Gilquin, G. 2021. 'Using corpora to foster L2 construction learning: A data-driven learning experiment'. *International Journal of Applied Linguistics* 31(2): 229-247. https://doi.org/10.1111/ijal.12317

Using corpora to foster L2 construction learning: A data-driven learning experiment

Abstract

This paper adopts a Construction Grammar approach to test the efficiency of data-driven learning (DDL) when applied to constructions. High-intermediate learners of English as a foreign language took part in an experiment consisting, for each construction, in (i) a pre-test, in which the learners had to produce as many sentences as possible illustrating the constructions, (ii) a DDL intervention, which involved various concordance-based activities, and (iii) several post-tests, similar to the pre-test, performed at regular intervals. The analysis looks at the number of (valid and invalid) constructions produced as well as the native-like nature of these constructions, as established through a comparison with native corpus data. The analysis also examines the evolution of these results over the different post-tests. Finally, it considers learners' attitudes towards the DDL interventions, as evaluated through a questionnaire.

Keywords: data-driven learning, Construction Grammar, MAKE causative construction, *way* construction, *into* causative construction

1. Introduction

Data-driven learning (DDL), i.e. the use of corpus data in the language classroom, is a concept that has been around almost since the advent of corpora. Yet, it is fair to say that, to date, it has not been exploited to the full. On the one hand, many publications on DDL have adopted a theoretical approach, discussing the method without actually applying it. On the other hand, most DDL studies have been item-based, focusing on lexicogrammatical and lexical aspects. Larger units such as constructions, by contrast, tend to be neglected in DDL.

This paper adopts a constructionist approach to DDL, first by relying on the tenets of Construction Grammar to demonstrate the potential of DDL, and second by presenting the results of an experiment that applied DDL to the learning of constructions. Three constructions that are regularly discussed in the Construction Grammar literature, namely the MAKE causative construction, the *way* construction and the *into* causative construction, were tested among high-intermediate learners of English as a foreign language. Their answers to a sentence production task were compared before and after DDL activities, and their attitudes towards this pedagogical method were evaluated, in an attempt to assess the efficiency of DDL in fostering construction learning. Not only learners' knowledge of the constructions will be considered, but also their ability to develop more general cognitive skills through the use of corpora.

The paper is organised as follows. Section 2 briefly introduces the method of DDL and its variant, 'data-driven construction learning', showing that it can make instructed construction learning in a foreign language more similar to the natural process of construction learning in first language acquisition. In Section 3 the experimental design is described, while in Section 4 the results are presented, with a distinction between the pre- and post-tests on the one hand, and the participants' evaluation of the DDL activities on the other hand. Section 5 discusses the results and Section 6 offers some concluding remarks.

2. Data-driven construction learning

DDL is a pedagogical method that involves the use of corpora for learning purposes, in the form of activities encouraging language learners to engage with corpus data in order to make discoveries about language. According to McEnery & Wilson (1997), the use of corpora for language teaching purposes started at Aston University in the late 1960s, that is, in the early days of corpus linguistics. As for the first paper on DDL, it is thought to have been published in 1980 (McKay, 1980; see Cobb & Boulton, 2015: 482). Boulton & Cobb (2017), who carried out a meta-analysis of empirical studies published before 2015 and investigating the efficiency of DDL, identified 205 such studies. This, in effect, corresponds to a relatively limited number of six publications per year on average – although, as pointed out by Boulton & Cobb (2017: 381), the number of empirical studies increased year after year. This does not mean that DDL has not been given due attention in the literature, but may suggest that many publications devoted to DDL have adopted a more theoretical perspective, discussing the method without actually applying it and/or testing it (see Gilquin & Granger (2010: 365) and the references therein on the need to validate the DDL approach empirically).

Most of the empirical DDL studies so far have focused on lexicogrammatical and lexical aspects, in that order (Boulton & Cobb, 2017: 380). Hunston (2002: 184) notes that "DDL materials (…) tend to deal with the minute details of the phraseology of *particular words*" (emphasis added). The dominant approach is thus item-based, starting from a word and asking students to examine its linguistic environment. Constructions (in the sense of phrasal patterns), on the other hand, tend to be neglected in DDL. One exception is Rapti (2010), who applies DDL to phenomena like the passive voice or the *would rather* and *had better* constructions. Yet, following the central tenets of the theories of Construction Grammar, whose focus is on constructions (in the broad sense of 'pairings of form and meaning', see Goldberg, 2006), we can hypothesise that DDL should be beneficial for the learning of constructions. As a usage-based framework, Construction Grammar claims that constructions emerge in the constructicon (i.e. the mental repertoire of constructions) through abstraction over a large number of instances of the construction. Thus, children acquiring their first language (L1) are said to generalise pairings of form and meaning from the input they receive and to gradually build their knowledge of the language on that basis (see, e.g., Tomasello, 2003). If we consider second language acquisition, with a focus on foreign language (L2) learners, construction learning is expected to work differently, however. The first reason for this is that foreign language learners usually receive much less exposure to the target language than children acquiring their mother tongue, because they live in an environment that does not use the target language for daily interactions. This absence of regular input may hinder the process of generalisation. Second, most of their input in the target language comes from classroom instruction, and the authenticity of this input may be questioned (see Gilquin & Paquot, 2008 on spoken input in the classroom). If learners' input lacks authenticity, they are unlikely to develop constructions in the same way as native speakers. Finally, construction learning tends to be explicit and deductive in the case of foreign language learners, as opposed to the implicit and inductive learning that characterises first language acquisition. In the foreign language classroom, for example, a construction like the passive voice is typically taught through a set of rules, rather than (or at least in addition to) exposure to instances of the construction. Since, as Boulton & Cobb (2017: 350) put it, "[r]ules are hard, patterns are easy", the task of foreign language learners studying the target language through formal instruction is assumed to be more difficult than that of children acquiring their L1.

Thanks to DDL, and in particular the examination of concordances showing multiple instances of a certain construction in context (see Figure 2 below for an example), L2 construction learning can become more like L1 construction learning (see also Torres-Martínez, 2017). This approach, which could be referred to as 'data-driven construction learning', presents several advantages in comparison to a more traditional type of construction learning, as summarised in Table 1. First, while traditional construction learning usually involves little exposure to instances of the construction, data-driven construction learning gives learners access to a potentially large number of instances of a construction, thus corresponding to "condensed exposure", as Gabrielatos (2005: 10) puts it. By going through a concordance extracted from a one-million-word corpus, one can receive an exposure to the construction equivalent to several weeks of reading. Second, unlike traditional construction learning which often relies on invented examples, in data-driven construction learning the instances that learners get to see are authentic corpus examples, representing naturally-occurring language. They may also represent a specific kind of language (speech, writing, academic English, etc.), provided an appropriate corpus is selected. Finally, contrary to the mainly explicit and deductive approach of traditional construction learning, the DDL approach can be inductive (it usually is) and implicit. Students doing DDL have often been described as detectives (cf. Johns's (1997: 101) slogan: "Every student a Sherlock Holmes"), who have to generalise about language use on the basis of their own observation of corpus data (inductive approach). DDL may also be implicit, "enabling learners to infer rules without awareness" (Ellis, 2008: 879). These three features, namely the large amount of exposure to the target language, the authenticity of the input, and the inductive and implicit approach, should contribute to making L2 construction learning more similar to L1 construction learning and hence, presumably, more natural and effortless.

