

DOES L2 ASSESSMENT MAKE A DIFFERENCE? TESTING THE EMPIRICAL VALIDITY OF APPLIED COGNITIVE LINGUISTICS IN THE ACQUISITION OF THE SPANISH/L2 PSYCH-VERB CONSTRUCTION

Beatriz Martín-Gascón

Universidad Complutense de Madrid, Spain

Reyes Llopis-García

Columbia University, USA

Irene Alonso-Aparicio

Columbia University, USA

Abstract

This article lies within the field of Applied Cognitive Linguistics (ACL) and presents empirical work that addresses overlooked effects of assessment typology in L2 learning. It examines whether pairing a cognitive instructional approach with matching assessment design results in greater learning outcomes over more pervasive notional-functional approaches. Although the last two decades have witnessed a proliferation of empirical research searching for evidence of the productivity of cognitive approaches, studies have only been partly fruitful in eliciting data that truly favors ACL. We argue that this is largely due to assessment design, which typically measures performance via correct-vs-incorrect tasks. To overcome this caveat, two studies addressing the complex Spanish psych-verb construction were conducted with a pretest/posttest/delayed posttest design for three research conditions (control, cognitive, and traditional): a pilot study ($n=59$) and a larger replication ($n=160$). Data collection entailed ACL-based assessment for interpretation and production tasks. Results showed that after instruction, the cognitive group significantly outperformed the traditional in both tasks. These findings lend support to the effectiveness of cognitive instruction and assessment for difficult grammatical constructions.

Keywords: *Spanish/L2; psych-verbs; Applied Cognitive Linguistics; L2 assessment; L2 acquisition; L2 pedagogy*

INTRODUCTION: APPLIED COGNITIVE LINGUISTICS AND INSTRUCTED L2 DEVELOPMENT

Cognitive Linguistics (CL), as an interdisciplinary approach to language, focuses on speakers' cognitive abilities and claims that the linguistic system –its underlying structure and how we acquire it– is to be understood in terms of cognitive processes such as attention, memory, perception, reasoning, or mental processing (Ibarretxe-Antuñano, Cadierno, and Castañeda, 2019). Furthermore, CL approaches are based on the experientialist view of human faculties and contend that language meaning is experiential, and thus embodied and symbolic (Ellis, 2019; Gibbs, 1996, 2005; Lakoff & Johnson, 2002). Departing from structuralist and generativist models, cognitivist approaches claim that frequency and meaning of forms, as usage-based properties, determine the mental representation of grammatical constructions (Tomasello, 2005).

The last two decades have witnessed a growing interest in the potential of CL to gain deeper insights into how language conceptualization and representation work, and as such, Applied Cognitive Linguistics (ACL) has emerged (Pütz, Niemeier, and Dirven, 2001; Tyler, Huang, and Jan, 2018) as a research field that seeks to contribute new pathways for the understanding of the L2 teaching-learning process (Achard & Niemeier, 2004; Cadierno & Eskildsen, 2015; Ellis & Cadierno, 2009; Ibarretxe-Antuñano, Cadierno, and Castañeda, 2019; Littlemore, 2009; Littlemore & Juchem-Grundmann, 2010; Robinson & Ellis, 2008; Tyler, 2008, 2012). Since then, significant empirical studies have addressed the relative effectiveness of L2 cognitive-based pedagogical approaches against more traditional didactic options that tend to rely on a structuralist view of the language (e.g., Author 2, 2010; Author 3 & Author 2, 2019; Boers & Lindstromberg, 2008; De Knop & De Rycker, 2008; De Knop, Boers, and De Rycker, 2010; Holme, 2009; Lam, 2009). Most studies, however, have only been partially successful in eliciting data favoring the cognitive condition versus traditional ones (Author 2, forthcoming; Bielak & Pawlak, 2011; Bielak et al. 2013; Kissling et al., 2018; Kohl-Dietrich et al., 2016; Piquer-Piriz et al., 2022; Reif, 2012; Suñer & Roche, 2019; Tyler, 2008; Tyler et al., 2011; Yasuda, 2010). These outcomes differ from the cognitive-based L2 classroom reality, where both students and instructors report positive views and experiences (Author 2, forthcoming; Suñer & Roche, 2019). It is contended here that an explanation of these results might be found in the type of assessment tools used for empirical research in L2 acquisition and pedagogy. As claimed by Author 2 and Author 3 (2018, 2019) and Author 2 (2021, forthcoming), these remain unchanged and favor traditional learning-teaching approaches in detriment of a more innovative, cognitive-based pedagogy.

In light of this evidence, the aim of this article is to explore whether a cognitive-based approach to teaching, but also to assessing, results in greater learning outcomes than a traditional one in the development of the Spanish/L2 psych-verb construction (PVCx) at the elementary level (e.g., *me gusta*, 'I like'), a well-known challenging form in the acquisition of Spanish/L2 (e.g., Author 1, 2020, 2021; Gascon, 1998; Marras & Cadierno 2008; Miglio et al., 2013).

This article is structured as follows. First, the PVCx will be described, as well as their presence in L2 teaching and learning. Second, an explanation of what traditional instructions are will be briefly addressed (i.e., notional-functional approaches), specifically for the case of the PVCx in Spanish/L2. Third, an ACL view of these constructions for the L2 classroom will be discussed, together with a pilot study that dives into the empirical challenge towards ACL research for the design and implementation of assessment materials. Then, the main study will be presented alongside its methodological aspects, and special attention will be paid to the assessment tasks for data collection. The results will be reviewed in light of our research questions, and finally, a discussion of findings and the newly found implications for empirical testing and assessment design will follow.

To date and to the best of our knowledge, no investigation has examined the pedagogical benefits of the ACL approach in the L2 instructional setting using ACL-based assessment, nor have Spanish psych-verbs been empirically addressed from this perspective. The

hypothesis posited here is that an ACL-based pedagogy will be more fruitful in enhancing students' L2 learning outcomes of our target linguistic construction when learning is coherently assessed with the same methodology as the instruction.

THE SPANISH PSYCH-VERB CONSTRUCTION: A THEORETICAL AND PEDAGOGICAL CHALLENGE FOR ANGLOPHONE LEARNERS

Psych-verbs, i.e., verbs that express mental states, changes of state or psychological processes (Jackendoff, 1990; Levin, 1993), have consistently been an intriguing topic for linguists and psycholinguists alike because of their pervasiveness and peculiar grammatical behavior (Croft 1986; Dowty, 1988, 1991; Jackendoff, 1990, 2007; Levin, 1993; Van Valin, 1990, 2005; Vázquez Rozas, 2006, 2012; Zaenen, 1993). They constitute a general class that comprises verbs of affection (e.g., 'like', 'bother'), perception (e.g., 'see', 'touch'), cognition (e.g., 'know', 'believe'), and evaluation (e.g., 'appreciate', 'respect') (Bossong, 1998). The most prominent subclass with cross-linguistically identifiable semantic and syntactic characteristics is that of verbs of affect (Talmy, 2000), and accordingly, the term 'psych-verbs' has been often used to design this subcategory. Due to the distinct and unsystematic ways in which they are represented, psych-verb constructions (PVCx) have been analyzed in a variety of languages including Italian (Belletti & Rizzi, 1988), Brazilian Portuguese (Cançado, 1996), Spanish (Fábregas, Marín, and McNally, 2012; Mayoral-Hernández, 2012), Navajo (Jelinek & Willie, 1996) English (Hartshorne et al., 2017) German (Klein & Kutscher, 2002, 2015), Mandarin Chinese (Liu, 2016), or Sign Language of the Netherlands (Oomen, 2016).

