ‘Pedro finished reading the book’.

(12) PT: *O Pedro acabou por ler o livro*

ES: *Pedro acabó por leer el libro* ‘Pedro end up (by) reading the book’.

For the purpose of this paper, it is not relevant to name the difference in meanings conveyed by the constructions shown above (4-12). We retain these structures as two distinct auxiliary verb constructions. However, for sake of clarity in the comparison and the reuse of this information in several applications, a tentative classification of the semantic/grammatical value of the auxiliary verb constructions has been proposed by different research (Baptista, Mamede & Gomes, 2010; Baptista & Mamede, 2016). We based our research on that proposal, so that, this work was revised and added to the basic meaning presented in (i)-(iii). In the case of the constructions in examples (1) and (2) above, they have been defined as *acabar de W* (VASP/terminative/0) and *acabar por W* (VASP/terminative/resultative). This information can then be used to explore and annotate or parse textual corpora.

In the present paper, we analyzed different auxiliary verb constructions according to a structural criteria; however, we also considered the classification in the *Nueva gramática de la lengua española* in order to specify more possibilities of classification and functional factors for auxiliary verb constructions in both languages.

We also must say that the *Nueva gramática de la lengua española* (2009) considers three types of auxiliary verb constructions: verbal periphrasis, verbal semiperiphrasis and verbal phrases. We observed this taxonomy according the examples found in the linguistic corpus:

- Verbal periphrasis: comply the conditions of syntactic operation and grammaticalization of the auxiliary verb.
- Verbalsemiperiphrasis: comply only a few conditions, but not all of them. According to the *Nueva gramática*, verbs as *conseguir, intentar, lograr, necesitar, tratar, querer+infinitive, prometer, demostrar, mostrar* and the *esquemas fraseológicos semiproductivos* such as *romper a+infinitive* which only operates with the verbs *aplaudir, llorar, reír; salir+gerundio; dar+infinitive; echar a + infinitive* and *echarse a+infinitive*.
- Verbal phrases: they are constructions which do not have two verbs, but also one verb and another grammatical category. Some examples are *tener en cuenta, echar en cara* or *dar pie*.

Methods

For this study we started with a list of Portuguese auxiliary verbs drawn from the theoretically-based concept of auxiliary verb construction as defined, among others, by Gross (1999), and built by Baptista, Mamede & Gomes (2010). Then, Baptista & Mamede (2016) provided a new review.

According to this theoretical framework, the existence of auxiliary verb constructions was manually confirmed for Portuguese in the publicly available CETEMPúblico corpus, especially for the less obvious examples. CETEMPúblico is a corpus of texts of the Portuguese newspaper *Público* under an agreement between *Público* and the Ministry of Science and Technology of Portugal (MCT).

The interface to access this corpus allows for the use of regular expressions to build very precise queries involving lemma, part-of-speech and grammatical features. Regarding Spanish, the corpus of Real Academia Española (CREA) was used. This is a linguistic corpus about Spanish expressions from Latin America and Spain from all kind of written texts. The corpus covers the last twenty five years. It was developed by the Real Academia de la Lengua Española.

We consider as a basis of our research analyzing the possibilities of both linguistic corpus, as long as those linguistic uses are contextualized into determined communication processes. The auxiliary verb constructions analyzed were chosen according to the definition of Baptista, Mamede & Gomes (2010). In this research, auxiliary verb constructions were chosen according their semantic values and their

frequency of occurrence in CETEMPúblico corpus and in CREA. We analyzed the most frequent. In order to improve the study, we combine research methods of corpus linguistic and introspective lexicon grammar methods (Paribakht & Weseche, 1999) based on the inference as a cognitive process to vocabulary acquisition. Linguistic corpus allowed analyzing each auxiliary verb construction according to a specific context, which allows students to carry out a contextualized inference process.

Each auxiliary verb construction (its internal structure, the auxiliary type (Vasp, Vtemp, Vmod) and its grammatical and semantic values (e.g. Durative, Imperative, etc.) was encoded in a tabular format. Then, for each Portuguese entry a tentative equivalent in Spanish was determined in order to establish an interpretative correspondence as exact as possible between both languages. Conversely, several Spanish constructions were found without any equivalent in Portuguese. We made sure these equivalents had the same etymon, otherwise, a new entry was used. At this point, we based our research on a historical criterion of the language.

