

Learners' Overgeneralization of Dative Clitics to Accusative Contexts: Evidence for Prototype Effects in SLA

Eve Zyzik

Michigan State University

1. Introduction

This paper investigates the acquisition of the dative case by L2 learners of Spanish whose native language is English. Specifically, I focus on the emergence and subsequent development of the third-person clitics *le(s)*, since these forms are the outward manifestation of the dative case in Spanish. There are several reasons why the dative case constitutes a valuable focus of investigation for SLA researchers. For those interested specifically in Spanish as a second language, it is clear that all verbal clitics (including *le/les*) are an important hurdle along the path to proficiency, as has been demonstrated by a number of studies in the past two decades. Previous research has established that English-speaking learners tend to misinterpret preverbal clitics as subjects (VanPatten 1984, VanPatten & Cadierno 1993, VanPatten & Sanz 1995) and accept nominative experiencers with verbs that require the dative (Montrul 1997). Furthermore, these and other studies point to the scant production of object clitics during oral production tasks (Licerias et al. 1997, Sanchez & Al-Kasey 1999).

In the broader context of SLA, the study of the dative case can provide intriguing empirical data concerning the development of two closely related language areas: morphosyntax and semantic categories. In this sense, the dative can be considered a complex semantic category with a simple morphological manifestation.¹ This paper provides initial support for the idea that learners approach the task of learning complex linguistic structures by creating prototype categories, which may or may not coincide entirely with the prototype of native speakers. As the data will illustrate, L2 learners frequently overgeneralize dative case marking to accusative contexts, especially when the referent is [+animate]. The goals of this paper are twofold: 1) to postulate stages of acquisition for dative case marking among L2 learners and 2) to suggest some possible reasons for the overgeneralization of dative clitics to accusative contexts.

2. The dative case in Spanish

At first glance, the dative case in Spanish seems to encode an almost unlimited variety of functions, ranging from the prototypical indirect object in ditransitive constructions to a number of less easily classified cases in which the dative clitic appears to form an idiomatic phrase with the verb (e.g., *le da igual* 'it's all the same to him/her'). Dative constructions in Spanish have traditionally resisted a unified account, given the fact that they involve different verb types (ditransitive, transitive, and intransitive) and display a range of semantic functions. The examples that follow serve to illustrate a sample of construction types:

* The data presented in this paper are part of a larger project on the acquisition of clitics in L2 Spanish (Zyzik 2004). I am grateful to Patricia Lunn for her thorough reading of a preliminary version of this paper. I also thank an anonymous reviewer for his/her comments and suggestions. All errors remain my own.

¹ The dative is complex semantically because it displays polysemy in its core functions. Morphologically, however, the dative is uniquely encoded by *le(s)*, which is linked to an optional prepositional phrase introduced by *a*.

- (1) *Le dieron otra oportunidad.*
'They gave him/her another chance.'
- (2) *A Miguel le encanta el chocolate.*
'Miguel loves chocolate.'
- (3) *Les llegó una carta del FBI.*
'They got a letter from the FBI.'
- (4) *Alberto le abrió la puerta.*
'Alberto opened the door for him/her.'
- (5) *La taza se le cayó.*
'The cup fell from his/her hands.'
- (6) *Le robaron el auto a Claudio.*
'They stole Claudio's car.'

As can be observed from these examples, dative constructions can involve ditransitive verbs as in (1), psych verbs (2), intransitive verbs (3), prototypical transitive verbs (4), and middles when the *se* is present (5). The last example, (6), is included here to illustrate that datives occur frequently with transitive verbs to signal possession, either inalienable or not.

Part of the difficulty in classifying the dative has stemmed from competing formal and functional criteria. Formal approaches (Alarcos 1980, Hernanz & Brucart 1987) emphasize the argument status of indirect objects in contrast with "superfluous" datives (Bello 1988 [1847]) and thus rely on a series of syntactic tests to validate such a distinction. Semantic approaches, on the other hand, call attention to the thematic roles expressed by the dative participant such as Goal, Origin, Beneficiary, and others. In spite of competing motivations, most classifications recognize the distinction between dative participants that are part of the valence of the verb and those that are not. More recent studies (c.f. Delbecque & Lamiroy 1996, Maldonado 2002) have attempted to unify both types of datives in their analyses, an approach that reflects a conceptualization of the dative as a grammatical resource that allows the verb to incorporate an additional argument (Gutiérrez Ordóñez 1999, Masullo 1992).

I adopt a model of the Spanish dative from Cognitive Grammar (CG), as outlined in Maldonado (2002). One of the principle advantages of the CG approach is that it conceptualizes all uses of the dative as part of a category that exhibits prototype effects. In this sense, certain dative constructions are central or best examples of the category, while others are more peripheral to the core representation (c.f. Taylor 1989). Maldonado (2002) proposes that all dative constructions are based on the transfer schema, which is described as the relocation of the theme (either concrete or abstract) from the dominion of the subject/agent to that of the dative participant. In the CG approach, the dative participant is best represented as an experiencer, which is characterized by its active status in the target domain (c.f. Langacker 1987). After having established the basic nature of the dative participant, Maldonado explains how the specific properties of dative constructions are determined by verb type (intransitive, transitive, ditransitive), directionality (whether the transfer proceeds in canonical fashion from source to goal) and finally, by the space the dative participant occupies in the stage model (on-stage or off-stage).²