In Spanish, the PVCx belongs to one of the areas of grammar that presents the most difficulties, especially for English native speakers (González, 1998; Mayoral-Hernández, 2012; Montrul, 1997a, 1997b; Marras & Cadierno, 2008). This acquisitional struggle is the result of a conflict between the natural level of prominence of the object (the *experiencer*) and that of the subject (or *stimulus*) in the Spanish PVCx (cursives in Talmy's 2000 terminology). Such a correspondence between the natural prominence of the *experiencer* and its syntactic function does not occur in English (Marras & Cadierno, 2008, p. 246). Let us consider, for instance, the psych-verb *gustar* 'to please', whose most frequent English equivalent is 'to like'. The English grammatical construction casts its arguments in a similar fashion to prototypical transitive constructions such as 'I drink coffee', that is, with a nominative subject (the *experiencer*) and an accusative direct object (the *stimulus*), both based on the semantic roles of the participants, as in 'I like Spanish Grammar' (Example 1). The Spanish equivalent *Me gusta la gramática española* (Example 2) casts the *stimulus*, i.e., the cause of the psychological state taking place in the mind of the experiencer (or as scholars refer to it, what "causes some emotional reaction" (Dowty, 1991, p.579; Talmy, 2000, p.101)) as the syntactic subject. The *experiencer*, or entity that undergoes a certain emotion, functions as the indirect object or dative, capitalizing, therefore, on the semantic roles of each (Miglio et al., 2013, p. 268). Structurally, then, the Spanish PVCx differs from that of English in how the two arguments semantically categorized as *experiencer* and *stimulus* surface as syntactic entities.

- (1) *I* exp/subj *like* psych-verb *Spanish grammar* stim/object (+freq.)
 (2) *Me* exp/object *gusta* psych-verb *la gramática española* stim/subj (+freq.)
 Me dat *likes* 3sg *the grammar Spanish* sg.nom
 ‘Spanish grammar pleases me’ (-freq.)

Spanish verbs of affect can also appear in transitive constructions with the *experiencer* as subject (e.g., *amar* ‘to love’, *odiar* ‘to hate’), yet most psych verbs appear only in the inverse construction explained above (e.g., *encantar* ‘to delight’, *fastidiar* ‘to be annoyed by’). These constructions exist in many languages and tend to express the cross-linguistic tendency to encode the experience of feelings in a fundamentally different way from the encoding of canonical transitive actions (Viola et al., 2013, p. 268). Hence, inverse constructions are typologically linked to experiential predicates that are less transitive (Bauer, 2000; Bossong, 1998; Dahl & Fedriani, 2012), and therefore express less volition, control, and affectedness of the object than agentive ones (Dahl & Fedriani, 2012, p. 343). The English language can also lexicalize the *experiencer* as the object, as in (5), but it is more common to find the *experiencer* argument as the subject (6).

- (5) *Spiders* stim/subj *please / disgust / frighten the children* exp/obj
 (6) *The children* exp/subj *like / hate / fear spiders* stim/obj

According to the Contrastive Analysis Hypothesis (Ellis, 1990, p. 25), the closer the correspondence of a linguistic structure of the L2 is to its equivalent in the L1, the easier it is learned, and vice versa. Since most frequent English PVCx make use of an *experiencer*-subject structure, while the equivalent expressions in Spanish (e.g., *gustar* ‘to like’, *encantar* ‘to delight’, *molestar* ‘to be bothered by’, *fastidiar* ‘to be annoyed by’, *interesar* ‘to be interested by’) recur to inverse constructions with *experiencer*-as-object, it is not surprising that English learners of Spanish struggle with these constructions (Gascon, 1998, p.70).

THE TRADITIONAL PEDAGOGICAL APPROACH TO THE SPANISH/L2 PVCx
 ‘Traditional instruction’ is not an easy concept to pin down, just as ‘Communicative Language Teaching’ is hard to define (Dörnyei, 2013). Both terms encompass a wide range of methodological criteria and classroom-based dynamics, but Larsen-Freeman (2015, p.263) hits core issues in her definition of traditional instruction as “grammar teaching centered on accuracy of form and rule learning, and with mechanical exercises seen as the way to bring about the learning of grammar.” This relates directly to the still-current notional-functional approach that largely portrays grammar as a “foreign entity”, which features the descriptive rules and lists of usage-and-context that are often found in most Spanish/L2 market textbooks (Author 2, Author 5, and Author 6, 2012; Author 2 & Author 4, 2019; Mayoral-Hernández, 2012). Author 1 (2020, 2021) explored the inclusion and usage of PVCx in 70 units from 36 Spanish/L2 textbooks. Proficiency levels ranked from beginner to low-intermediate (A1, A2 and B1, according to the *CEFR*). Results from quantitative and qualitative analyses showed an unsystematic inclusion of

these constructions and great variation in the contents assigned to each level. Additionally, negative emotions conveyed by psych-verbs were relegated to higher levels of proficiency, with only positive verbs at the A1-A2 levels (usually *gustar -to like-*, *encantar -to love-* or *interesar -to interest-*). These studies also confirmed that instruction of the Spanish/L2 PVCs tends to follow this pedagogical approach outlined by Larsen-Freeman: a rule of syntactic structure, usage contexts, and practice exercises of correct vs. incorrect options. Most textbooks mainly focus on formal aspects and taxonomies through the syntactic properties of psych verbs in constructions created *ad hoc* to the teaching unit at hand (e.g., Brewer, 1970; Comrie, 1989; Dahl & Fedriani, 2012). These are based on pattern recognition, syntactic combinations and the placement of the various elements involved in the construction (subject, verb and object pronoun) in the correct order. Assessment tasks then include grammaticality judgements, fill-in-the-blank exercises, multiple choice, or true/false statements, all within a correct/incorrect dichotomy. In sum, and as the textbook analysis points out, attention to the relationship between the linguistic structure and its meaning in order to prioritize communicative intent has not been the focus of traditional instruction.

ACL: A PRODUCTIVE FRAMEWORK FOR THE PEDAGOGY OF THE SPANISH/L2 PVCx

From the perspective of cognitive semantics, the way speakers conceptualize an emotional state affects how they construe it linguistically (Talmy, 1985, 2000). In this regard, it is important to understand that many issues in L2 learning derive from these specificities in the conceptualization of each language, i.e., how every discourse community experiences and perceives the world, and how they communicate about it). This idea of linguistic conceptualization as being universal but not equal across languages is emphasized by Taylor (1993) and Littlemore and Taylor (2014), who recommend that language instruction acknowledge that formal differences in construal between languages result from conceptual differences. From this perspective, a contrastive pedagogical strategy is to be advocated, and since learners tend to transfer from their L1 to the L2 (Boers, 2013; Cadierno, 2008; VanPatten, 2004), raising awareness of the grammatical variances between languages presents a pedagogical advantage for a more meaningful assimilation of these difficult constructions. In teaching, then, L2 learners should be made aware of the conceptual image or representation associated with a grammatical construction (such as the PVCx) in both languages.