Traditional construction learning	**Data-driven construction learning**
Little exposure to target language	"Condensed exposure" (Gabrielatos, 2005: 10)
Lack of authenticity	Authentic data
Mainly explicit and deductive	Can be inductive and implicit

Table 1. Main differences between traditional construction learning and data-driven construction learning

In an attempt to test the efficiency of DDL in fostering the learning of constructions (in the narrow sense of phrasal patterns), an experiment was conducted that involved, for each of three constructions, a pre-test, a DDL intervention and several post-tests. The three constructions that were focused on are constructions whose existence is well-established in the Construction Grammar literature: the MAKE causative construction (see, e.g., Gilquin, 2010), the *way* construction (see, e.g., Luzondo Oyón, 2013) and the *into* causative construction (see, e.g., Kim & Davies, 2015). Besides determining the effect of data-driven construction learning on learners' knowledge and skills development, a second aim of the study was to evaluate their attitudes towards this type of pedagogical approach. In what follows, the experiment is described and the results of the pre- and post-tests and of the evaluation of the DDL interventions are presented.

3. The experiment

3.1. Participants

The participants who took part in this experiment were a group of high-intermediate learners of English as a foreign language. They studied English as a major in a Belgian, French-speaking university, and were in their second year. Most of them were native speakers of French. They were all students in a course on English syntax and stylistics taught in English by the author, and formed a group of 40-50 students regularly attending the course.

3.2. Design

Table 2 shows the three constructions at the core of the experimental design, with examples taken from the British National Corpus. The students were expected to be familiar with the MAKE causative construction, but less so with the *way* and *into* constructions, which do not really have any equivalent in French.

Construction	**Structure**	**Example**
MAKE causative construction	MAKE + NP + V_{inf}	*She tried to make him look at her.*
way construction	V + Poss. Det. + *way* + PP	*She elbowed her way through the crowd.*
into causative construction	V1 + NP + *into* + $V2_{ing}$	*He had blackmailed her into marrying him.*

Table 2. Constructions focused on in the experiment

As appears from Figure 1, which shows the timeline for the whole experiment, it started with a pre-test, followed by the DDL intervention, and then a number of post-tests. The first post-test took place about an hour after the DDL intervention, and additional post-tests were organised every week after that, for up to three weeks. The experiment ended with a final post-test that took place as part of the exam for the course, one month after the last class.

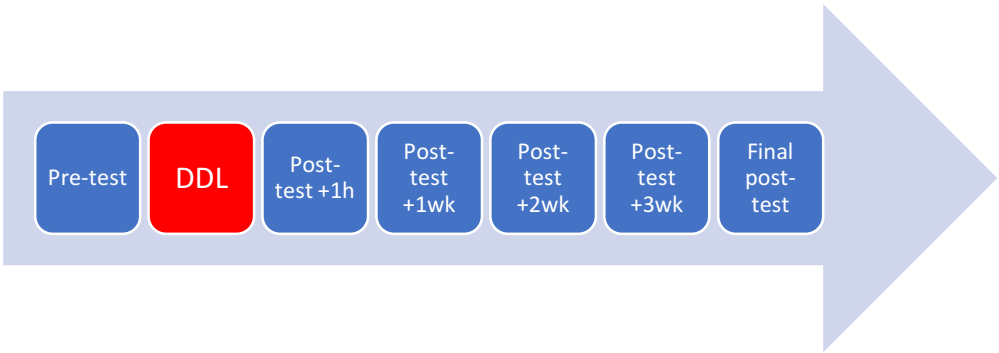

Figure 1. Timeline of the experiment, with pre-test, DDL intervention and post-tests

The main experiment (minus the exam) was spread over a period of five weeks. Only one DDL session was organised per week, which means that it was not possible to perform all the post-tests for each construction, but there were at least three post-tests per construction: one an hour

after the DDL intervention, one a week later, and the final post-test. Table 3 summarises the different stages of the experiment according to the construction. The DDL activity on the MAKE causative construction involved the highest number of post-tests (five), followed by the *way* construction (four post-tests) and the *into* causative construction (three post-tests).

	Week 1	**Week 2**	**Week 3**	**Week 4**	**Week 5**	**Exam (Week 9)**
MAKE	Pre-test	DDL & post-test +1h	Post-test +1wk	Post-test +2wk	Post-test +3wk	Final post-test
WAY	Pre-test	–	DDL & post-test +1h	Post-test +1wk	Post-test +2wk	Final post-test
INTO	Pre-test	–	–	DDL & post-test +1h	Post-test +1wk	Final post-test

Table 3. Different stages of the experiment per construction

The DDL sessions took 40 to 45 minutes, while the pre- and post-tests took 10 to 15 minutes each. An evaluation of students' attitudes towards the DDL interventions was organised at the end of the last class (Week 5), before the exam and its final post-test.

3.3. DDL interventions

The DDL interventions relied on a so-called indirect approach to DDL, in which students themselves do not have access to the corpus and corpus query software (see, e.g., Yoon & Jo, 2014 on the difference between direct and indirect approaches to DDL). Instead of actually searching corpora on a computer, the participants were given printed handouts with selected corpus materials. Figure 2 shows part of the concordance that they received for the *into* causative construction. As appears from the figure, the pattern was highlighted by means of angle brackets in each sentence. The concordances were extracted from the British National Corpus, a balanced corpus of native English representing different written and spoken genres. The sentences were mostly taken in their original form, but occasionally they had to be slightly adapted, for example to make the sentences shorter and/or simpler. For each construction, the handout was a double-sided sheet including 80 concordance lines. However, the students were advised to first focus on one side of the sheet, and to only look at the other side if they had time left. In addition to the concordance handout, they received a worksheet with various tasks related to the concordances (see Appendices 1 to 3). The tasks required them, for example, to underline elements within the construction, examine the features of these elements, check certain words in a dictionary, or rely on their observations to describe the meaning of the construction. The students had to carry out the tasks individually and autonomously.

1	He's very good at < talking people into doing > things.
2	Another anonymous woman, "a looker-on" claimed that "the men have recently shown a curious and novel diligence in the way of coaxing, bribing or even < fighting girls into joining > a trade union".
3	She sensed his waiting for her confidence, but knew he < would never pressure her into saying > things she would rather not.
4	The agony of Othello's realization that he < has been duped into killing > Desdemona is one of the two or three most intense sequences in Shakespeare.
5	How could there be any intent to < deceive people into thinking > it is French champagne?
6	There are few things worse than < being bludgeoned into reading > a book you hate.
7	Holyoak believes that the council < has been bamboozled into adopting > LET's monolithic scheme which has had outline planning permission for the last year.
8	They < should never be seduced into thinking > their policies had been so effective that there was no risk of relapse.
9	She was also instrumental in < nudging him into giving > 25 per cent of the positions in the Socialist Party to women.
10	They complained about < being "bounced" into accepting > last-minute decisions.
11	The claim is so shocking -- a paradox and even a horror, which we < may easily be lulled into taking > too lightly -- that only two views of this man are possible.
12	Martin Edwards insisted that he would still not < be rushed into selling > his majority shareholding, writes Stephen Bierley.
13	She was staring at him, with her great golden eyes, and her expression was so sad and apprehensive that the sight of it < shocked him into asking > the question he had been trying to ignore.
14	The aims of this new coalition were to < press the British government into acting > more forcefully against terrorism.
15	We need to look at what < triggers people into expressing > grief, at what moment they realize they have lost something.