Maldonado's (2002) model emphasizes that the meanings obtained in a wide variety of dative constructions all involve an affected participant in the target domain. The central cases (i.e., prototype) of the dative involve five interrelated properties: 1) a transference, 2) a trajectory, 3) a target domain that involves locative properties, 4) a non-initiative participant actively involved in some experience, and 5) a reference point from which to locate an entity in the participants' dominion (Maldonado 2002:12). The notion of dominion, defined as "a virtual area to which a participant has access for

² For reasons of space, I do not provide further details regarding Cognitive Grammar and the canonical event model, which serves as the foundation for Maldonado's distinction between objective (on-stage) and subjective (off-stage) datives. The interested reader is referred to Langacker (1987, 1991).

interaction” (Maldonado 2002:26), is essential because it accounts for the use of the dative to signal possession, and furthermore, demonstrates that inanimate objects can also be affected participants. This last point deserves further clarification. Although Spanish is a language in which inanimates can be marked dative, corpus data indicate that the overwhelming majority of datives are [+animate] and [+human] (Silva-Corvalán 1984, Vázquez Rozas 1995). This correlation between animacy and dative case is not coincidental, given the characterization of datives as active participants.

3. Methodology

3.1 Participants

The participants were 50 L2 Spanish learners enrolled in a variety of courses at the University of California, Davis. Together, these 50 learners comprised four proficiency groups, based on their institutional status and prior exposure to Spanish: beginning, intermediate, high intermediate, and advanced. The criteria for each grouping are listed in Table 1.

Proficiency group	Number	Course Enrollment	Time Abroad
Beginner	12	Spanish 3—end of 1 st year	None
Intermediate	12	Spanish 22—end of 2 nd year	None
High Intermediate	12	Spanish 100—end of 3 rd year	2-3 months
Advanced	14	any upper division course	1 year

Table 1: Participant profiles

The group of 50 L2 learners included 38 females and 12 males, representing an age range from 17 to 24, with an average age of 20.5. In addition to their institutional status, the participants were accepted for the study only after being screened for a number of background factors: L1, study abroad, and formal instruction in other Romance languages. All participants were native English speakers who reported English as the only language spoken in their household; none had formally studied another Romance language. The participants in the beginner and intermediate groups had studied Spanish exclusively in the U.S. setting. The high intermediate and advanced students had all studied abroad, with an important difference between the two groups: the former had been on short-term study abroad programs (i.e., 3 months or less), whereas the latter group consisted solely of Spanish majors/minors who had returned from an academic year abroad.

3.2 Tasks

Oral production data were gathered from the participants by means of four different elicitation procedures: a picture book narration, a structured interview task, and two video narration tasks. The first task, the storybook narration, was based on a popular children’s picture book: dePaola’s (1978) *Pancakes for Breakfast*. In order to assist them with the storybook task, the participants were given five minutes of planning time to look at the pictures and to consult a list of eighteen vocabulary items. The second task was a structured interview consisting of a series of comprehension questions related to the storybook. The questions were intended to establish a particular object as discourse topic, which would then encourage learners to respond with an anaphoric clitic pronoun. This structured interview format has been used in previous research focusing on Spanish clitics (Sanz 1997, VanPatten & Sanz 1995). The two remaining tasks involved a video narration procedure in which participants watched two silent video clips that depicted a variety of two- and three-participant events. The participants were given a list of English-Spanish equivalents before viewing each video clip. The video clips were shown twice and the participants were asked to narrate the events that were happening on the screen during the second viewing (i.e., in real time). Data collection took place individually with each participant during two half-hour sessions. The oral data were recorded, transcribed, and then analyzed with specific attention paid to the use of the dative clitics *le(s)*.

3.3 Data analysis

The analysis of *le(s)* in the oral production of L2 learners focuses on the functional distribution of these forms. In a distributional analysis (cf. Pienemann 1998), functions performed by the morpheme(s) in the target language are included, as well those functions unique to interlanguage. This kind of analysis is believed to be more meaningful than a simple accuracy measure because it reveals which functional and/or structural contexts learners are sensitive to in their use of a particular morpheme. For the purposes of this study, the categories included in the functional distribution are based largely on Maldonado (2002), but I draw on some terminology from Relational Grammar as well (c.f. Bickford 1985). These categories, which correspond to semantic notions, are listed below.

- (a) Recipient
- (b) Source
- (c) Experiencer
- (d) Beneficiary
- (e) Possessor
- (f) Patient (m.)
- (g) Patient (f.)
- (h) Other

The first functional category, Recipient, is the prototype of dative constructions in Maldonado (2002) and involves a ditransitive verb in which the agent transfers the theme (either concrete or abstract) to the indirect object. The next category, Source, is the reversal of the prototypical transfer event, and is meant to account for directionality of transfer as a possible variable in acquisition. The Experiencer function is best illustrated by psych verbs like *gustar* or *encantar*, in which the dative participant experiences some psychological sensation (i.e., pleasure, pain, sadness). Next, the semantic role of Beneficiary occurs in the context of a potentially unlimited number of transitive and intransitive verbs, where the dative is a grammatical means by which the verb can incorporate an additional argument that is not originally part of its valence. Finally, the Possessor role refers to the use of the dative to profile a whole/part relation.