Slobin (1996, p. 76), following the foundational tenets of CL, conceives language and its grammar as one, and as such, only through form-and-meaning symbolic units can speakers communicate their perspective on a given situation. According to this author, children learn particular linguistic patterns when acquiring their native language, which means that they learn to focus on specific dimensions of experience that are, in turn, embedded in the grammar of their L1. This has been termed *thinking-for-speaking* and links the conceptualization of sensorimotor experiences and encyclopedic knowledge to grammatical structure. Author 2 (2011) emphasizes the impact of *thinking-for-speaking* in the L2 teaching-learning process and argues that the idea of language as a reflection of how a linguistic community perceives the world, humanizes and motivates the language

classroom. Cognitive semantics, as opposed to objectivist semantics, defends the idea that languages are not neutral coding systems of an objective reality but “a subjective experience to the world of human experience” (Slobin, 1996, p. 91). Talmy (1985, 2000) argues that languages differ in terms of the type of information presented in the foreground and in the background, an idea also put forward in cognitive grammar by Langacker as ‘profiling’ (2008, 2016). Talmy (2000, p. 128) refers to this cognitive phenomenon as ‘salience’, i.e., “the degree to which a component of meaning, due to its type of linguistic representation, emerges into the foreground of attention or, on the contrary, forms part of the semantic background where it attracts little direct attention.” In this sense, languages can vary in the lexicalization of different facets of emotions and profile different participant roles according to their conceptualization (Langacker, 1999). This is consistent with the *experiencer* and *stimulus* view of the PVCx, which brings semantic roles to L2 instruction over morpho-syntactic constraints.

ACL provides an interesting framework to study how PVCx conceptualize and encode emotions (likes and dislikes, annoyance, love, etc.) in the lexicon and grammar of a given language. As suggested in previous theoretical studies, *experiencer* and *stimulus* are the two major roles involved in psych-verb constructions (e.g., Dowty, 1991; Jackendoff, 1990, 2007; Levin, 1993; Talmy, 1985, 2000). When describing an affective situation, Talmy (2000) highlights the lexical differences in verbs of affect, which are motivated by distinct semantic and referential roles and whose use varies according to the salience or focus of attention, i.e., if the salient aspect is a quality of the *stimulus* (see Figure 1) or if, otherwise, it is the state of the *experiencer* (see Figure 2). When speakers utter the sentence *Me fastidia tu alarma de las 5:00 am cada mañana*, ‘I am annoyed by your 5:00 am alarm every morning’, their aversion refers to the characteristics and implications of an external *stimulus* (e.g., the alarm and its unwelcome repetitive noise, leading to sleep interruption or deprivation). Differences in lexical usage also vary according to the origin or cause of the emotional state (see arrows in Figures 1 and 2). Thus, when the projection is targeted from an *experiencer* onto an external entity (the *stimulus*/alarm), we encounter a prototypical grammatical construction (Subj.>V.>Obj.). However, if the emotional state results from the semantic protagonism of a *stimulus*/alarm, it gives rise to an inverse construction (Obj.>V.>Subj.).

[FIGURE 1 NEAR HERE]

[FIGURE 2 NEAR HERE]

Up to this section, we have presented the pedagogical challenges outlined so far, as well as the differences in the mapping of the *experiencer* in different semantic and syntactic roles in English and Spanish. Given these considerations, ACL in general, and cognitive semantics in particular, are presented as a productive framework for teaching and learning the expression of emotion in Spanish/L2 through these PVCx, starting at the elementary level. This approach can offer a systematic explanation of the relationship between the semantic/conceptual structure and the formal/linguistic representation, aiming at fostering motivated, form-meaning connections for the learners. To our knowledge, no

published studies have explored the PVCx in Spanish/L2 from a cognitive perspective in order to examine the benefits of the ACL approach and with the use of ACL-based assessment tests. The following section addresses these questions.

THE ACL EMPIRICAL CHALLENGE: TOWARDS NEW AVENUES IN ASSESSING LEARNING OUTCOMES

In the intersection of ACL with L2 Pedagogy, there has been a growing number of empirical studies on the effectiveness of a cognitive linguistic approach over a more traditional methodology of teaching (Lam, 2009; Author 2 & Author 3, 2019; Martín-Gilete & Piquer-Píriz, 2021; Tyler, Mueller, and Ho, 2011; among many others). Published papers, however, are a fraction of the number of studies that are presented at conferences but not beyond. The main reason seems to be that their findings and empirical validity do not meet the hypotheses posed, i.e., the superiority of ACL. This incongruity, however, does not match the literature on cognitive-based L2 pedagogy (Author 2, Author 5, and Author 6, 2012; Castañeda Castro, 2014; Ibarretxe-Antuñano, Castañeda Castro, and Cadierno, 2019), nor the classroom experience of L2 instructors at large (Author 2, forthcoming; Holme, 2009).

Classroom-based empirical research in general is very complex because there are many intervening factors at play: disparity of proficiency levels within the same level, student attendance, cognitive fatigue during posttests, the novelty of a researcher who teaches the instruction part of a study, or simply, the needs of a classroom setting with a curriculum to cover in a limited amount of time. At the curricular level, additionally, metaphorical competence and other cognitive-based notions are not part of either the *CERF* or the *ACTFL* (American Council on the Teaching of Foreign Languages) descriptors, so their absence in mainstream classroom textbooks and materials hinders their pedagogical potential (Nacey, 2017).

We posit here that the main hurdle that prevents statistically significant results for ACL, however, stems from a key methodological aspect: the design and implementation of assessment tests, as contended by Author 2 and Author 3 (2018, 2019), Suñer and Roche (2019), Arnett and Suñer (2019), and Author 2 (2021, forthcoming). To our knowledge, all studies conducted in the comparison of an ACL approach against more traditional methods (notional-functional instruction with correct-vs-incorrect testing), have implemented pre- and posttests that measured learner performance via traditional tasks: grammaticality judgements, fill in the blanks, true vs. false, or multiple-choice exercises, all with one correct answer only. This presents a major disadvantage for the cognitive-based group, since their instruction is based on semantic interpretations and communicative intent and not on right vs. wrong choices. Additionally these tasks are what students are familiar with, and what presents less of a challenge for them in the assessment of instructional effects. These types of tasks are the constant companions of learners in regular classroom testing, and students complete them during their day-to-day instruction. ACL instruction, however, focuses on embodiment, semantic motivation, image-based form-and-meaning pairings, or the saliency of communicative intent. These pedagogical notions veer greatly from the more automated answer choices of traditional

tasks, so assessing the effects of a novel, brand new instruction with traditional tests stacks the odds against the cognitive groups every time.

From this point of view, when studies report no statistical differences between the traditional and the cognitive groups, instead of a failure for the cognitive instruction, what should be inferred is a tremendous success of the latter. The reason for this is that while the cognitive instruction usually introduces completely new ways of processing and understanding target forms, the (traditional) assessment used does not evaluate them at all. And yet, the performance of the cognitive group usually equals that of the traditional block, making the cognitive instruction effective despite the lack of coherent assessment. With the two studies presented here, we put forward the notion that novel instructional methods require evaluation tasks that match the new way of learning and may therefore more effectively measure the effects of their innovative approach.