Besides, we also considered for the constitution of the lexicon-grammar entries the preposition (if any) linking the auxiliary to the main verb and the type of non-inflected form of the main verb (infinitive, gerund or past participle).

Only fully equivalent structures (same etymon of Vaux, same Prep, same V) were considered. In case there was no precise equivalent, a new entry was created and the remainder of that line was left empty.

Results

In this section we present and comment on the results of the formal comparison and study of the set of auxiliary verb constructions found in the Portuguese/Spanish language pair.

Perfect matches

Perfect match occurs when the Aux, the preposition and the non-finite form

of the main verb were exactly the same meaning in both languages (notwithstanding any translation/orthographic issues). Seventy-nine auxiliary verb constructions were found with a perfect match between the two languages. Table 1 shows the breakdown per *Vaux* type and their respective number. These corresponded to about 80% of all the auxiliary verb constructions described so far.

Table 1
Perfect matches of auxiliary verb constructions between Portuguese and Spanish. Breakdown per Vaux type and aspect/modality values

VauxType	Aspect/modality	Count
Vasp	Durative	21
Vmod	Epistemic	29
Vmod	Deontic	6
Vasp	Past/Perfective	2
Vasp	Terminative	10
Vasp	Inchoative	11
Total		79

Source: self made.

An example of these constructions is *acabar por Vinf*, functioning as an aspectual auxiliary with a terminative value:

- (13) PT: *O Pedro acabou por ler o livro*
(PT) ES: *Pedro acabó por leer el libro* 'Pedro ended up reading the book'

In all other cases where a correspondence exists, there may be some slight differences worthy of being described in view of foreign language teaching/learning. These will be highlighted next.

Lexical mismatches

In other cases, we observed lexical mismatches, that is, an auxiliary verb construction that only exists in one of the languages. We have found 13 PT constructions without ES equivalent, and 11 ES auxiliaries without PT equivalent.

For example, for the PT, *dar em V+inf* construction, there seems to be no equivalent form in SP, even though the verb *dar* ‘give’ exists in both languages:

(14) PT: *O Pedro deu em ler este tipo de livros*

‘Pedro ended up reading/started to read this kind of books’

In Spanish case, we found in Corpus ES TenTen18 of Sketch Engine:

Evoca así el caso de los tiranos que dilapidan con ligereza sus riquezas, hasta el punto de dar en cometer abusos incluso extremos con tal de no caer en la indignancia.

The translations issues raised in the English gloss indicated the particular modality besides the inchoative aspect expressed by this auxiliary verb construction. Namely, it expresses a negative/depreciative attitude of the speaker towards the content of the proposition and in particular towards the subject. In most cases, this auxiliary verb construction was applied to sentences with human-agentive subjects. Passive sentences seemed incompatible with this auxiliary verb construction:

(15) PT: *Este tipo de livros foi lido pelo Pedro* ‘This kind of books were read by Pedro’

**Este tipo de livros deu em ser lido pelo Pedro*

‘This kind of books ended up being read/started to be read by Pedro’

Lexical mismatches can also be the result of a non-extant lexical equivalent. For example, the Portuguese aspectual-intensive auxiliary verb constructions *desmanchar-se/escangalhar-se a Vinf* have no Spanish counterpart.

(16) PT: *O Pedro desmanchar-se/escangalhar-se a rir*

‘Pedro started laughing a lot/strongly’ (lit: ‘Pedro teared himself apart/brokedown laughing’)

Besides, a construction exists that is synonym of these two, PT: *desfazer-se a Vinf*,

(17) PT: *O Pedro desfazer-se a rir* ‘Pedro started laughing a lot/strongly’

(lit: ‘Pedro teared himself apart laughing’)

The *Aux desfazer-se* has an ES equivalent in *deshacerse* which is a non auxiliary verb construction **deshacerse a Vinf* and seems to exist in Spanish.

Anyway, the *Nueva gramática* considers the auxiliary verb construction *dar a+infinitive* as a semiperiphrasis verbal, however, *dar en+infinitive* has no equivalence.

On the other hand, another inchoative construction exists in Spanish with no Portuguese equivalent: *echar(se)a+Vinf*. The non-pronominal use of this auxiliary also seems to be in use. In fact, in the CREA corpus, the pronominal construction *echarse a Vinf* occurs 968 times while the

non-pronominal *echarse a + Vinf* appears 1,607 times. The *Nueva gramática* considers this construction as an *esquema fraseológico semiproductivo* which means that this construction has a few of verbal periphrasis factors but not all of them.