The next two categories, (f) and (g), pertain to learner language (and to certain dialects of Spanish). The thematic role ‘Patient’ refers to two-participant events in which the non-agentive participant is marked with dative case. As the data will show, there is a tendency for L2 learners to overgeneralize *le(s)* to contexts that are normally reserved for accusative clitics. This feature of learner language shares some common properties with *leísmo* in the Peninsula, but there are important differences as well. The ‘Patient’ category was subdivided (masculine and feminine) in order to determine if the gender of the antecedent is a relevant variable. For example, if learners use *le(s)* to refer to masculine referents only, this would be an indication they have constructed a pronoun system based on gender rather than case. However, if learners use *le(s)* to refer to both masculine and feminine antecedents with more or less equal frequency, then their use of the dative must be conditioned by other factors. Specifically, I explore the possibility that learners construct a prototype of dative participants based on animacy.

Finally, the ‘Other’ category has been included in order to deal with any ungrammatical learner utterances that are not easily classifiable within the existing framework. These include utterances like the following, produced by one of the advanced learners:

- (7) *La muchacha le está viéndose en el espejo.*
 ‘The girl *le* is looking at herself in the mirror.’

It is clear that this utterance is semantically incoherent; the learner has used a dative clitic as well as the co-referential *se* in a single clause. Fortunately, these idiosyncratic uses of the dative clitics constitute a minimal percentage of all tokens.

4. Results

4.1 Group totals

The total number of dative clitics (*le/les*) produced by the 50 L2 learners was 309. These tokens are distributed disproportionately among the four proficiency levels; the beginning learners produced a negligible number of dative clitics, while the advanced learners outperformed all the remaining groups with 143 tokens. The descriptive statistics for all four groups are presented below in Table 2.

	Beginner (A)	Intermediate (B)	High Intermediate (BC)	Advanced (CD)
Total tokens	5	68	93	143
Mean	.42	5.67	7.75	10.21
SD	.79	4.5	7.39	5.82

Table 2: Descriptive statistics for dative clitic use

Results from a one-way ANOVA show that there are significant differences among the four groups ($F(3, 46) = 7.94, p < .001$). A post-hoc Tukey test reveals significant differences between the beginners and the three other groups ($p < .001$ for all three comparisons), as well as between the intermediate and advanced learners ($p < .05$). These results are presented in the top row of Table 2 (following standard statistical notation, a shared letter indicates that the difference between groups failed to reach significance).

As shown in Table 2, the beginner group displayed minimal use of dative clitics. Due to the low number of tokens, a functional analysis for this group of learners would be questionable. Three of the tokens occurred in the context of *gustar/encantar* ‘like,’ one token of *le* appeared in a two-participant event (*besar* ‘kiss’) and one learner used a dative clitic with the ditransitive *dar* ‘give.’

It would be premature to draw any conclusions based on five tokens of use, but it is interesting that all of these uses, with the exception of *le besa* ‘he kisses her,’ could be viewed as lexicalized chunks. Clearly the verbs *gustar* and *encantar* are taught repeatedly in first-year language classes, and holophrastic utterances like *me gusta* and *le gusta* are commonly used by students at the beginning level. There is some evidence to indicate that the *le gusta* sequence is an unanalyzed chunk, such as the fact that learners rarely mark subject-verb agreement in these constructions (see example 8 below). This tendency is found in more experienced language learners as well. For the reader’s convenience, all samples of learner language from the corpus are marked for participant number (1-50) and task type (ST = storybook, SI = structured interview, V1/V2 = video narrations).

- (8) *Porque le encanta los panqueques.* (P7) (ST)
 ‘Because he loves pancakes.’

Table 3 presents the functional distribution for the remaining three groups of learners. The columns are split in order to present the number of tokens in each semantic category (e.g., Recipient, Beneficiary) as well as the percentage of the total number of tokens produced by each group.

	Intermediate (68)		High Intermediate (93)		Advanced (143)	
Recipient	29	42.6%	37	40%	48	33.5%
Source	0	0	4	4.3%	12	8.4%
Experiencer	14	20.6%	13	14%	8	5.6%
Beneficiary	5	7.4%	3	3.2%	7	4.9%
Possessor	0	0	1	1%	0	0
Patient (m.)	6	8.8%	25	26.8%	39	27.3%
Patient (f.)	9	13.2%	10	10.7%	27	18.9%
Other	5	7.4%	0	0	2	1.4%

Table 3: Distributional analysis

The distributional analysis for the intermediate, high intermediate and advanced learners indicates several interesting trends in the acquisition of the dative case. First, it is the Recipient function that

remains strong throughout all levels, accounting for a large proportion of all tokens of *le(s)*. The ditransitive verb *dar* ‘give’ is the most frequently used in this category, but *decir* ‘tell,’ *traer* ‘bring,’ and *servir* ‘serve’ are also found with some regularity (refer to example 9). The Source function, which was inexistent at the beginning and intermediate levels, emerges in the speech of some high intermediate learners, as illustrated in (10).

- (9) *Ella come la cena y luego Pablo le sirve un café.* (P19) (V1)
 ‘She eats the dinner and later Pablo serves her a coffee.’
- (10) *Carmen va a bañar su perro Tico y le quita el collar.* (P32) (V2)
 ‘Carmen is going to bathe her dog Tico and she takes off his collar.’

With respect to the Experiencer role, the results indicate that this function appears early in the acquisition process. However, caution must be used in interpreting the results, given that many of these tokens of *le(s)* only superficially correspond to a dative experiencer. Learners from all proficiency groups consistently employ the verb *gustar* with the appropriate dative clitic, but rarely does the utterance conform to the intransitive structure of Spanish. Consider the following examples from intermediate and high intermediate learners:

- (11) *Pero Camila no le gusta tirar su chaqueta en el armario.* (P16) (V1)
 ‘But Camila (nom.) doesn’t like to throw her jacket in the closet.’
- (12) *Y el perrito le gusta la comida.* (P28) (ST)
 ‘And the dog (nom.) likes the food.’