The question to pose, then, becomes: how would a cognitive linguistic group perform when given the chance to demonstrate new learning through cognitive-based assessment tasks? Since ACL approaches to L2 learning concentrate on embodied meaning, perspective, metaphorical mappings, and speakers' stance (among others), L2 learning should be assessed following the same methodology of the approach at stake.

The Pilot study

Prior to the main study, a pilot study was conducted to address these concerns and examine the feasibility of the methodology intended for a larger-scale iteration. Students from a North American-based university taking online elementary Spanish/L2 courses participated in the study. The initial pool consisted of 59 learners from different course sections and was randomly assigned to one of the three research conditions: a control or comparison group (CG) ($n = 15$), a traditional instruction group (TIG) ($n = 22$) and a cognitive instruction group (CIG) ($n = 22$). Students were asked to complete three tests (pretest, posttest and delayed test) of 20 minutes each. These tests were delivered online via Wufoo and included two types of tasks: interpretation and production at a very basic level to ensure very targeted form-meaning connections. On the basis of two criteria (i.e., achieving scores of less than 60% in the pretest and attending all experimental sessions), the pool was reduced to a final group of 40 participants (CG = 6, CIG = 18, and TIG = 16). The CIG experimental groups were taught using an ACL-based approach, while a more traditional teaching method consistent with mainstream textbooks (Author 1, 2020, 2021) was used for the TIG. The CG received no instruction on the target form. Despite the small-sized samples ($n < 20$), departure from normal distribution was found in most of the subsamples. Thus, nonparametric tests were selected for statistical analyses. In the Interpretation task, results from Friedman tests yielded a significant difference in test scores across the three testing situations only for the CIG ($\chi^2(2) = 32.118, p = .000$) and a Wilcoxon signed-rank test revealed a statistically significant increase in scores ($Z = -3.732, p = .000$), with a medium effect size ($r = .62$). In turn, results from Kruskal-Wallis tests showed a statistically significant difference between groups in the posttest ($H(2) = 31.527, p = .000$) and the delayed test ($H(2) = 28.366, p = .000$). Additionally, results from Mann-Whitney U tests revealed differences between experimental treatments on the

immediate posttest ($Z = -4.913, p = .000$) and on the delayed posttest ($Z = -4.721, p = .000$).

In the Production task, Friedman tests yielded a significant difference in test scores across the three testing situations for the TIG ($\chi^2(2) = 11.220, p = .004$) and the CIG ($\chi^2(2) = 26.638, p = .000$). In fact, Wilcoxon signed-rank tests revealed significant increases in scores for the TIG ($Z = -2.523, p = .012$), with a medium effect size ($r = .44$), and the CIG ($Z = -3.627, p = .000$), with a medium effect size ($r = .60$). On the other hand, results from Kruskal-Wallis tests showed a statistically significant difference between groups in the posttest ($H(2) = 13.618, p = .001$) and the delayed test ($H(2) = 16.872, p = .000$). Also, results from Mann-Whitney U tests revealed differences between experimental treatments on the immediate posttest ($Z = -2.669, p = .007$) and on the delayed posttest ($Z = -3.812, p = .000$).

The promising results obtained in the pilot study led to a larger-scale replication in which all conditions were held constant for all variables involved. The following section addresses the main study.

THE MAIN STUDY

Research Questions and Hypotheses

In order to extend our previous research on the effectiveness of an ACL-inspired pedagogical approach, the following research questions were explored:

1. (RQ1). What is the relative effect of an ACL-based approach and a traditional approach to teaching psych-verbs when knowledge is measured by means of ACL-based assessment interpretation tasks?
2. (RQ2). What is the relative effect of an ACL-based approach and a traditional approach to teaching psych-verbs when knowledge is measured by means of ACL-based assessment production tasks?

In light of the previously reviewed literature, the hypotheses were: H1. An ACL-based pedagogical approach will render better results than a traditional pedagogical approach when knowledge of psych-verbs is measured by means of ACL-based interpretation assessment tasks; H2. An ACL-based pedagogical approach will render better results than a traditional pedagogical approach when knowledge of psych-verbs is measured by means of ACL-based production assessment tasks.

Participants

Participants were recruited from an initial pool of 160 undergraduate students enrolled in eleven intact sections and attending their first semester of an online elementary Spanish/L2 course as part of their core curriculum at a university in North America. The eleven sections were randomly assigned to one of the three treatment groups: (a) a CG ($n = 41$); (b) a TIG ($n = 53$); and (c) a CIG ($n = 66$).

The same two criteria as with the pilot study were in place for sample selection. First, only those participants who scored below 60% in a pretest were included. Second, participants who were absent for one or more sessions during the experimental treatment

were excluded. After these selection criteria, the initial pool of 160 students was reduced to 140 participants, distributed as follows: (a) CG ($n = 32$); (b) TIG ($n = 49$); and (c) CIG ($n = 59$).

Procedure

In order to examine linguistic learning when both instructional and assessment procedures align from an ACL perspective, a pretest/posttest/delayed posttest design was implemented. The experimental phase occurred during three regular class sessions (75 minutes each) and spanned approximately one week and a half. Session 1 involved the completion of a consent form as required by the committees of human research protection regarding the design and duration of the study, plus the administration of pretests. No explicit information about the nature of the experiment was offered to the participants. Session 2 took place two days after Session 1 and involved both the instructional phase and the administration of the first posttest. During the instructional phase, both experimental groups were exposed to the target construction while the control group worked on an unrelated linguistic form. Posttests were administered immediately after the instruction. Session 3 was delivered four days after Session 2 and entailed the implementation of the delayed posttests. Participants were allowed a maximum of 20 minutes for the completion of each test.

Instructional materials

Two instructional packages were designed and implemented: a traditional one (see Appendix A) and a cognitive counterpart (see Appendix B). The traditional teaching package was based on most market-ready textbooks and it involved the presentation of the target form via tables, followed by a sequence of activities that required students to manipulate it in order to proceduralize the new content. The presentation of the target form was carried out inductively by means of a text flooded with the target structure (Figure 3). Students were required to read and discuss the topic at hand through some comprehension checks. Next, the target form was presented through tables depicting morphosyntactic roles and with the conjugation of psych-verbs (see also Figure 3). These materials also incorporated the metalanguage used with the CIG as well, with the goal to ascertain that the TIG would not be at a total disadvantage during the assessment phase. The presentation and discussion of the target form spanned about 20 minutes, and afterwards, students were asked to work on the target form to think of and produce sentences with the PVCx under study. Practice of the target form entailed activities focused on morphosyntactic cues (e.g., conjugating psych-verbs based on given grammatical subjects and objects), as well as free sentence formulation.