Furthermore, there are several types of predicates that can be combined with this auxiliary: (a) sueño: *dormir, dormir*; (b) manifestation of intense feelings: *reír* (468 cases) *llorar* (298 cases); (c) language verbs: *hablar* (7 cases), *gritar* (3 cases); (d) verbs about move: *correr* (431), *andar* (502), *volar* (99), *recorrer* (2 cases), *nadar* (3 cases), *caminar* (58 cases), *pedalear* (0 cases); (e) other verbs: *perder* (319 cases), *morir* (8 cases), *faltar* (17 cases), *sudar* (1 case).

In Portuguese, though, this seems to be a special word combination forming perhaps a collocation between the verbs *desmanchar-se*, *escangalhar-se* and *desfazer-se*, on the one hand, and the verb *rir* 'laugh', on the other hand. In fact, almost all combinations of these

verbs in the pronominal construction appear only with verb *rir* (except for one occurrence of *desmanchar-se a chorar*).

A somewhat similar situation occurs with the pairs PT: *começar*/ES: *comenzar, empezar*. The pair PT: *começar*/ES: *comenzar* is not only equivalent but the two verbs are cognates (< LAT: *COMINITIARE, from INITIARE), so a perfect match equivalence has been established. On the other hand, though the *Dicionário da Língua Portuguesa Contemporânea* (DLPC) indicates the verb *empezar*, it seems that the verb is no longer use unless perhaps a Brazilian regionalism (Rio Grande do Sul). This corresponds to an auxiliary verb construction equivalent to *começar a Vinf* and is supposed to be derived from Castilian *empezar*. Since the verbs of the pair PT: *começar*/ES: *empezar* are not cognate, we kept them as non-equivalent constructions. Furthermore, each construction both in PT and in ES has exactly the same formal structures: *Aux a Vinf*, *Aux E Ger* and *Aux por Vinf*.

List of specific verbal periphrasis in Portuguese

Deitar-se	a	W	Vmod	<i>O Pedro deitou-se a ler este tipo de livros</i>
Findou	por	W	Vasp	<i>O Pedro findou por dizer que estava cansado</i>
Hesitou	em	W	Vmod	<i>O Pedro hesitou em ler o livro</i>
Permanecer	a	W	Vasp	<i>O Pedro permaneceu a ler o livro</i>
Principiou	a	W	Vasp	<i>O Pedro principiou a ler o livro</i>
Principiou	por	W	Vasp	<i>O Pedro principiou por ler o livro</i>
Tencionar	<E>	W	Vmod	<i>O Pedro tenciona ler o livro</i>
Terminar	a	W	Vasp	<i>O Pedro terminou a ler o livro</i>
Tocar	a	W	Vasp	<i>Toca a levantar-se, meninos</i>
Ver	a	W	Vasp	<i>O Pedro veio a ler o livro</i>
Ver	de	W	Vasp	<i>O Pedro veio de ler o livro</i>
Viver	a	W	Vmod	<i>O Pedro vive a ler o livro</i>
Voltar	a	W	Vasp	<i>O Pedro voltou a ler o livro</i>

List of specific verbal periphrasis in Spanish

Acostar	a	W	Vasp	<i>Pedro se acostó a leer este tipo de libro</i>
Acostumbrar	a	W	Vasp	<i>Pedro acostumbra a leer ese tipo de libros</i>
Arrancar	a	W	Vasp	<i>Pedro arrancó a aplaudir</i>
Dar	por	W	Vasp	<i>A Pedro le dio por leer este tipo de libros</i>
Echarse	a	W	Vasp	<i>Pedro se echó a leer este libro</i>
Empezar	a	W	Vasp	<i>Pedro empezó a leer el libro</i>
Empezar	<E>	G	Vasp	<i>Pedro empezó por leer el libro</i>
Estar	por	W	Vmod	<i>Pedro estaba por leer ese libro</i>
Ir	a	W	Vtmp	<i>Pedro va a leer el libro</i>
Romper	a	W	Vasp	<i>Pedro rompió a llorar</i>
Tocar	<E>	W	Vmod	<i>Toca levantarse, chicos</i>

A more complex/subtle situation concerns the use of verb ES: *acostar* ‘lay down’. While this is not exactly a lexical mismatches, the Portuguese verb PT: *acostar* ‘dock’ is mostly used in a nautical sense:

