Gradient competence
at the Syntax-Discourse Interface

Roumyana Slabakova,* Jason Rothman
and Paula Kempchinsky*
*University of Iowa / University of Florida

In this article, we present additional support of Duffield's (2003, 2005) distinction between Underlying Competence and Surface Competence. Duffield argues that a more fine-grained distinction between levels of competence and performance is warranted and necessary. While underlying competence is categorical, surface competence is more probabilistic and gradient, being sensitive to lexical and constructional contingencies, including the contextual appropriateness of a given construction. We examine a subset of results from a study comparing native and learner competence of properties at the syntax-discourse interface. Specifically, we look at the acceptability of Clitic Right Dislocation in native and L2 Spanish, in discourse-appropriate context. We argue that Duffield's distinction is a possible explanation of our results.

1. Introduction: Gradient competence

A basic tenet of generative grammar since its beginnings has always been the distinction between the underlying system of linguistic knowledge – competence – and the use of this system – performance. While competence has been predicated of an ideal speaker-listener in a completely homogeneous speech-communication who knows his/her (the speech community’s) language perfectly, performance has been defined as applying this knowledge of language in actual communication situations. To the extent that competence is a mental state of linguistic representation, it is fixed, whereas performance is arguably affected by grammatically irrelevant conditions such as memory limitations, processing complexities, distractions, shifts of attention and interest (Chomsky, 1965: 3).

However, Duffield (2003, 2005) argues that a more fine-grained distinction between levels of competence and performance is warranted and necessary. As it relates to the macro-category of competence, Duffield suggests a division between
what he calls “underlying competence” and “surface competence.” Although both types of competence are implicit, underlying competence is *categorical* while surface competence is more *probabilistic* because it encodes lexical and constructional contingencies, including the contextual appropriateness of a given construction. Below are Duffield’s definitions of surface and underlying competence:

“By hypothesis, underlying competence (UC) is categorical, and consists of formal (phonological and syntactic) principles, autonomous from the lexicon. It is plausible to think of UC as innate. Surface competence (SC), by contrast, is intimately determined by the interaction of contextual and specific lexical properties with the formal principles delivered by UC: as a consequence, SC generates gradient effects. SC is largely language-specific learned knowledge.”

(Duffield, 2003: 101)

Dividing competence into two types as Duffield suggests is not a mere terminological recapitulation of “performance”, as it might seem at first glance. The notion of surface competence differs from performance in that it is *always* constrained faithfully by underlying forms; however, such forms themselves are sensitive to frequency, context and other variables giving rise to gradient judgments. Even and especially for native speakers (NSs), categorical judgments of properties that are infrequent and/or highly constrained by information structure are difficult precisely because the underlying representations of such properties themselves are sensitive to probabilistic contingencies. Thus, surface competence is closer to the performance of actual utterances than underlying competence, because the notions of gradience and variability are central to it, as they also are for linguistic performance. This dual competence model endeavors to explain (among other linguistic findings) the cases in L2 acquisition where near-native L2 speakers “outperform” native speakers with respect to some expected set of acceptability judgments.

Sorace and Keller (2005) make a very similar distinction: that between “soft” and “hard” linguistic constraints. They define the two types of constraints in the following way:

“If a constraint violation induces strong unacceptability and fails to show context effects and developmental optionality, then it can be classified as a hard constraint. If a constraint triggers only mild unacceptability and is subject to contextual variation and developmental optionality, then the constraint is soft.” The authors further speculate that “hard constraints are purely structural (i.e., syntactic) in nature, while soft constraints are at the interface between syntax and other domains (i.e., semantics or pragmatics”).

(Sorace & Keller, 2005: 1521)

Duffield’s (2003, 2005) proposal together with Sorace and Keller’s (2005) proposed distinction offers an interesting and theoretically important way of explaining
interlanguage findings. In this article, we present in detail a further example of findings that support this distinction in competence, from a study comparing native and L2 learner competence of properties at the syntax-discourse interface. Specifically, we present data from NSs of Spanish and English-native L2 learners of Spanish. The grammatical domains under investigation in the larger study, from which the data we focus on here is derived, include the discourse-sensitive domains of dislocates, namely clitic left and clitic right dislocations (CLLD and CLRD) as well as focus fronting. In light of the discussion we endeavor to provoke, our focus will primarily be on the results of near-native speaker (NNS) group as compared to the NS controls for CLRD, where the L2 acquirers seemingly outperform the NSs. We will argue that Duffield's dual competence model can explain our findings for CLRD.

2. Duffield's (2003, 2005) competent gradience and the competence paradox

Duffield starts his argument by pointing out that linguistic competence includes not only categorical knowledge of grammaticality (intuitions like “this sentence is ungrammatical while this sentence is perfectly fine”), but also knowledge of gradient grammaticality. He gives examples cited in Kluender (1992), but originally provided by Chung and McCloskey (1983) (where the symbol ‘>’ indicates ‘more acceptable than’):

(1) a. This is the paper that we really need to find someone who understands >
   b. This is the paper that we really need to find a linguist who understands >
   c. This is the paper that we really need to find the linguist who understands >
   d. This is the paper that we really need to find his advisor who understands >
   e. This is the paper that we really need to find John, who understands.

These and similar examples illustrate the phenomenon of gradient acceptability of sentences, the opposite of categorical grammaticality (e.g., canonical word order violations, lack or use of overt subjects in contexts where presence or absence are obligatory as in expletive contexts).¹ Such knowledge of gradient acceptability can be demonstrated by native speakers and second language learners alike when the test design allows the test-takers to express gradience; that is, when magnitude estimation or a numerical scale of acceptability is used. When a putative NS shows gradient judgments for a given set of sentences but a putative L2 learner
(e.g., a NNS) has categorical judgments of the same set, then by hypothesis the competence of the L2 learner diverges from the competence of the NS. To be more concrete, a learner who categorically accepts a construction that is judged as less acceptable by native speakers could be considered as “more competent than the native speaker” (Duffield, 2003: 100). By the measure of L2 competence converging on NS competence, that same learner could also be evaluated as less competent than the NS. Both approaches would be somehow unsatisfactory, however, since they simplify over a theoretically interesting distinction. This type of divergence is one of the possible scenarios proposed by Duffield for the relationship between native and learner competence.