[FIGURE 3 NEAR HERE]

Alternatively, the cognitive-based instructional package was inspired by ACL tenets with regard to PVCx. Parallel to the traditional package, the CL-based instruction also involved the presentation and practice of the target structure, as well as the morphosyntactic metalanguage present for the TIG. The didactic sequence started by

comparing psych-verbs to other verbs that follow the prototypical grammatical construction, also present in their L1 (Subj.>V.>Obj.) (Figure 4). For the next 20 minutes, the target form was explicitly presented through explanations that included cognitive aspects such as embodiment, motion, experientially-based metalanguage (i.e., *experiencer* and *stimulus/main idea*), the use of GIFs to represent the *stimuli*, or attention to the motivation behind the grammatical structure (see also Figure 4). In the final stage, learners were asked to practice what they had learned by creating a) sentences with verbs, *stimuli* and *experiencers* provided to them; and b) by coming up with new sentences themselves, just like the TGI.

[FIGURE 4 NEAR HERE]

Data elicitation instruments

Data elicitation and assessment tools involved a pretest (Appendix C), a posttest (Appendix D), and a delayed posttest (Appendix E) designed and delivered via Wufoo, a web application for the design of SurveyMonkey online forms. Each test included two tasks, an interpretation task and a production task. In turn, each task contained multiple items, both used traditional and ACL-based concepts (e.g., subject and idea/performer), and both showed the answer option “Skip it” to prevent students from guessing the answers (see samples in Figures 5 and 6). Additionally, all exercises had all prompts and content in English and Spanish in order to ensure availability of the elementary learner’s very limited processing resources for the form-meaning connections sought (following studies like DeKeyser & Sokalski, 1996 or Gudmestad, 2006) because, “simply put, fewer unknown words in the input string requires less processing of novel lexical items that in turn means releasing resources to process grammatical form.” (VanPatten, 2004, p. 12). The exercises in the interpretation task (see samples in Figure 5) were designed to elicit very basic form-meaning-image connection where learners had to match images to linguistic items by focusing on form, in order to apprehend the meaning of each construction, also in line with tenets for processability from VanPatten’s Processing Instruction (2004; Lee & VanPatten, 2003). Items included the use of ACL-based metalanguage (*experiencer/performer*, *stimulus/main idea*), as well as the traditional, more opaque terminology (indirect object/dative, subject/predicate). Furthermore, the exercises did not focus on the concept of “correct/incorrect” for the answers, but rather on the representational understanding of the meaning behind the grammatical construction.

[FIGURE 5 NEAR HERE]

The production task, on the other hand, included items at the sentence level and required students to perform a simple, but crucial task: to identify the semantics of “who does what to whom” in the PVCx. All activities linked the different morphosyntactic components (subject, verb, indirect object) with their semantic value and asked students to differentiate *experiencer* vs. verb of affection vs. *stimulus/main idea*, depending on the item (see samples in Figure 6). The aim was thus to pair meaning and form as one

teachable unit at a sentence level, as this would allow our beginner learners to focus their efforts on form-meaning connections based on communicative intent.

[FIGURE 6 NEAR HERE]

RESULTS

Interpretation task

The first RQ examined the relative effectiveness of an ACL approach and a traditional approach to teaching psych-verbs when knowledge is measured by means of ACL assessment tools in interpretation tasks. Table 1 displays the means, standard deviations, and confidence intervals of the interpretation scores for all three conditions.

[TABLE 1 NEAR HERE]

The statistical analyses consisted of nonparametric tests as normality criteria were not met for all subsamples. To assess within-group differences across pretest, immediate posttest, and delayed posttest scores, Friedman tests were used. Results indicated significant differences in test scores across the three testing situations for the CG ($\chi^2(2) = 11.322, p = .003$), the CIG ($\chi^2(2) = 92.327, p = .000$), and the TIG ($\chi^2(2) = 65.742, p = .000$). Post-hoc analyses with Wilcoxon signed-rank tests were then conducted. For the TIG, results revealed a statistically significant increase in test scores ($Z = -5.298, p = .000$), with medium effect size ($r = .54$). Indeed, the median score increased from 4.00 in pretest to 6.00 in posttest, to further 8.00 in delayed test. For the CIG, results also revealed a statistically significant increase in test scores ($Z = -6.703, p = .000$), with medium effect size ($r = .62$). The median score increased from 4.00 in pretest to 10.00 in posttest and remained 10.00 in delayed test.

Kruskal-Wallis tests were used to assess between-group differences in the three testing situations. Results showed no statistically significant difference in test scores between baseline, traditional and cognitive conditions in the pretest ($H(2) = 3.064, p = .216$). However, a statistically significant difference in test scores between the three conditions was found both on the immediate posttest ($H(2) = 100.374, p = .000$) and on the delayed test situation ($H(2) = 84.580, p = .000$). Further, to examine the relative effectiveness of the ACL-inspired approach and the traditional approach, Mann-Whitney U tests were run in the three testing situations. Results revealed no significant difference between both experimental conditions on the pretest ($Z = -1.116, p = .265$). There was a significant difference between conditions on the posttest ($Z = -8.345, p = .000$) and on the delayed posttest ($Z = -5.675, p = .000$).

Production task

The second RQ surveyed the relative effectiveness of an ACL approach and a traditional approach to teaching PVCx when knowledge is measured by means of ACL assessment

tools in production tasks. Table 2 shows the means, standard deviations, and confidence intervals of the interpretation scores for all three conditions.

[TABLE 2 NEAR HERE]

The statistical analysis consisted of nonparametric tests as subsamples were not normally distributed. Friedman tests were used to evaluate within-group differences across pretest, immediate posttest, and delayed posttest scores. Results indicated significant differences in test scores across the three testing situations for the CG ($\chi^2(2) = 14.957, p = .001$), the CIG ($\chi^2(2) = 89.876, p = .000$), and the TIG ($\chi^2(2) = 51.158, p = .000$). Next, post-hoc analyses with Wilcoxon signed-rank tests were conducted. For the TIG, results revealed a statistically significant increase in test scores ($Z = -5.755, p = .012$), with medium effect size ($r = .58$). The median score increased from 1.00 in pretest to 5.00 in posttest and remained 5.00 in delayed test. For the CIG, results also revealed a statistically significant increase in test scores ($Z = -6.695, p = .000$), with medium effect size ($r = .62$). Indeed, the median score increased from 1.00 in pretest to 9.00 in posttest and remained 9.00 in delayed test.

Kruskal-Wallis tests were run to test differences between conditions in the three testing situations. Results showed no statistically significant difference in test scores between groups in the pretest ($H(2) = 2.487, p = .288$). However, a statistically significant difference in test scores between the three conditions was found on the immediate posttest ($H(2) = 66.235, p = .000$) and on the delayed test situation ($H(2) = 57.821, p = .000$). Further, Mann-Whitney U tests were run in the three testing situations in order to examine the relative effectiveness of the two teaching approaches. Results revealed no significant difference between both experimental conditions on the pretest ($Z = -0.610, p = .542$). There was a significant difference between conditions on the posttest ($Z = -4.895, p = .000$) and on the delayed posttest ($Z = -5.168, p = .000$).