Figure 2. Concordance for the *into* causative construction

3.4. Pre- and post-tests

In order to determine the effect of the DDL interventions, several tests were carried out that made it possible to compare the students' knowledge of the constructions before and after the activities. This was done both on a short term and long term basis. The pre- and post-tests, with the exception of the final one, were carried out by means of Wooclap, a classroom response system ('clicker') that enables students to answer the teacher's questions via their mobile phones or laptops. The last post-test was done on paper as part of the course exam. The tests mainly consisted in sentence production tasks: the students had three minutes to produce as many sentences as they could illustrating the construction under study, except for the final post-test, in which the students had to produce only one sentence with no time limit. The first post-test also included a multiple-choice question, which aimed to test students' understanding of the construction.

All the students who were attending a DDL session did the activities organised as part of that session (worksheet and/or test). Since not all students attended each and every session, however, in effect some students took tests on constructions for which they had not participated in the DDL intervention itself. Data produced in one of the post-tests by students who had not benefited from the corresponding DDL session were excluded from the analysis, and the students who had produced them were not taken into account. For the sentence production tasks, for example, while in total 2,102 sentences were produced as part of the pre- and post-tests, only 1,620 sentences could be taken into account in the analysis. Table 4 shows, per construction, the distribution of the data that were actually analysed and that lie at the basis of the results presented in Section 4 (number of students and number of sentences produced during the sentence production tasks; the number of answers to the multiple-choice questions corresponds to the number of students in the post-test +1h).

	Pre-test	Post-test +1h	Post-test +1wk	Post-test +2wk	Post-test +3wk	Final post-test
MAKE	45 students 107 sentences	47 students 220 sentences	32 students 152 sentences	34 students 151 sentences	36 students 175 sentences	49 students 49 sentences
WAY	45 students 82 sentences	38 students 129 sentences	25 students 66 sentences	32 students 100 sentences	–	39 students 38 sentences
INTO	45 students 46 sentences	42 students 141 sentences	41 students 112 sentences	–	–	55 students 52 sentences

Table 4. Number of participants and sentences produced for each test per construction

3.5. Evaluation

During the last class of the course (i.e. after all the DDL interventions and all the post-tests except the final one), students had to fill out a questionnaire asking their opinions about the DDL activities (see Appendix 4). The questionnaire included both statements to rate along a Likert scale and open questions. The students had about ten minutes to complete it, and did so on paper and anonymously.

4. Results

4.1. Pre- and post-tests

In order to evaluate the effect of the DDL interventions on students' knowledge of the three constructions, we can first compare the average number of sentences produced by each participant before and after the activities. As appears from Figure 3, the DDL interventions clearly have a positive effect in this respect, since the number of sentences produced increases significantly after the DDL session. On average, students produce at least one or two more sentences per construction. Interestingly, a week or two after the session, the number of sentences drops, but it tends to go up again later on, which suggests that it might take students a little while to integrate knowledge about the constructions beyond the immediate post-DDL test.

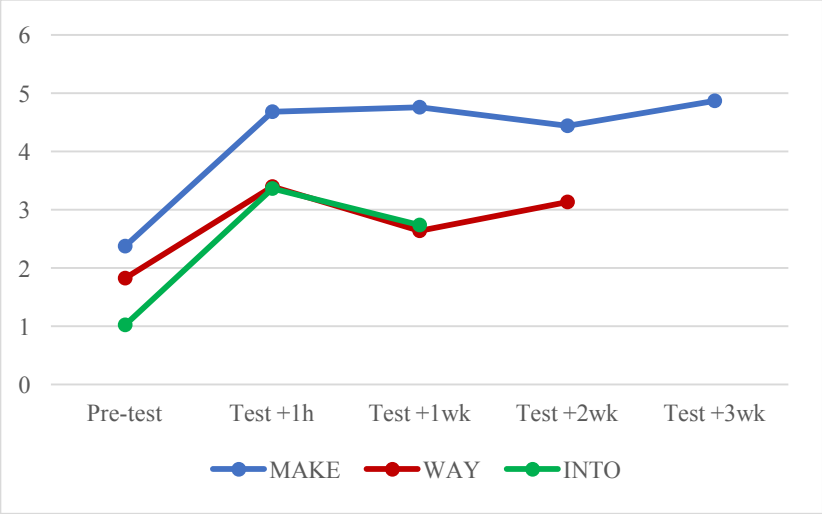

Figure 3. Average number of sentences produced per participant

Figure 3 is based on the total number of sentences produced by the students, regardless of whether they are valid examples of the constructions or not. However, some students produce sentences that do not correspond to the expected constructions (e.g. 'He made a good cake to eat' for the MAKE causative construction, 'I asked my way to the destination' for the *way* construction, and 'My friend put me into this fight' for the *into* causative construction). Figures 4 to 6 show the percentage of sentences that do correspond to a valid construction. Again, the effect of the DDL intervention is clearly visible, in that the proportion of valid sentences is much higher after the session than it is before. The difference is particularly impressive in the case of the *way* construction, which starts with less than 30% of valid sentences and, right after the DDL session, displays a proportion of almost 90%. As is the case with the number of sentences produced (see Figure 3), one week after the DDL session, the percentages appear to drop slightly, but then go up again, as if it took students some time to truly benefit from the long-term effects of the DDL interventions. The *into* causative construction is an exception, in that the percentage of valid constructions stabilises at 90% right after the DDL session (the number of data points is smaller for this construction, though, since it was the last one to be tested). The U-shaped development could partly be related to the repetition of post-tests and the fact that, after one or two post-tests, the students have become more familiar with the construction. It should however be mentioned that the students were not allowed to keep the DDL handouts (concordances and worksheets) and that no further examples of the construction were provided before or during a new post-test.