In a series of graphs, Duffield represents the possible logical states of convergence and divergence of native and learner competence. In Figure 1, convergence is illustrated: both native and non-native speakers demonstrate sensitivity to underlying and surface competence.

An example of a study documenting such full convergence could be Martohardjono’s (1993) already classic experiment, which investigated knowledge of the contrast between “weak” island violations such as those in (2) and “strong” island violations such as those in (3).

(2) a. ?? Which book did John hear a rumor that you had read?
   b. ?? Which book did John wonder who had read?

(3) a. * Which book did John meet a child who read?
   b. * Which soup did the man leave the table after the waiter spilled?

Whatever the theoretical accounts of these violations, the data speak for themselves: NSs reject the strong violations with 99% accuracy while they reject the weak violations only 78% of the time. Keeping in mind that these sentences are ungrammatical and presumably do not appear in the input, learners of English whose native language was Chinese rejected the violations 75% and 44%, respectively.
Indonesian-native learners of English rejected them 88% and 46% of the time, respectively, both learner groups clearly demonstrating gradient competence. However, it is also logically possible that the judgments of NSs and non-NSs reflect diverging competencies, as illustrated in Figures 2 and 3. Duffield argues that this situation might hold even when NSs and non-NSs share the same judgments, because these shared judgments could be a reflex of either only underlying competence or only surface competence.

What would be examples of findings illustrating these two types of NS/non-NS divergences? Examples cited by Duffield (2003) show that it is possible for the judgments of two participant groups (L1 vs. L2) to reflect accessing of different competences, that is, surface or underlying competence. An example of a research finding conforming to the situation in Figure 2 comes from Duffield, Sabourin and Curtin (1998), a study of L2 lexical access and priming effects. NS results from Marslen-Wilson et al. (1994) documented that only words that could be decomposed into a shared stem plus a legitimate affix are considered to be related, as estimated by the effects of priming. However, morphological relatedness was a necessary, but not a sufficient condition for lexical relatedness; in addition, the
two forms had to be semantically related. Thus, for the NSs, govern was shown to prime government (and vice versa), but depart failed to prime department, since the latter pair do not share any meaning. The advanced learners in Duffield, Sabourin and Curtin (1998) replicated the NS behavior. Their intermediate learner group, however, showed priming effects for morphologically related pairs irrespective of semantic relatedness. These learners “better approximated” the categorical behavior predicted by the standard theoretical models of lexical access, as compared to the NSs.

Type 2 disjoint convergence is exemplified by the behavior of NSs and learners on knowledge of pronominal clitic placement in French (Duffield & White, 1999; Duffield et al., 2002). In this test, the French NSs were tested with an off-line grammaticality judgment task and an on-line sentence-matching task, the latter presumed to be a more unobtrusive measure of grammaticality. The examples under discussion involve infinitives and object clitics in French:

(4) a. je veux le voir.  
   I want 3sg to.see  
   ‘I want to see him.’

b. *je le veux voir.

c. je le fais chanter.  
   I 3sg make to.sing  
   ‘I make him sing.’

d. *je fais le chanter.

Equivalent examples to (4b), in which the clitic pronoun thematically associated with the infinitive appears before the conjugated verb, are possible in other Romance languages such as Italian and Spanish. Various linguistic analyses propose that in these languages, verbs such as Italian volere ‘to want’ (so-called restructuring verbs) appear in a monoclausal structure. As is well-known, modern French lacks “restructuring”, given the ungrammaticality of (4b). Conversely, in the case of causative structures (4c, d), the structure is proposed to be biclausal (and in all the Romance languages, the clitic associated with the causee must appear with the causative verb and not the infinitive).

Duffield cogently points out that these syntactic analyses actually contradict naïve inspection of the French facts: the clitic le ‘him’ is placed between the putative restructuring verb and the infinitive (cf. 4a, b) but before the causative verb (cf. 4c, d); hence it would be expected in the latter case that there is “restructuring”, i.e., a monoclausal structure. Duffield argues that the syntactic analyses of these two structures lead to different predictions of grammaticality: while (4b) is expected to be ungrammatical at all levels of representation, (4d) is expected to be grammatical only at the surface level, because the clitic is attached to the verb.
with which it is thematically associated. In the Duffield et al. (2002) experiment, a sentence matching task uncovered that the native speakers treated “ungrammatical causatives” differently from “ungrammatical restructuring” sentences. In this respect, they were sensitive to the contrast predicted by the linguistic literature, judging (4b) as ungrammatical but (4d) as grammatical. On the other hand, both English and Spanish-speaking learners of French showed a grammaticality effect for both restructuring and causative contexts in the sentence-matching task, in contrast to the French native speakers. Duffield argues that the standard model of treating a unitary competence separate from performance cannot provide a satisfactory account of these findings.

Theoretical explanations of the second type of parallel divergence, illustrated in Figure 3 and exemplified in the experiment described above, would be positions such as the Fundamental Difference Hypothesis (Bley-Vroman, 1989, 1990), the Failed Functional Features Hypothesis (Hawkins & Chan, 1997) and the Interpretability Hypothesis (Hawkins & Hattori, 2006; Tsimpili & Dimitrakopoulou, 2007). In this acquisition situation, learners are imitating, faking or learning in the true sense of the word target-like L2 behavior, having necessarily arrived at it primarily through observation of surface patterns.

The second type of divergence has been the subject of much debate in L2 research. However, it is the first type of parallel divergence which is the focus of discussion in Duffield (2003, 2005) and in this article. This is the situation in which learners may seem to outperform the native speakers. “What this model is intended to account for are cases in which the L2 learner has attained the target grammar, but where his/her judgment patterns nevertheless diverge systematically from those of the native-speaker” (Duffield, 2003: 103).

Gradient acceptability is even more important at the syntax-discourse interface where acceptability depends not only on purely syntactic factors but also on contextual variables such as “topic” and “focus”. In the next section, we discuss one such set of data.

