

# **Error analysis, contrastive linguistics and learner corpora, or how to use current linguistic tools to improve the level in SFL class: the case of change-of-state verbs**

Lydia Fernández Pereda, Kris Buyse & Katrien Verveckken

*KU Leuven*

## **1. Introduction**

This study reflects the intention to verify, corroborate and rectify (if necessary) an intuition based on daily work in the classroom: experienced SFL (Spanish as a Foreign Language) teachers in Flanders have noticed repeatedly that students frequently make errors when using the change-of-state verbs<sup>1</sup> and do not have at their disposal either knowledge or well-elaborated tools to solve these problems. Indeed, change-of-state verbs are a group of Spanish verbs which do not fully correspond to their French or Dutch counterparts. Further, the degree of interference between the languages has not been investigated objectively yet. This topic is usually overlooked in SFL books and barely treated by SFL teachers. Besides, the latter have not always received the necessary theoretical background or training to teach the intricacies of this topic<sup>2</sup>.

Therefore, the five main objectives of this paper can be summarized as follows:

1. to examine the learner corpus *Aprescrllov* (*Aprender a Escribir en Lovaina* ‘Learning to write in Leuven’) and to create a new corpus with compositions by French-speaking students from the UCL (Louvain-la-Neuve) and by Dutch students of Spanish from the Radboud Universiteit (Nijmegen, the Netherlands);

---

<sup>1</sup> For a definition and brief description of these verbs, see Section 3.

<sup>2</sup> The problem of the lack of theoretical training for SFL teachers was once again confirmed by the testimonials of teachers during the presentation of this paper at the BKL Taaldag, Leuven, May 11<sup>th</sup> 2013.

2. to analyze not only the errors, but also the target-like use of change-of-state verbs by SFL students in order to determine the parameters in the selection of an appropriate equivalent;
3. to check more objectively the influence of French (L1 of Walloon students, L2 of Flemish students and L3-4-5 of Dutch students) and of Dutch (L1 of Flemish and Dutch students, L2-3-4 or even 5 of Walloon students) on these errors. By incorporating a contrastive perspective in the study of the interference of other languages in the acquisition of SFL, we obtain an explicative and at the same time predictive tool that helps to detect systematic interferences of French and Dutch;
4. to elaborate, on the basis of the data extracted from the analysis of the extended learner corpus, a new didactic tool that includes a contrastive and cognitive perspective;
5. to check the impact of this presentation on the output quality of the learning process of the change-of-state verbs by SFL learners.

## **2. Theoretical framework**

In the thesis project the following three theoretical frameworks come together: Error Analysis and Interlanguage (section 2.1), Cognitive Linguistics (section 2.2) and Collection, description and analysis of the learner corpora (section 2.3).

### *2.1. Error Analysis and Interlanguage*

The methods of Error Analysis (henceforth EA) and Interlanguage (IL) already exist since the 1960s-1970s, yet lost their focus on SFL at the end of the last century. The field of English as a second/foreign language (ESL/EFL), however, has continued to adopt both these frameworks in various studies. In addition, other research domains also show a renewed interest in these methods thanks to the impulse of current fruitful research about the compilation, description and analysis of learner corpora (cf. 2.3).

EA became relevant since the pioneer work of Corder (1967, 1971, 1981). It is essentially a scientific procedure that allows to determine the nature of the errors made by non-native speakers. Moreover, the final objective of EA is to draw conclusions from the identification, description and explanation of the errors in the interlanguage of non-native speakers and to propose didactic procedures which help to avoid them. Therefore, EA is closely related to Interlanguage (IL, Selinker 1972).

Also related to IL is the dilemma of referring to the L1 and the L2 in the foreign language classroom and its impact on the learning process (Galindo Merino

2011). As far as we know, this project would be a pioneer in studying the influence of more than one language in second language acquisition through corpus analysis, with the ultimate goal of determining more objectively the possible interferences between one language and another. By adding other texts written by Walloon (French-speaking) and Dutch students to the set of Spanish texts written by Flemish (Dutch-speaking) students, we expect variations regarding the interference of French and Dutch, explaining those errors. The underlying working hypothesis is that French-speaking students will make more errors than Dutch students due to the influence of French L1, while Flemish students will be situated in the center of the scale, between French-speaking and Dutch students, due to the different status of French as a FL in Flanders and the Netherlands.

