

Article

Feature Matching Does Not Equal Convergence: Acquisition of L2 French Accusative Pronouns by L1 Spanish Speakers

Elena Shimanskaya * and Tania Leal *

Department of World Languages and Literatures, University of Nevada, Reno, NV 89557, USA

* Correspondence: eshimanskaya@unr.edu (E.S.); taniel@unr.edu (T.L.)

Abstract: Our study aims to determine whether formal similarity between two languages (operationalized via the Feature Reassembly Hypothesis) allows adult L2 learners of French (Spanish native speakers; NSs) to straightforwardly acquire third-person singular accusative clitics in their L2. Additionally, we examined the role of surface similarity, since French and Spanish overlap and diverge in several ways. In terms of formal similarity, third-person accusative clitic pronouns in Spanish are almost perfect analogues of their French counterparts. In terms of surface similarity, however, while the feminine accusative pronouns are identical (“la” [la]), the masculine ones differ in Spanish (“lo” [lo]) and French (“le” [lə]). Participants included French NSs ($n = 26$) and Spanish-speaking L2 French learners ($n = 36$). Results from an offline forced-choice picture selection task and an online self-paced reading task did not support the Feature Reassembly Hypothesis because learners showed considerable difficulty with the interpretation and processing of these pronouns, revealing that, unlike French NSs, their interpretations and processing are guided by the feature [\pm Human] and, to a lesser degree, by gender, which might be due to the surface-level similarity between feminine accusative clitic pronouns in both languages.



Citation: Shimanskaya, Elena, and Tania Leal. 2021. Feature Matching Does Not Equal Convergence: Acquisition of L2 French Accusative Pronouns by L1 Spanish Speakers. *Languages* 6: 144. <https://doi.org/10.3390/languages6030144>

Academic Editor: Aarnes Gudmestad

Received: 8 June 2021

Accepted: 17 August 2021

Published: 26 August 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: L2 acquisition; clitic pronouns; feature reassembly; French; Spanish; accusative

1. Introduction

It has long been discussed in the literature on bilingualism/multilingualism that first (L1) and second (L2) languages are interconnected in the mind of a speaker. To explain such relationships, several frameworks have made different proposals. From a usage-based perspective, abundant evidence shows that both languages in a bilingual’s mind remain available and activated when processing both written and spoken language. This parallel activation of the lexicon in both languages has been shown to operate when surface-level similarity exists between isolated words (Dijkstra et al. 1998; Marian and Spivey 2003) and between sentence structures (Baten et al. 2011; Duyck et al. 2007).

In contrast to usage-based perspectives, generative approaches have traditionally focused on understanding speakers’ underlying knowledge of syntactic and morpho-syntactic properties. Numerous studies conducted within this framework investigate whether knowledge of L1 grammar affects how the L2 is represented and used (the question of ‘L1 transfer’) and whether bilinguals are able to derive the same meanings and arrive at the same linguistic judgments in their L2 as native speakers (NSs).

While usage-based research demonstrates the important role surface-level similarity plays in language acquisition and processing, generative research has provided crucial evidence of the importance of abstract-level phenomena in L2 acquisition and processing. Therefore, one important open question is the role that both abstract-level similarity (e.g., similarity in structural or syntactic function) and surface-level similarity (e.g., similarity in how two forms look or sound) play in the acquisition and use of a second or subsequent (Ln) language. In this study, we investigate how abstract- and surface-level similarity between a speaker’s L1 and L2 affect L2 comprehension and incremental L2/Ln language processing.

We do so by investigating L2 acquisition of French pronominal clitics by NS of Spanish. To the best of our knowledge, no previous study has investigated situations wherein both formal- and surface-level similarities can be established between grammatical forms in two languages. One could, thus, argue that a situation like the one we are considering here constitutes a situation where both usage-based and generative approaches can provide valuable insights into the minds of bilingual speakers.

The rest of the paper is organized as follows: the next section details theoretical background, describes the linguistic property under investigation, and briefly discusses some previous research relevant for the current study. The section ends with the presentation of the research questions. Sections three and four present the two experiments and their results followed by a discussion section and a brief conclusion.

2. Background

2.1. Theoretical Background

The phenomenon of parallel activation has received ample empirical support in usage-based research. For example, in an experiment conducted by [Schwartz and Kroll \(2006\)](#), a group of Spanish-English bilinguals living in Spain read English sentences that contained words such as “pan,” which are inter-lingual homographs: words that look the same but have different meanings in English and Spanish (Spanish *pan* = bread). The bilinguals, who resided in Spain and used Spanish regularly, made more errors naming words like “pan” in English sentences as opposed to English words that do not exist in Spanish.

Within the generative SLA framework, a particularly useful conceptualization of L1 influence—especially as it concerns formal similarity—is offered by the *Feature Reassembly Hypothesis* (FRH; [Lardiere 2009](#)), which has operationalized L1 transfer as the process of mapping abstract feature bundles (realized in the L1 as words or morphemes) to forms in the L2. Lardiere adopts a Minimalist view of language architecture ([Chomsky 1995](#)) to propose that L2 learners transfer abstract units of meaning—linguistic features—expressed by L1 words and morphemes. Importantly, this view assumes that all lexical items (words and morphemes) are physical realizations of features, whether functional (grammatical) or semantic. These features (which include number and, crucially for our investigation, gender) are essential for mapping sound-meaning pairings. However, while the atomic units of meaning or linguistic features are assumed to be universal, the departing notion is that languages vary in how they organize features into words and morphemes, such that learners will have to reassemble features if the L1 and L2 differ in their feature encoding. Additionally, features not present in the L1 but present in the L2 will need to be added from the Universal Grammar feature inventory.

As we will see in more detail in the next section, Spanish and French accusative clitic pronouns should not require reassembly because they encode the same feature sets. The FRH does not predict difficulties for forms with different surface realizations if the corresponding abstract features can be mapped perfectly from the L1 to the L2 since “learners will look for morpholexical correspondences in the L2 to those in their L1, presumably on the basis of semantic meaning or grammatical function (the phonetic matrices will obviously differ)” ([Lardiere 2009](#), p. 191). The situation regarding surface-level similarity is more complex because Spanish has a clitic form that looks exactly like the French accusative masculine clitic “le,” although it is not its functional analogue. Spanish “le” [le] can refer to masculine and feminine antecedents and is used as a dative (not accusative) pronoun. Therefore, if the parallel activation phenomenon observed with lexical words applies to functional words, surface-level form similarity might cause Spanish-French bilinguals some trouble mapping “le” in their L2-French.

2.2. Linguistic Property: Gender and [\pm Human] Encoding in Accusative Clitic Pronouns in French and Spanish

We examine how Spanish-French bilinguals process French third-person singular accusative clitic pronouns. These forms, both in French and Spanish, have been of interest to linguists in different L1–L2 language combinations from a variety of theoretical perspectives. However, investigations focusing on Spanish speakers learning French remain relatively scarce—a situation we attempt to improve with the present study. Importantly, personal pronouns are a good testing ground for the Feature Reassembly Hypothesis because they constitute phonetic realizations of features such as person, number, case, and, in this particular language combination, gender.

Let us have a closer look at the French and Spanish paradigms for accusative and dative singular pronouns. On one hand, with respect to the similarity in terms of abstract features, the Spanish paradigm aligns perfectly with the French one. Both languages have pronominal clitics in their inventory (Roberge 1990; Zagona 2002). Compare this to languages like English, where all pronouns are strong. Additionally, in Spanish and French, clitics occupy different positions from a full Determiner Phrase (DP): While full DPs occur after the verb, clitics are positioned pre-verbally (with finite verbs).¹ Examples (1) and (2), (3) and (4) show this contrast for French and Spanish, respectively.

(1)	Je		verrai		le parc/Pierre.	French
	I		will-see		the park/Pierre	Full DP
	'I will see the park/Pierre.'					
(2)	Je		le		verrai.	French
	I		3.sg.acc.masc		will-see	Accusative Clitic
	'I will see it/him.'					
(3)	Yo	veré	el parque	/a	Pedro.	Spanish
	I	will-see	the park	DOM ²	Pedro	Full DP
	'I will see the park/Pedro.'					
(4)	Yo		lo		veré.	Spanish
	I		3.sg.acc.masc		will-see	Accusative Clitic
	'I will see it/him.'					