3. **Clitic left dislocation, clitic right dislocation and focus fronting in Spanish**

In this experimental study, we looked at the acquisition of dislocation structures in L2 Spanish by L1 English speakers. These structures are acceptable in certain discourse contexts and unacceptable in others. Furthermore, there are semantic constraints on the felicitous use of these constructions that may make them even harder to acquire. The syntactic analysis assumed as the basis of the study is that of López (2009), which we briefly present below.
López (2009) offers a new theoretical approach to information structure, proposing a *pragmatic* module of computation which yields an intermediate level of representation between the syntactic structure and the discourse structure. At certain points in the derivation (for López, at the end of each phase, following Chomsky, 2001) certain constituents in the syntactic structure are assigned pragmatic features by this pragmatic module, yielding “the information structure of a syntactic object Σ”, represented as Σ[p] (López, 2009: 22). These pragmatic structures are then assembled into longer units of text in the discourse module, in the form of Discourse Representation Structures. The specific positions where the pragmatic features are assigned are positions at the edge of the two phases, vP and CP.

According to López, “topic” and “focus” are not primitives of the grammar, as many previous models have assumed, but are at most descriptive terms. The crucial information structure notions are discourse anaphor and contrast, represented in terms of the primitive features [±a(naphor)] and [±c(ontраст)], assigned to elements that have undergone dislocation. Elements which move to the edge of the vP and agree with a doubling clitic are assigned the feature [+a]; elements which move to the left periphery are assigned the feature [+c]. Non-dislocated elements receive no pragmatic features:

| Table 1. Features bundles and constructions in López’s (2009) model |
|-------------------|-------------------|
|                   | +contrast          | −contrast         |
| +anaphor          | Clitic Left Dislocation (CLLD) | Clitic Right Dislocation (CLRD) |
| −anaphor          | Fronted Focus (FF)   | Rheme             |

We shall illustrate these features with Spanish examples, where all capitals stand for focus intonation. The feature [+anaphoric] reflects the fact that the phrase moved from its underlying word order position, known as the dislocate, is obligatorily related to an antecedent in the discourse. Clitic Right Dislocation (CLRD) and Clitic Left Dislocation (CLLD) are [+anaphoric] constructions. However, the conditions on the anaphoric relationship differ (Villalba, 2000). In CLRD, the dislocate must be semantically equivalent to the antecedent, as shown by the contrast in (5) and (6) below.

(5) [Context: Where did you buy that dress?]
Lo compré en El Corte Inglés, *ese vestido*.
‘I bought it at El Corte Inglés, *that dress’

(6) [Context: Where did you buy those clothes?]
#Lo compré en El Corte Inglés, *el vestido*.
‘I bought it at El Corte Inglés, *the dress’
In the case of CLLD, the dislocate may be semantically identical to the discourse antecedent; however, unlike CLRD, the dislocate can also be in a subset, superset, or part/whole relationship with the antecedent, as in (7) and (8) below:

(7) [Context: What did the movers do with the furniture?]
Las sillas las dejaron en el pasillo, pero no sé dónde están las mesas.
‘The chairs they left them in the hallway, but I don’t know what they did with the tables’ [subset-set]

(8) [Context: What can we do with this table? It doesn’t fit through the doorway]
Mira, las patas las podemos doblar.
‘Look, the legs we can fold them’ [part-whole]

With respect to [±contrast], Fronted Focus (FF) and CLLD are positively marked. The [+c] feature encodes the idea that the dislocated element identifies a missing element (in semantic terms, a variable) within a presupposition. In (9) below, the fronted [+c] element is la alfombra ‘the rug’ which has undergone focus fronting, while in (10) the fronted [+c] element is the CLLD dislocate las sillas ‘the chairs’. In both cases, the preceding discourse established the presupposition “John bought x”. In (9), the identity of x does not need a discourse antecedent; Focus Fronting yields the combination [–a, +c]. In (10) the identity of x must have a discourse antecedent (CLLD yields [+a, +c]), where the relationship between the antecedent and the CLLD dislocate can be one of those enumerated above. It is in this sense that Arregi (2003) refers to CLLD as “contrastive topicalization”:

(9) [Context: John bought the furniture]
LA ALFOMBRA compró (no los muebles) FF
THE RUG he bought (not the furniture)

(10) Las sillas sí que las compró, pero las mesas … CLLD
The chairs he did indeed buy, but the tables …

Finally, the Rheme construction, which is used as a foil in our experiment, is neutral in the sense that it only introduces new information in answer to a preceding question which opens up a variable position to be satisfied.

(11) ¿Qué compró Juan?
What did John buy?
Juan compró muebles.
John bought furniture.
4. Learning tasks for native speakers of English

Of the four constructions illustrated in the previous section, English allows FF with prosodic emphasis on the dislocate, the so-called focus intonation (see for example the English translation of (9)), and simple Rheme with SVO word orders (as in the English translation of (11) above).

However, as we will see below, Spanish presents syntactic options for rheme constructions which are not available in English, namely VOS and VSO word orders. Since English does not have clitics, CLLD and CLRD are by definition new constructions to the English-native learners of Spanish.5 Syntactic knowledge of clitics was a condition for inclusion of participants in our experimental groups. In addition to syntactic knowledge of fronting operations, the discourse appropriateness of the clitic-doubled dislocations is the next property that the learners need to acquire, and that is a property at the syntax-discourse interface. That includes knowledge that a dislocate in an anaphoric relationship with another previously mentioned constituent should be clitic-doubled. Finally, the semantic constraints on the antecedent-dislocate relationship have to be learned, contrasting CLLD and CLRD with respect to the freedom of association between antecedent and dislocate.

Note also that CLLD and CLRD are constructions with differing frequency and distribution. As noted by López (2009) CLRD is much less frequent in Spanish than, for example, Catalan as well as much less frequent in Spanish as compared to CLLD. He attributes this difference to the availability of p-movement (Zubizarreta, 1998) in Spanish, which provides a syntactic mechanism for placing the direct object in VP-final position without dislocation.6 Catalan lacks p-movement, and hence CLRD provides the mechanism for giving heavier stress to the direct object in a non-contrastive context. As we discuss below, these differences will become very important in interpreting the experimental results.