## 2.2. Cognitive Linguistics

The visual component is highly important in Cognitive Linguistics (henceforth CL). It goes without saying that, in developing a method of second language teaching (SFL in our case) that bears in mind the correlation between the visual and the verbal realities, metaphors, schemes and comparisons with real life are highly useful. Likewise, the cognitive approach takes into account the symbolic character of language: via metaphorization, languages allow to transfer information from sensory experiences. While languages can present remarkable differences between one another, the world situations to which they refer are more or less equivalent within broad geographical borders.

This is why Cognitive Linguistics is gaining both in importance and relevance in second language teaching (a.o. *Handbook of Cognitive Linguistics and Second Language Acquisition*, cf. references), and why our research also includes several key concepts taken from this theory, namely *construal*, *categorization*, *encyclopaedic knowledge*, *metaphor*, *metonymy*, *embodiment*, and *motivation*.

In addition to establishing a cognitive relation between the learner and the linguistic form and to allowing a process based on the understanding of the forms, the major contribution of CL in language teaching is not only to help the students acquire the language, but also to teach them how to use in a conscious way those mechanisms that native speakers use unconsciously. In doing so, the cognitive approach motivates students to think about why the target language expresses things the way it does and to reflect on the “difficulties” of the target language, which, according to Langacker (2008: 73), is “a far more natural and enjoyable process than sheer memorization”.

Thus, by approaching the Spanish change-of-state verbs from a contrastive perspective, we meet again with one of the principles of cognitivism: every

language expresses its perception of reality and the world in a different way. This principle will become obvious when we analyze Spanish change-of-state verbs in a *contrastive* comparison to Dutch, as we will see in section 3.

Finally, we would like to stress the link in applied linguistics between the cognitive and contrastive dimension on the one hand and second language teaching on the other hand. This connection sustains one of the three pillars of the theoretical framework of our project:

“one of the main contributions that cognitive linguistics can make to theories of language learning and teaching is in the area of contrastive analysis (...) Findings from cognitive linguistics can thus complement and extend earlier approaches to contrastive analysis (...)” (Littlemore 2009: 6)

### 2.3. *Collection, description and analysis of the learner corpora*

This section emphasizes the importance of applying corpus linguistics to teaching SFL. The fundamentals of the latter can no longer be exclusively intuitive, but have to be grounded in reliable data that only corpora can offer.

The architecture of our corpus *Aprescrilov* and its analysis are based on the theoretical frameworks of EA (for more information, cf. Santos Gargallo 2008), on the one hand, and of learner corpora (cf. Díaz-Negrillo and Fernández Domínguez 2006) on the other hand. The important utility of these frameworks has been illustrated, among others, in Gilquin, Granger & Paquot (2007), but is normally limited to the interference between one language and another (Buyse et al. 2009).

This project intends to set a step forward in applying corpus linguistics to SFL teaching (Cruz Piñol 2012) and to include the perspective of the interference of more than one language (Native Language, but also other L2) in foreign language learning.

## 3. **State of the art**

In Morimoto & Pavón Lucero (2007), the change-of-state verbs are considered a subgroup of pseudo-copula verbs. This subgroup includes *hacerse*, *volverse*, *quedarse*, *ponerse*, *llegar a ser*, *convertirse (en)* and other, less frequent items such as *tornarse (en)*, *meterse a*, and *trocarse*.<sup>3</sup> Their main characteristics, according to Morimoto & Pavón Lucero (2007), are the following:

1. loss of lexical meaning;
2. nexus role: the attribute is the semantic core of the predicate;

---

<sup>3</sup> The most frequent translation for these nine verbs is *to become*, *to turn (oneself) into*.

3. attribute is obligatory;
4. attribute pronominalization is not possible (*Juan se volvió loco/ \* Juan se lo volvió*);
5. higher semantic load than copula verbs (*ser* and *estar*).

The space dedicated to this issue in Spanish grammars is usually limited: apart from Morimoto & Pavón Lucero (2007), it has been studied by Coste & Redondo (1965), Fente (1970), Pountain (1984), Eberenz (1985), Alba de Diego & Lunell (1988), Porroche Ballesteros (1988), Bermejo Calleja (1990), Eddington (1999), Rodríguez Arrizabalaga (2001), and Bybee & Eddington (2006).

In Dutch, change-of-state is usually expressed by only one verb, i.e. *worden*, which does not have a direct, specific counterpart in Spanish. As a consequence, its translation is not easy for Flemish learners of SFL. Interestingly, they try to solve this problem by using a literal and non-existing translation from its French counterpart (*\*devenir*) or by risking to use one of the Spanish counterparts at random.