Other similarities include that, dative and accusative clitics in both languages cannot be separated from the verb (except by other clitics, in which case both languages have specific ordering criteria) (van Riemsdijk 1999). Additionally, clitics cannot stand alone and, therefore, cannot be the answer to a *wh*-question (Cardinaletti and Starke 1999). Similarly, neither Spanish nor French clitics can be either stressed or conjoined (Cardinaletti and Starke 1999).

Similarities between French and Spanish extend to the encoding of gender and the [\pm Human] feature because both French and Spanish accusative clitics encode grammatical gender (masculine/feminine) but not the feature [\pm Human]. Unlike English, where the pronominal paradigm lexically encodes biological gender of human referents (*she* vs. *he*, *her* vs. *him*), both French and Spanish lexically encode grammatical gender, meaning that clitics match the gender of both human and, importantly, non-human antecedents. This difference between French/Spanish and English can be observed in examples (2) and (4), where the English pronoun *him* marks [+human] (while the pronoun *it* is reserved for [−human] referents). Thus, both French (2) and Spanish (4) could be translated into English as “I will see it [masculine]” if the referential expression associated with the pronoun is [−human] instead.³ Regarding the syntactic analysis of clitics, we assume, following Cardinaletti and Starke (1999), that clitics in French and Spanish are deprived of the projection that realizes referential, human, and case features (unlike the case for strong pronouns). This analysis assumes that this syntactic deficiency explains their inability to stand alone and to receive stress. Additionally, we follow Jakubowicz et al. (1998, p. 120) in their assumption that

these clitics are underspecified for the [±Human] feature, which allows them to be merged in the functional domain of the verb (unlike full DPs.)

Regarding gender features, French and Spanish nouns are posited to bear grammatical gender features that can be specified or unspecified. Masculine forms are *unspecified* because they are the default and are assumed to lack lexical marks (e.g., Harris 1991; Lumsden 1992). In contrast, feminine nouns are (lexically) *specified* as feminine since these are not the default forms. An interesting correlate of this premise is that there are asymmetries in terms of feature clashes: Masculine determiners or adjectives do not clash when they co-occur with feminine nouns because [masculine] bears default, unspecified agreement (see Licerias et al. 2008; White et al. 2004). On the contrary, feminine determiners and adjectives clash with masculine nouns because the feminine is specified for gender.

Finally, French and Spanish are also comparable in terms of the resolution of pronouns when these co-refer with referential expressions (R-expressions). This process requires for a dependency between the pronoun and the referent to be established. In the case of these languages, this dependency relies on gender matching, such that φ -features between the referent in the discourse and the features on the pronouns themselves must match. Once these features are co-indexed, relevant uninterpretable features can be checked and then deleted in the computation. In short, French and Spanish accusative clitic pronouns form a perfect analogue in terms of abstract similarity.

In terms of surface-level similarity, French and Spanish pronoun paradigms share commonalities, to different degrees, in writing and pronunciation. The feminine accusative clitic has the same spelling and pronunciation in both French and Spanish: “la” [la]. The masculine clitics differ, however: “lo” [lo] in Spanish and “le” [lə] in French. The parallels do not end there, either, because the French *accusative* masculine clitic has the same written form as the Spanish *dative* singular clitic—“le”—although it is pronounced differently in the two languages: [lə] in French and [le] in Spanish. Finally, the French dative clitic has a unique spelling and pronunciation that shares no similarity with Spanish forms: lui [lɥi]. Table 1 illustrates the extent to which the paradigms in French and Spanish parallel each other in abstract-level features and feature bundles expressed by individual forms. It also shows the extent to which these surface forms are similar or different in the two languages in writing and pronunciation.

Table 1. Third-person singular accusative and dative object clitics in French and Spanish.⁴

Case	Gender	±Human	French	Spanish
Accusative	Masculine	[+human] [−human]	“le” [lə]	“lo” [lo]
	Feminine	[+human] [−human]	“la” [la]	
Dative	Masculine	[+human] [−human]	“lui” [lɥi]	“le” [le]
	Feminine	[+human] [−human]		

As an interim summary: in terms of abstract functional features (the learning task), there is nothing to reassemble for L2 learners of French who are Spanish NSs when it comes to accusative clitics; the paradigms in both languages align perfectly. Thus, based on the predictions of the Feature Reassembly Hypothesis, we would expect that acquisition should proceed straightforwardly. In terms of surface-level similarity, however, the learning task differs: The fact that some forms align perfectly (e.g., both feminine accusative clitics are the same in writing and pronunciation), while others do not (“le” in French and Spanish, as shown in Table 1), might interfere with how these accusative clitics are judged or processed in L2 French.

2.3. Previous Studies: The Acquisition of Gender and French Accusative Clitics

Overall, there is consensus in the L2 literature that, relative to subject clitics, the acquisition of French object clitics lags behind in terms of overall clitic production (e.g., Schlyter 2003) but also in terms of accuracy regarding object clitics' distributional properties (e.g., placement) and gender agreement (Prévost 2009). Hawkins (2001) has proposed that the acquisition of the distributional properties of object clitics proceeds in stages, with learners first placing them in postverbal position, then omitting them, placing them in the intermediate position, before reaching target-like placement. Placement seems to be unproblematic for Spanish-French bilinguals, however. Bruhn de Gavarito and Montrul (1996) have shown that Spanish-speaking learners of French have no issues learning placement except in cases where Spanish (but not French) allows clitic climbing, i.e., Spanish-French bilinguals failed to consistently reject ungrammatical sentences like **Je le veux voir*, ("I want to see him").

While French and Spanish clitic systems are very similar, there is evidence that even in situations with two closely related languages wherein L2 learners can successfully transfer some L1 processing strategies to the L2, acquiring gender might not be effortless. For instance, Sabourin and Haverkort (2003) studied the acquisition of gender agreement with German NSs who were L2 learners of Dutch. The authors used a Grammaticality Judgment Task (GJT) and a second GJT paired with ERP versions of the experiments. They found that learners were successful with definite Noun Phrases (NPs), which are very similar in both languages, but performed poorly with indefinite NPs, which feature important differences between the two languages.

Overall, previous research, mainly from generative and variationist perspectives, presents a complex picture when it comes to the L2 acquisition of gender. While some studies show that L2 learners can be quite accurate in their judgments and use of gender-marked determiners and adjectives (e.g., Bruhn de Garavito and White 2002; Edmonds and Gudmestad 2018; Gabriele et al. 2013; Gudmestad et al. 2019; Montrul et al. 2008), other studies have documented extensive difficulties when realizing gender agreement (e.g., Franceschina 2005; Grüter et al. 2012). Additionally, it appears that speakers whose L1 has grammatical gender have an easier time than those whose L1 does not have grammatical gender (e.g., Foucart and Frenck-Mestre 2011; Sabourin and Stowe 2008). This would mean that, in theory, our learners should have an easier time with the acquisition of gender distinctions in L2 French because Spanish, like French, encodes grammatical gender.

An ancillary consideration is that a disconnect may exist between how grammatical gender is processed/computed and learners' lexical knowledge of gender. For instance, Hopp (2013) investigated the acquisition of gender with L2 learners of German and found that only those learners who had stable (lexical) knowledge of gender could use gender as a predictive cue. In our investigation, we tried to disentangle the lexical knowledge of gender from how it is processed or computed by always providing the correct gender marking in the determiner (e.g., *sa* "his/her"/*la* "the" or *son* "his/her"/*le* "the"). In other words, we were interested in studying how gender was processed and interpreted rather than whether learners knew the lexical gender of a given NP.

A final factor to consider is how the feature [\pm Human] may affect L2 performance. As with the acquisition of gender, however, the research on this semantic feature has produced divergent results. Some researchers have proposed that the processing of animate nouns is more "cognitively demanding" (Sagarra and Herschensohn 2010, p. 106). Other researchers have suggested that gender cues can facilitate the processing of both [+human] and [−animate] pronouns (Irmen and Knoll 1999). In the case of our learners, we would expect positive transfer of this semantic feature because it works in equivalent ways in both languages. The question of the acquisition of gender for Spanish-French bilinguals, however, is an experimental one. In the following section, we describe the two experiments that we conducted to answer the following research questions:

1. RQ1-Does the acquisition of the features associated with French accusative clitics proceed straightforwardly when there is nothing to reassemble (e.g., when the formal similarity coincides), as predicted by the Feature Reassembly Hypothesis?, and
2. RQ2-What role, if any, does surface-level similarity play?