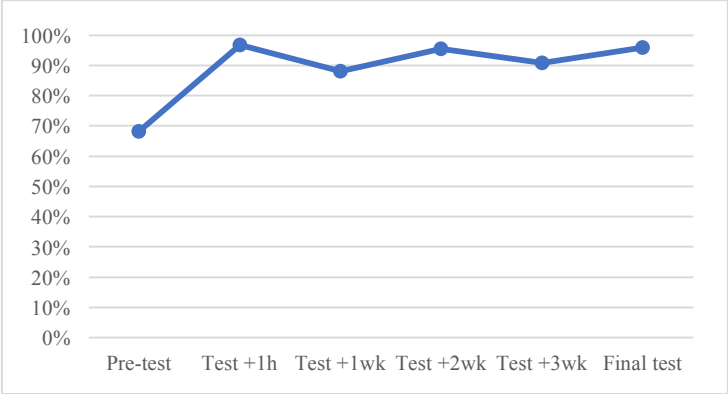

Figure 4. Proportion of valid MAKE causative constructions produced by the participants

Figure 5. Proportion of valid *way* constructions produced by the participants

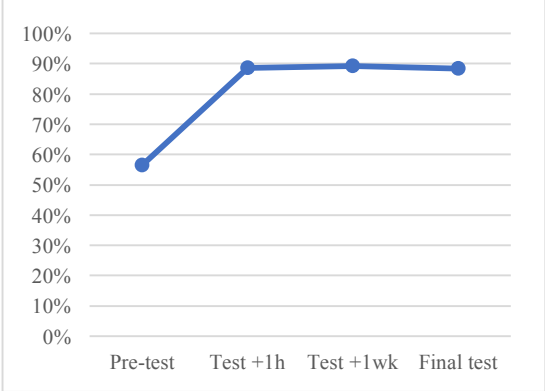

Figure 6. Proportion of valid *into* causative constructions produced by the participants

In addition to the validity of the constructions produced, we can also examine their native-like character. This will be illustrated here for the MAKE causative construction, which involves the largest number of post-tests. A corpus analysis based on a 10-million-word sample of the British National Corpus (Gilquin, 2010) reveals that most causative constructions with MAKE include a non-volitional infinitive, which refers to something happening independently of the subject's will (e.g. *feel*, *laugh*, *sound*), and that about half of the constructions have an inanimate subject. The coercive meaning, with MAKE being used as a synonym of FORCE, is quite rare in naturally-occurring language. In the present experiment, quite a few constructions produced before the DDL session have an animate subject and a volitional verb, and express a coercive meaning, as in (1) and (2). After the DDL session, more of the constructions produced by the participants are typical of native English, including a non-volitional verb, cf. (3) and (4), sometimes combined with an inanimate subject (4).

(1) She made her brother leave the room.
(2) She made her son finish his plate of vegetables.
(3) You made me realize that I was better off without you.
(4) That recipe made my mom's pancakes taste better.

Figure 7 shows the different combinations of subject and infinitive in the MAKE causative constructions, and their proportions in the concordance lines from the worksheet (which correspond to authentic instances from the British National Corpus), as well as in the pre-test and each of the five post-tests. Although none of the tests mirrors the proportions of the worksheet, it can be noticed that right after the DDL session, the combination of animate subject and volitional verb, which mainly corresponds to a coercive meaning, becomes less common, from 42% in the pre-test to 21% in the post-test +1h (to be compared with 15% in the worksheet). As for the combination of inanimate subject and non-volitional verb, which is predominant in the worksheet (41%), it becomes more common, from only 6% in the pre-test to 29% in the post-test +1h. Inanimate subjects also emerge timidly in combination with volitional verbs, from being non-existent in the pre-test to representing a small 3% in the post-test +1h (as against 11% in the worksheet). These effects, despite being positive, turn out not to be long-lasting, unfortunately, since the proportions in the final post-test are very similar to those in the pre-test, with only a 9% difference for inanimate subjects with non-volitional verbs (15% in the final post-test vs 6% in the pre-test) and for animate subjects with non-volitional verbs (43% in the final post-test vs 52% in the pre-test). This could possibly be explained by the context of the final post-test, which took place as part of the course exam (a high-stakes and stressful event for students) and involved the production of one single sentence, which might have led to a play-it-safe strategy and the preference for a grammatically correct, but not necessarily idiomatic construction. Yet, it should be noted that all the delayed post-tests, not only the final one, are more similar to the pre-test than to the post-test +1h (and the worksheet).

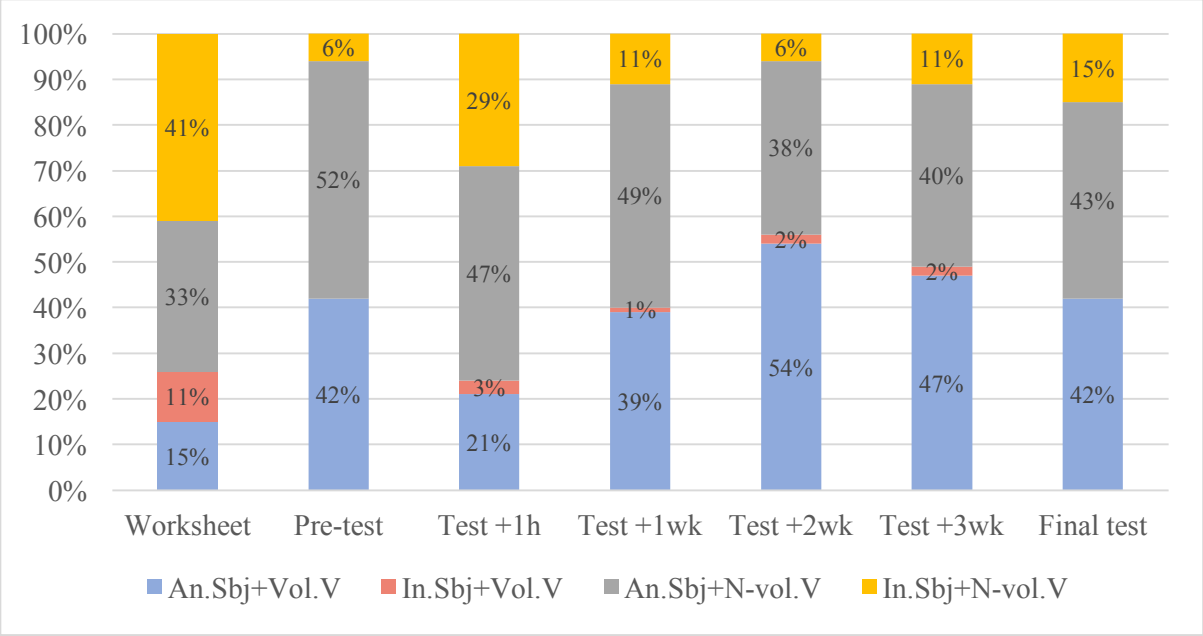

Figure 7. Features of the MAKE constructions in the worksheet, pre-test and post-tests

While this lack of long-lasting effect on the quality of the constructions is a slightly disappointing outcome, it seems as if the students have at least been able to infer what the main meaning of the constructions is. In the case of the MAKE causative construction, 78% of them were able to correctly select 'to cause something to happen to somebody independently of their will' as the meaning most typically associated with the construction, as opposed to only 15% for the coercive meaning 'to force somebody to do something' (see Figure 8).