Our predictions with respect to the acquisition of CLRD reflected the anecdotal claims of its being rare and dialectally varied. However, our native speaker results constitute the first experimental evidence of NS judgments on CLRD; thus we could not incorporate them in our predictions. Since we knew CLRD was rare in the input, we predicted lower acceptance rates for learners, as compared to CLLD, due to infrequent evidence for it in the input.
5. Experimental study

5.1 Participants and test materials

Eighty-eight individuals took an online battery of tests, posted on WebSurveyor, an online survey service. Twenty-one were native speakers of Spanish from different Spanish-speaking countries (1 Argentina, 4 Chile, 4 Colombia, 1 Costa Rica, 2 Cuba, 3 Spain, 5 Mexico, 1 Venezuela) and 67 were adult learners of Spanish. An anonymous questionnaire with personal questions ascertained that the learners had started acquiring Spanish after puberty. They reported Spanish as their second language. The large majority were classroom learners. They also reported education levels higher than high school (at least some college). The learners’ names were not known to the investigators, to protect their identity.

The test materials included a proficiency test based on the DELE (official exam of proficiency administered by the Spanish Ministry of Education) that is standardly used by generative L2 Spanish research (e.g., White, Valenzuela, Kozlowska-MacGregor & Leung, 2004; Montrul, 2004; Rothman & Iverson, 2008). The test contains 50 items (distributed between a close test format and multiple choice). The latter was used to divide the 67 learners into proficiency levels: 21 individuals fell in the near-native range (score range 47–50), 22 were classified as advanced learners (score range 40–47), and 23 as intermediate learners of Spanish (score range 30–39). In the battery of tests, we also included a test on the syntax of clitics. This test was necessary to make sure that learners had knowledge of the correct placement and obligatory presence of clitics in some Spanish sentences. There is no point in testing learners’ knowledge of the discourse and semantic constraints of clitics if they lack the prerequisite syntactic knowledge. The test included ten forced multiple-choice items. Each sentence had five options to choose from, for a total of fifty points. Points were given for correctly choosing the grammatical options and for correctly rejecting the ungrammatical choices. The cutoff point for inclusion in the rest of the study was set at 35, or 70% accuracy. All advanced and near-native participants scored over 45, or were 90% accurate, with a single exception, an advanced-proficiency individual who scored 40, or 80%. All intermediate learners scored above 35 points. Below are two examples from the clitic test.

(12) Dulce: ¿Te gustó algo de la tienda?
   Did you like something from the store?
   Julia: Me encanta esa chaqueta, ______ pronto.
      CL love this jacket, ______ soon
      ‘I love this jacket, ______ (I-want to buy it) soon’
The main experimental task was the Felicity Judgment Task. It included 40 context-test sentences combinations: 10 CLLD and 10 CLRD constructions, 5 with and 5 without identity between antecedent and dislocate. The test also included 5 Fronted Focus constructions and 5 Rheme constructions where the object was new information. Ten fillers were also included. A sample test item is presented in (14) below.

Each test item had the following structure: after a brief context, a short dialogue was presented which contained two alternatives, one with and one without a clitic. The 5 Rheme constructions had the object as new information, and the two alternatives to evaluate for acceptability were VSO vs. VOS word orders. In Spanish, SVO, VSO and VOS are all grammatical orders; however, the discourse contexts in which each of these is appropriate differ. In particular, the research on Spanish clausal syntax converges on the generalization that VOS is appropriate only when the subject represents new information. In contrast, when the object represents new information, both SVO and VSO are possible, although SVO is preferred (we will return to this in the discussion). We hypothesized that if the SVO word order were included, this would make the task of the learners too easy. Since SVO is their native default word order, they would not have to access any new knowledge in order to evaluate these sentences. In particular, we would not be able to determine whether rejection of the pragmatically infelicitous SOV reflected problems at the syntax-discourse interface or lack of syntactic knowledge, and therefore we would not have been able to tell whether the native language or the L2 is the source of SVO acceptance.

The context stories and test sentences were presented both aurally and visually to all participants. The aural presentation of test items was crucial for distinguishing CLLD constructions, which lack emphatic stress on the fronted constituent, and FF constructions, where the fronted constituent must be stressed.
and a (doubling) clitic is ungrammatical. It was also important for the testing of CRLD items, which require an intonational break between the end of the clause and the right dislocated element. Participants were asked to judge sentences as felicitous or infelicitous in the context of the story on a scale of 1 to 4, but they could also choose the answer “I don't know.” A negligible number of answers displayed this option. A screenshot of a test item presentation appears in Appendix A.

Below we give examples from the CLRD construction. The context and the question mention furniture, which is the discourse antecedent. Sofia’s answer contains reference to a table, which acts as a subset of all possible furniture. Thus answer A with a clitic and a right-dislocated constituent was expected to be inappropriate (marked with a hachure #) because the relationship between the antecedent and the dislocate is not one of equivalence. On the other hand, answer B without a clitic is simply ungrammatical, and marked with a star (*).

(14) Context: Juan y Mónica quieren comprar un sofá nuevo pero no están seguros si la tienda de muebles que está cerca de su casa es buena y entonces le preguntan a Sofía.

Juan and Monica want to buy a new sofa but they are not sure whether the furniture in their neighborhood store is good, and so they ask Sofia:

Juan: ¿Crees que los muebles aquí son buenos?

Do you think that the furniture here is good?

A. Sofía: # Claro que sí, lo compré ahí, mi sofá.

Of course, it I-bought there, my sofa

B. Sofía: * Claro que sí, compré ahí, mi sofá.

Of course, I-bought there, my sofa

On the other hand, in Example (15), where there is equivalence between the antecedent and the dislocate (both are ‘the homework’), answer A is acceptable and answer B is still ungrammatical due to the lack of clitic-doubling.

(15) Context: Denise, la hija de Juan, le dice a su padre que quiere salir a jugar. Eso sí, por regla no le permiten salir hasta que termine su tarea. Por lo tanto, Juan le dice a su hija:

Denise, Juan’s daughter, asks her father permission to go out and play. However, she is only allowed to play after she finishes all her homework. So, Juan asks his daughter:

Juan: Depende ¿terminaste tu tarea?

It depends, have you finished your homework?

A. Denise: Sí. Ya la hice, la tarea. ¿Puedo salir a jugar?

Yes. Already it I-did, the homework. Can I go out and play?