3. Experiment 1—Picture Selection Task

3.1. Participants

We tested two groups of participants: NSs of French who provided a baseline of behavior and who resided and were tested in Paris ($n = 26$); and NSs of Spanish who were L2/L3 learners of French ($n = 36$), tested in Mexico (Puebla). Table 2 includes the demographic information. All participants reported learning languages other than their native ones. NSs of French all reported speaking English ($n = 26$) at various levels, while half reported having learned Spanish ($n = 13$). A few reported learning German ($n = 5$), Italian ($n = 3$), or other languages (Russian, Portuguese, or French Sign Language ($n = 3$)).

Within the L2 learner group (Spanish NSs), all learners reported learning French and English ($n = 34$).⁵ A few reported learning German ($n = 9$), while others reported learning Portuguese ($n = 2$) and Croatian ($n = 1$). The L2 group also completed a proficiency test that consisted of a C-test containing 50 blanks (50 points total) (Renaud 2010). The test contained two short authentic texts (74 and 97 words); the first clause of each was left intact. Starting at the second clause, the second half of every other word was deleted to target different categories (e.g., nouns, verbs, or prepositions). In terms of scoring, each correct answer received one point. Answers with minor spelling errors or unidiomatic phrasing received half a point. The instructions encouraged participants not to take more than five minutes per text to fill in the blanks, but they were not timed. Average scores are included in Table 2.

Table 2. Demographic information for both participant groups.

Demographic Information	French NSs	Spanish NSs (French L2 Learners)
<i>n</i>	26	36
Mean age (SD)	20.67 (2.24)	21.35 (6.04)
Female (Male)	25 (1)	24 (10)
Proficiency, max 50 (SD)	n/a	34.77 (8.14)

3.2. Materials and Procedure

3.2.1. Procedure

Our forced-choice picture selection task tested participants' interpretations of clitic pronouns that differed in terms of the encoding of the feature [\pm Human] and gender. The task was delivered via Qualtrics and completed in a quiet lab under supervision. This test, including the proficiency task and the background questionnaire, took L2ers an average of 41 min to complete.

3.2.2. Materials

The picture selection task followed a 2×2 design, with gender (masculine/feminine) and [\pm Human] (human/nonhuman) as factors. The task included a total of 46 items: 12 experimental and 34 fillers. Items were distributed into two lists so that each participant only saw one version of each item (masculine/feminine), judging a total of 46 items. Participants were introduced to a scenario: They were to help Nicolas, a schoolboy, understand who or what his friends are talking about because he always arrives late. For each item, participants read a short discourse context where three potential referents were introduced (either two masculine referents and one feminine or vice versa). The test sentence following the context contained a clitic pronoun and was followed by four choices/pictures, only one of which was felicitous. Option 1 correctly matched the gender of the referent, while Option 2 referred to the opposite gender. Option 3 allowed participants to choose both, while

Option 4 represented a distractor that was thematically compatible with the context but had gone unmentioned in the discourse. These options were randomized per participant. Figure 1 shows a sample item. Contexts and test sentences were presented in French. Spanish translations were not provided.

Every Wednesday, Nicolas and his friends play in the park. Today, Luc, Claire, and Anne are playing hide-and-seek. Nicolas is late. When Nicolas arrives at the park, he sees Anne and a policeman, but he sees neither Luc nor Claire. When Nicolas comes up to Anne, Anne says to the policeman:

Je la cherche partout.
 I 3.sg.acc.fem look.for everywhere
 "I am looking for her/it everywhere."

Donc, Anne cherche...
 "So, Anne is looking for..."

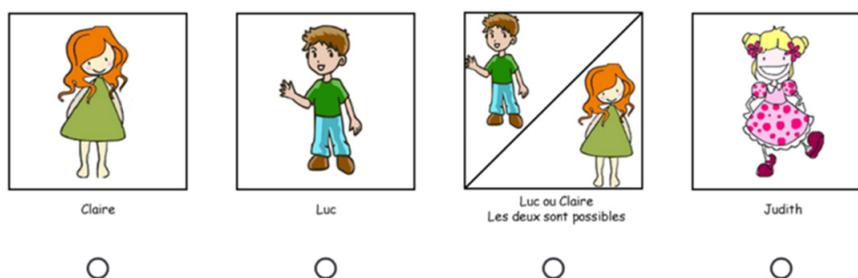


Figure 1. Sample item, picture selection task.

3.3. Results

Our task investigated the interpretations participants assigned to third-person accusative French clitic pronouns. We designed the task to determine the effects of the [\pm Human] feature plus grammatical gender to determine whether lack of reassembly would lead to straightforward acquisition. Participants were presented with four options:

- Correctly gender-matched referent;
- Incorrect: gender-mismatched referent;
- Incorrect: both genders; and
- Incorrect: distractor (gender-matched).

Table 3 presents the proportion of responses (counts in parentheses), arranged by answer type for the L1-French NSs.

Table 3. Picture selection task results by gender and the [\pm Human] feature, L1 group. (# = number).

Feminine	Human (#)	Nonhuman (#)	Total
Correct	97.5% (79)	95.1% (77)	96.3% (156)
Incorrect: Gender Mismatch	-	-	-
Incorrect: Both Genders	-	4.9% (4)	2.5% (4)
Incorrect: Distractor	2.5% (2)	-	1.2% (2)
Masculine	Human (#)	Nonhuman (#)	Total
Correct	98.8% (80)	88.9% (72)	93.8% (152)
Incorrect: Gender Mismatch	-	-	-
Incorrect: Both Genders	-	4.9% (4)	2.5% (4)
Incorrect: Distractor	1.2% (1)	6.2% (5)	3.7% (6)

Overall, Table 3 shows that the NSs behaved as predicted in the literature. Note that the error type “Gender Mismatch” did not have a single incidence: gender-matching was particularly salient for NSs. In the [+Human] conditions, the only type of error was “distractor,” which did not make sense in terms of the story but matched the gender of the referent. Below, Table 4 shows the proportion of responses for the L2 group, whose accuracy scores were much lower, their error types more varied. Recall that this task was not performed under time pressure—even if learners were unfamiliar with the gender of a nonhuman NP, they could look at the story and retrieve the gender marked in the determiner.

Table 4. Picture selection task results by gender and the [\pm Human] feature, L2 group.

Feminine	Human (#)		Nonhuman (#)		Total	
Correct	76.6%	(82)	63%	(68)	69.8%	(150)
Incorrect: Gender Mismatch	4.7%	(5)	3.7%	(4)	4.2%	(9)
Incorrect: Both Genders	8.4%	(9)	26.9%	(29)	17.7%	(38)
Incorrect: Distractor	10.3%	(11)	6.5%	(7)	8.3%	(18)
Masculine	Human (#)		Nonhuman (#)		Total	
Correct	57.9%	(62)	60.7%	(65)	59.3%	(127)
Incorrect: Gender Mismatch	6.5%	(7)	13.1%	(14)	9.8%	(21)
Incorrect: Both Genders	20.6%	(22)	19.6%	(21)	20.1%	(43)
Incorrect: Distractor	15.0%	(16)	6.5%	(7)	10.8%	(23)

We analyzed our data through a series of pairwise comparisons because our dependent variable was categorical (selection responses). To set up the comparisons, we collapsed results over binary categories using logistic mixed-effects models via the *glmer* function of the *lme4* package (Bates et al. 2015) in the R environment (R Core Team 2020). Dependencies were captured in maximal random effect structures (RES). We report on separate models per group (L1 vs. L2). This decision follows Dekydtspotter et al. (2006), who argue that observed isomorphy between groups should not be taken to mean that L1 and L2 acquisition are fundamentally different. In doing so, we also avoid the comparative fallacy (Bley-Vroman 2009). Thus, we focus on finding out whether learners can make significant distinctions *within* their own grammar, rather than whether they exactly mirror their L1-French counterparts. Models included the following RES: random intercepts for participants and items, as well as random slopes. When the model did not converge with the maximal RES, we ran it again with the next maximal RES until convergence.