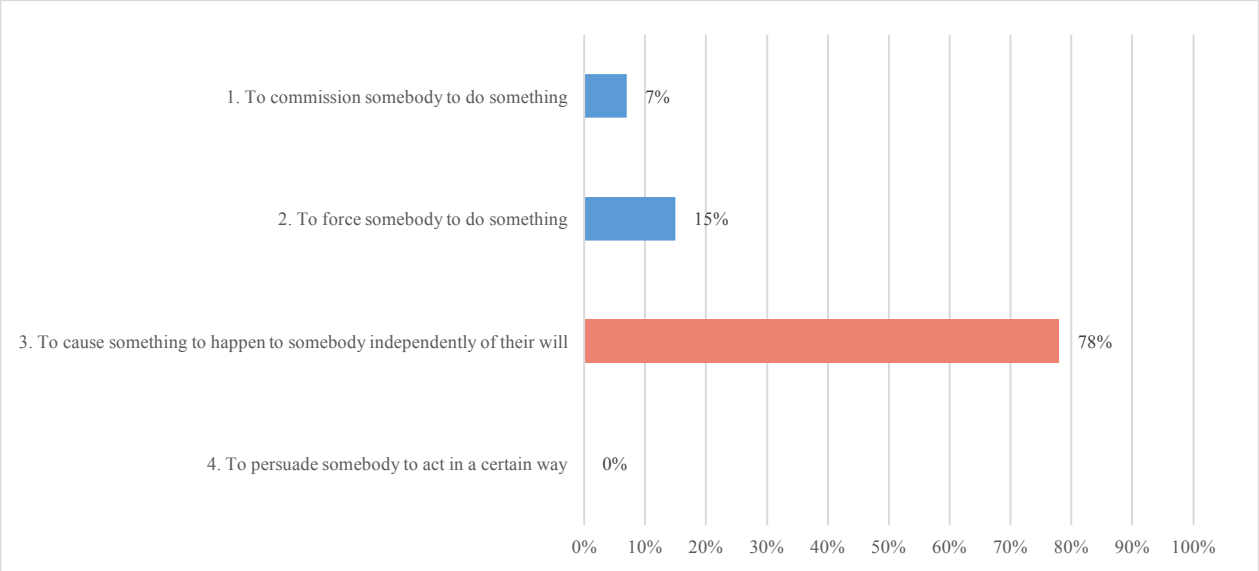

Figure 8. Distribution of the answers to the multiple-choice question 'Choose the meaning that best describes the pattern MAKE + Noun Phrase + Infinitive'

For the *way* construction, 60% of the students chose the right combination of semantic roles, namely 'Subject = agent; Verb = manner; Prepositional Phrase = direction' (see Figure 9). If

we disregard the semantic role of the subject, which is less relevant here and may have to do with the students' ignorance of the right terminology, we even reach a proportion of almost 70% of the students who discovered that the verb in this construction expresses a manner and the prepositional phrase a direction. This very good result is to be contrasted with the small proportion of valid sentences that were produced by the students in the pre-test (see Figure 5).

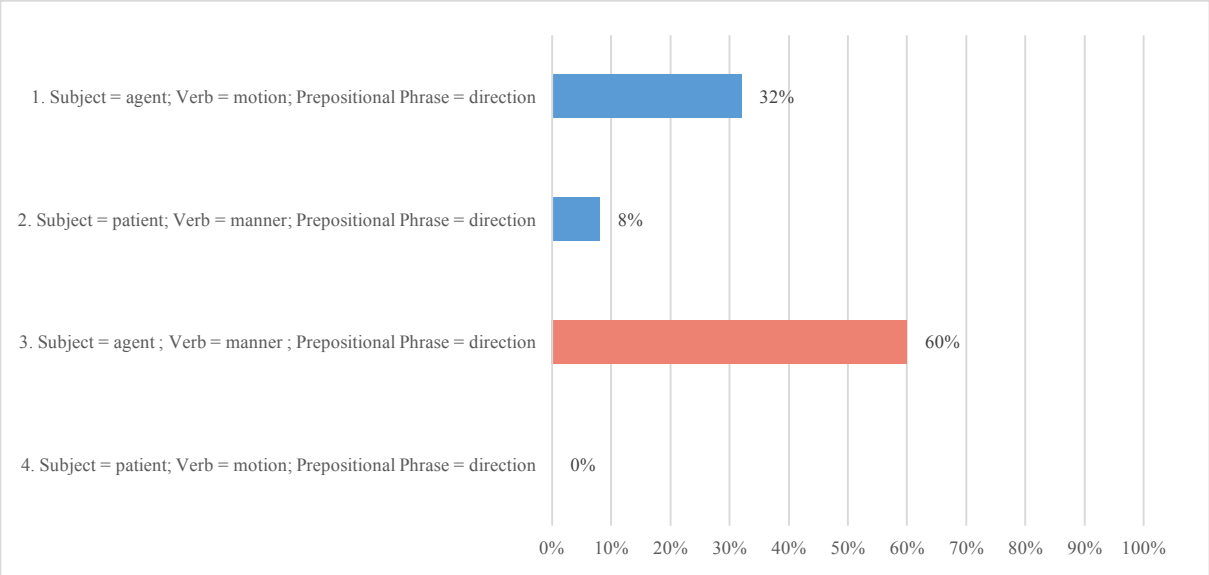

Figure 9. Distribution of the answers to the multiple-choice question 'What are the semantic roles of the different elements in the pattern Subject + Verb + Possessive Determiner + *way* + Prepositional Phrase?'

Finally, for the *into* causative construction, it is a staggering 83% of the students who were able to identify the right pattern describing the construction, namely 'MAKE NP V2 by $V1_{ing}$' (see Figure 10).

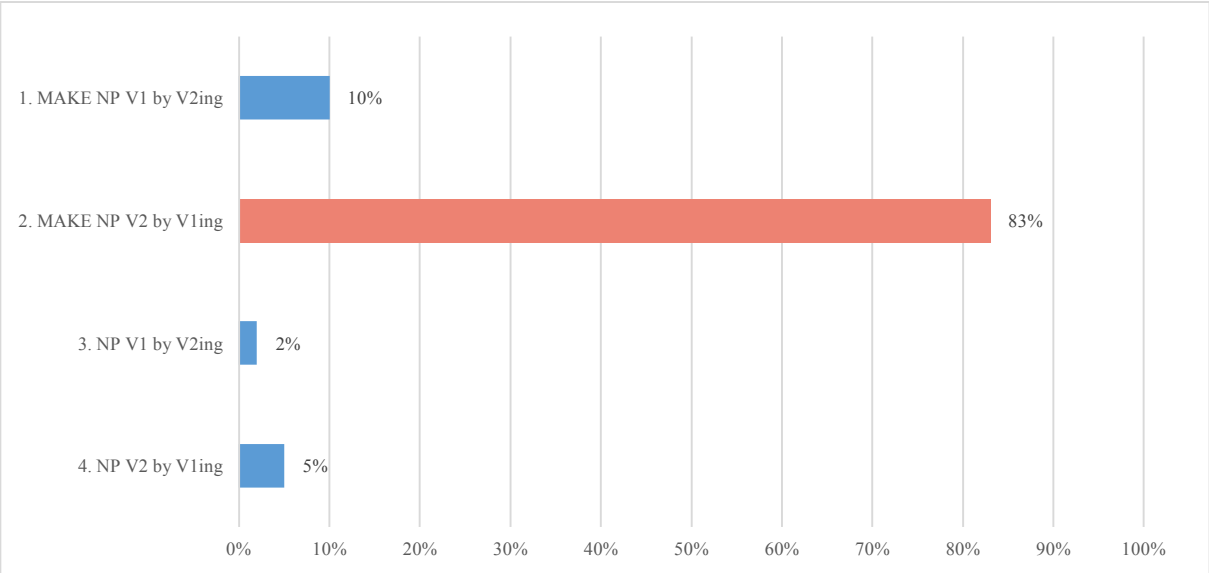

Figure 10. Distribution of the answers to the multiple-choice question 'Which of these patterns corresponds to the meaning expressed by the V1 + Noun Phrase + *into* + $V2_{ing}$ construction?'