These examples provide further evidence for the hypothesis that *le gusta* is initially learned as a chunk and that learners reinterpret the structure to match the transitive frame of their L1 (i.e., nominative experiencer). Most learners, even at high intermediate and advanced levels, continue to use *gustar* according to English syntactic frames for the verb ‘to like.’³

The Beneficiary function of the dative case accounts for a small percentage of *le(s)* usage at all three proficiency levels. This could be a task effect in that there may not have been enough contexts in which to use this particular function of the dative. However, another possibility is that English-speaking learners rely on prepositional phrases to signal an affected participant, as illustrated in (12).

- (13) *Y Pablo pone la manta sobre Camila y le besa.* (P13) (V1)
 ‘And Pablo puts the blanket over Camila and kisses her (dat.).’

Note how this intermediate learner chooses the prepositional phrase in the first clause (instead of *le pone la manta*) while making use of the dative in the two-participant event. The use of prepositional phrases like [*para* + NP] and [*sobre* + NP] may partially account for the low incidence of the Beneficiary function in learner language.

The use of *le(s)* to indicate possession was almost nonexistent in the corpus. In general, two tendencies can be observed in the context of whole/part relations: some learners rely on the use of possessive determiners (conforming to the L1 pattern) while others use the definite determiner and no corresponding clitic to indicate the affected participant. The following examples illustrate both tendencies:

- (14) *Y empezó a secar el Tico con una toalla. Y seca su cabeza y su cuerpo.* (P25) (V2)
 ‘And she began to dry Tico with a towel. And she dries his head and his body.’

³ This can only be confirmed when the learner uses the full NP and/or a plural subject. On the contrary, an utterance like *le gusta el café* makes it impossible to determine the status of the clitic pronoun in interlanguage. It could be an indication that the learner has appropriately used the dative experiencer construction, or that he/she has simply reinterpreted *le* as subject (‘he likes coffee’).

- (15) *Ella usa un secador para secar el pelo, y las orejas, y el cuerpo.* (P26) (V2)
 ‘She uses a dryer to dry the hair, the ears, and the body.’

These examples are representative of how the dative is absent from learner language in possessive contexts, even among advanced learners who have spent an academic year abroad.

4.2 Extension of the dative to accusative contexts

The next two categories in the distributional analysis (masculine and feminine patients) point to an overgeneralization of the dative to accusative contexts. This trend begins at the intermediate level and shows a steady increase as learners gain exposure to the target language. A particularly interesting fact about this phenomenon is that it is not limited to masculine referents, as one might expect.⁴ Table 3 indicates that intermediate learners actually use *le* more frequently in the context of feminine direct objects. Example (16) illustrates this usage:

- (16) *Y él le tapa con la manta para que ella pueda dormir.* (P14) (V1)
 ‘And he covers her (dat.) with the blanket so that she can sleep.’

Since *le(s)* is used in accusative contexts with both masculine and feminine referents, the data were further analyzed to determine if another factor—animacy—constrains this type of overgeneralization. The four elicitation tasks used in the present study provided the learners with multiple opportunities to make reference to both animate and inanimate participants. While the storybook and structured interview tasks contained a higher proportion of inanimate referents, both video tasks focused on animate referents in non-agentive roles. For the intermediate group, 12 of the 15 tokens of *le/les* in accusative contexts referred to animate participants (80%). For the high intermediates, 30 out of 35 tokens occurred in the context of animate referents (85.7%). At the advanced level, where the use of *le(s)* in two-participant situations predominates, 61 of 66 tokens (92.4%) were used to mark animate participants. Table 4 summarizes this information.

	Intermediate	High Intermediate	Advanced
% of overgeneralization dative to accusative	22%	37.6%	46.2%
% of <i>le(s)</i> marking [+animate] referents in accusative contexts	80%	85.7%	92.4%

Table 4: Overgeneralization of dative clitics to accusative contexts

These results suggest the presence of two important trends. First, the tendency to overgeneralize the dative to accusative contexts grows stronger as learners gain exposure to the target language. Second, the extension of *le(s)* is not random, but rather learners are highly sensitive to the animate/inanimate distinction. There were only a handful of isolated cases in which *le(s)* was used to signal an inanimate referent. In fact, it is likely that learners specifically avoid the use of the dative with inanimate referents, despite the fact that *le(s)* is their preferred form in two-participant situations. Self-corrections, like the one in (17), are revealing because they show the hesitation that many learners experience during oral production when forced to choose between dative and accusative clitics. Note how this high-intermediate learner uses the dative *le* on two occasions (despite the fact that the first use occurs in an accusative context), then hesitates between *le* and *lo* when referring to *el abrigo*, finally settling on *lo*.

⁴ In *leísta* dialects, the dative clitic *le* is used in accusative contexts, but primarily when the referent is masculine, human and singular. The use of *le* for feminine referents is rare (cf. Klein Andreu 2000 for a detailed treatment of these dialectal phenomena).

- (17) *Camila llega a su casa después de un día de trabajo y su marido le abraza, y luego le quitó el abrigo y le, lo tiró en en sillón.* (P37) (V1)
 ‘Camila returns home after a work day and her husband hugs her (dat.), and later takes off her (dat.) coat, and throws it (dat.) it (acc.), on the chair.’

