The Acquisition of Differential Object Marking in Spanish[♦]

Miguel Rodríguez-Mondoñedo
University of Connecticut - February 2006

Abstract.- In this paper, I show that a constraint-based approach to DOM, such as the OT system proposed by Aissen 2003, leads to the conclusion that the child will entertain a number of grammars different from the target grammar, before acquiring the final ranking of constraints, even under the learnability assumptions of the OT framework (such as those made by Tesar and Smolensky 1998, 2000). The data examined here, from Spanish-speaking children in the CHILDES database, clearly shows that children master Spanish DOM with a performance virtually errorless. This raises doubts regarding the capacity of the OT framework to explain a key aspect of human language, namely, the process of acquisition.

0. Introduction

The main goal of this paper is to study the use of direct object in Spanish-learning children with the longitudinal corpora in CHILDES database (MacWhinney 2002), in order to check the acquisitional predictions made by Aissen's 2003 Optimality Theory (OT) explanation of Differential Object Marking (DOM)—which is expressed in Spanish by the use of dative preposition A with some direct objects. This model lets us to anticipate several types of errors in the use of the A-marker, given the adjustments the child must make to rank the constraints correctly.

Spanish DOM constructions have attracted the attention of several grammarians and linguists in the last two hundred years (arguably, it is one of the most discussed topics in Hispanic linguistics). The syntactic literature dealing with Spanish DOM is huge—see Isenberg 1968, Luján 1978, Demonte 1987, Brugé and Brugger 1996, Torrego 1998, Aissen 2003, Leonetti 2003, just to mention a few works.[1] Nevertheless, the bibliography from an acquisitional

[♦] I would like to thank William Snyder and Diane Lillo-Martin for their many suggestions, which greatly improved this paper. Many thanks also to Andrea Calabrese and Bosook Kang for their advice and support. All remaining errors are mine.

[1] For an extensive review of the literature in different frameworks, see Pensado 1995b. Bossong (1985, 1991) deserves special mention, since he coined the term "Differential Object Marking" and presented a cross-linguistic study of the phenomenon in more than 300 languages for the first time.

point of view is scarce or null—there is, however, some important literature about the effects of bilingualism on DOM; see Silva-Corvalán 1994, Luján 1996, Sánchez 2003, Montrul 2004. To my knowledge, this is the first attempt to describe the acquisition of this phenomenon in Spanish by monolingual children.

The paper is organized in the following way. In Section 1, I will present a brief description of Spanish DOM, formulating a generalization that encodes the vast majority of cases. Section 2 expounds Aissen's 2003 proposal to account for DOM, based on ranked constraints. Section 3 presents the type of acquisitional process that Aissen 2003 leads us to expect, in particular the idea that the child has to undergo a period during which s/he does not know what the right proper ranking of the constraints is, which should be obvious in his/her production. Section 4 presents the analysis of the case from CHILDES database. Section 5 offers the conclusions.

1. Spanish DOM

In general, Spanish Objects that are [+specific] and [+animate] are marked with the preposition A, and other objects are unmarked:[2]

(1) a. Juan mató **a** María [+animate, +specific]
John killed PREP Mary
John killed Mary

b. *Juan mató María
John killed Mary
John killed Mary

(2) a. María quiere **a** un abogado [+animate, +specific]
Mary wants PREP a lawyer
Mary wants a (specific) lawyer

b. María quiere un abogado [+animate, -specific]
Mary wants a lawyer
Mary wants a lawyer (any lawyer)

(3) a. *Juan destruyó a la ciudad [- animate, +specific]
John destroyed PREP the city
John destroyed the city

[2] A note of caution is necessary here. Let me use Farkas' words: "The notion of specificity in linguistics is notoriously non-specific" (2002:19). This paper will not contribute anything to this complex issue. For discussion of the notion of specificity (and other semantic factors involved) in the context of DOM languages see Kliffer 1982, King 1984, Leonetti 2003, 2004, Heusinger and Kaiser 2003, Farkas and Heusinger 2003, among others.

b. Juan destruyó la ciudad
John destroyed the city
John destroyed the city

(4) a. *Juan destruyó a una ciudad [-animate, -specific]
John destroyed PREP a city
John destroyed a city

b. Juan destruyó una ciudad
John destroyed the city
John destroyed the city

The exact semantic conditions that require the accusative to have this preposition A are not totally clear, however. Putting aside some details, from the previous cases we may propose this scenario of distribution for Spanish DOM structures—along the lines of Heusinger and Kaiser (2003: 53), Zagona (2002: 140), among others:

(5)

Marked Object	[+animate] [+specific]
Unmarked Object	[+animate] [-specific]
	[-animate] [+specific]
	[-animate] [-specific]

There are several problems with the generalization in (5). First, a negative quantifier like ***nadie*** (nobody) receives a mandatory A but it is [-specific]:[3]

[3] Notice that we cannot entertain the idea that [a nadie] is some kind of idiom, or that ***nadie*** is somehow [+specific]. Under existential sentences (subject to the Definiteness Effect), where ***nadie*** is also an object (Suñer 1982, and others), this element cannot receive the preposition A (for discussion on Spanish existential constructions and their relation with DOM, see Leonetti 2003, 2004 and Rodríguez-Mondoñedo 2006a):