DISCUSSION

The results of the statistical tests showed no significant differences in the pretest for any of the conditions in the interpretation or the production tasks. However, after the instruction was implemented, results revealed that the cognitive group (CIG) performed significantly better than the traditional (TIG) and the control (CG) blocks for both tasks in both posttests. These findings are in line with the pilot study and are also consistent with the two hypotheses posited, since students in the main study also performed more effectively in interpreting and producing the PVCx after a cognitive teaching session than those who received the notional-functional instruction. The latter still showed improved performance in both tasks and both tests after the intervention, yet this improvement in learning outcomes was not superior to the CIG, even when the assessment also included the terminology that they had been exposed to during their instruction session. It is worth remembering that, as reported in previous sections, all empirical studies conducted to date have implemented 'correct-vs-incorrect only' tests for data elicitation, and that the cognitive-based students, even without having received traditional instruction, still performed as well as the traditional group.

The results presented here confirm that a cognitive-based pedagogy, when followed by a coherent assessment that prioritizes form-meaning connections and meaning motivation, becomes a productive method for teaching and learning difficult grammatical constructions because it yields statistically significant learning gains for all students, but more so for the cognitive groups. Our findings are in agreement with the existing literature on ACL-based L2 teaching (Author 2 et al., 2012, Castañeda Castro, 2014; Ibarretxe-Antuñano et al., 2019; Nacey, 2017) and with the reported classroom experience of L2 instructors (Author 2, forthcoming), who advocate for and directly observe the more memorable and effective teaching-and-learning processes afforded by the ACL paradigm.

More research is needed in order to replicate and corroborate the results here reported, i.e., the pedagogical advantages of teaching with a cognitive approach, as well as the need to assess student learning with the same techniques, instead of using the pervading traditional tasks that highlight a correct-vs-incorrect approach but do not necessarily measure *understanding* of the target form. Future studies should expand the types of tasks used to assess student knowledge beyond the very basic form-meaning connections presented here, and create more “classroom-like” teaching materials for more realistic instruction. Regarding limitations to the studies and recommendations for further research, due to the situation derived from the COVID-19 pandemic, both the pilot and the main study had to be conducted in language courses that were taking place in an online setting. Empirical investigations with face to face, in-class conditions could help corroborate our findings. Additionally, more studies on Spanish and other L2s are desirable in order to assess the validity of the ACL approach across different languages, cementing thus its idoneity for a more communicative and meaning-based L2 classroom.

CONCLUSION

Psych-verbs have been a challenging topic for linguists and L2 instructors alike, due to their uneven and unsystematic representation in different languages. For the case of Spanish, PVCx have been widely studied because of their complex acquisition by L2 learners, especially by native speakers of languages that differ in terms of patterning and order of syntactic and semantic roles, such as English (e.g., Gascon, 1998; Marras & Cadierno, 2008; Miglio et al., 2013). An additional difficulty to the L2 learning of PVCx lies in their irregular inclusion in textbooks (Author 1, 2020, 2021) and the prevailing notional-functional approach in instructional settings. This approach can become rather challenging for English native-speakers in the US, for it relies heavily on the usage of morphosyntactic metalanguage (e.g., subject, direct/indirect object) that is to a certain extent unfamiliar to them.

The study presented here addresses some of these caveats and contributes data to the growing number of studies of ACL and L2 learning in languages other than English. The purpose of this article was to explore whether pairing a cognitive instructional approach with matching assessment design could result in greater learning outcomes over the more pervasive notional-functional approaches described here. The underlying methodological aim was to motivate the expression of emotion already in the early stages of L2 learning

through the use of embodiment, perspective, or the concept of salience, as well as to replace opaque morphosyntactic metalanguage with experience-based and representational roles of the *experiencer* and the *stimulus* (Talmy, 1985, 2000). The other aim attended to with this study was to break with the pervading typology of assessment typically employed by empirical research in L2 teaching and learning.

Both the pilot and the main study presented confirm that an ACL-based instruction of the PVCx in Spanish/L2 creates access to systemic and motivated explanations that are based on the embodiment and real-life perceptions of grammatical structure, and not on obscure morphosyntactic criteria presented as a property of the linguistic system alone. Focusing on these meaningful aspects lends validity to Langacker's (2008) claim that "learning grammar does not have to be the soulless internalization of arbitrary restrictions." (p. 78). Additionally, assessment design has typically focused on tests that were coherent only with traditional instruction, featuring mechanical, correct-vs-incorrect tasks. Assessing learners who receive innovative and novel instructional approaches only with traditional tests treats them unfavorably when measuring learning gains. Departing from a traditional conception of grammar associated with lists of rules and contexts of usage, a cognitive-based approach proposes a more accessible understanding of PVCx. Our two studies render solid evidence in support of a continued exploration of ACL approaches to L2 teaching and learning, especially when assessment design considers cognitive-inspired techniques (i.e., form-meaning-image pairings, motivated understanding of meaning, or semantic identification of linguistic roles), while also including traditional terminology. This way, all instructional groups can make learning gains, with the cognitive group in the statistical lead.

Finally, the work presented here adds to the growing body of research conducted with L2s other than English. To the best of our knowledge, this is the first empirical study to successfully examine the effects of a CL-based methodology in both instruction and assessment design. The positive results for both interpretation and production of a complex construction by elementary L2 learners offer further support to the field of ACL at large. They also contribute to the ecological validity of its methods, opening up new avenues of research for Applied Cognitive Linguistics in the L2 classroom.

REFERENCES

- Author 1. (2020).
- Author 1. (2021).
- Author 2. (2010).
- Author 2. (2011).
- Author 2. (2021).
- Author 2. (Forthcoming).
- Author 2. & Author 3. (2019).
- Author 2., & Author 3. (2018).
- Author 2., & Author 4. (2019).
- Author 2., Author 5., & Author 6. (2012).
- Author 3., & Author 2. (2019).

- Achard, M., & Niemeier, S. (Eds.). (2004). *Cognitive Linguistics. Second Language Acquisition, and Foreign Language Teaching*. Vol. 18. The Hague: De Gruyter Mouton <https://doi.org/10.1515/9783110199857>.
- Arnett, C., & Suñer, F. (2019). Using Cooperation Scripts and Animations to Teach Grammar in the Foreign Language Classroom. *Yearbook of the German Cognitive Linguistics Association*, 7(1), 31–50. <https://doi.org/10.1515/gcla-2019-0003>.
- Bauer B, L. M. (2000). *Archaic Syntax in Indo-European: the Spread of Transitivity in Latin and French*. The Hague: Mouton de Gruyter.
- Belletti, A., & Rizzi, L. (1988). Psych-verbs and θ -theory. *Natural Language & Linguistic Theory*, 291-352.
- Bielak, J., & Pawlak, M. (2011). Teaching English tense and aspect with the help of cognitive grammar: An empirical study. *Studies in Second Language Learning and Teaching*, 1(3), 365. <https://doi.org/10.14746/sslit.2011.1.3.4>
- Bielak, J., Pawlak, M., & Mystkowska-Wiertelak, A. (2013). Teaching the English active and passive voice with the help of cognitive grammar: An empirical study. *Studies in Second Language Learning and Teaching*, 3(4), 581. <https://doi.org/10.14746/sslit.2013.3.4.7>
- Boers, F. (2013). Cognitive Linguistic approaches to teaching vocabulary: Assessment and integration. *Language Teaching*, 46(2), 208-224.
- Boers, F., & Lindstromberg, S. (Eds.). (2008). *Cognitive linguistic approaches to teaching vocabulary and phraseology*. Berlin: Mouton de Gruyter.
- Bossong, G. (1998). Le marquage différentiel de l'objet dans les langues d'Europe. In J. Feuillet (Ed.), *Actance et valence dans les langues de l'Europe* (pp. 193-258). Berlin: Mouton de Gruyter.
- Brewer, W B. (1970). Extent of Verbal Influence and Choice between *le* and *lo* in Alphonsine Prose. *Hispanic Review*, 38(2), 133-146.
- Cadierno, T. (2008). Learning to talk about motion in a foreign language. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 249-285). New York, NY: Routledge. DOI:10.4324/9780203938560-20
- Cadierno, T., & Eskildsen, S. W. (Eds.). (2015). *Usage-based perspectives on second language learning*. Berlin, Germany: De Gruyter Mouton.
- Cançado, M. (1996). Análise descritiva dos verbos psicológicos do português brasileiro. *Revista de Estudos da Linguagem*, 4(1), 89-114.
- Castañeda Castro, A. (2014). Cognitive Linguistic Approaches to Second Language Spanish. In Geeslin, K. (Ed.), *The Handbook of Second Language Acquisition* (pp. 96-110). Oxford: Wiley & Sons.
- Comrie, B. (1989). *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: University Press.
- Cook, V. (1985). Language functions, social factors, and second language learning and teaching. *IRAL*, 23(1-4), 177-198.
- Cook, V. (2016). *Second language learning and language teaching*. Routledge.
- Croft, W. A. (1986). Surface subject choice of mental verbs. Paper presented at the Annual Meeting of the Linguistic Society of America, New York.