Clearly, case distinctions are problematic for these learners, given that their native language does not encode the accusative/dative contrast morphologically. What is significant, however, is that learners tend to overgeneralize in only one direction: the data indicate that learners are much more likely to err in favor of the dative than vice versa. While there were 116 tokens of dative clitics in accusative contexts, there were only 18 tokens of accusative clitics in dative contexts (i.e., *laísmo* and *loísmo*). Furthermore, in the few instances when this occurred, the learner typically corrected himself and supplied the appropriate dative clitic. This is illustrated in (18).

- (18) *La pregunta si quiere entrar a su casa. Le pregunta.* (P36) (V1)
 ‘He asks her (acc.) if she wants to go into his house. Asks her (dat.).’

5. Discussion

5.1 Acquisition of dative functions

The results presented above clearly indicate that the dative case is not acquired by L2 learners as a homogeneous structure. The emergence of the clitics *le(s)* in a learner’s oral production does not mean that he/she has acquired the full range of dative functions. On the contrary, the distributional analysis demonstrates that certain functions emerge prior to others. For example, directionality seems to play an important role: the use of the dative to mark the Recipient is prevalent among all learners past the beginner level, while the Source function appears later. This delayed use of the dative to indicate Source may be due to the lack of equivalence with the indirect object category in English, where directionality is limited to transfer from the source domain to the target domain (i.e., English indirect objects can only receive). Equally important is the finding that certain functions, such as Beneficiary and Possessor, are almost entirely absent from the repertoire of these L2 learners. With respect to possession, these English-speaking learners often rely on possessive determiners (a likely result of L1 transfer).

The results of this study also underscore the fact that dative clitics are practically inexistent in the spontaneous production of first-year learners. Despite multiple opportunities to produce the target structure, these learners displayed only isolated uses of *le(s)* in the context of chunks like *le gusta*. It is only at the intermediate level when the dative appears with some regularity, especially in the context of high frequency ditransitive verbs like *dar* ‘give’ and *decir* ‘say/tell.’ This suggests that the grammatical system of these intermediate learners is becoming tuned to particular L2 word sequences (Ellis 1996). In Spanish, the verbs *dar* and *decir* have a high statistical probability of appearing adjacent to *le(s)*, given that indirect objects are reduplicated freely in all contexts and all dialects of Spanish. It seems that the intermediate learners have seized onto this co-occurrence pattern of *le(s)* with *dar* and *decir*, and it has subsequently surfaced in their oral production. In fact, VanPatten (1987:65) had put forward as similar hypothesis in order to explain learners’ more accurate performance with indirect objects (as opposed to direct):

It does not seem to go beyond reason to suggest that perhaps learners first acquire *le* and *les* as unanalyzed chunks that go along with *dar* and *preguntar*, particularly when the two verbs are highly frequent in most language classrooms. In fact, the subjects in this study, as learners of Spanish, may have expected a clitic pronoun to accompany these verbs, although they are not quite sure what the *le* and *les* do as surface features of the sentence.

The oral production data from the current study certainly corroborate VanPatten's (1987) hypothesis with regard to the kind of item-learning (i.e., chunking) that sets in motion the acquisition of dative clitics.⁵

The effects of chunking are also readily noticeable with the verb *gustar*, which was frequently used in combination with *le(s)* by L2 learners. As an intransitive psychological predicate, the argument structure of *gustar* is unlike that of *dar* and *decir*. It does, however, promote use of the clitics *le(s)* for much the same reason. The verb *gustar* appears obligatorily with a dative clitic to signal the presence of an affected participant. As a result, L2 learners appropriately attach the dative clitic *le* to *gustar*, but they reanalyze the verb according to the L1 transitive frame of 'to like.' The resulting hybrid structure (cf. examples 11 and 12) is a combination of two mechanisms: chunking and L1 transfer in the syntactic domain.

With respect to the remaining functions of dative clitics in learner language, it has been demonstrated that learners tend to utilize the dative in two-participant events, which would normally call for accusative case. It should be emphasized, however, that this type of overgeneralization is not a categorical rule but rather a tendency that appears in the speech of L2 learners of Spanish early in the acquisition process and grows stronger as exposure to Spanish increases. It could be argued that study abroad is a factor here, especially with respect to the learners in the advanced group. However, many of the tokens of *le(s)* in accusative contexts were produced by learners who had studied in Latin America, as well as by learners who had only been exposed to classroom Spanish.⁶ Certainly, exposure to Spanish from *leísta* speakers is likely to reinforce a learner's tendency to use the dative clitics in accusative contexts. It cannot be the only factor, however, given the learners' extension of *le* to feminine referents. More importantly, we would not expect to see this phenomenon among learners who have not been abroad. As means of an explanation for why L2 learners forgo case distinctions and overgeneralize the dative to accusative contexts, I propose that they construct a semantic prototype of the dative based on frequency patterns in the input.

5.2 Constructing a dative prototype

SLA researchers have called attention to the role of internal and external factors that guide a learner's linguistic development, and such factors must be present in the acquisition of case marking as well. Internal factors refer to possible UG-derived knowledge and L1 influences, while external factors include L2 input data and explicit instruction. For the purposes of this paper, I will consider only L1 influences and L2 input data. Although explicit instruction probably has some effect on learners' use of dative clitics, this study was not designed to measure what such effects might be. As for UG-derived knowledge, it is unclear how it would contribute to learners' ability to make case distinctions, or to their acquisition of L2 morphology in general. White (1989:30-1) states that "UG does not have anything to say about morpheme acquisition as such; morphemes are lexical items; they are language specific and have to be learned."