More generally, it is interesting to notice that students seem to have generalised from the corpus examples rather than simply reproducing them. Most of the sentences produced during the post-tests are, obviously, different from those found on the worksheets, but most importantly, some of the verbs used by the students in the *way* and *into* constructions are new verbs, which were not presented to the students as being acceptable in these constructions. These new verbs represent 17% of the *way* constructions and 12% of the *into* causative constructions. While some of these verbs resulted in awkward constructions, like *to tell somebody into doing something*, most of the time the students ended up producing perfectly fine and even innovative constructions, as illustrated in the following sentences:

(5) My mother bribed me into cleaning my room.
(6) George tortured Donald into giving up intel.
(7) He hopped his way towards class.
(8) I laughed my way through the oral exam.
(9) I crawled my way into his heart.

4.2. Evaluation

Although the main criterion to decide whether a method is pedagogically relevant or not is probably its capacity to foster students' development (whatever form this development may take), it is also interesting to consider students' attitudes towards the method, as positive attitudes may contribute to enhancing this capacity. After all the DDL sessions, but before the exam and its final post-test, the participants in this experiment were asked to fill out an evaluation form about their experience. They did so anonymously, so that they could express their honest opinions freely. The answers provided insights into the students' attitudes towards data-driven construction learning.

The first question of the evaluation form required the students to write down the first word that came to their minds in relation to the DDL activities. Table 5 lists all the answers, in decreasing order of frequency. The most frequent word, with 11 occurrences, is *interesting*. This positive word, however, is followed by less positive ones: *repetitive* (5), *difficult* (4) and *long* (3). The term *research* (2) shows that a couple of students perceived the exploratory nature of the DDL tasks, which they equated with some kind of research activity. Among the answers occurring only once, we also find both positive words, like *good*, *helpful*, *useful* or *ludique* (a French word meaning 'playful'), and less positive ones, like *boring*, *redundant* or *time-consuming*.

N	**Answer**
11	interesting
5	repetitive
4	difficult
3	long
2	research
1	boring, challenging, constructions, correction, different, discovering, drilling exercise, good, grammar, helpful, interactive, linguistic analysis, list, ludique, redundant, related to the course, remember, repetition, statistics, time-consuming, too much, unclear, useful

Table 5. Answers to the question 'What is the first word that comes to your mind in relation to the tasks?' (n = number of times the answer was given)

Despite some negative terms among the answers to the first question, the students' attitudes towards the experiment seemed to be globally positive, as appears from the scores given to a number of evaluative statements, which the students had to rate from 1 'I strongly disagree' to 4 'I strongly agree' (see Figure 11). The statement that obtained the lowest score is actually one that asked about the students' previous experience with DDL exercises ('I had already done similar exercises before', score=1.49). For most of the students, this type of pedagogical approach was obviously a first. Not all the students agreed that the tasks were fun (score=2.09), but overall they enjoyed them (score=2.65), and they considered them interesting (score=2.97) and useful (score=2.91). They also thought that the tasks helped them learn more about the constructions (score=2.93), and perhaps even remember more about them than if they had been taught by the lecturer (score=2.44). More generally, the students felt that they became more efficient at the tasks after each session (score=2.65), which suggests that they not only learned about specific constructions, but also acquired special skills in the analysis of concordances.

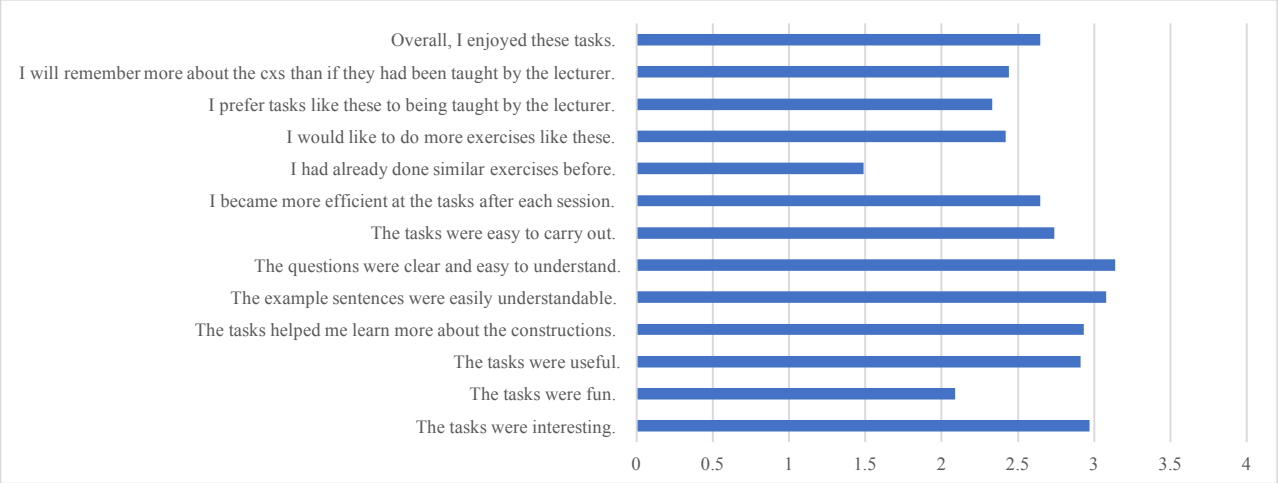

Figure 11. Students' rating of evaluative statements concerning the DDL activities

The students were also asked what they liked about the tasks and what they saw as the main advantages of this kind of approach. Their answers, some of which are reproduced (as is) below, underline that the tasks improved their understanding of the constructions, which became more *logical* (10). They also point to the research aspect of the task and to the fact that it is a different kind of activity which makes them more active (11). One of the students describes a process which actually corresponds to pattern-finding (Tomasello, 2003: 28) (12), and another one highlights the creative nature of the sentence production task (13).

(10) What I liked is that it made me able to really understand the construction, and semantic role of the verb in the construction. Something that for me was just a 'locution', 'a way of speaking' turned out to be very logical.

(11) I felt like a researcher, discovering things, trying to understand by myself. In this sense, these tasks were very interesting. Moreover, it brought a little changement to the usual activities and allowed us to be ourselves, more active.

(12) It's useful because looking at the corpus, we see that each sentence has a similar construction that we might have not noticed without it, so it makes it easier to remember and notice them later.

(13) You get to be creative with the different sentences you write.

When asked what they did not like about the tasks and what they saw as the main disadvantages of the approach, several students mentioned the fact that the tasks were long, annoying and boring, that there was not enough time and too many examples, so that the students sometimes felt a bit lost. They also underlined the lack of interaction. These ideas are visible in the following comments:

(14) In my opinion, the tasks were too long, we had about 30 minutes to fill the sheets and it becomes quickly a bit annoying and so after 20 minutes or so I would be fed up and stop working conscientiously.
(15) We didn't have enough time. For those exercises, I prefer to take my time to complete the tasks properly. I suggest students should do one of the three tasks at home.
(16) The disadvantages of such task are the lack of interaction and the fact that sometimes it can be boring looking for the same expression all over again.
(17) There are maybe a bit too many examples and then we are lost.