B. Denise: * Sí. Ya hice, la tarea. ¿Puedo salir a jugar?

Yes. Already I-did, the homework. Can I go out and play?
5.2 Results

In this article, we will focus on the acceptability judgments of the CLRD construction, and briefly summarize the CLLD results for comparison purposes only. Figure 4 presents the mean acceptability on the CLRD construction, on a scale from 1 to 4. The first two columns present the acceptability judgments on the test items where the dislocate and the antecedent were semantically equivalent; the next two columns present acceptability judgments on the test items where the dislocate and the antecedent are in subset-set relationship. The first thing to notice in the graph is that the native speakers don’t really accept the CLRD construction that is expected to be rare but still acceptable, in accordance with López (2009). On the other hand, the near-natives and the advanced learners rate the acceptable CLRD with a clitic to be more acceptable than the non-clitic variant, i.e., the variant that should be ungrammatical.

Table 2 presents the intra-group comparisons based on the post-hoc tests of a repeated measures ANOVA with group, semantic equivalence, and presence of clitic as factors. When the semantic relationship between the antecedent and the dislocate is one of equivalence, hence the construction is purportedly acceptable, the native speakers do not demonstrate a significant contrast in acceptability, while the near-native and advanced groups do. When the semantic relationship between dislocate and antecedent is one of subset to set, the native speakers do show a contrast, but in the opposite to the expected direction. They like the variant without a clitic better, probably reanalyzing the construction as p-movement or ignoring altogether the pause before the dislocated object. It is the near-native group that behaves according to the expectations in the literature: they accept the

![Figure 4. Mean acceptability judgments on the CLRD construction](image-url)
clitic-doubled dislocations and reject the non-clitic-doubled ones. Furthermore, they consider the unacceptable and the ungrammatical CLRD to be equally bad. The advanced learners also show a significant contrast between the clitic-doubled and the non-clitic-doubled test items. Their judgments are in the expected direction, unlike the native speakers. However, they still prefer CLRD with a clitic to one without a clitic in the subset-set condition, demonstrating that the semantic constraints on CLRD play no role in their grammar. Finally, intermediate learners have acquired neither the discourse requirements nor the semantic constraints of CLRD, and thus have neither the discourse-based nor the semantic contrasts in their grammar.

In order to ascertain that group results are not hiding large individual variation, we looked at individual results. We calculated them in the following way: If a participant had a mean felicity scores on unacceptable and acceptable test sentences differing by a whole point or more, in the right direction, we considered them aware of the expected distinction. As the reader can see in Table 3, even some individuals in the native speaker group make the cut. For CLRD in the subset-set condition, having a contrast is not expected, and the fact that we have individual learners making a distinction that the natives as a group do not make, is interesting. A good percentage of the near-native and advanced learners, and even some intermediate learners, also exhibit this sensitivity in their individual grammars.

Table 2. Intra-group comparisons between mean ratings of sentences with clitics and w/o clitics on the CLRD construction, significance set at the 0.05 level

<table>
<thead>
<tr>
<th></th>
<th>Equivalence ✓</th>
<th>Subset-set #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natives (n = 21)</td>
<td>p = .96 ns</td>
<td>p = .006*</td>
</tr>
<tr>
<td>Near-natives (n = 22)</td>
<td>p &lt; 0.0001</td>
<td>p = .062 ns</td>
</tr>
<tr>
<td>Advanced (n = 22)</td>
<td>p &lt; 0.0001</td>
<td>p = .0001</td>
</tr>
<tr>
<td>Intermediate (n = 23)</td>
<td>p = .28 ns</td>
<td>p = .1 ns</td>
</tr>
</tbody>
</table>

* These means are opposite to the expected direction.

Table 3. Number of subjects with a numeric difference of 1 between sentences with clitics and sentences without clitics on the CLRD construction, in the right direction (%)

<table>
<thead>
<tr>
<th></th>
<th>Equivalence ✓</th>
<th>Subset-set #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natives (n = 21)</td>
<td>6 (20%)</td>
<td>0</td>
</tr>
<tr>
<td>Near-natives (n = 22)</td>
<td>15 (68%)</td>
<td>8 (36%)</td>
</tr>
<tr>
<td>Advanced (n = 22)</td>
<td>16 (72%)</td>
<td>10 (45%)</td>
</tr>
<tr>
<td>Intermediate (n = 23)</td>
<td>7 (30%)</td>
<td>7 (30%)</td>
</tr>
</tbody>
</table>
This behavioral pattern contrasts very much with learners and natives’ acceptance of CLLD. As mentioned above, CLLD is a frequent construction in spoken Spanish of all varieties. When asked to judge the acceptability of left dislocations, native, near-native and advanced participants were equally accurate in contrasting the clitic-doubled from the clitic-less variants. In other words, natives and up to advanced learners demonstrated a reliable contrast with the CLLD construction. Intermediate learners display no contrast, just as with the CLRD construction, see Table 4.

Finally, we present the group results on FF and Rheme in Figure 5 and Table 5. All groups are aware of the fact that when you front a constituent with a focus intonation, no clitic is needed, and in fact, the presence of a clitic is ungrammatical.

<table>
<thead>
<tr>
<th></th>
<th>Equivalence √</th>
<th>Subset-set #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natives (n = 21)</td>
<td>p &lt; 0.0001</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>Near-natives (n = 22)</td>
<td>p &lt; 0.0001</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>Advanced (n = 22)</td>
<td>p = .04</td>
<td>p = .02</td>
</tr>
<tr>
<td>Intermediate (n = 23)</td>
<td>p = .61 ns</td>
<td>p = .9 ns</td>
</tr>
</tbody>
</table>

Figure 5. Mean acceptability judgments on the FF and Rheme constructions
6. Discussion

Let us start by summarizing the results relevant to the present discussion. This experimental study looks at the acceptability of four constructions at the syntax-discourse interface. In the two dislocations with anaphoric relations to the preceding context, CLLD and CLRD, clitic-doubling is obligatory. In the FF dislocation marked by intonation, clitic-doubling is ungrammatical. This is the only construction where transfer from English is possible since English uses the same device for contrastive focus marking. Finally, in the unmarked (object) Rheme construction, we tested two word orders, one pragmatically acceptable and one pragmatically unacceptable in Spanish; neither one is syntactically grammatical in English (see also discussion below). In addition, we tested semantic constraints on the antecedent-dislocate relationship within the two dislocations.