With respect to L1 influence, the native English speaker brings to the task of SLA a linguistic system in which case distinctions are not encoded morphologically.⁷ English object pronouns have

⁵ A 'chunk' can be operationalized as any sequence of language that is produced or recalled as a whole (Weinert 1995). MacWhinney (2002) identifies chunking as one of the nine main cognitive processes involved in second language learning.

⁶ The learners in the high-intermediate and advanced groups studied in Spain as well as in various Latin American countries, including Honduras, Ecuador, Chile, and Argentina. It should be noted that certain dialects of Latin American Spanish, particularly in the Andean region, also display blurring of the *le/lo* distinction (c.f. Suñer & Yépez 1988). It is possible that this may have influenced the speech of one learner who spent a year in Ecuador. Furthermore, as suggested by a reviewer, it is impossible to know the extent to which the classroom dialect(s) to which these learners were exposed have influenced their use of dative and accusative clitics.

⁷ There is evidence to suggest that the syntax of English blurs this distinction as well: some datives can be promoted to primary object (i.e., dative shift) and can be passivized. It can be argued that English speakers know the conceptual difference between dative and accusative, even though the grammar does not encode it. I leave this broader question aside and focus instead on the issue of obligatory grammatical distinctions.

fused together accusative and dative case into one set of forms ('me,' 'you,' 'her,' 'him,' 'us,' and 'them'). In the third person singular, only the gender distinction is obligatory (e.g., 'I saw him' vs. 'I saw her'). To summarize, the dative/accusative distinction, which is absent in English, has been grammaticized in Spanish within the clitic pronoun system. Assuming that making new grammatical distinctions is problematic for adult learners (Slobin 1996), choosing the appropriate clitic in two-participant situations is guaranteed to cause some uncertainties on the part of the learner. Two-participant situations are bound to be more difficult in this respect because there are few syntactic clues that might indicate case.⁸

Turning now to the learner-external factor, L2 input, it becomes immediately clear that the dative clitics *le(s)* usually refer to animate participants. Researchers and grammarians alike have pointed out that the dative is the case used to signal animacy (Fernández-Ordóñez 1999) and corpus data confirm this hypothesis (Silva-Corvalán 1981, Vázquez Rozas 1995). Although inanimate referents are not excluded from dative case marking, the overwhelming majority of dative participants in Spanish are animate and human. The L2 production data presented here indicate that learners are sensitive to the correlation between dative case and animacy. The strong frequency effect in the input (90%) is bound to have an effect on a learner's developing grammar (Ellis 2002).

Moreover, animacy is a concept that exists independently of the linguistic forms that mark it. As Slobin (1996:94) has pointed out, "some grammaticized categories may be obvious on non-linguistic grounds." For example, Slobin suggests that the notion of plurality is obvious to all human beings, even to a speaker of a language that lacks a plural marker. It is likely that animacy is also one of these salient notions that all human beings are aware of, probably even at pre-linguistic stages. In other words, the animate/inanimate distinction is much more primary, transparent, and salient than the distinction between accusative and dative case. Accordingly, it is not surprising that L2 learners of Spanish focus on the animate/inanimate distinction when choosing which clitic pronoun to use.

The construction of any prototype depends crucially on the abstraction of features from the input. In their work on L1 acquisition of irregular past tense forms, Bybee and Slobin (1982) suggest that children create schema in order to make connections among the rote-learned forms they already have. In other words, morphological knowledge may proceed as follows: (1) rote-learned forms, (2) a schema for connecting the rote-learned forms to one another, and (3) a categorical rule. The progression from (1) to (2) takes place when the learner abstracts features from the rote-learned forms.⁹

The data indicate that the same progression described by Bybee and Slobin (1982) may be at work in the acquisition of the Spanish dative by L2 learners. Recall that the earliest appearance of the clitics *le(s)* occurred in the context of high frequency verbs like *dar* and *gustar*. Evidence was presented to indicate that learners were using *le(s)* with these verbs as unanalyzed chunks. Although these two verbs have radically different argument structures, the learners begin to notice that the *le* refers to an animate (usually human) participant. This is especially salient in three-argument verbs like *dar*, where there is a sharp contrast between the animate recipient and the inanimate theme. This abstraction of the [+animate] feature forms the basis for the prototype schema. Accordingly, the learners begin to use dative clitics in the context of numerous two-participant verbs such as *mirar*, *besar*, *bañar*, whenever the non-agentive participant is animate. This usage is reinforced in several ways: (1) exposure to input from *léista* speakers and (2) input strings in which these same verbs subcategorize for a third

⁸ Note that both direct and indirect objects appear with the object marker 'a' in Spanish. In most cases, the presence of the accusative *a* is conditioned by a set of properties of the object: proper, human (animate), concrete, count, referential, and definite (see Weissenrieder 1990 for a review). Although the presence of *a* before an NP would not help learners disambiguate direct and indirect objects, the frequency of clitic doubled constructions may be a syntactic clue: it is clear that indirect objects are reduplicated freely in all dialects of Spanish, while direct object doubling is more restricted (Silva-Corvalán 1981).