Moreover, *worden* presents another problem: in Dutch, it is primarily used to form the passive. Therefore, when learners find themselves in need of using such a syntactic structure in Spanish, *worden* – whose possibilities of translation into Spanish are perceived as problematic by learners – is also “activated” in their minds.

In the Dutch-speaking domain, only Hanegreefs (2004) and Meulman (2003) have, as far as we know, paid attention to the problem of finding an accurate equivalent in Spanish to the Dutch verb *worden* as well as to the intrinsic difficulty of this group of change-of-state verbs. Both have done so from the point of view of the Dutch-Spanish translation, on the basis of a corpus of translations made by native speakers. Hanegreefs also points out that the context in which *worden* appears could be helpful when translating this verb into one of the Spanish change-of-state verbs, although she does not specify which textual analysis criteria would be of use for this purpose. This is thus a possible research path to be explored.

#### 4. Hypotheses

In a nutshell, the hypotheses of this project can be summarized as follows:

1. “*\*devenir*” is the most frequently used change-of-state structure by French-speaking students in the first place – but also, though to a lesser extent, by Flemish students.

2. The quantitative imbalance between the Dutch structure (*worden*, one verb only) and the Spanish verbs (as mentioned above, at least nine verbs) makes it very difficult for the students to choose the appropriate verb and thus triggers the errors.
3. As a consequence, SFL students tend to avoid these change-of-state verbs, which are perceived as problematic and lacking an exact equivalent, both in their Native Language (NL) and in their L2.
4. If the presentation of this topic is done contrastively “L1-L2-NL” and based on real and frequent problems revealed by the error analysis, students would be able to use these structures properly and avoid the typical obstacles of interferences between Spanish, French and Dutch.

## 5. Methodology

The starting point is the *Aprescilov*-corpus, which perfectly fits the three-fold theoretical framework described in Section 2 as well as the theoretical principles of collecting, describing and analyzing learner corpora presented in 2.3. The present section outlines its main characteristics.

*Aprescilov* (“*Aprender a Escribir en Lovaina*”, see Buyse 2011 and Buyse & González 2013) is an online corpus composed of more than 2700 texts written in the academic years from 2004-2005 until 2009-2010 by students of Spanish Linguistics and Literature at the Faculty of Arts of KU Leuven, and by students of Applied Linguistics at Lessius Hogeschool Antwerpen (now part of KU Leuven as well). The compositions were written by 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> Bachelor students of both institutions and have been digitally marked with a customized version of the Markin<sup>4</sup> software, whose “button set” allows systematic marking of problems or ‘errors’ in the texts. The annotations cover all components of writing – from spelling to discourse structure, checking punctuation, morphology, morphosyntax, pragmatics, lexicon, etc. The online corpus also includes a qualitative and quantitative description of each component (number of compositions, number of words per text) as well as a search engine which allows

---

<sup>4</sup> Markin, elaborated by Creative Technology, allows noting down compositions digitally. “It is a Windows program which runs on the teacher’s computer. It can import a student’s text for marking by pasting from the clipboard, or directly from an RTF or text file. Once the text has been imported, Markin provides all the tools a teacher needs to mark and annotate the text. When marking is complete, the teacher can export the marked text as an RTF file for loading into a word-processor, or as a web page so that students can view the marked text in a web browser. Marked work can even be emailed directly back to the student, all from within the Markin program.” (<http://www.cict.co.uk/software/markin/>, last accessed 14-04-2014)

to search examples and their contexts by using criteria such as the type of problem, the course, the academic year or the institution.

The current corpus will be the starting point of a pilot study that will help to refine our research hypotheses as well as the parameters (type of subject:  $\pm$  animated,  $\pm$  agentive; type of process: voluntary, spontaneous, permanent, transitory, abrupt, gradual; context elements both in Spanish and in Dutch; cotext etc.) for the analysis of change-of-state verbs.

Starting from the *Aprescrllov* corpus, we will create a second corpus, called *Aprescrllov II*, that integrates texts written from 2010 onwards by Dutch-speaking students of the same academic degrees. The corpus will also contain texts in Spanish written by Dutch students from the Radboud Universiteit (Nijmegen, the Netherlands) and by French-speaking Walloon students from the UCL (Louvain-la-Neuve, Belgium) in the same period and on comparable subjects. These texts will be gathered in an online corpus with the same interface as *Aprescrllov*.