When we analyzed the natives’ response behavior statistically, we collapsed data over Correct/v/Incorrect (all error types collapsed). Results from a logistic mixed-effects model revealed no main effects or interactions, suggesting that the natives’ responses, which were overwhelmingly accurate, did not show a relationship with either the [\pm Human] feature or gender. The maximal converging RES (Table 5) included random effects by participant and by item.

Table 5. Linear mixed-effects model, picture selection task, L1 group.

Fixed Effects	Estimate	SE	z-Value	Pr(> z)
[\pm Human]	0.8591	0.957	0.898	0.369
Gender	−1.620	1.479	−1.095	0.273
Gender*[Human]	2.562	1.637	1.565	0.118
Random Effects			Variance	SE
By-Participant Intercept			4.616	2.149
By-Item Intercept			2.675	1.636

For the model for L2 learners (Table 6), we also collapsed results over binary categories. Results from a logistic mixed-effect model revealed main effects for [\pm Human] and gender ($p = 0.0103$ and $p = 0.001$, respectively), as well as a gender* $[\pm$ Human] interaction ($p = 0.0304$). The maximal RES that converged (Table 6) included random effects by participant and by item. Significant results are marked with an asterisk (*).

Table 6. Linear mixed-effects model, picture selection task, L2 group.

Fixed Effects	Estimate	SE	z-Value	Pr(> z)
[\pm Human]	0.897	0.350	2.567	0.010 *
Gender	1.191	0.351	3.393	0.001 *
Gender* $[\pm$ Human]	-1.032	0.477	-2.165	0.030 *
Random Effects			Variance	SE
By-Participant Intercept			1.866	1.366
By-Item Intercept			0.226	0.475

We followed up with pairwise comparisons (paired-samples t -tests, *emmeans* in R), which revealed a significant contrast between human and nonhuman referents when the gender was feminine ($t = -2.567$, $p = 0.0103$), whereby the learners were more accurate with [+human] referents (Tukey HSD). Comparisons also unveiled a significant distinction between masculine and feminine ($t = -3.393$, $p = 0.0007$) when selecting among [+human] referents (learners were more accurate with feminine). No other contrasts were significant.

Finally, we examined the data by proficiency, treated as a continuous variable (Leal 2018). We ran a generalized linear mixed-effect regression with (centered) proficiency as a fixed-effect predictor and the maximal RES that would converge on the data set (by-subject and by-item slopes, plus their intercepts), collapsed over Correct/ v /Incorrect responses. To better visualize the relationship, we plotted proficiency against accuracy in Figure 2. The model yielded a main effect of proficiency ($\beta = -0.137$, $SE = 0.0228$, $z = -5.98$, $p = 0.000$), suggesting that accuracy on the task increased with higher proficiency.

The results of the picture selection task revealed several important findings. First, as expected, the L1 group was extremely accurate in their selections (96.86% overall), which were not influenced by whether the referent was [\pm Human] or masculine/feminine. The L2 learners were much less accurate (64.57% overall), showing a two-way interaction (gender*Human), whereby they were more accurate with [+human] referents when the referent was feminine. Thus, although the French and Spanish clitic systems have similar feature bundles, this formal similarity did not automatically translate into an advantage in their L2 French. Finally, we found a positive effect of proficiency on overall accuracy.

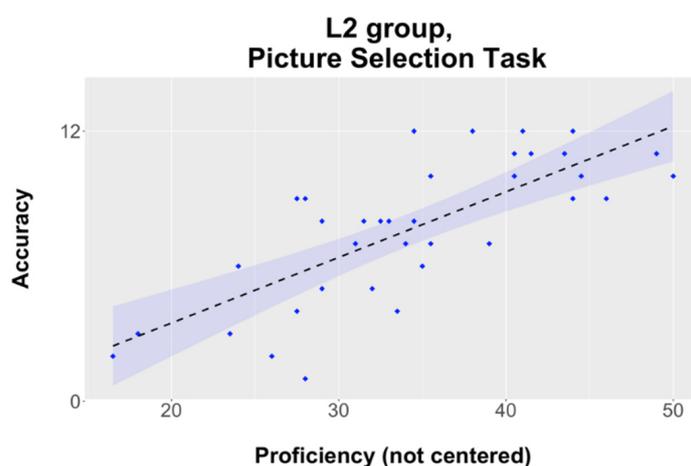


Figure 2. Scatterplot of proficiency/ v /overall accuracy, picture selection task (L2 group).

4. Experiment 2—Self-Paced Reading Task (SPR)

4.1. Participants

The same participants completed both tasks. For the SPR, we had two criteria for exclusion. We excluded one participant because their overall mean Reading Times (RTs) were more than 2.5 SDs from the mean. We excluded five additional participants whose accuracy in the comprehension questions was lower than 70%. All exclusions came from the L2 group.

4.2. Materials and Procedure

4.2.1. Procedure

Participants read sentences in French in a non-cumulative moving-window display (Just et al. 1982) as presented by the software Linger (Rohde 2003). Each trial included a non-moving discourse context followed by a sentence with non-space characters replaced by dashes. To read sentences, participants were asked to press the space bar at whatever pace suited them. The software recorded the duration between space-bar presses in milliseconds. After participants read a moving segment, the word changed back to dashes so segments could not be reread. To verify comprehension (and to test attentiveness), each sentence was followed by a true/false question, for which half of the answers were true, half false. Additionally, half of the questions focused on the context and half on the moving-window sentences. After each sentence, participants received automatic feedback on their accuracy on the comprehension questions. Instructions were presented in French and Spanish (on different pages) for the learners, and in French for the NSs. To familiarize the participants with the task, we included five practice items (participants received feedback). Trials were randomized per participant and presented in a Latin square; a given participant saw only one version, out of four conditions, per lexicalization/item.

4.2.2. Materials

Excluding practice items, participants read 58 sentence/context combinations, of which 24 were experimental sentences and 34 were fillers. The context included at the outset of each item was vital to foreground the information in the test sentence; that way, participants would have the potential antecedents for each clitic pronoun (e.g., whether the antecedent was feminine/masculine, human/nonhuman). The inclusion of the context, however, did make the experiment a long one: L2 Participants completed this portion of the study, on average, in 40 min (NSs completed the task in less than 30 min, on average). Participants were encouraged to take optional breaks, which occurred every 20 sentences.

The SPR task had a $2 \times 2 \times 2$ design; the 24 experimental items (six items per cell) manipulated gender (masculine/feminine), [\pm Human] (human/nonhuman), and match (match/mismatch) as independent variables. Figure 3 illustrates a sample item (English translations were not present). Note that the “match” options come from the match or mismatch of gender features between the referent in the discourse (le passeport “passport” [masculine]) and the features in the accusative clitic (“le” [masculine]). In this sense, the mismatches used in the experiment were soft (discourse) violations, rather than instances of ungrammaticality. Based on the psycholinguistic literature, we expected that mismatched responses would be read more slowly, but we did not expect longer RTs based on the [\pm Human] feature, since French and Spanish clitics are perfect analogues regarding this formal similarity. However, because French and Spanish have a complete overlap of function and form for the feminine clitic [la], we expected RTs to be a bit slower for masculine referents if surface similarity played a role.

Context	Il y a deux mois Claire a renouvelé son passeport et sa carte bancaire. Quelques jours plus tard elle a réalisé que le nouveau passeport était introuvable, mais pas la carte bancaire. <i>Two months ago, Claire renewed her passport [masculine] and her debit card [feminine]. Some days later, she realized that she couldn't find her passport, but she did find her debit card.</i>						
Item	Voilà pourquoi Claire le cherche avec acharnement. <i>That is why Claire searches for it relentlessly.</i>						
CONDITION	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6	Region 7
nonhuman, masculine, match	Voilà	pourquoi	Claire	le	cherche	avec	acharnement.
	That's	why	Claire	it [acc.masc]	searches	with	frenzy

Figure 3. Schematization of an SPR item with English translations.