The data uncovered highly differentiated behavior on the part of the native speakers. The CLLD construction was accepted with high reliability. On the other hand, the CLRD construction was found mildly unacceptable by the native speakers. Ratings of native speakers (roughly, 2.5) were exactly in the middle of the range of possible answers (1 to 4). We hypothesize that the natives were not sure how to process this construction, due to a conflict between their pragmatic preferences and the syntactic and phonological properties of the test items. As observed earlier, CLRD is infrequent in Spanish, due in part to the availability of p-movement. For López (2009), both CLRD and p-movement have the effect of moving a constituent to Spec,vP, where it can be marked [+a] by the pragmatics module. However, in PF a direct object in Spec,vP which agrees with a clitic – that is, a CLRD direct object – will be linearized after all material in the verb phrase, whereas constituents which have moved to Spec,vP by p-movement without an agreeing clitic, which includes direct objects, will be linearized preceding other material in the verb phrase, by Kayne’s (1994) Linear Correspondence Axiom. Suppose now that the verb phrase contains a direct object as well as another complement XP. Then the order XP > DO can result either from p-movement of

<table>
<thead>
<tr>
<th>Fronted Focus (Clitic vs. No Clitic)</th>
<th>Rheme (VSO vs. VOS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natives (n = 21)</td>
<td>$p &lt; 0.0001$</td>
</tr>
<tr>
<td>Near-natives (n = 22)</td>
<td>$p = 0.00059$</td>
</tr>
<tr>
<td>Advanced (n = 22)</td>
<td>$p = .05$</td>
</tr>
<tr>
<td>Intermediate (n = 23)</td>
<td>$p = .003$</td>
</tr>
</tbody>
</table>

Table 5. Intra-group comparisons between mean ratings of grammatical and ungrammatical sentences in the FF and Rheme constructions, significance set at the 0.05 level.
the XP to Spec,vP (and consequent marking of the XP as [+a]), or from CLRD movement of the direct object to Spec,vP, in which case there must be an agreeing clitic present and there must be a pause boundary between the XP and the direct object, which is the [+a] marked element.

Note that in these two cases the discourse status of the direct object is quite different. In the case of CLRD, the dislocated direct object must have a discourse antecedent. In the case of p-movement of another constituent in the verb phrase, the direct object will be in the complement domain of v, which by default (see López, 2009) is marked [−a]. This means only that the direct object need not have a discourse antecedent, not that it cannot have one. In effect, therefore, in a discourse context where the direct object has a discourse antecedent, it may be in its original position, or it may be the dislocate in a CLRD construction. López speculates that given the availability of p-movement in Spanish, the use of CLRD tends to be limited to the one case where p-movement does not achieve the desired result: correction of the focused part (the non-dislocated part) of the verb phrase. That is, the discursively ideal CLRD sentence would be one such as the following:

(16) Context: Ana left her purse in the entryway.
    No, lo dejó en la mesa de la cocina, su bolso.
    ‘No, she left it on the kitchen table, her purse’

This perspective on the role of CLRD in Spanish can shed some light on our results. In our sample there were five semantically felicitous examples of CLRD, in terms of the identity condition on the antecedent-anaphor relation for CLRD. However, in four of these examples, there is no overt material between the verb and the (dislocated) direct object, and in the fifth the only other overt material in the VP is an adverbial phrase (muy bien ‘very well’) which even without dislocation would intervene linearly between the verb and the direct object due to V raising to T. Given the lack of overt material to receive corrective focus, and the fact that the discourse context of the items did not provide that context, the low ratings given by our native speakers to these items may reflect their rejection of the discourse appropriateness of CLRD in these contexts.12

On the other hand, the near-native and advanced groups of learners demonstrate reliable knowledge of the syntax-discourse interface phenomena, namely, that topic dislocations have to be marked by a clitic while focus dislocations must not be clitic-doubled. These two groups show significant contrasts between the acceptable and unacceptable variants. Around 70% of individuals in these groups behave as expected in the linguistics literature. If we assume that behavioral tests are indicative of linguistic knowledge, where can this knowledge have come from?
This is the question that interests us in this article. Why are the NNSs accepting CLRD? Our speculative answer would be that the learners have not acquired p-movement, or have not acquired it to a sufficient extent, so that they would have ‘specialized’ CLRD in the way that López suggests for Spanish. In this acquisition situation, we argue that the NNSs are accessing a grammatical structure because they have no alternative in their grammar, while the NSs are confused in choosing between two legitimate alternatives, because unfortunately our test does give them the opportunity to tease apart the two possible analyses (CLLD and p-movement).\textsuperscript{13}

Of course, there is always the possibility that López’s description of CLRD in Spanish is simply incorrect. However, we maintain that it is premature to make such a claim for both theoretical and acquisition reasons. From a theoretical standpoint, consider the fact that the discourse features López proposes and their combinations meet descriptive and explanatory adequacy for three out of four constructions (see Table 1). This is supported by the fact that all groups (with the exception of the intermediate group) demonstrated a contrast between acceptable and unacceptable test sentences as predicted by this feature bundle theory. In view of this, it seems prudent to proceed with caution in not abandoning what his theory would predict for the 4th property in Table 1.

Alternatively, one immediate appeal of rejecting López’s claims for CLRD would be that doing so would explain the native speaker performance without further conjecture or stipulations. Nevertheless, such a claim would also effectively leave the NNSs and advanced L2 learners’ performances unexplained. Let us keep in mind as well, that the NNSs not only showed knowledge of the syntax-pragmatics of CLRD itself as described by López, but also were sensitive to the semantic imports that distinguish CLRD from CLLD in the exact way that López describes. Coincidence? We think not. Unless one is prepared to accept that the data of the L2 learners in this study converge almost perfectly on López’s theoretical description by sheer happenstance, which we are not prepared to do, the alternative we highlighted in the last paragraph seems even more reasonably plausible.