(18) PT: *O acidente ocorreu quando o pequeno petroleiro [...] se preparava para acostar ao terminal*

‘The accident occurred when the small oil tanker [...] was preparing to dock to the terminal’

In Spanish, however, the current meaning hints of this being still a full verb, as it is confirmed by its distribution in the corpus (4,468 occurrences), which features:

- (a) A typical locative construction with a subject and a locative (destination) complement:

Nobj0 V Prep-loc1 Nloc1 (with prep = *a, en* ‘to/in’) ‘motion’, e.g.:

(19) ES: *COSME: ¿Y qué pensabas hacer en la desfloración de Purificación? ¿Acostarte a la izquierda? ¿Como un pederasta? ¡Ridículo! ‘Cosme: And what did you plan to do in Purification’s desfloration? Lay on the left? Like a pedophile? Ridiculous’.*

(20) ES: **Se la dijo y ella le reprochó por la tardanza. Se acostó a su lado ‘S/he told her and she reproached him because of his delay’.*

(21) ES: *Luego anunció que él se acostaría a este lado de la cama king size porque era el lado derecho y solo duerme bien ‘Then s/he announced that he would lie down on this side of the king size bed because it was the right side and only sleeps well’.*

(22) ES: *Lo hemos tenido que acostar en su cuarto ‘We had to top him up in his room’*

Its agentive-causative counterpart, where the object is also preceded by preposition *a*:

(23) ES: *La noche antes, después de haber acostado a sus dos hijas, sus mellizas de apenas tres años ‘The night before, after having put her two daughters in bed, their twins just three years old’.*

Preposition *a* ‘to’ is slightly more frequent (138) than preposition *en* ‘in’ (112). In Portuguese, this corresponds to the locative full verb *deitar* ‘lay down’.

- (b) An intransitive construction, often with a time complement, in the sense of 'lay down, going to sleep':
- (24) ES: *Yo me acosté a las doce, igual que de costumbre* 'I went to bed at twelve, as usual'
- (c) When in combination with an infinitive introduced by preposition *a*, 37 out of 48 were instances of *acostar a dormir*:
- (25) ES: *¡Con cuánto placer me acostaría a dormir siesta!* 'With how much pleasure I would lie down to take a nap!'

However, other verbs were found like *chupar* 'suck', *leer* 'read', *pensar* 'think' and *ver* 'see'. Besides, though some of these verbs show very similar or semantically close meaning to *dormir* 'sleep': *descansar* 'rest', and *echar la siesta* 'take

an afternoon nap'. Therefore, we conclude that this particular construction is likely to be in a process of grammaticalization with an inchoative value with verb *dormir* 'sleep', or at least, it is acquiring a co-location status.

In Portuguese, the equivalent auxiliary verb construction involves the verb *deitar* (lit., 'lay-down') with an intensive value: *O Pedro deitou a ler o livro* 'Pedro start reading the book in earnest'

2. This situation with *acostar* is to be distinguished from the ES constructions with *dar*, on the one hand, and *echar-se*, on the other hand, because these are no longer interpreted as full/lexical/distributional verbs and are more clearly auxiliary verb constructions (with inchoative value):

dar	por	W	Vasp	Inchoative	<i>A Pedro le dio por leer libros</i>
echarse	a	W	Vasp	Inchoative	<i>Pedro se echó a reír</i>

In the case of *romper+a+* infinitive, we found 318 cases but for the construction *arrancar+a+* infinitive only 35 cases were found:

- (26) ES: *... no lograba romper a leer y a escribir/borrar de la memoria* 'I could not break to read and write / erase from memory'
- (27) ES: *Y un día, de buenas a primeras, te arrancaste a leer como un procurador* 'And one day, suddenly, you started to read like a solicitor'

According to the *Nueva gramática*, this construction is an *esquema fraseológico semiproductivo* and in Spanish it occurs with verbs *aplaudir*, *llorar* y *reír* and other verbs with expressive

values. According to our research, this construction operates in this way. We did not find other special possibilities, although the *Nueva gramática* does not consider the example *leer* specifically.