The SPR was not designed to test the participants' lexical knowledge of gender, which was always unambiguously marked on the determiner (e.g., *sa* "his/her"/*la* "the" vs. *son* "his/her"/*le* "the") in the (non-moving) context. As in Figure 3, the context in experimental items introduced a main character and two other potential referents (human or non-human, depending on the condition) of different gender. To control for phonological weight, Regions 4 and 6 all had the same number of syllables. Verbs (Region 5) were one- ($n = 5$) and two-syllable ($n = 7$) core transitive verbs in the third-person singular present tense (*regarde* "looks", *cherche* "looks for", *comprend* "understands", *choisit* "chooses", *dessine* "draws", *décrit* "describes", *déteste* "hates", *touche* "touches", *pousse* "pushes", *laisse* "leaves", *surveille* "monitors", *voit* "sees"). Each verb was repeated twice.

4.3. Results

To account for differences in word length and individual reading speeds, we length-adjusted RTs using Fine et al.'s (2013) procedure. Then, because RTs were positively skewed (typical of RTs elicited with this type of task), we log-transformed the residuals. This procedure has been used successfully in previous psycholinguistic investigations (e.g., Kim 2018). We then analyzed these length-adjusted log-transformed RTs (henceforth logRTs) using mixed models for each group (L1 vs. L2), based on the reasoning offered earlier. For each model, we included three within-subjects fixed factors: whether the referent was [\pm Human] (human/non-human); whether the clitic constituted a match with the discourse antecedent (match/mismatch); and whether the gender of the clitic was masculine or feminine.

Because our experimental design was repeated-measures, we included random effects by participant and by item. Following Barr et al. (2013), we fit the maximal converging random-effects structure using a top-down stepwise method. As a departure point, we included all possible random by-subject and by-item slopes, plus their intercepts and interactions. When models did not converge, we removed the random slope that accounted for the least variation. This was an iterative process that continued until the model converged. Alpha was set at 0.05. Once again, we used the *lmer* package (Bates et al. 2015).

Figures 4 and 5 show the logRTs for NSs (L1 French) and L2 learners of French (L1 Spanish), respectively, plotted by gender, [\pm Human], and match (error bars represent standard error). We expected the effects to show up at the critical region (Region 5), which corresponded to the verb. As is typical in self-paced reading investigations, however, we also expected to analyze subsequent regions for spill-over effects. This is especially common for L2 learners, who typically read more slowly than natives (Dekydtspotter et al. 2006).

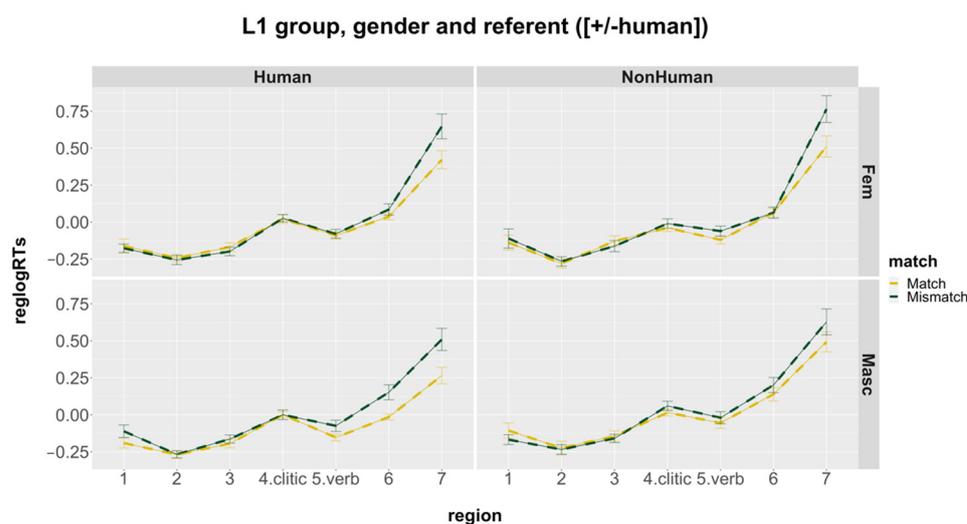


Figure 4. Length-adjusted, log-transformed reading times by region, L1 group.

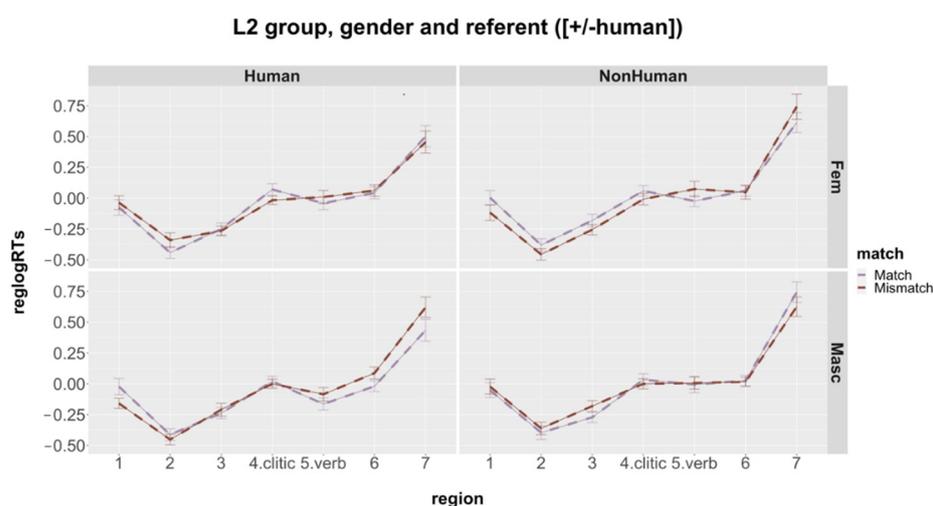


Figure 5. Length-adjusted, log-transformed reading times by region, L2 group.

At the critical region (Region_5-verb), the model for the NSs revealed no main effects, but it did unveil a gender*[Human] interaction (Table 7). Follow-up pairwise comparisons using paired-samples *t*-tests in *emmeans* in R (Tukey HSD) revealed that [+human] referents were read faster than [−human] referents, but only in the masculine conditions ($t = -2.150$, $p = 0.0409$). No other contrasts were significant. The model for L2 learners did not have any main effects or interactions (Appendix A); L2 participants did not make significant distinctions between matched and mismatched referents, irrespective of whether the referent was human, nonhuman, masculine, or feminine.

At the post-verbal region (Region 6), the model for NSs (Appendix B) unveiled an overall marginal human*gender interaction ($p = 0.0827$), indicating that the differences in logRTs between human and nonhuman referents varied according to gender (masculine, feminine). Because the interaction was marginal, we did not conduct post-hoc pairwise comparisons, although inspecting Figure 4 (Region 6) reveals that masculine referents were read faster when they were human and slower when non-human, while this trend was reversed for feminine referents. The model for L2 learners revealed no main effects or interactions (Appendix C).

Table 7. Linear mixed-effects model, postverbal region (Region 5-SPR), L1 group.

Fixed Effects	Estimate	SE	df	t-Value	Pr(> t)
Match	0.007	0.051	71.90	0.145	0.885
[±Human]	−0.034	0.047	116.95	−0.725	0.470
Gender	−0.066	0.046	131.83	−1.439	0.152
Match*[Human]	0.054	0.062	545.03	0.872	0.384
Match*Gender	0.078	0.062	549.56	1.267	0.206
[Human]*Gender	0.137	0.062	546.02	2.203	0.028 *
Match*[Human]*Gender	−0.106	0.088	546.04	−1.213	0.226
Random Effects				Variance	SE
By-Participant Intercept				0.007	0.086
By-Participant Slope over Match				0.005	0.073
By-Participant Slope over [Human]				0.002	0.039
By-Participant Slope over Gender				0.000	0.011
By-Item Intercept				0.003	0.054
By-Item Slope over Match				0.005	0.072
By-Item Slope over [Human]				0.003	0.053
By-Item Slope over Gender				0.002	0.049

Finally, although it is not typical to analyze the last region in SPR investigations, Figures 4 and 5 show a pronounced delay in the onset of effects. Our sentences were relatively short, so a delayed effect would not be unusual. However, we realize that disentangling wrap-up effects, which are not completely understood, is impossible with our design. We will address the potential consequences of this design feature in the discussion of our limitations. The model for the L1 group is presented in Table 8, while the model for L2 learners is presented in Table 9.

Table 8. Linear mixed-effects model, Region 7-SPR, L1 group.