(i) No hay (*a) nadie en el parque.
Not there-is PREP nobody in the park
There is nobody in the park

(6) No mató *(a) nadie
NEG killed PREP nobody
He didn't kill anybody

Secondly, sometimes, the preposition A seems to occur just to avoid confusion about which nominal is the object (given the relatively free word-order in Spanish) if the subject is also [-animate]:[4]

(7) La calma sigue a la tormenta
The calm follows PREP the storm

Third, in some objects, the preposition seems to be optional (Laca 1995, Aissen 2003). For instance, nonhuman [+animate] nominals can drop the preposition even if they are definite (and, then, [+specific]):

(8) Mató (a) la mosca
Killed PREP the fly
He killed the fly

It is necessary to point out that the optionality here is not so clear. It has been suggested that the factor involved in Spanish DOM is not really referentiality, which is the traditional property of [specific] (see Heusinger 2002 for a discussion with respect to DOM), but topicality (Weissenrieder 1990, Leonetti 2003, 2004) and even telicity (Torrego 1998).[5] Under certain circumstances, for instance, even non animate objects can have A:

(9) a. WITH PROPER NOUNS (SPECIALLY SHIPS)
Vio al *Huáscar*
He-saw PREP +the *Huáscar*
He saw the *Huáscar*

b. UNDER CLITIC DOUBLING (ESPECIALLY IN INFORMAL SPEECH)
Las compramos a las cortinas en Pátzcuaro
Them we-bought PREP the curtains in Pátzcuaro
The curtains, we bought them in Pátzcuaro
[From CHILDES, Montes database (9Jun81), uttered by the mother]

[4] This is a very controversial explanation. Although some researches (Zubizarreta 1994, Zagona 2002: 14) argue that the preposition A is actually a way to mark the object when both arguments are animate or both are inanimate, there are counterexamples:

(i) El huracán causó la tormenta
The hurricane caused the storm

See Weissenrieder 1985 and Martin 1999 for a discussion on this issue.
[5] This is also recognized by Aissen (2003: footnotes 2 and 24)

This could mean that the presence of the preposition is not really optional, but is controlled by a different factor. Since [specific] is a very vague concept, that can have different interpretations (see Farkas 2002 for an exploration of its possibilities), I see no reason why we could not redefine [specific] to include those other factors—in fact Leonetti 2003 attempts precisely this.[6] If this task is completed successfully, the generalization in (5) will hold (other problems notwithstanding), and DOM would not be optional in Spanish. Here, however, I will leave the issue open.

Aissen's 2003 proposal deals with the cases of optionality, as we will see, but the other cases (which are not exhaustive, see Weissenrieder 1985, 1990 and Torrego 1999 for more cases) are problematic for her account too. Here I will assume that these are produced by some intervening factor, and I will ignore them; after all, Aissen 2003 accounts for the vast majority of Spanish direct objects. Notice that this does not affect our goal, since we are trying to check if Aissen's proposal makes the right prediction for the process of DOM acquisition. In the next section, I will discuss her explanation.

2. The mechanism for DOM in OT Syntax

Aissen 2003 presents an explanation for DOM that relies on the harmonic alignment of the two hierarchies or scales involved in the distribution of DOM. Harmonic Alignment (Prince and Smolensky 1993, cited by Aiseen 2003) is an operation to combine two scales in order to obtain a set of constraints. One of the scales must be binary. The operations connect the highest element on the binary scale with all the elements in the other scale, from the highest to the lowest; in addition, it connects the lowest element on the binary scale from the lowest to the highest. For instance, suppose we have two hierarchies S_1 and S_2 with these elements:

(10) a. S_1: M > N
b. S_2: m > n > o

Given that S_1 is binary, we can align S_1 and S_2. The first part of the operation will relate *M* with *m*, *n* and *o* (in this order), producing the harmonic scale (11a); the second part will associate *N* with *o*, *n* and *m* (that is, in the reverse order), producing the harmonic scale (11b):

(11) a. H_M: M/m $\succ$ M/n $\succ$ M/o
b. H_N: N/o $\succ$ N/n $\succ$ N/m

[6] Of course, it is not important if we must change the name "specific" for a more suitable one. See also King 1984 for the idea that all instances of A-objects share the same semantic ground.

The scales in (10) are hierarchies of prominence, with the more prominent element at the top of the hierarchy (symbolized by ">"). (11), on the other hand, contains hierarchies of harmony (or markedness), with the most harmonic (i.e. the less marked) element at the top (symbolized by "≻"). That is, according to (11), if an element has the property *M*, it is more likely for this element to have also the property *m* than the property *o*. Conversely, if an element has the property *N*, it is more likely for this element to have the property *o* than the property *m*.[7]

In the OT framework, Harmonic scales can be translated straightforwardly to constraint alignments by inverting their order, and interpreting them as Avoid constraints (whose ranking is symbolized by ">>"). For (11), the constraint alignments will be:

(12) a. C_M: * M/o **>>** * M/n **>>** * M/m
b. C_N: * N/m **>>** * N/n **>>** * N/o

The constraints in (12) mean that, for instance, elements that are both M/o (that is, *M* and *o*) or N/m (that is, *N* and *m*) will be penalized (this is symbolized by the "*"); in other words, (12) will reward, for instance, the candidates that are M/m or N/o (given that * M/m and * N/o are in the lower position in the ranking)—remember that in the OT framework, all constraints can be violated; however, the violation of higher constraints is considered worse than the violation of the lower ones.