⁹ Determining exactly how this happens may be the key to understanding language acquisition in general. Ellis (1996) argues that much of this process is dependent upon analysis of sequence information. However, there is some evidence that children are much more successful at abstracting regularities from input data than late learners (Newport 1990), which may be one of the primary differences between first and second language acquisition. Clearly this issue is beyond the scope of the present study.

participant (e.g., *El hombre le besó la mano*. ‘The man kissed her hand.’). Assuming that sequence learning is a powerful mechanism in both L1 and L2 acquisition, learners may store *le besa* as a chunk in exactly the same way as they had previously stored *le da* or *le gusta*.

How does the native-speaker prototype (Maldonado 2002) compare to the one constructed by L2 learners? To the extent that the transfer construal involves an Experiencer that is [+animate], the prototypes are extremely similar. However, the prototype of the L2 learners begins to diverge from that of the native speakers when we take a look at less central cases. The most obvious context is that of possession, which involves an Experiencer and some object in his/her dominion. Here, the L2 learners in all proficiency groups consistently avoided using the dative, relying instead on possessive adjectives or definite articles. The same is true for most cases encompassing the semantic role of Beneficiary, where the L2 learners attempt to encode meaning via prepositions in contexts where Spanish native speakers are more likely to use the dative.

In summary, it seems that the two prototypes exhibit an enormous amount of overlap because of the fact that Experiencers are usually [+animate] participants. Nevertheless, they differ in terms of their breadth. The native speaker’s prototype for the dative is broader than the L2 learner’s in the sense that it can easily be applied to less central cases of participant affectedness. On the other hand, the L2 dative prototype is more extensive in a different direction; it can mark all animate participants in the target domain, regardless of their status as Patients or Experiencers. The NS prototype is more in tune to the property of affectedness (which covers a wider variety of functions), while the L2 prototype is focused almost exclusively on animacy.

A final observation regarding the potential convergence of NS and L2 prototypes is needed. The present study cannot predict whether or not the learners’ prototype will eventually acquire the same breadth as that of native speakers. With more experience and time on task, it is possible that some advanced L2 learners will become increasingly sensitive to the property of affectedness, thereby expanding the contexts of use of the dative. A similar proposal was put forth by Andersen and Shirai (1994), who traced the development of tense-aspect marking in SLA. The authors demonstrate that prototypical past events (those which are [+punctual], [+telic], and [+result]) are generally acquired earlier than other types of past situations, but they suggest that more proficient learners are able to free themselves from the prototypical uses of tense-aspect markers.

6. Conclusion

The data presented in this paper suggest a possible sequence of acquisition for the dative clitics *le* and *les* that is based, at least partially, on the semantic role of the non-agentive participant. Two equally important factors in this acquisition process include chunking and overgeneralization. The first of these two mechanisms promotes the use of dative clitics with high-frequency verbs, especially in the initial stages of acquisition. Overgeneralization, on the other hand, is manifested in the extension of dative clitics to accusative contexts. This tendency to utilize the dative in two-participant events has largely been overlooked in the SLA research on clitics (with the notable exception of Licerias et al. 1997), and suggests a potentially important role for prototypes in adult language acquisition. Specifically, I have argued that the L2 learners in this study constructed a mental schema of the prototypical dative participant, one that contains the essential feature [+animate]. This prototype coincides with the frequency effects present in the input, which ascribe an important role to the animate/inanimate distinction. As Taylor (1989:53) points out in his discussion on the basis for prototypes, “certain attributes might be particularly salient, e.g. because they are especially important in a society, with the result that these attributes cluster in prototypes.” It appears that for the L2 learners in this study, animacy was a highly salient feature that they were able to abstract from the range of dative constructions in the input.

References

- Alarcos, Emilio. 1980. *Estudios de gramática funcional del español*. Madrid: Gredos.
- Andersen, Roger W., and Yasuhiro Shirai. 1994. Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16.133-56.