Fixed Effects	Estimate	SE	df	t-Value	Pr(> t)
Match	0.229	0.113	56.74	2.026	0.048 *
[±Human]	0.089	0.103	104.71	0.868	0.388
Gender	−0.157	0.096	203.57	−1.637	0.103
Match*[Human]	0.030	0.132	545.05	0.230	0.818
Match*Gender	0.012	0.132	548.53	0.091	0.927
[Human]*Gender	0.137	0.131	543.23	1.042	0.298
Match*[Human]*Gender	−0.139	0.186	546.30	−0.743	0.458
Random Effects				Variance	SE
By-Participant Intercept				0.043	0.208
By-Participant Slope over Match				0.029	0.170
By-Participant Slope over [Human]				0.021	0.145
By-Participant Slope over Gender				0.010	0.098
By-Item Intercept				0.001	0.023
By-Item Slope over Match				0.035	0.186
By-Item Slope over [Human]				0.014	0.117
By-Item Slope over Gender				0.003	0.050

The model for NSs at Region 7 revealed a significant main effect of match ($p = 0.0475$); no other effects or interactions reached significance. Thus, native French speakers were able to distinguish between matched and mismatched referents, although that decision was not influenced by the status of the referent as [\pm Human] or by the clitic's gender. The model for the L2 learners revealed only a significant three-way interaction (match*[Human]*gender). Follow-up pairwise comparisons using paired-samples t -tests (*emmeans* in R; Tukey HSD) revealed a significant contrast between human and nonhuman referents when matched for gender ($t(30.4) = -2.703$, $p = 0.0111$) and a marginal effect when there was a gender mismatch ($t(32.2) = -1.931$, $p = 0.0624$). Similarly, there was a significant contrast between human and nonhuman referents for masculine ($t(31.9) = -2.046$, $p = 0.0490$) and feminine ($t(31.1) = -2.584$, $p = 0.0147$). No other contrasts were significant. Significant results are marked with an asterisk (*).

Table 9. Linear mixed-effects model, Region 7-SPR, L2 group.

Fixed Effects	Estimate	SE	df	t -Value	Pr(> t)
Match	−0.050	0.127	88.38	−0.392	0.696
[\pm Human]	0.114	0.110	296.21	1.038	0.300
Gender	−0.073	0.113	134.16	−0.643	0.522
Match*[Human]	0.187	0.153	578.43	1.223	0.223
Match*Gender	0.244	0.154	579.71	1.582	0.114
[Human]*Gender	0.205	0.153	577.77	1.345	0.179
Match*[Human]*Gender	−0.493	0.218	578.08	−2.268	0.024 *
Random Effects				Variance	SE
By-Participant Intercept				0.091	0.302
By-Participant Slope over Match				0.070	0.264
By-Participant Slope over [Human]				0.012	0.108
By-Participant Slope over Gender				0.002	0.047
By-Item Intercept				0.019	0.138
By-Item Slope over Match				0.022	0.147
By-Item Slope over [Human]				0.002	0.043
By-Item Slope over Gender				0.012	0.111

Finally, we examined the data by proficiency, treated as a continuous variable, to determine whether it had an impact on match. We ran a generalized linear mixed-effects regression with (centered) proficiency as a fixed-effect predictor and the maximal RES that would converge on the data set (by-subject and by-item slopes, plus their intercepts). This model did not yield any effects related to proficiency (Appendix D), suggesting that accuracy on the task did not increase with higher proficiency.

In summary, the SPR unveiled several important findings. First, as expected, effects were delayed, especially for the L2 group. Second, the NSs made a relevant distinction between matched and mismatched referents, although this result was not affected by the other factors (gender/[\pm Human]). The L2 learners, on the other hand, evinced a complex three-way interaction (gender*Human*match), whereby they read human and nonhuman referents at different speeds (to different degrees) when referents were matched for gender. They made this distinction within both masculine and feminine clitics. In short, while NSs made the expected distinctions between matched and mismatched referents, the L2 learners were guided by the [\pm Human] feature, even though their L1 and L2 operate in analogous ways.

5. General Discussion

As we mentioned at the outset, the goal of our study was two-fold. First, we aimed to determine whether formal similarity between two languages (operationalized via the Feature Reassembly Hypothesis, which is couched in the generative tradition) would allow adult L2 learners of French (Spanish NSs) to straightforwardly acquire third-person singular accusative clitics in their L2 (RQ1). We focused on Spanish-French bilinguals because, in terms of formal similarity, both languages encode the same features in analogous ways regarding accusative clitics. Additionally, following insights from usage-based research on bilinguals, we tried to determine the role played by surface similarity because French and Spanish overlap and diverge in interesting ways; we also aimed to determine whether the difference in the actual form of the masculine (but not feminine) clitics could affect results (RQ2).

To answer these questions, we used an offline picture selection task and an online self-paced reading task. The results of both tasks showed that, although French and Spanish encode features in similar ways, Spanish-French bilinguals evinced non-trivial difficulties in their interpretation and processing of French accusative clitics. With regard to our research questions, the data from both tasks showed that the answer to RQ1 was negative: Although there was nothing to reassemble, the acquisition of these clitics did *not* proceed straightforwardly, against the predictions of the Feature Reassembly Hypothesis.

The picture selection task showed that the overall accuracy of the learners was relatively low (64.57% overall, compared to 96.86% for the L1 group), while also showing that learners were more accurate with feminine clitics, but only where these were [+human]. Alternatively, learners were more accurate with [+human] referents, but only when these were feminine. This result lends partial support to the notion that surface-level similarity may play a role, since the feminine clitics share the same form in both languages (“la” [la]), while the masculine clitics differ both in written form and sound (“le” [lə] French, “lo” [lo] in Spanish). Additionally, we did observe proficiency effects over overall accuracy, which means the developmental trajectory is positive.

The results of the self-paced reading task also showed that learners did not process accusative clitics in the same way as natives, and this difference was not only quantitative. Although the learners evinced a delayed effect, compared to NSs, they were also influenced by the gender of the clitic and whether the referent was [\pm Human] such that they made a distinction between [+human] and [−human] referents within the matched conditions but also for both masculine and feminine referents. The native French speakers’ processing of these clitics, however, was not guided by either gender or the feature [\pm Human] because the only relevant distinction was between (gender-) matched or mismatched referents. These results again challenge the Feature Reassembly Hypothesis, although they do not offer unequivocal support to the notion that surface-level similarity matters: This issue was present in both the feminine and masculine conditions.

Together, the results from both experiments showed that Feature Reassembly is not the only factor in the acquisition of object clitics: Even when there is no need for reassembly, that does not entail convergence (RQ1). In terms of surface-level similarity (RQ2), we have partial support from our offline task that there might be a facilitatory effect when formal and surface-level similarity coincide. However, our investigation does leave several open questions, the most pressing of which is perhaps this: Given the overwhelming similarity between French and Spanish feminine accusative clitics, why do we not observe convergence?

One interesting feature of both participant groups is their degree of multilingualism. As shown earlier, all participants, irrespective of group, reported learning English. This is not surprising because in Mexico, exposure to English is, at least nominally, compulsory under a program known as *El Programa Nacional de Inglés* (National English Program) sponsored by the Ministry of Mexican Public Education (*Secretaría de Educación Pública*). In practice, however, Spanish-English bilingualism is far from universal. This is relevant because, as mentioned in the introduction, English differs from both Spanish and French in terms of encoding the [\pm Human] feature, since the pronominal system does encode

[±Human] (e.g., *he/she* vs. *it*). Previous research with Anglophone learners of French revealed differences in reading time patterns depending on the value of the [±Human] feature (Shimanskaya and Slabakova 2017). Thus, we cannot rule out that the L2 learner group could be influenced by the other languages in their repertory. This would necessitate, however, that they had actually acquired the English pronominal system, which is far from an easy task. Given the ubiquity of English as an L2 in Mexico, future research should control for learners' knowledge of the English pronominal system.

An additional open question from the picture selection task is the higher accuracy rates with feminine referents, especially when these are [+human]. We noted earlier the potential role of surface-level similarity, but there might be additional explanations. Recall that, in terms of gender, French NPs bear a gender feature that can be specified as [feminine] or else left unspecified, which means it receives a [masculine] valuation by default. This means that [feminine] values are contingent on it being distinguished by the feature [Gender]. [Masculine], on the other hand, does not have to be individuated because it is not assumed to be represented in the feature matrices. Thus, it is possible that the lower accuracy with masculine clitics is related to its underspecification of gender [Gender: ∅]. Given our particular language combination, this is a possibility we cannot discard.