- Dahl, E., & Fedriani, C. (2012). The argument structure of experience: experiential constructions in Early Vedic, Homeric Greek, and Early Latin. *Transactions of the Philological Society*, 110(3), 342-362.
- De Knop, S., & De Rycker, T. (Eds.). (2008). *Cognitive approaches to pedagogical grammar: A volume in honour of René Dirven* (Vol. 9). Berlin: Walter de Gruyter.
- De Knop, S., Boers, F., & De Rycker, A. (Eds.). (2010). *Fostering language teaching efficiency through cognitive linguistics* (Vol. 17). Berlin: Walter de Gruyter.
- Deignan, A. (2003). Metaphorical expressions and culture: An indirect link. *Metaphor and symbol*, 18(4), 255-271.
- DeKeyser, R., & Sokalski, K. (1996). The Differential Role of Comprehension and Production Practice. *Language Learning*, 46(4), 613-642.
- Dowty, D. (1988). On the semantic content of the notion 'thematic role'. In G. Chierchia, B. H Partee & R. Turner (Eds.), *Properties, types, and meaning* (Vol. 2) (pp. 69-130). Dordrecht: Kluwer.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3), 547-619.
- Ellis, N. C. (2019). Essentials of a theory of language cognition. *The Modern Language Journal* 103, 39-60.
- Ellis, N. C., & Cadierno, T. (2009). Constructing a second language: Introduction to the special section. *Annual Review of Cognitive Linguistics*, 7(1), 111-139.
- Ellis, R. (1990). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2012). *Language Teaching Research and Language Pedagogy*. NJ: John Wiley & Sons.
- Ellis, R. (2015). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Fábregas, A., Marín, R., & McNally, L. (2012). From psych verbs to nouns. In V. Demonte & L. McNally (Eds.), *Telicity, change, and state: A cross-categorial view of event structure* (pp. 162-184). Oxford: Oxford University Press.
- Gascon, C. D. (1998). The Spanish Psych Verb Construction: Beginning and Intermediate Learners' Patterns of Usage. *Texas papers in foreign language education*, 3(3), 69-84.
- Gibbs Jr, R. W. (1996). Why many concepts are metaphorical. *Cognition*, 61(3), 309-319.
- Gibbs Jr, R. W. (2005). *Embodiment and cognitive science*. Cambridge: Cambridge University Press.
- González, L. (1998). Dative/accusative alternations in gustar-type verbs. *Spanish Applied Linguistics*, 2, 137-167.
- Gudmestad, A. (2006). L2 Variation and the Spanish Subjunctive: Linguistic Features Predicting Mood Selection. In Klee, C. & Face, T. (Eds.), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Language*. Cascadilla Proceedings Project, Massachusetts. Págs. 170-184.

- Hartshorne, J. K., O'Donnell, T. J., Sudo, Y., Uruwashi, M., Lee, M., & Snedeker, J. (2017). Psych verbs, the linking problem, and the acquisition of language. *Cognition*, 157, 268-288.
- Holme, R. (2009). *Cognitive linguistics and language teaching*. Basingstoke: Palgrave Macmillan.
- Ibarretxe-Antuñano, I., Cadierno, T., & Castañeda, A. (Eds.). (2019). *Lingüística cognitiva y español LE/L2*. New York, NY: Routledge.
- Jackendoff, R. (1990). On Larson's account of the double object construction. *Linguistic Inquiry*, 21(3), 427-456.
- Jackendoff, R. (2007). *Language, consciousness, culture: Essays on mental structure*. Cambridge: MIT Press.
- Jelinek, E., & Willie, M. A. (1996). Psych verbs in Navajo. *Athabaskan Language Studies*. University of New Mexico Press, Albuquerque, 15-34.
- Kissling, E., Tyler, A., Warren, L., & Negrete, L. (2018). Reexamining *por* and *para* in the Spanish foreign language intermediate classroom: A usage-based, cognitive linguistic approach. In Tyler, A., Huang, L., & Jan, H. (Eds.). *What is applied cognitive linguistics? Answers from current SLA research*. Berlin: Mouton de Gruyter.
- Klein, K., & Kutscher, S. (2002). Psych-verbs and lexical economy. *Theorie des Lexikons Nr. 122. Arbeiten des Sonderforschungsbereichs*, 282, 1-84.
- Klein, K., & Kutscher, S. (2015). *Lexical economy and case selection of psych-verbs in German*. Institut für Deutsche Sprache, Bibliothek.
- Kohl-Dietrich, D., Juchem-Grundmann, C., & Schnotz, W. (2016). Conceptual motivation as a tool for raising language awareness in the English as a foreign language classroom – Does it enhance learning outcomes? Insights from an empirical study. *Yearbook of the German Cognitive Linguistics Association*, 4(1), 193–210.
<https://doi.org/10.1515/gcla-2016-0013>
- Lakoff, G. & Johnson, M. (2002). Why cognitive linguistics requires embodied realism. *Cognitive Linguistics*, 13(2), 245-263.
- Lam, Y. (2009). Applying cognitive linguistics to teaching the Spanish prepositions *por* and *para*. *Language Awareness*, 18(1), 2-18.
- Langacker, R. W. (1999). Assessing the cognitive linguistic enterprise. *Cognitive linguistics: Foundations, scope, and methodology*, 13-59.
- Langacker, R. W. (2008). *Cognitive Grammar. A basic introduction*. Oxford University Press.
- Langacker, R. W. (2016). Working toward a synthesis. *Cognitive Linguistics*, 27(4), 465-477.
- Lee, J.F., & VanPatten, B. (2003). *Making Communicative Language Teaching Happen*. McGraw-Hill: New York.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.
- Littlemore, J. (2009). *Applying cognitive linguistics to second language learning and teaching*. New York: Springer.
- Littlemore, J., & Grundmann, C. J. (Eds.). (2010). Applied cognitive linguistics in second language learning and teaching. *AILA Review (Special Issue)*, 23(1).