The construction *dar+por+* infinitive occurred 337 times in the corpus. In this construction, besides the inchoative aspect, there is also some connotation of 'unexpectedness, impulsiveness' which is very difficult to capture with the categories we have used so far to classify auxiliaries. Besides, this is an impersonal construction, since the semantic subject of the infinitive appears as a dative complement of *dar* 'give':

- (28) ES: *Se volvió raro después, al morir mamá, cuando le dio por leer*

librotes y cartearse con medicos extranjeros 'S/he became weird later, when mom died, when s/he took to reading books and correspondence with foreign doctors'

- (29) ES: *Me dio por leer en francés a Kafka y Rimbaud sin haber leído ni a Gracián ni a Góngora* 'I read French in Kafka and Rimbaud without having read either Gracián or Góngora'

An exactly equivalent Portuguese auxiliary verb construction exists with the verb *dar*: *Deu-lhe para ler o livro* (lit: 'Give to-him to read the book', 'He had the sudden urge of reading the book'.

Finally, in the case of *echar+a+infinitive* (1,383 instances), some remarkable combinations were found, namely 502 instances with the verb *andar* 'walk' and 468 with *reír* 'laugh', *perder* 'lose' with 120 instances, which may have a collocational status:

- *-andar* 'walk': ES: *Señaló como meta de su gobierno el echar a andar la economía* 'S/he indicated as the goal of her/his government to start the economy'
- *-reír* 'laugh': ES: *Se ha vuelto a echar a reír sin control* 'S/he has laughed again without control'
- *-perder* 'lose': ES: *Croacia debe ser paciente ahora para no echar a perder la iniciativa de Estado* 'Croatia must be patient now in order to not spoil the State initiative'

Naturally, other verbs can also occur:

- (30) ES: *En realidad, su interés se multiplicó y posiblemente se echó a leerlo todo, sin siquiera sentarse,*

apoyando los codos en los tablones 'Actually, her/his interest multiplied and possibly s/he read everything, without even sitting down, resting her/his elbows on the boards'

This combination of ES: *echar* with *reír* 'laugh' is similar to the one we have found in the Portuguese *escangalhar-se arir*. This construction also is considered as an *esquema fraseológico semiproductivo* by the *Nueva gramática*.

3. The ES aspectual-frequentative auxiliary verb *acostumbrar* 'use to' shows two constructions, one where it is directly linked to the infinitive (521 instances) and another one linked with preposition *a* 'to' (2,308):

- (31) ES: *Caído tanto de ellas como de los four tracks que acostumbra correr en compañía de sus amigos* 'Fallen off them as much as the four tracks that s/he usually runned in the company of her/his friends'
- (32) ES: *Volvió a demostrar esa garra y esa fuerza que acostumbra a poner en los partidos* 'S/he returned to demonstrate that force that s/he used to put in matches'

In Portuguese, however, the verb *acostumar* 'get used to' is a full verb:

- (33) PT: *O Pedro acostumou-se a fazer isso* 'Pedro got used to doing that';
- (34) PT: *O Pedro acostumou-se a que a Ana lhe fizesse isso* 'Pedro got used to Ana did him that'.

On the other hand, an auxiliary verb construction exists the same frequentative value but with the verb *costumar* 'use to':

(35) PT: *O Pedro costuma fazer isso*
‘Pedro uses to do that’.

This would correspond approximately to the auxiliary verb constructions of *acostumbrar* in Spanish. There is also another Portuguese construction with *costumar* but with linking preposition *de* ‘of’ which seems specific or at least more frequent in Brazilian Portuguese:

(36) PT: *O Pedro costuma de fazer isso*
‘Pedro uses of to do that’.

5. In Spanish, the auxiliary *estar+por+infinitive* (502 instances):

(37) ES: *Pedro está por leer un libro*
‘Pedro was supposed to but he has not yet started to read a book’

This construction seems to have no direct and exact counterpart in Portuguese. In this language, the *estar+por+infinitive* only operates on passive or passive-like constructions:

(38) PT: *O livro está por ler/ser lido*
‘The book has not yet been read (though it was supposed to)’; cp. *O

Pedro está por ler o livro (lit: ‘Pedro is for reading the book’). However, the passive form is also available in Spanish:

(39) ES: *...llevado la cruz, porque ya habrás visto, mamá, que ellos ni se acercaban a tu alcoba.*

‘...carried the cross, because you’ve already seen, Mom, that they did not even get closer to your bedroom’.

(40) ES: *Estoy por entrar en ella ahora mismo, vencer ese miedo estúpido que me embarga* ‘I’m about to enter right now, overcome that stupid fear that overwhelms me’.

So there is only a partial equivalence with the auxiliary operating on passive forms while the construction with an active subject is only available in Spanish.