- Bello, Andrés. 1988 [1847]. *Gramática de la lengua castellana destinada al uso de los americanos*, ed. by Ramón Trujillo. Madrid: Arco/Libros.
- Bickford, J. Albert. 1985. Spanish clitic doubling and levels of grammatical relations. *Lingua* 65.189-211.
- Bybee, Joan, and Dan Slobin. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58.265-89.
- Delbecque, Nicole, and Béatrice Lamiroy. 1996. Towards a typology of the Spanish dative. *The dative*, vol. 1: descriptive studies, ed. by William Van Belle and Willy Van Langendonck, 73-117. Amsterdam: Benjamins.
- dePaola, Tommie. 1978. *Pancakes for breakfast*. San Diego, CA: Voyager Books/Harcourt Brace & Company.
- Ellis, Nick. 1996. Sequencing in SLA: phonological memory, chunking, and points of order. *Studies in Second Language Acquisition* 18.91-126.
- Ellis, Nick. 2002. Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24.143-88.
- Fernández-Ordóñez, Inés. 1999. *Leísmo, laísmo y loísmo*. *Gramática descriptiva de la lengua española*, vol.1, ed. by Ignacio Bosque and Violeta Demonte, 1317-97. Madrid: Espasa Calpe.
- Gutiérrez Ordóñez, Salvador. 1999. *Los dativos*. *Gramática descriptiva de la lengua española*, vol.2, ed. by Ignacio Bosque and Violeta Demonte, 1855-930. Madrid: Espasa Calpe.
- Hernanz, M. Lluisa, and José M. Brucart. 1987. *La sintaxis*. Barcelona: Editorial Crítica.
- Klein-Andreu, Flora. 2000. *Variación actual y evolución histórica: los clíticos le/s, la/s, lo/s*. Munich: Lincom Europa.
- Langacker, Ronald. 1987. *Foundations of Cognitive Grammar*. Stanford, CA: Stanford University Press.
- Langacker, Ronald. 1991. *Concept, image, and symbol: the cognitive basis of grammar*. Berlin: Mouton de Gruyter.
- Liceras, Juana M.; Denyse Maxwell; Biana Laguardia; Zara Fernández; Raquel Fernández; and Lourdes Díaz. 1997. A longitudinal study of Spanish non-native grammars: beyond parameters. *Contemporary perspectives on the acquisition of Spanish*, vol. 1: developing grammars, ed. by Ana Teresa Pérez-Leroux and William R. Glass, 99-132. Somerville, MA: Cascadilla Press.
- MacWhinney, Brian. 2002. Extending the Competition Model. *Bilingual sentence processing*, ed. by Roberto Heredia and Jeanette Altarriba, 31-58. Amsterdam: North Holland/Elsevier.
- Maldonado, Ricardo. 2002. Objective and subjective datives. *Cognitive Linguistics* 13.1-65.
- Masullo, Pascual. 1992. *Incorporation and Case-theory in Spanish: a cross-linguistic perspective*. Seattle: University of Washington dissertation.
- Montrul, Silvina. 1997. On the parallels between diachronic change and interlanguage grammars: the L2 acquisition of the Spanish dative case system. *Spanish Applied Linguistics* 1.87-113.
- Newport, Elissa. 1990. Maturational constraints on language learning. *Cognitive Science* 14.11-28.
- Pienemann, Manfred. 1998. *Language processing and second language development: Processability Theory*. Amsterdam: John Benjamins.
- Sanchez, Liliana, and Tamara Al-Kasey. 1999. L2 acquisition of Spanish direct objects. *Spanish Applied Linguistics* 3.1-32.
- Sanz, Cristina. 1997. Experimental tasks in SLA research: amount of production, modality, memory, and production processes. *Contemporary perspectives on the acquisition of Spanish*, vol. 2: production, processing, and comprehension, ed. by William R. Glass and Ana Teresa Pérez-Leroux, 41-55. Somerville, MA: Cascadilla Press.
- Silva-Corvalán, Carmen. 1981. La función pragmática de la duplicación de pronombres clíticos. *Boletín del Instituto de Filología de la Universidad de Chile* 31.561-70.
- Silva-Corvalán, Carmen. 1984. Semantic and pragmatic factors in syntactic change. *Historical Syntax*, ed. by Jacek Fisiak, 555-74. Berlin: Mouton de Gruyter.
- Slobin, Dan. 1996. From "thought and language" to "thinking for speaking." *Rethinking linguistic relativity*, ed. by John J. Gumperz and Stephen C. Levinson, 70-98. Cambridge, UK: Cambridge University Press.
- Suñer, Margarita, and María Yépez. 1988. Null definite objects in Quitoño. *Linguistic Inquiry* 18.511-19.
- Taylor, John. 1989. *Linguistic categorization: prototypes in linguistic theory*. Oxford: Oxford University Press.
- VanPatten, Bill. 1984. Learners' comprehension of clitic pronouns: more evidence for a word order strategy. *Hispanic Linguistics* 1.57-67.
- VanPatten, Bill. 1987. Classroom and naturalistic language acquisition: a comparison of two case studies in the acquisition of clitic pronouns. *Language and language use: studies in Spanish*, ed. by Terrell A. Morgan, James F. Lee, and Bill VanPatten, 241-62. Landham, MD: University Press of America.
- VanPatten, Bill, and Teresa Cadierno. 1993. Explicit instruction and input processing. *Studies in Second Language Acquisition* 15.225-43.

- VanPatten, Bill, and Cristina Sanz. 1995. From input to output: processing instruction and communicative tasks. *Second Language Acquisition theory and pedagogy*, ed. by Fred R. Eckman, Diane Highland, Peter W. Lee, Jean Mileham, and Rita Rutkowski Weber, 169-85. Mahwah, NJ: Erlbaum.
- Vázquez Rozas, Victoria. 1995. *El complemento indirecto en español*. Santiago de Compostela: Universidad de Santiago.
- Weinert, Regina. 1995. The role of formulaic language in second language acquisition: a review. *Applied Linguistics* 16.181-205.
- Weissenrieder, Maureen. 1990. Variable uses of the direct object marker *a*. *Hispania* 73.223-31.
- White, Lydia. 1989. *Universal grammar and second language acquisition*. Amsterdam: John Benjamins.
- Zyzik, Eve. 2004. *Encoding meaning with polyfunctional forms: the acquisition of clitics in L2 Spanish*. Davis: University of California, Davis dissertation.

Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages

edited by Carol A. Klee and Timothy L. Face

Cascadilla Proceedings Project Somerville, MA 2006

Copyright information

Selected Proceedings of the 7th Conference on the Acquisition
of Spanish and Portuguese as First and Second Languages
© 2006 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 1-57473-409-1 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Zyzik, Eve. 2006. Learners' Overgeneralization of Dative Clitics to Accusative Contexts: Evidence for Prototype Effects in SLA. In *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, ed. Carol A. Klee and Timothy L. Face, 122-134. Somerville, MA: Cascadilla Proceedings Project.

or:

Zyzik, Eve. 2006. Learners' Overgeneralization of Dative Clitics to Accusative Contexts: Evidence for Prototype Effects in SLA. In *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, ed. Carol A. Klee and Timothy L. Face, 122-134. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1280.