We noted that researchers such as Hopp (2013) have proposed that lexical knowledge of gender might be distinguished from how gender is processed online. In theory, then, this could constitute an explanation for our results. In our materials, we did not match the gender of the referents because the gender was unambiguously marked in the determiner in the context for both tasks. Moreover, after examining our materials, we note that most NPs were of the same gender in both languages. In the SPR, only six referents (out of the 48 introduced) did not match in gender: With one exception (*grille-pan* [masc.French], *tostadora* [fem.Spanish] "toaster"), these referents were all feminine in French but masculine in Spanish. In the picture selection task, only two of the referents (*le paplemousse* "the grapefruit", *le sac à dos* "the backpack") were not matched, since they were masculine in French but feminine in Spanish (*la toronja* "the grapefruit", *la mochila* "the backpack"). Thus, even if learners had missed the gender in the context, this is an unlikely explanation for our results.

No study is without limitations, and ours is no exception. As mentioned earlier, our study was very long overall (almost two hours of testing), and we tried to shorten the duration by including shorter test sentences in the SPR. This meant the delayed effects occurred in the last region, which is not typically analyzed because wrap-up effects are not well understood (Just et al. 1982). We will note, however, that we are not the first to analyze this region (Tokowicz and Warren 2010); moreover, our data (Figures 4 and 5) show that the effects were quite delayed, which is not surprising because our regions were very small, often including a single word. Future research should take this into account by ensuring that sentences have a sufficient number of regions after the critical one.

6. Conclusions

Our study aimed to determine whether formal similarity between two languages (operationalized via the Feature Reassembly Hypothesis) would allow adult L2 learners of French (Spanish NSs) to straightforwardly acquire third-person singular accusative clitics in their L2. Additionally, as suggested by usage-based accounts, we examined the role of surface similarity, since French and Spanish overlap and diverge in several ways. Results from a forced-choice picture selection task and a self-paced reading task could not support the Feature Reassembly Hypothesis, as our learners showed considerable difficulty with the interpretation and processing of these pronouns, revealing that, unlike French NSs, their interpretations and processing are guided by the feature [±Human] and, to a lesser degree, by gender, which might be due to the surface-level similarity between feminine accusative clitic pronouns in both languages.

Author Contributions: All authors were involved in the conceptualization, methodology, analysis, and writing. All authors have read and agreed to the published version of the manuscript.

Funding: Funding provided by an NSF Dissertation Improvement Grant (1452413).

Institutional Review Board Statement: Approved by the University of Iowa’s IRB (ID# 201302704).

Informed Consent Statement: Informed consent was obtained from all subjects involved.

Data Availability Statement: Data will be available upon request without undue reservation.

Acknowledgments: We are thankful to Roumyana Slabakova for her help with data collection. We are also indebted to the *Alianza Francesa campus Puebla* for their help.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Linear mixed-effects model, postverbal region (Region 5-SPR), L2 group.

Fixed Effects	Estimate	SE	df	t-Value	Pr(> t)
Match	0.053	0.080	114.74	0.665	0.507
[±Human]	0.020	0.086	58.69	0.229	0.820
Gender	−0.119	0.082	69.98	−1.459	0.149
Match*[Human]	0.045	0.103	596.28	0.432	0.666
Match*Gender	0.026	0.103	595.97	0.255	0.799
[Human]*Gender	0.138	0.103	596.26	1.339	0.181
Match*[Human]*Gender	−0.113	0.146	596.16	−0.773	0.440
Random Effects				Variance	SE
By-Participant Intercept				0.000	0.000
By-Participant Slope over Match				0.012	0.110
By-Participant Slope over [Human]				0.012	0.108
By-Participant Slope over Gender				0.000	0.019
By-Item Intercept				0.007	0.081
By-Item Slope over Match				0.008	0.088
By-Item Slope over [Human]				0.020	0.143
By-Item Slope over Gender				0.016	0.126

Appendix B

Table A2. Linear mixed-effects model, postverbal region (Region 6-SPR), L1 group.

Fixed Effects	Estimate	SE	df	t-Value	Pr(> t)
Match	0.047	0.060	66.64	0.776	0.440
[±Human]	−0.027	0.055	139.39	0.489	0.625
Gender	−0.049	0.054	200.02	−0.902	0.368
Match*[Human]	−0.046	0.073	573.50	−0.632	0.527
Match*Gender	0.112	0.073	571.25	1.544	0.123
[Human]*Gender	0.013	0.073	572.70	1.738	0.083
Match*[Human]*Gender	−0.064	0.103	574.02	−0.624	0.533
Random Effects				Variance	SE
By-Participant Intercept				0.000	0.000
By-Participant Slope over Match				0.003	0.046
By-Participant Slope over [Human]				0.002	0.041
By-Participant Slope over Gender				0.003	0.056
By-Item Intercept				0.000	0.000
By-Item Slope over Match				0.011	0.106
By-Item Slope over [Human]				0.004	0.059
By-Item Slope over Gender				0.003	0.046

Appendix C

Table A3. Linear mixed-effects model, postverbal region (Region 6-SPR), L2 group.

Fixed Effects	Estimate	SE	df	t-Value	Pr(> t)
Match	0.019	0.067	98.29	0.286	0.776
[±Human]	0.015	0.064	246.48	0.225	0.822
Gender	−0.065	0.066	162.78	−0.994	0.322
Match*[Human]	−0.029	0.090	588.26	−0.324	0.746
Match*Gender	0.089	0.089	587.67	0.991	0.322
[Human]*Gender	0.032	0.090	588.20	0.354	0.724
Match*[Human]*Gender	−0.087	0.127	588.08	−0.689	0.491
Random Effects				Variance	SE
By-Participant Intercept				0.012	0.108
By-Participant Slope over Match				0.004	0.061
By-Participant Slope over [Human]				0.002	0.046
By-Participant Slope over Gender				0.001	0.028
By-Item Intercept				0.004	0.062
By-Item Slope over Match				0.004	0.060
By-Item Slope over [Human]				0.001	0.031
By-Item Slope over Gender				0.003	0.058

Appendix D

Table A4. Linear mixed-effects model, Region 7, Proficiency and Match, L2 group.

Fixed Effects	Estimate	SE	df	t-Value	Pr(> t)
Match	0.004	0.057	0.062	0.753	0.451
Prof_c	0.051	0.010	0.385	0.501	0.61
Match*[Prof_c]	0.0172	0.008	6.207×10^2	1.320	0.187
Random Effects				Variance	SE
By-Participant Intercept				0.099	0.314
By-Item Intercept				0.010	0.101

Notes

- ¹ In Spanish, clitics can be attached to an infinitive or a gerund after a reconstruction verb. However, this discussion is outside of the scope of our investigation.
- ² DOM stands for “Differential Object Marker,” which is used in Spanish to mark objects when these are [+specific] and [+human].
- ³ Although our study does not focus on dative pronouns, we will point out that the situation is analogous for dative clitics because both French and Spanish use forms that are not inflected for gender (*lui*-French, *le*-Spanish; see Table 1).
- ⁴ Our presentation leaves aside the phenomenon known as *leísmo*—the use of the dative clitic *le* instead of the (etymological) accusative pronouns (*lo/la*). This is because Mexican Spanish is not considered a *leísta* dialect and instances of *leísmo* are viewed as representing intensifying verbal affixes instead (Cacoullos and Hernández 1999). While our French NSs might have come in contact with some *leísta* varieties due to proximity with Spain, only a few were Spanish learners and, crucially, we test only their native French.
- ⁵ Following our IRB protocol, participants could skip questions. Two participants (L2 group) did not answer all demographic questions. Our background questionnaire did not allow us to determine whether learners learned English or French/Spanish first. The information from two participants is missing in some categories, which is why the totals do not reach 100%.