Beyond its plausibility, however, the admittedly speculative explanation we laid out is definitely testable in future research and as such should not be considered as vacuously ad hoc. As we mentioned above, we were surprised by the fact that CLRD was not processed correctly by the native speakers, although we were aware of its low frequency and dialectal variation. If we are on the right track in our explanation of the acquisition pattern, and NNSs did not show the same resistance to CLRD because they have not acquired p-movement in part or at all, then this should be testable by creating a test in which p-movement and CLRD
can be teased apart. If indeed, NNSs show lack of or incomplete knowledge of p-movement but still accept CLRD in its appropriate contexts, we would then have empirical support for the possibility we laid out above. Equally, in subsequent experiments we could resolve the ambiguity in our CLRD exemplars that we ourselves highlighted might not permit (or at least encourage) the NSs to tease apart two legitimate competing analyses (i.e., CLRD and p-movement). Such improvements could yield better results from the NS group and uncover a principled grammatical reason for why the NNSs performed so well on the present tasks.

At this point, let us come back to the definitions of underlying and surface competence from Duffield (2003, 2005). Although both types of competence are considered implicit, underlying competence is categorical while surface competence is more probabilistic because it encodes lexical and constructional frequencies, among other factors. It appears that Duffield’s concepts of underlying and surface competence go a long way towards explaining the findings of our study. That is, the behavioral patterns of our native speakers and learners can be captured by Duffield’s parallel divergence type 1, illustrated in Figure 2. The NSs show sensitivity to the low frequency of the CLRD, while the learners demonstrate categorical knowledge of the acceptability of the right dislocation. However, we do not think that the NSs are unaware of the underlying acceptability of CLRD. Rather, we submit that their awareness of alternative ways of capturing the same message in a discourse-sensitive way (p-movement) interferes with their categorical underlying competence and results in wider surface competence. Recall that Catalan has only CLRD while Spanish allows CLRD and p-movement in competition. In a sense, the learners develop a ‘Catalan grammar’ so to say, with respect to these properties. In such learning situations, other test instruments are warranted that are more sensitive to both types of knowledge and test both constructions. These are left for further research, expanding and differentiating the test items of the present test.

Another finding supporting this proposal is the NS patterns of acceptance on the Rheme construction. As we noted in Section 5.1, VOS is infelicitous in this context. In theory, both SVO and VSO should be equally possible, given that the object is in sentence-final position to receive nuclear stress (Zubizarreta, 1998), and we chose to not include SVO for the reasons stated in that section. However, it has been noted in the literature (see for example Zubizarreta, 1998; Ordóñez, 2000; Sheehan, 2010) that in cases where the entire sentence represents new information, many speakers find VSO sentences pragmatically odd without a fronted locative or temporal adverbial. In our test sentences, only the object represents new information, but this pragmatic preference could explain the low acceptability of VSO in our native speaker group. Even though they do show a reliable
contrast between VSO and VOS word orders, the native speakers are much more “grudging” in their acceptance of VSO, giving it an average score of 2.32 out of 4, while all the learners give it scores higher by a whole digit. It must be kept in mind that the English native speakers are distinguishing in acceptability between word orders that are both ungrammatical in their native language. In choosing VSO as acceptable, they demonstrate successful acquisition, and in a much more categorical way than the native speakers. In fact, the learners treat the transferable FF construction and the non-transferable Rheme construction in the same way. Thus, we find advanced learners’ and NNSs’ acceptance of CLRD and VSO to be parallel.

Coming back to the more general question, could it be the case that our experimental data do not tell us anything about the competence of the non-native speakers, since the native speakers’ behavior does not confirm that the CLRD is an acceptable construction? After all, it is standard practice in the generative second language acquisition literature that the property under investigation has to be supported with native data first, before we look at its L2 acquisition. We would like to propose that this practice is overly restrictive and may obscure interesting findings about learners’ competence. Rather, we should be asking the question: How did learners come to possess this categorical knowledge of CLRD, if the construction is so rare in the input? The answer that we tentatively submit to this question is that advanced learners have generally acquired the knowledge that in Spanish dislocations, when there is a discourse antecedent, the dislocate must be doubled by a clitic. Furthermore, even if infrequent, CLRD must have sufficient presence in spoken language in order to support its successful acquisition.

7. Conclusions

In this article, we discuss a subset of experimental data which seem to point to the conclusion that native speakers do not readily accept a specific construction, CLRD, while near-native and advanced speakers behave as expected and distinguish between clitic-doubled and non-clitic-doubled dislocated objects. We argued, using the concepts of surface competence and underlying competence, that in fact the native speakers are experiencing competition from another construction in Spanish, p-movement. In this sense, they have additional acceptable ways in their grammar of ensuring that an object gets nuclear stress and they are more sensitive to the gradient acceptability of CLRD, therefore demonstrating gradient, or surface competence. On the other hand, advanced and near-native learners of Spanish demonstrate the categorical underlying competence, which requires a dislocated topicalized object to be clitic-doubled, whether it is moved to the left or
the right edge. Our findings support Duffield's (2003) distinction between surface and underlying competence. Without this distinction, our data, and in particular, why learners are accurate on the CLRD construction while the natives are not, would have remained unexplained.

Notes

1. A reviewer argues that frequency of constructions and probabilistic constraints may be constantly at play in establishing the acceptability of a sentence, not only at surface competence but at underlying competence, too. We submit that this is certainly a plausible point of view (see e.g., Yang, 2002). However, we believe that there are certain core syntactic properties that would be close to categorically ungrammatical, such as expletive omission in English or expletive suppliance in Spanish weather sentences. Probabilistic constraints certainly come to play in the acquisition of such constructions, but once established, they are not variable.

2. There is much debate on the precise syntactic analysis of sentences like the Italian equivalent to (4b); “restructuring” has been used as a label to refer to this phenomenon since Rizzi (1982), even in accounts which do not presuppose an actual reconstruction process. The lack of this construction in French has been linked to the non-licensing of null subjects in this language (see for example Kayne, 1989; Roberts, 1997).

3. Take for example, Rizzi’s (1997) analysis of the left periphery, where Topic and Focus are assumed to be functional heads.

4. Examples (5) to (11) are Spanish translations/adaptations of corresponding Catalan examples in López (2009).

5. English of course has topicalization without a resumptive pronoun, and Romance CLLD is often seen as discourse-equivalent to topicalization. However, this is a simplification of a complex phenomenon. Minimally, we can observe that CLLD is distinct from topicalization both in its syntactic distribution (see Haegemann, 2006) and in the discourse properties of the left dislocate (see Bianchi & Frascarelli, 2010, among others).