Table 2 shows the main distinction of auxiliary types. It can be seen that both languages have a similar distribution of these types with aspectual auxiliaries totalling 58% of all the structures found. Next, the modal auxiliaries (39%) and temporal auxiliaries are just 3%.

Table 2
Results of the comparisons between Portuguese and Spanish auxiliary verb constructions by Vaux Type

	PT	ES	%PT	%ES
Vasp	56	52	0.57	0.57
Vmod	40	37	0.40	0.40
Vtmp	3	3	0.03	0.03
Total	99	92		

Source: self made.

Results show that Portuguese had more possibilities and examples of auxiliary verb constructions for every category and subcategory, except for VASP inchoative. Spanish had more possibilities and constructions for this Vasp category. The auxiliary verb constructions that only Spanish language had were *arrancar a*+infinitive (W) and *romper a*+infinitive (W), both inchoative VASP constructions. Both auxiliary verb constructions are considered *esquemas freseológicos semiproductivos* by the *Nueva gramática* (2009: 35): “Así presenta restricciones romper a+infinitivo que solo se construye con los verbos *aplaudir, llorar y reír* y unos pocos más que expresan manifestaciones expansivas”.

Besides, Spanish had also the construction *quedar en*+ infinitive (W) which expresses a modal value (VMOD).

On the other hand, Portuguese had the auxiliary verb constructions *acabar a*+ infinitive (W), *andar a* + infinitive (W), *continuar a* + infinitive (W), *escangalhar-se a* + infinitive (W), *estar a* + infinitive (W), *ficar de* + infinitive (W), *tardar a* + infinitive (W), *tencionar* + infinitive (W), *ter de* + infinitive (W), and *terminar a* + infinitive (W).

In some cases, the Spanish equivalent was found in the reference

corpus though with a very low frequency. This is the case of *continuar a Vinf* for which only two instances were found. In this corpus, we found the examples *continuar a ser un hombre* and *continuar a ver*. Because of that we included this auxiliary verb construction in the Spanish list.

Another particular point is the construction *estar a Vinf*, which was a very low usual auxiliary verb construction. We only found the example *estar a rebosar* two times in a list of 130 examples of *estar a* + personal tense verb (1.5%) and the example *está a ver* as a single one:

- (42) ES: *Lagente del mundo de la cultura está a ver si consigue [obtener] gaita* ‘The people of culture world are to see if it gets [get] money’

The construction *andar a Vinf* only was found in the next example:

- (43) ES: *Pedro se despertó e iba a andar a tomarlas* (las medicinas) ‘Pedro woke up and was going to walk to take them (the medicines)’

It was one example among 151 *andar* + personal tense verb (0.66%).

Table 3
Results of the comparison between Portuguese and Spanish auxiliary verb constructions: Vaux Type and Values

VauxType	Value	#PT	#ES	Difference
Vasp	Durative	25	22	3
Vasp	Inchoative	16	20	-4
Vasp	Terminative	15	10	5
Vmod	Deontic	7	6	1
Vmod	Epistemic	32	30	2
Vmod	Imperative	1	1	0
Vtmp	Future/Imperfective	1	1	0
Vtmp	Past/Perfective	2	2	0
-	-	13	20	-7
Total		112	112	

Source: self made.

Discussion

Topic 1: Number of structures for each vaux/vmod/vtem type in each language

This study showed Portuguese language, at least in comparison with Spanish, has more examples of auxiliary verb constructions. Otherwise, Spanish has more examples in the subcategory inchoative. Anyway, results can provide more evidence about the issues of the way tense, mood and aspect which are expressed in these romances languages (Hassler, 2016). According to second languages learners, we think is interesting for learners of both languages observe the way both languages understand the time related to aspectual factors. For future research, it would be interesting to study

the specific use of some periphrasis according to a lexical selection. This is the case of *de arrancarse a* + infinitive or *romper a* + infinitive.

Topic 2: Advantages of current methodology, particularly the combination of corpus linguistics and introspective methods of lexicon-grammar

On the other hand, this study shows that the empirical methodology of corpus linguistics can be a complementary approach for the traditional hypothetic-deductive research. It also can provide ways of analysis from direct contexts and communicative situations (Aparecida, 2005; Callou, 2017).

In order to know the internal functioning of a language, this kind of research must be necessary for the contrastive linguistic (Callou, 2013)

and its applications in foreign languages learning.