- Littlemore, J., & Taylor, J. R. (Eds.). (2014). *The Bloomsbury companion to cognitive linguistics*. Bloomsbury Publishing.
- Liu, M. (2016). Emotion in lexicon and grammar: lexical-constructional interface of Mandarin emotional predicates. *Lingua Sinica*, 2(1), 1-47.
- Marras, V., & Cadierno, T. (2008). Spanish *gustar* vs. English like: A cognitive analysis of the constructions and its implication for SLA. *Language in the context of use: Discourse and cognitive approaches to language*, 233-254.
- Martín-Gilete, M. (2021, September 9-10). Aplicación de enfoques lingüístico-cognitivos para la enseñanza de verbos con partículas para el aula ILE [Conference presentation]. AALiCo online Conference.
- Mayoral-Hernández, R. (2012). Teaching Subject Position in Spanish: *Gustar* Verbs, Unaccusatives and the Impersonal Passive. *Journal of Language Teaching and Research*, 3(2), 321-330.
- Miglio, V. G., Gries, S. T., Harris, M. J., Wheeler, E. M., & Santana-Paixão, R. (2013). Spanish *lo* (s)-*le* (s) clitic alternations in psych verbs: A multifactorial corpus-based analysis. In *Selected proceedings of the 16th Hispanic linguistics symposium* (pp. 268-278). Somerville, MA: Cascadilla.
- Montrul, S. (1997a). On the parallels between diachronic change and interlanguage grammars: The L2 acquisition of the Spanish dative case system. *Spanish Applied Linguistics*, 1, 87-113.
- Montrul, S. (1997b). Spanish *gustar* psych verbs and the unaccusative *se* construction: The case of dative experiencers in SLA. In A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish* (pp. 189-207). Somerville, MA: Cascadilla Press.
- Nacey, S. (2017). Metaphor comprehension and production in a second language. In Semino, E., & Demjén, S. (Eds.). *The Routledge Handbook of Metaphor and Language* (pp. 503-515). New York: Routledge.
- Nunan, D. (1999). *Second Language Teaching & Learning*. Heinle & Heinle Publishers, 7625 Empire Dr., Florence, KY 41042-2978.
- Oomen, M. (2017). Iconicity in argument structure: Psych-verbs in Sign Language of the Netherlands. *Sign Language & Linguistics*, 20(1), 55-108.
- Piquer Píriz, A. M. (2021, September 9-10). Un recorrido por algunas de las principales aportaciones de la lingüística cognitiva a la enseñanza de lenguas extranjeras [Conference presentation]. AALiCo online Conference.
- Piquer-Píriz, A. M., Castellano-Risco, I., & Martín Gilete, M. (2022, June 28). *Cognitive Linguistics-oriented teaching activities in real L2 classrooms: Design, implementation and assessment*.
- Pütz, M., Niemeier, S., & Dirven, R. (Eds.). (2001). *Applied cognitive linguistics: Language pedagogy* (Vol. 2). Walter de Gruyter.
- Reif-Hülser, M. (2012). *Making progress simpler? Applying cognitive grammar to tense-aspect teaching*. Peter Lang.
- Robinson, P., & Ellis, N. (Eds.). (2008). *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Routledge/Taylor & Francis.

- Slobin, D. I. (1996). From "thought and language" to "thinking for speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge University Press. (Reprinted in modified form from "Pragmatics," 1, 1991, pp. 7–26)
- Suñer, F., & Roche, J. (2019). Embodiment in concept-based L2 grammar teaching: The case of German light verb constructions. *IRAL (International Review of Applied Linguistics in Language Teaching)*, 59(3,27), 421-447. <https://doi.org/10.1515/iral-2018-0362>
- Suñer, F., & Roche, J. (2021). Embodiment in concept-based L2 grammar teaching: The case of German light verb constructions. *International Review of Applied Linguistics in Language Teaching*, 59(3), 421–447. <https://doi.org/10.1515/iral-2018-0362>
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. *Language typology and syntactic description*, 3(99), 36-149.
- Talmy, L. (2000). *Toward a cognitive semantics, Vol. II: Typology and process in concept structuring*. Cambridge, MA: MIT Press.
- Taylor, J. R. (1993). Some pedagogical implications of cognitive linguistics. *Conceptualizations and mental processing in language*, 3, 201-226.
- Tomasello, M. (2005). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.
- Tyler, A. (2008). *Cognitive linguistics and second language instruction* (pp. 466-498). Routledge.
- Tyler, A. (2008). Cognitive linguistics and second language instruction. In P. J. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 456–488). Routledge.
- Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. New York: Routledge.
- Tyler, A. (2015). Cognitive Linguistics and SLA. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (1. publ. in paperback, pp. 88–90). Routledge.
- Tyler, A. Huang, L. and Jan, H. (Eds.). (2018). *What is applied cognitive linguistics: Answers from current SLA research* (Vol. 38). Walter de Gruyter GmbH and Co KG.
- Tyler, A., Mueller, C., & Ho, V. (2011). Applying Cognitive Linguistics to Learning the Semantics of English to, for and at: An Experimental Investigation. *Vigo International Journal of Applied Linguistics*, (8), 180-205.
- Tyler, A., Mueller, C., & Ho, V. (2011). Applying Cognitive Linguistics to Learning the Semantics of English to, for and at: An Experimental Investigation. *Vigo International Journal of Applied Linguistics*, 8, 180–205.
- Van Valin Jr, R. D. (1990). Semantic parameters of split intransitivity. *Language*, 66(2), 221-260.
- Van Valin Jr, R. D. (2005). *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- VanPatten, B. (2015). Foundations of processing instruction. *International Review of Applied Linguistics in Language Teaching*, 53(2), 91-109.

- VanPatten, B. (Ed.). (2004). *Processing instruction: Theory, research, and commentary*. Mahwah, NJ: Lawrence Erlbaum
- Vázquez-Rozas, M. V. (2006). Gustar-type verbs. In J. Clancy Clements & J. Yoon (Eds.), *Functional Approaches to Spanish Syntax* (pp. 80-114). Palgrave Macmillan, London.
- Vázquez-Rozas, M. V. (2012). Construyendo emociones: sintaxis, frecuencia y función comunicativa. In T. Jiménez, B. López Meirama, M. V. Vázquez Rozas, & A. Veiga (Eds.), *Cum corde et in nova grammatica. Estudios ofrecidos a Guillermo Rojo*. Santiago de Compostela: USC Editor, pp. 841-854.
- Zaenen, A (1993). Unaccusativity in Dutch: Integrating syntax and lexical semantics. In J. Pustejovsky (Ed.), *Semantics and the lexicon* (pp. 129-161). Dordrecht: Kluwer.

APPENDICES

All appendices have been submitted for review along with the manuscript and appear anonymously in the open-access website 'Open Science Framework' in the following link: https://osf.io/ufm64/?view_only=264780abda9340dd8ab7967b4d0eaa02