References

- Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68: 255–78. [\[CrossRef\]](#)
- Baten, Kristof, Fabrice Hofman, and Tom Loeys. 2011. Cross-linguistic activation in bilingual sentence processing: The role of word class meaning. *Bilingualism, Language, and Cognition* 14: 351–59. [\[CrossRef\]](#)
- Bates, Douglas, Martin Maechler, Ben Bolker, Steven Walker, Rune Haubo Bojesen Christensen, Henrik Singmann, Bin Dai, Fabian Scheipl, and Gabor Grothendieck. 2015. *Package 'lme4'*. Vienna: R Foundation for Statistical Computing.
- Bley-Vroman, Robert. 2009. The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition* 31: 175–98. [\[CrossRef\]](#)
- Bruhn de Gavarito, Joyce, and Silvina Montrul. 1996. Verb movement and clitic placement in French and Spanish as a second language. In *Proceedings of the 20th Annual Boston University Conference on Language Development*. Edited by Andy Stringfellow, Dalia Cahana-Amitay, Elizabeth Hughes and Andrea Zukowski. Somerville: Cascadilla Press, pp. 61–78.
- Bruhn de Garavito, Joyce, and Lydia White. 2002. The L2 acquisition of Spanish DPs. The status of grammatical features. In *The Acquisition of Spanish Morphosyntax: The L1/L2 Connection*. Edited by Ana Teresa Pérez-Leroux and Juana Liceras. Dordrecht: Kluwer, pp. 151–76.
- Cacoullos, Rena Torres, and Esteban Hernández. 1999. A trabajarle: La construcción intensiva en el español mexicano. *Southwest Journal of Linguistics* 18: 79–100.
- Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In *Clitics in the Languages of Europe*. Edited by Henk van Riemsdijk. New York: Mouton de Gruyter, pp. 145–233.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Dekydtspotter, Laurent, Bonnie Schwartz, and Rex Sprouse. 2006. The comparative fallacy in L2 processing research. In *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*. Edited by Mary Grantham O'Brien, Christine Shea and John Archibald. Somerville: Cascadilla, pp. 33–40.
- Dijkstra, Ton, Henk Van Jaarsveld, and Sjoerd Ten Brinke. 1998. Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition* 6: 51–66. [\[CrossRef\]](#)
- Duyck, Wouter, Eva Van Assche, Denis Drieghe, and Robert J. Hartsuiker. 2007. Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology* 33: 663–79. [\[CrossRef\]](#) [\[PubMed\]](#)
- Edmonds, Amanda, and Aarnes Gudmestad. 2018. Gender marking in written L2 French: Before, during, and after residence abroad. *Study Abroad Research in Second Language Acquisition and International Education* 3: 58–83. [\[CrossRef\]](#)
- Fine, Alex, Florian Jaeger, Thomas Farmer, and Ting Qian. 2013. Rapid expectation adaptation during syntactic comprehension. *PLoS ONE* 8: e77661. [\[CrossRef\]](#)
- Foucart, Alice, and Cheryl Frenck-Mestre. 2011. Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition* 14: 379–99. [\[CrossRef\]](#)
- Franceschina, Florencia. 2005. *Fossilized Second Language Grammars: The Acquisition of Grammatical Gender*. Amsterdam: John Benjamins Publishing.
- Gabriele, Alison, Robert Fiorentino, and José Alemán-Bañón. 2013. Examining second language development using event-related potentials: A cross-sectional study on the processing of gender and number agreement. *Linguistic Approaches to Bilingualism* 3: 213–32. [\[CrossRef\]](#)
- Grüter, Theres, Casey Lew-Williams, and Anne Fernald. 2012. Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research* 28: 191–215. [\[CrossRef\]](#)
- Gudmestad, Aarnes, Amanda Edmonds, and Thomas Metzger. 2019. Using variationism and learner corpus research to investigate grammatical gender marking in additional-language Spanish. *Language Learning* 69: 911–49. [\[CrossRef\]](#)
- Harris, James. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22: 27–62.
- Hawkins, Roger. 2001. *Second Language Syntax*. Amsterdam: John Benjamins Publishing.
- Hopp, Holger. 2013. Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research* 29: 33–56. [\[CrossRef\]](#)
- Irmen, Lisa, and Jochen Knoll. 1999. On the use of the grammatical gender of anaphoric pronouns in German. A comparison between Finns and Germans. *Sprache Kognition* 18: 123–35. [\[CrossRef\]](#)
- Jakubowicz, Celia, Lea Nash, Catherine Rigaut, and Christophe-Loïc Gerard. 1998. Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition* 7: 113–60. [\[CrossRef\]](#)
- Just, Marcel, Patricia Carpenter, and Jacqueline Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology* 111: 228–38. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kim, Hyunwoo. 2018. Second language processing of Korean floating numeral quantifiers. *Journal of Psycholinguistic Research* 47: 1101–19. [\[CrossRef\]](#) [\[PubMed\]](#)
- Leal, T. 2018. Data analysis and sampling: Methodological issues concerning proficiency in SLA research. In *Critical Reflections on Data in Second Language Acquisition*. Edited by Amanda Edmonds and Aarnes Gudmestad. Amsterdam: John Benjamins, pp. 63–88.
- Lardiere, Donna. 2009. Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25: 173–227. [\[CrossRef\]](#)

- Liceras, Juana, Raquel Fernández-Fuertes, Susana Perales, Rocío Pérez-Tattam, and Kenton Spradlin. 2008. Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua* 118: 827–51. [CrossRef]
- Lumsden, John. 1992. Functional categories in the lexicon. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 37: 219–40. [CrossRef]
- Marian, Viorica, and Michael Spivey. 2003. Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism Language and Cognition* 6: 97–115. [CrossRef]
- Montrul, Silvina, Rebecca Foote, and Silvia Perpiñán. 2008. Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning* 58: 503–53. [CrossRef]
- Prévost, Philippe. 2009. *The Acquisition of French: The Development of Inflectional Morphology and Syntax in L1 Acquisition, Bilingualism, and L2 Acquisition*. Amsterdam: John Benjamins Publishing.
- R Core Team. 2020. R: A Language and Environment for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 25 August 2021).
- Renaud, Claire. 2010. On the Nature of Agreement in English-French Acquisition: A Processing Investigation in the Verbal and Nominal Domains. Unpublished Ph.D. dissertation, Indiana University, Bloomington, Indiana.
- Roberge, Yves. 1990. *Syntactic Recoverability of Null Arguments*. Montreal: McGill-Queen's Press.
- Rohde, Doug. 2003. Linger: A Flexible Platform for Language Processing Experiments. Available online: <http://tedlab.mit.edu/~dr/Linger> (accessed on 25 August 2021).
- Sabourin, Laura, and Marco Haverkort. 2003. Neural substrates of representation and processing of a second language. In *The Lexicon-Syntax Interface in Second Language Acquisition*. Edited by Roeland van Hout, Aafke Hulk, Folkert Kuiken and Richard Towell. Amsterdam: John Benjamins, pp. 175–96.
- Sabourin, Laura, and Laurie Stowe. 2008. Second language processing: When are first and second languages processed similarly? *Second Language Research* 24: 397–430. [CrossRef]
- Sagarra, Nuria, and Julia Herschensohn. 2010. The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua* 120: 2022–39. [CrossRef]
- Schlyter, Suzanne. 2003. Development of verb morphology and finiteness in children and adults acquiring French. In *Information Structure and Dynamics of Language Acquisition*. Edited by Christine Dimorth and Marianne Starren. Amsterdam: John Benjamins, pp. 15–44.
- Schwartz, Ana, and Judith Kroll. 2006. Bilingual lexical activation in sentence context. *Journal of Memory and Language* 55: 197–212. [CrossRef]
- Shimanskaya, Elena, and Roumyana Slabakova. 2017. Re-assembling objects: A new look at the L2 acquisition of pronominal clitics. *Bilingualism: Language and Cognition* 20: 512–29. [CrossRef]
- Tokowicz, Natasha, and Tessa Warren. 2010. Beginning adult L2 learners' sensitivity to morphosyntactic violations: A self-paced reading study. *European Journal of Cognitive Psychology* 22: 1092–106. [CrossRef]
- van Riemsdijk, Henk. 1999. Clitics: A state of the art report. In *Language Typology: Clitics in the European Languages*. Edited by Henk van Riemsdijk. Berlin: Mouton de Gruyter, pp. 1–30.
- White, Lydia, Elena Valenzuela, Martyna Kozłowska-Macgregor, and Yan-Kit Ingrid Leung. 2004. Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25: 105–33. [CrossRef]
- Zagona, Karen. 2002. *The Syntax of Spanish*. Cambridge: Cambridge University Press.