Topic 3: Complex cases of reduced constructions involving auxiliary verbs that look like (but are not) full verbs plus a subclause

About the concept of periphrasis, we consider that many constructions which traditionally were not classified as such should see reviewed in order to enhance an optimal definition of this concept. The category auxiliary verb construction allows review the traditional concept of periphrasis and its possibilities of classification and translation among languages.

An example is the Spanish the auxiliary verb construction *esforzarse para Vinf*. Normally, some grammars (Alarcos Llorach, 1994, p. 318; Gómez Torrego, 1988, p. 114) do not include this construction within the list of auxiliary verb constructions they describe. The grammarians consider this construction not as a periphrasis but as a full verb (*esforzarse*) plus a final (infinitive) subclause.

Notwithstanding, we construe this structure as an auxiliary verb construction because we do not change the subject in many ways (*Pedro se esforzó para leer el libro*) and the meaning of the expression is a global sense. Ej.:

- (44) ES: *Pedro se esforzó para que su hijo fuera a la universidad* 'Pedro struggled to get his son to university'
 (45) PT: *O Pedroi esforçou-se para entrar na Universidade* 'Pedro struggled to get into the university'
 (46) PT: *O Pedro esforçou-se para [que o seu filho fosse para a Universidade]= O Pedro esforçou-se*

[a fazer algo] a fim de que/para que o seu filho fosse para a Universidade 'Pedro struggled to get his son to university'

- (47) PT: * *O Pedro esforçou-se a o João fazer algo* 'Pedro struggled to get Joao to do something'
 (49) PT: *O Pedro conseguiu fazer isso* 'Pedro was able to do that'
 (48) PT: *O Pedro conseguiu que o João fizesse isso= O Pedro conseguiu [fazer algo] [para/a fim de] que o João fizesse isso* 'Pedro got Joao to do that'

Other examples are:

- (49) *Largarse a + infinitive*:
 ES: *Pedro se largó a llorar* 'Pedro left to cry'
 (50) *Largarse + gerundio*:
 ES: *Pedro se largo corriendo* 'Pedro left running'

Topic 4: Introducing the topic of auxiliary chains (not considered, for the most part in this paper)

Andar a + infinitive (W) like in the example:

- (51) ES: *Pedro se despertó e iba a andar a tomarlas (las medicinas)* 'Pedro woke up and he was going to take them (the medicines)'. We only found one single example in CREA like this on the construction *estar a + infinitive* (W).

Estar a+infinitive. We found the example:

- (52) ES: *Las consultas van a estar a rebosar en el mismo corpus* 'The consultations are going to be overflowing in the same corpus'

However, the auxiliary verb construction *a rebosar* could be considered as a phrasal verb as long as it refers *totalmente o en situación de total lleno* (Seco, Andrés & Ramos, 2004).

Conclusion

As a conclusion, we must say actually many research are based on traditional concept of verbal periphrasis instead of auxiliary verbal construction.

This approach has shown two ways to work both systems. Spanish has more inchoative meanings, and Portuguese has more possibilities as auxiliary verb constructions. These results allow investigating new paths for the translation research and for the study of ES and PT as second languages. According to this approach, the *Nueva gramática* considers many of results of matching as *esquemas fraseológicos semiproductivos*. This is particularly interesting for the translation and the equivalent meaning between both languages and the semantic and functional nature of their auxiliary verb constructions. The way how auxiliary verb constructions operate in both languages by *esquemas fraseológicos semiproductivos* contributes to classify with better precision the nature of auxiliary verb constructions in different contexts. This can help students of PT and ES as second languages.

From our point of view, the linguistic research tradition has based on applying reductionist or atomistic approach to a complex system so that the results have always been partial and segmented. We have recognized the plans of Saussure and the overcoming of them but we have not

overcome the methodologies of positivist research based on deterministic theses and reductive and mechanistic approaches to explain the laws of operation of a system.

In addition, given that our study has focused on the Spanish of Spain and the Portuguese of Portugal, we must consider this option for future research contemplating variation as a development of the systems (coming from a common system, Latin) according to the development of their own rules. This approach has shown two ways to work both systems. Spanish has more inchoative meanings, and Portuguese has more possibilities as auxiliary verb constructions.

For future research, more empirical investigations must be done to know the way of both languages work beyond of their traditional definitions by traditional grammars.