Between the lines, we can also read from the students' comments that DDL is really a matter of taste, with one student explaining that she is *the kind of person who needs to do it to remember* and another one admitting that she prefers *more traditional ways of learning* and that she would rather *have the lecture on the theory aspects and then some examples and exercises*. In the general evaluation of the course, several students mentioned the DDL activities, either among the aspects of the course that they liked the most or among those that they liked the least. The students, therefore, seemed to either love this approach or loathe it. These results could explain why other studies have reported both positive and negative attitudes towards DDL among students (see Gilquin & Granger, 2010: 365).

5. Discussion

The results of this data-driven construction learning experiment are mixed, both in terms of students' progress and their attitudes. The number of sentences produced increases and their quality improves right after the DDL activity, but this is not necessarily a long-lasting effect. And while most students recognised the usefulness of the method, many also thought that it was boring and repetitive. However, a number of factors having to do with the experimental setting of the study could partly account for these mixed results.

First, for the sake of the experiment and to avoid the influence of the teacher's intervention, her role was limited to that of "materials-provider" (Hunston, 2002: 184). In practice, however, the teacher should be more involved by providing feedback, giving additional explanations, discussing issues raised among students, etc. In fact, many students asked for such intervention by the teacher and felt insecure about not knowing what the 'correct answers' were. The DDL activities could also be more interactive in the sense that they could involve discussion between the students, who would thus have a chance to share their discoveries with their classmates and learn from them as well. In addition, the students might benefit more fully from the DDL interventions if the sessions were shorter and spread over a longer period, with regular exposure to different sets of concordance lines. Letting the students keep the worksheets and concordance handouts would also make it possible for them to continue their explorations at home and to memorise some of the typical features of the constructions, while asking them to insert these constructions in longer stretches of text would enable them to go beyond the isolated sentences of the experiment. It should also be borne in mind that DDL is not meant to

be used on its own, but should come as a complement to more traditional teaching methods, so that alternating between DDL and traditional sessions and combining them into a coherent pedagogical sequence would be useful. Such adjustments would make it possible to turn the controlled experiment of this study into real teaching practice, which would arguably have the effect of enhancing its efficiency.

A second point worth underlining is that sentence production tasks like the one used here tend to elicit a certain type of language which might differ from naturally-occurring language. This is due to the "inherently different nature of single sentence production and connected discourse sentence production" (Roland & Jurafsky, 2002: 330) and the fact that what is salient in the mind, and hence most readily accessible, does not necessarily coincide with what is frequent in language (Gilquin, 2008). The qualitative differences between the constructions produced by the students and those coming from the British National Corpus, as described in Section 4.1 for the MAKE causative construction, may thus not only reflect a lack of intake on the part of the students, but also partly result from the nature of the task itself, so that students' actual intake might be more significant than implied by the results of the sentence production. In fact, even native speakers have been shown to produce causative constructions expressing coercion more often in an experimental setting than in naturally-occurring language (Gilquin, 2010: 160-162). Yet, students' ability to produce MAKE causative constructions displaying features of naturally-occurring native language right after the DDL intervention (see Figure 7) suggests that the influence of the experimental setting is limited – although it might become stronger over time, which could explain the decrease in native-like features in further post-tests. Students' answers to the multiple-choice questions also seem to be indicative of real progress in their understanding of the constructions.

Finally, the ratio between the time spent on the DDL activities (3 x 45 minutes) and the students' progress may give the impression that the outcome is not worth the investment. Indeed, if we merely take the students' knowledge of the three constructions into account, the method is clearly unrewarding and impracticable given the limited amount of time that is available in the language classroom. However, one should keep in mind that, as pointed out by Boulton & Cobb (2017: 349), "[a]lthough the approach itself may seem rather time consuming for a single word, it is the processes that are important, and DDL may help learner development if it leads, for example, to increased language sensitivity, noticing, induction, and ability to work with authentic data". The results of the experiment suggest that long-lasting skills were also acquired by the students: they were able to make generalisations from the corpus instances (cf. multiple-choice questions and production of innovative constructions), they claimed to have become more efficient at the task, and they felt like researchers making discoveries about the language. Hopefully, the activities will also have made the students aware of the existence of a resource – the corpus – which they could use whenever they need information about language use.

6. Conclusion

The aim of this paper was to apply the pedagogical method of data-driven learning to constructions. Relying on the theoretical framework of Construction Grammar, it was shown that the method had the potential to make instructed L2 construction learning more similar to the natural process of L1 construction learning. So-called 'data-driven construction learning' was therefore expected to achieve some efficiency among foreign learners of English. An experiment applying DDL to three different constructions revealed that this type of approach

is not a panacea: it is time-consuming, its effects in terms of language acquisition are not necessarily long-lasting, and it does not suit all learning styles. However, "condensed exposure" (Gabrielatos, 2005: 10) to authentic examples of the constructions does have a positive impact on students' knowledge of these constructions, to the point where they can actually produce new, creative instances of the constructions, whose features are more similar to those of native language. Even more importantly, the students seem to have developed inductive skills, which allow them to make generalisations (e.g. about the construction's meaning) and to form patterns, just like someone acquiring their L1. In that sense, DDL is a pedagogical method that brings L2 acquisition a little closer to L1 acquisition – although the two processes will of course always remain distinct from one another.

While most DDL interventions described in the literature so far have been item-based, the present study has demonstrated that the method is also suitable for constructions. Data-driven construction learning basically presents the same advantages and disadvantages as item-based DDL, and its results are equally mixed, both in terms of progress and participants' attitudes (see Gilquin & Granger (2010) on the general strengths and weaknesses of DDL). Yet, the fact that the method appears to be applicable to different types of linguistic phenomena confirms, if confirmation were needed, that DDL deserves a place in today's language classroom, on a par with – and as a complement to – the other pedagogical practices adopted by teachers to the benefit of their students.

References

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, *67*(2), 348-393.

Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 478-497). Cambridge: Cambridge University Press.

Ellis, R. 2008. *The study of second language acquisition. Second edition*. Oxford: Oxford University Press.

Gabrielatos, C. (2005). Corpora and language teaching: Just a fling or wedding bells? *TESL-EJ*, *8*(4), 1-35.

Gilquin, G. (2008). What You Think Ain't What You Get: Highly polysemous verbs in mind and language. In J.-R. Lapaire, G. Desagulier & J.-B. Guignard (Eds), *Du fait grammatical au fait cognitif. From gram to mind: Grammar as cognition. Volume 2* (pp. 235-255). Pessac: Presses Universitaires de Bordeaux.

Gilquin, G. (2010). *Corpus, cognition and causative constructions*. Amsterdam: John Benjamins.

Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359-370). London: Routledge.

Gilquin, G., & Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction, 1*(1), 41-61.

Goldberg, A. E. (2006). *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Johns, T. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100-115). London: Longman.