The expansion of the corpus, leading to *Aprescrllov II*, is necessary for several reasons:

1. to determine in a more objective way the interference degree of French in the acquisition of Spanish as a foreign language by students who have this language as their NL or as a L2, L3-4-5..., (i.e. Walloon, Flemish, and Dutch students of Spanish, respectively). We expect Flemish students to be at the center of the interference scale in Figure 1:

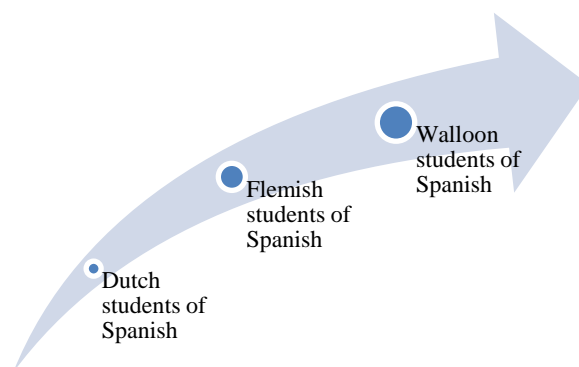


Figure 1. Scale of interference degree of French in the learning process of Spanish for students with different L1s and L2s.

2. to achieve the highest possible degree of systematicity in analyzing the errors as to change-of-state verbs in texts of the same academic years and on the same subject, in Flanders, Wallonia and the Netherlands.

3. to simultaneously focus on students' successes when choosing one of the change-of-state verbs correctly. The inclusion of this aspect of the language production is highly important because it helps to determine which parameters lead to errors in this matter. In order to do so, all correct choices of these verbs will be annotated – while this has not been done yet in a systematic way in *Aprescrllov*.

Once the analysis of both errors and target-like occurrences of change-of-state verbs in both corpora will be finished, a didactic tool that systematically and thoroughly addresses the matter of change-of-state verbs for Flemish SFL students will be designed. Finally, the efficiency of this new didactic material will be checked through a pre- and post-test system with parallel student groups.

## 6. Conclusions

The proposed project is innovative in the sense that the conclusions are based on the EA of a large corpus and that the results will be very useful, both for SFL teachers and learners in Flanders. At the same time, it fills an existing gap in SFL books, viz. the topic of change-of-state verbs. The study is also highly relevant from a didactic point of view, since it takes into account two major problems SFL teachers are constantly confronted with:

- a. how to cope with the interference of NL and/or another FL/L2 in the foreign language classroom.
- b. how to deal with errors in class and how to choose effective didactic tools to explain highly complicate issues which are probably influenced by L1 and L2, as is the case for the change-of-state verbs in Spanish.

The new didactic tool that will be created after a rigorous and scientific reflection and on the basis of the analysis of errors made by students of Spanish will prepare teachers to present the theory in such a way that their students will make less errors when facing the presented morphosyntactic issues.

## References

- Alba de Diego, V. & Lunell, K. (1988) 'Verbos de cambio que afectan al sujeto en construcciones atributivas'. In P. Jauralde, J. Sánchez Lobato, P. Peira & J. Urrutia, eds, *Homenaje a Alonso Zamora Vicente. Volumen I: Historia de la Lengua: El español contemporáneo*. Castilia, Madrid, 343-359.