Acknowledgments

Research for this paper has been partially funded by public national funds by FCT – Fundação para a Ciência e a Tecnologia (Portugal) UID/CEC/50021/2013.

Bibliography

- Academia das Ciências de Lisboa (2001). *Dicionário da Língua Portuguesa Contemporânea*. Verbo.
- Alarcos, E. (1994). *Gramática de la lengua española*. Gredos.
- Aparecida, C. (2005). *Diferencias de usos gramaticales entre español y português*. Editorial Edinumen.
- Bacelar do Nascimento, M.F.; Bettencourt, J.; Marrafa, P.; Ribeiro, R.;

- Veloso, R. & Wittmann, L. (1998). LE-PAROLE: do corpus à modelização da informação lexical num sistema-multifunção. Actas do XIII Encontro Nacional da Associação Portuguesa de Linguística, *Vol. II*, 115-134.
- Baptista, J.; Mamede, N. & Gomes, F. (2010). Auxiliary verbs and verbal chains in European Portuguese. *Proceedings of PROPOR'2010. LNCS/LNAI 6001*. Springer, 110-119.
- Baptista, J. (2012). ViPEr: A Lexicon-Grammar of European Portuguese Verbs. *Proceedings of the 31th International Conference on Lexis and Grammar*. Nové Hradý (Czech Republic), September 19-22, 2012, 10-16.
- Baptista, J. (2013). ViPEr: Uma base de dados de construções léxico-sintáticas de verbos do Português Europeu. In Silva, F.; Falé, I. & Pereira, I. (orgs.). *Textos Seleccionados, XXVIII Encontro Nacional da Associação Portuguesa de Linguística*, Coimbra, APL, 2013, 111-129.
- Baptista, J. & Mamede, N. (2018). *Dicionário Gramatical de Verbos do Português*. U.Algarve.
- Callou, D. & Avelar, J. (2013). *Ter/haver* constructions and verbal agrément. *Journal of Portuguese Linguistics*, 12(2), 187- 208.
- Callou, D. (2017). A propósito do uso de *ter* e *haver* em lexias verbo-nominais: análise contrastiva português-espanhol. *Linguística*, 33(1), 33-46.
- Coseriu, E. (1996). *El sistema verbal románico*. Siglo XXI.
- Es-Ten-Ten. Recuperado de <http://www.sketchengine.co.uk/documentation/wiki/Corpora/esTenTen> el 12 de febrero de 2020.
- Gelbukh, A. & Kolesnikova, O. (2013). Semantic analysis of verbal collocations with lexical functions. *Studies in Computational Intelligence*, 414, 1-146.
- Gómez Torrego, L. (1998). *Perífrasis verbales. Sintaxis, semántica y estilística*. Arco Libros.
- Gross, M. (1999). Compound Tenses in English. *Linguisticae Investigationes*, 22, 71-122.
- Hassler, G. (2016). *Temporalität, Aspektualität und Modalität in Romanischen Sprachen*. Berlin: De Gruyter.
- Kratochvilova, D. & Jindrova, J. (2017). Ingressive verbal periphrasis in Spanish and Portuguese. *Linguística Pragmática*, 27(1), 38-56.
- Markič, J. (2011). Acerca de las perífrasis verbales ir y venir + gerundio en español, portugués y gallego. *Verba Hispanica*, XIX, 129-141.
- Paribakht, T.S. & Wesche, M. (1999). Lectura y adquisición de vocabulario L2 incidental: un estudio introspectivo de la inferencia léxica. *Studies in second language acquisition*, 21(2), 195-224. Doi: www.jstor.org/stable/44486436
- Real Academia Española: Banco de datos (CREA) [en línea]. *Corpus de referencia del español actual*. Recuperado de <http://www.rae.es> el 19 de junio de 2019.
- Real Academia Española y Asociación de Academias de la Lengua Española, (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Santos, D. & Rocha, P. (2001). Evaluating CETEMPúblico, a free resource for Portuguese. *Proceedings of the 39th Annual Meeting of the*

- Association for Computational Linguistics* (Toulouse, 9-11 July 2001), 442-449.
- Seco, M.; Olimpia, A. & Ramos, G. (2004). *Diccionario fraseológico documentado del español actual*. Aguilar.
- Torijano, J.A. (2014). El análisis de errores de estudiantes lusófonos sobre los verbos españoles (I): presente y futuro. *Revista Española de Lingüística*, 44(1), 145-178.