6. López rejects prosodic considerations as the motivation for the movement, but retains the name ‘p-movement’.

7. We aimed at testing NSs from a variety of countries to get away from the possible dialectal differences in judging clitic constructions. Since we did not control for country of origin and the groups are too small, it is impossible to ascertain now whether there are significant dialectal differences between the NSs from the different countries. We chose to test speakers of different dialects because they represent who L2 learners in a classroom setting have as teachers, over the course of the several years they are learning Spanish. If CLRD is severely underrepresented in some dialect and almost common in others, there is no guarantee that learners would be exposed to one or the other dialect with any consistency. The fact that the near-natives and advanced speakers converge on target-like knowledge of CLRD attests to their having overcome the input of teachers from various Spanish dialects.
8. As pointed out by one anonymous reviewer, including the CLLD data with the same level of treatment could be beneficial. We, of course, do not disagree with such a suggestion a priori. However, given space limitations, but more importantly, the scope of our argumentation and the particular relevance of the subset of our data we focus on, namely the CLRD results, we will not be able to provide an equally detailed treatment of CLLD. We believe that the general details we provide of the CLLD results are sufficient to make the comparison. To include the CLLD and then not the FF and Rheme data would seem less than thorough, but adding all of the conditions would take up valuable space that, we believe, would also lead the present article astray from its goals. That said, we provide sufficient details of all the conditions to help orient the reader and also as a way of making the CLRD data stand out that much more.

9. If the native speakers were really hearing the pause, one would expect judgments of ungrammaticality, because *Ya hice#la tarea is just impossible.

10. Unfortunately, we cannot back observations of construction frequency with data from spoken language corpora because we are not aware of such counts. However, informal polls of NSs, both of Latin American and Peninsular Spanish, support our observations. Still, there were six individuals among the NSs who did observe the expected contrast, which suggests that it is not a figment of linguists’ imagination.

11. By López’s account, this linearization results from phonological constraints on the alignment of intonational phrases. In general, the nature of p-movement is controversial in the literature on Spanish linguistics. For example, as noted above, López (2009) does not assume p-movement to be prosodically motivated. In Zubizarreta’s (1998) account, the object moves to the right edge of the sentence when it is the constituent which must be final in order to receive nuclear stress; for example, when the direct object represents new information. For López, on the other hand, for whom p-movement is movement of some constituent to the left of the object, this movement will not necessarily happen in the same contexts predicted by Zubizarreta, because for him what triggers the movement is some unspecified feature which may or may not appear in the numeration. What is definitely the case is that virtually all speakers of Spanish would accept as perfectly grammatical an order like V – PP – Direct Object (unless the direct object is very light, like a bare noun), e.g., Puse en la cocina los platos ‘I put in the kitchen the plates’. This construction yields an order similar to the result of English Heavy NP Shift, but unlike English, the object doesn’t have to be very heavy. In sum, it is very likely that Spanish speakers do in fact have something in their grammar like p-movement as presented by López (2009): movement of constituents to (outer) Spec,vP, even though the precise motivations for this movement are still not clear.

12. One reviewer suggests that the similar low acceptance of VSO and CRLD on the part of the native controls shows a general tendency toward dispreference of right-displaced objects. However, it is important to note that syntactic analyses of Spanish VSO sentences, while they disagree on the syntactic position of the subject, do converge on the general assumption that the object remains in the vP-VP field.

13. Another, and also speculative, explanation of the behavior of the NSs would be that they (at least some of them) were not paying attention to the pause, and therefore in that case the test sentences would sound just out-and-out ungrammatical to them. However, how do we reconcile the fact that the NSs would not be hearing the pause with the fact that the NNS seem to be hearing it? NNSs are all very successful L2 learners, and they can be much more careful
and self-monitoring when participating in formal studies. They also feel that they have much more to prove, being in the 'hot seat of judgment' whereas NSs need not prove anything. Thus we do not find it completely implausible that (some) natives did not register the pause before the dislocated object while most of the near-natives did. We were aware of the problems with the pause in previous research (Valenzuela, 2005) and tried to make it prominent in the recordings. Unfortunately, we have no means of distinguishing a posteriori whether the pause in our test items was sufficiently prominent to notice.

14. As a reviewer suggests, one could think of the learners adopting a conservative strategy: once CLRD is acquired, they don’t need p-movement to capture the same meaning.

References


Hawkins, R., & Hattori, H. 2006. Interpretation of English multiple wh-questions by Japanese
301.
Kluender, R. 1992. Deriving Island Constraints from Principles of Predication. In Island Con-
Marslen-Wilson, W., Tyler, L. K., Waksler, R. & Older, L. 1994. Morphology and meaning in the
Martohardjono, G. 1993. Wh-movement in the Acquisition of a Second Language: A Cross-lingu-
istic Study of Three Languages with and without Movement. Unpublished doctoral dissert-
ation, Cornell University.
Rizzi, L. 1997. The fine structure of the left periphery. In Elements of Grammar, L. Haegeman
L2 knowledge of aspectual phrasal semantics. Language Acquisition: A Journal of Develop-
Sheehan, M. 2010. ‘Free’ inversion in Romance and the null subject parameter. In Paramet-
ric Variation: Null Subjects in Minimalist Theory, T. Biberauer, A. Holmberg, I. Roberts &
Valenzuela, E. 2005. L2 ultimate attainment and the syntax-discourse interface: the acquisition
of topic constructions in non-native Spanish and English. Ph.D. dissertation. McGill Uni-
versity.
de Barcelona.
Press.
Appendix A

Screen shot of a Felicity Judgment Test presentation

1) Juan y Mónica invitaron a María a comer. La cena se sirvió en la terraza y todo estaba muy rico. María felicitó a Juan por la sopa que había hecho. Cuando Mónica escucha esto, responde:

Mónica: La carne la preparó Juan, no la sopa. (Recuerde presionar el botón “play”)

☐ 4 (perfecta) ☐ 3 ☐ 2 ☐ 1 (muy extraña) ☐ No sé

2) Mónica: La carne preparó Juan, no la sopa. (Recuerde presionar el botón “play”)

☐ 4 (perfecta) ☐ 3 ☐ 2 ☐ 1 (muy extraña) ☐ No sé