Kim, J.-B., & Davies, M. A. (2015). The INTO-CAUSATIVE construction in English: A construction-based perspective. *English Language and Linguistics*, *20*(1), 55-83.

Luzondo Oyón, A. (2013). Revisiting Goldberg's semantic constraints on the 'way' construction. *Revista Española de Lingüística Aplicada*, *26*, 349-364.

McEnery, T., & Wilson, A. (1997). Teaching and language corpora (TALC). *ReCALL*, *9*(1), 5-14.

McKay, S. (1980). Teaching the syntactic, semantic and pragmatic dimensions of verbs. *TESOL Quarterly*, *14*(1), 17-26.

Rapti, N. (2010). A study of classroom concordancing in the Greek context: Data-driven grammar teaching and adolescent EFL learners. PhD thesis, University of Nottingham.

Roland, R., & Jurafsky, D. (2002). Verb sense and verb subcategorization probabilities. In S. Stevenson & P. Merlo (Eds.), *The lexical basis of sentence processing: Formal, computational, and experimental issues* (pp. 325-346). Amsterdam & Philadelphia: John Benjamins Publishing Company.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Torres-Martínez, S. (2017). Working out multiword verbs within an Applied Cognitive Construction Grammar framework. *European Journal of Applied Linguistics*, *5*(1), 55-86.

Yoon, H., & Jo, H. (2014). Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning & Technology*, *18*(1), 96-117.

Appendix 1. Worksheet for the DDL activity on the MAKE causative construction

Look at the examples of the <MAKE + NP + INFINITIVE> construction and answer the following questions.

1.1. Circle the infinitive (MAKE + NP + **INFINITIVE**) in all the examples.

1.2. Identify the infinitive that is the most frequent in the examples: _____

1.3. What kind of verb is this? 0 intransitive 0 transitive 0 copular 0 other: _____

1.4. Can you find other verbs from the same syntactic category (1.3) among the infinitives used in the examples? List them: _____

_____

1.5. What are some of the other verbs that are frequently used in the examples? _____

_____

1.6. Can you think of a semantic description that would apply to most of the infinitives found in the construction? _____

_____

2.1. Underline the subject (**NP** + MAKE + NP + INFINITIVE) of all the constructions.

2.2. Do the subjects mostly refer to: 0 animate entities 0 inanimate entities 0 both

3. On the basis of your observations, what can you say about the meaning of the <MAKE + NP + INFINITIVE> construction? (be as specific as you can)

_____

_____

_____

_____

_____

4. What other discoveries have you made about the construction? (continue overleaf if necessary)

_____

_____

_____

_____

_____

_____

Appendix 2. Worksheet for the DDL activity on the *way* construction

Look at the examples of the <Verb + Possessive Determiner + *way* + Prepositional Phrase> construction and answer the following questions.

1.1. Read the first example. Identify the two Prepositional Phrases: _____

_____

1.2. Paraphrase the sentence in your own words: _____

_____

2.1. Now read the second example. Try to guess the meaning of "worm one's way into" by relying on the meaning of the noun *worm* (Fr. "ver [de terre]"). Write it down here: _____

_____

2.2. Once you have answered the preceding question (only then!), use a dictionary to find the meaning of the phrase "worm one's way into". Find a definition that corresponds to the context of the first example, and write it here, together with the name of the dictionary you have used: _____

_____

2.3. The phrase "worm one's way into" can also be used in another, more metaphorical meaning. Use a dictionary to find the definition corresponding to this abstract use and write it here, together with an example and the name of the dictionary you have used: _____

_____

2.4. Explain how the concrete meaning (2.2) is related to the more abstract meaning (2.3):

_____

_____

3.1. Examine the other examples and in each of them highlight the verb of the construction (**Verb** + Poss. Det. + *way* + PP). If necessary, consult a dictionary to find the meaning of the verb.

3.2. Write down some of the verbs that you think are interesting/surprising/funny/etc. in the construction: _____

_____

_____

4. On the basis of your observations, what can you say about the meaning of the <Verb + Possessive Determiner + *way* + Prepositional Phrase> construction? Be as specific as you can, and in particular specify what the semantic role of the verb is.

_____

_____

_____

_____

5. What other discoveries have you made about the construction? (answer overleaf)

Appendix 3. Worksheet for the DDL activity on the *into* causative construction

Look at the examples of the <Verb + Noun Phrase + *into* + Ving> construction and answer the following questions.

1.1. Read the first sentence. Try to paraphrase it in your own words: _____

_____

1.2. Once you have answered the preceding question (only then!), use a dictionary to find the meaning of the phrase "talk someone into doing something". Write it here, together with the name of the dictionary you have used: _____

_____

1.3. If your dictionary mentions the opposite of "talk someone into doing something", please write it here, together with an example. (Tip: this involves using a different preposition.) _____

_____

2.1. Now read the second example. Identify the three coordinated verbs (**Verb** + Noun Phrase + *into* + Ving): _____

2.2. Examine the other examples and in each of them highlight the first verb of the construction (**Verb** + Noun Phrase + *into* + Ving). If necessary, consult a dictionary to find the meaning of the verb.

2.3. Write down some of the verbs that you think are interesting/surprising/funny/etc. in the construction: _____

_____

_____

_____

3. On the basis of your observations, what can you say about the meaning of the <Verb + Noun Phrase + *into* + Ving> construction? Be as specific as you can, and in particular specify what the semantic role of the first verb is.

_____

_____

_____

_____

4. What other discoveries have you made about the construction?

_____

_____

_____

_____

_____

Appendix 4. Evaluation questionnaire

Over the last few weeks, you have completed a number of corpus-based tasks to help you learn about three grammatical constructions. Please answer the following questions about these tasks.

1. Which of these sessions have you attended?
   - ○ MAKE + Noun Phrase + Infinitive
   - ○ Verb + Possessive Determiner + *way* + Prepositional Phrase
   - ○ Verb + Noun Phrase + *into* + Ving

2. What is the first word that comes to your mind in relation to the tasks? _____

3. Rate the following statements according to this scale:
1 = I strongly disagree; 2 = I disagree; 3 = I agree; 4 = I strongly agree

The tasks were interesting.	1	2	3	4
The tasks were fun.	1	2	3	4
The tasks were useful.	1	2	3	4
The tasks helped me learn more about the constructions.	1	2	3	4
The example sentences were easily understandable.	1	2	3	4
The questions were clear and easy to understand.	1	2	3	4
The tasks were easy to carry out.	1	2	3	4
I became more efficient at the tasks after each session.	1	2	3	4
I had already done similar exercises before.	1	2	3	4
I would like to do more exercises like these.	1	2	3	4
I prefer tasks like these to being taught by the lecturer.	1	2	3	4
I think I will remember more about the constructions than if they had been taught by the lecturer.	1	2	3	4
Overall, I enjoyed these tasks.	1	2	3	4

4. What aspects of grammatical constructions (and of language in general) do you think can best be learned on the basis of such tasks?

_____

_____

5. What have you liked about the tasks? What do you think are the advantages of such tasks?

_____

_____

_____

6. What haven't you liked about the tasks? What do you think are the disadvantages of such tasks?

_____

_____

_____

7. How do you think these tasks could be improved to make them more useful?

_____

_____

_____