- Bermejo Calleja, F. (1990) 'Verbos de cambio o devenir en español'. In S. Montesa Peydro & A. Garrido Moraga, eds, *Actas del Segundo Congreso Nacional de ASELE. Español para extranjeros: Didáctica e investigación*. Moraga, Málaga, 47-60.
- Buyse, K. (2011) 'El corpus en línea. ¿Fuera de línea en la clase de ELE?' Congreso Internacional de ASELE. Universidad de Valladolid, Spain, 21st-24th September 2011.
- Buyse, K., Delbecque, N. & Speelman, D. (2009) *Adquisición de la expresión escrita. Trampas y pistas concretas en el español del neerlandófono*. Edelsa/Averbode, Madrid.
- Buyse, K. & González Melón, E. (2013, i.p.) 'El corpus de aprendices Aprescrilov y su utilidad en la Bélgica multilingüe'. *ASELE, actas del XXIII Congreso Internacional*, Gerona, 18th-22nd September 2012.
- Bybee, J. & Eddington, D. (2006) 'A usage-based approach to Spanish verbs of 'becoming''. *Language*, vol. 82, No. 2, 323-355.
- Cifuentes Honrubia, J.L. (1996) *Gramática cognitiva: fundamentos críticos*. Eudema, Madrid.
- Corder, S.P. (1967) 'The significance of learner's errors'. *International Review of Applied Linguistics* 5, 161-169.
- Corder, S.P. (1971) 'Idiosyncractic dialects and error analysis'. *International Review of Applied Linguistics* 9, 149-159.
- Corder, S.P. (1981) *Error analysis and Interlanguage*. Oxford University Press, Oxford.
- Coste, J. & Redondo, A. (1965) *Syntaxe de l'espagnol moderne*. Société d'édition d'enseignement supérieur, Paris.
- Cruz Piñol, M. (2012) *Lingüística de corpus y enseñanza del español como 2/L*. Arco/Libros, Madrid.
- Díaz-Negrillo, A. & Fernández Domínguez, J. (2006) 'Error Tagging Systems for Learner Corpora'. *Revista española de lingüística aplicada* 19, 83-102.
- Doquin de Saint Preux, A. (2012) 'El español y el francés, ¿lenguas hermanas? El análisis contrastivo como herramienta clave para la enseñanza del español a francófonos', available from <<http://www.encuentro-practico.com/paris/pdf/12/doquin.pdf>> (last accessed 06-05-2014).
- Eberenz, R. (1985) 'Aproximación estructural a los verbos de cambio en Iberorromance'. *Linguistique comparée et typologie des langues romanes* 2, 460-475.
- Eddington, D. (1999) 'On 'becoming' in Spanish: A corpus analysis of verbs expressing change of state', available from <<http://linguistics.byu.edu/faculty/eddingond/become.pdf>> (last accessed 06-05-2014).

- Fente, R. (1970) 'Sobre los verbos de cambio o devenir'. *Filología Moderna* 38, 157-172.
- Galindo Merino, M. M. (2011) 'La lengua materna en el aula de ELE'. *Monografías ASELE 15*. Sta. Cruz de Tenerife, ASELE.
- Gilquin, G., Granger, S. & Paquot, M. (2007) 'Learner corpora: the missing link in EAP pedagogy'. In P. Thompson, ed, *Corpus-based EAP Pedagogy. Journal of English for Academic Purposes* (IRAL Special issue) 6, 319-335.
- Hanegreefs, H. (2004) 'The Dutch change-of-state copula *worden* and its Spanish counterparts. A matter of aspect and voice'. *Annual Review of Cognitive Linguistics* 2. John Benjamins, Amsterdam, 1-29.
- Langacker, R. N. (2008) 'Cognitive grammar and language instruction'. In P. Robinson & N. Ellis, eds, *Handbook of Cognitive Linguistics and Second Language Acquisition*. Routledge, New York, 66-88.
- Liceras, J.M. (1992). La adquisición de lenguas extranjeras: hacia un modelo de análisis de la interlengua. Madrid: Visor.
- Littlemore, J. (2009) *Applying cognitive linguistics to second language learning and teaching*. Palgrave Macmillan, Hampshire.
- Meulman, T. (2003) *Worden y su traducción. ¿Cómo se traducen los diferentes valores del verbo neerlandés worden al español y qué matices se pierden o se añaden en la traducción?*. Unpublished dissertation, Universiteit Leiden.
- Morimoto, Y. & Pavón Lucero, M.V. (2007) *Los verbos pseudo-copulativos del español*. Arco/Libros, Madrid.
- Porroche Ballesteros, M. (1988) *Ser, estar y verbos de cambio*. Arco Libros, Madrid.
- Pountain, C. (1984) 'La noción de 'devenir' en español'. *Boletín de la Asociación Europea de profesores de Español* 31, 111-116.
- Rodríguez Arrizabalaga, B. (2001) *Verbos atributivos de cambio en español y en inglés contemporáneos*. Universidad de Huelva, Huelva.
- Santos Gargallo, I. (2008) 'El análisis de errores en la interlengua del hablante no nativo'. In Sánchez Lobato, J. & Santos Gargallo, I., eds, *Vademécum para la formación de profesores: enseñar español como segunda lengua (L 2) / lengua extranjera (LE)*. SGEL, Madrid.
- Selinker, L. (1972) 'Interlanguage'. *International Review of Applied Linguistics in Language Teaching* 10:3, 209-232.
- Suby, J. & Asención-Delaney, Y. (2009) 'El uso del español del profesor en las clases de principiantes'. *Hispania*, vol. 92, n. 3, 593-607.