In order to explain DOM, Aissen uses two sets of Prominence scales (from Croft 1988):

(13) RELATIONAL SCALE
Subject > Object

(14) a. ANIMACY SCALE
Human > Animate > Inanimate

[7] Markedness hierarchies are built to indicate the expectation that an element that has the property *M* could also have the property *m*. They are not intended to identify bad or ungrammatical elements. Given (11), if an element has both properties *M* and *m*, it will be less marked (or more harmonic); on the other hand, if and element has both *M* and *o*, it will be more marked (or less harmonic), but it could still be grammatical. Let me provide an example outside of the linguistic field. If we know that an animal has its youngsters fed with milk (=*M*), we don't expect it to have been born from eggs (=*o*); nevertheless, platypus and echidnas (which are both *M* and *o*) do exist; they are "grammatical" but marked. As will be apparent in the following discussion, a markedness hierarchy constitutes the essential mechanism behind DOM, according to Aissen 2003.

b. DEFINITENESS SCALE
Pronoun > Name > Definite > Indefinite Specific > Non Specific

Aissen does not present any definition for these concepts. Notice that in the Animacy Scale "Animate" actually means "Non Human Animate" (otherwise we would oppose "Animate" to "Human"). With respect to the Definiteness Scale, I assume that a phrase like ***the cat*** is "Definite" and also "Specific", and a phrase like ***a cat*** is "Indefinite" but it could be "Specific" or "Non Specific" (see also footnote 2).

After harmonizing (13) with (14a) and with (14b), we obtain these sets of Harmonic scales (respectively):

(15) HARMONIC ANIMACY SCALES

a. Subject/Human ≻ Subject/Animate ≻ Subject/Inanimate

b. Object/Inanimate ≻ Object/Animate ≻ Object/Human

(16) HARMONIC DEFINITENESS SCALES

a. Subject/Pronoun ≻ Subject/Name ≻ Subject/Definite ≻
Subject/Specific ≻ Subject/Non Specific

b. Object/Non Specific ≻ Object/Specific ≻ Object/Definite ≻
Object/Name ≻ Object/Pronoun

The Harmonic scales express the markedness reversal between subject and object: what is marked for objects is unmarked for subjects and vice versa. For instance, we expect an object to be inanimate and non specific, whereas a subject is expected to be a human pronoun. In other words, they show the likelihood for a member of the Relational Scale (13) to be placed high or low in the Animacy Scale (14a) or in the Definiteness Scale (14b): the more prominent an element is in (13), the more we expect it to be prominent with respect to (14) too (and vice versa). In that sense, they are harmonic (or markedness) hierarchies. As we will see, this mechanism is essential to produce DOM, whose fundamental property, according to Aissen, can be stated in this way: "The higher in prominence a direct object, the more likely it is to be overtly case marked" (2003: 436), where "prominence" must be understood as the position of the object with respect to the scales in (14). I will come back to this immediately.

As mentioned, markedness hierarchies can be Avoid constraints, if we reverse their rankings:

(17) ANIMACY CONSTRAINTS
   a. *Subject/Inanimate **>>** *Subject/Animate **>>** *Subject/Human
   b. *Object/Human **>>** *Object/Animate **>>** *Object/Inanimate

(18) DEFINITENESS CONSTRAINTS
   a. *Subject/Non Specific **>>** *Subject/Specific **>>** *Subject/Definite **>>** *Subject/Name **>>** *Subject/Pronoun
   b. *Object/Pronoun **>>** *Object/Name **>>** *Object/Definite **>>** *Object/Specific **>>** *Object/Non Specific

These constraint hierarchies imply that a sentence will be penalized if its subject or its object does not comply with the markedness reversal. Notice that DOM objects are "marked" objects, that is, objects that do not comply with the markedness reversal—for Spanish, [+animate] and [+specific] objects, as we saw in (5). For Aissen, the gist of DOM is that "marked" objects are also morphologically marked. We need to be aware that two senses for *marked* are being used here. In the sense of the markedness reversal, "marked object" means that the characteristics of the object are unexpected (for an object). In the morphological sense, it simply means that there is an overt mark (preposition, case ending, etc). This implies that DOM is an iconic procedure: we mark with an expression what is marked in the content. It is necessary to come up with a way to encode this iconicity.

As the first step to achieve this goal, Aissen (2003: 447) proposes this iconic constraint:

(19) STAR ZERO
   *$Ø_C$: Penalizes the absence of a value for the feature CASE.

It is apparent that (19) by itself cannot deliver what we want. *$ø_C$ is compatible with a situation where all objects have a morphological mark of Case. But we need to induce Case-marking just for the "marked" objects. In other words, it is necessary (i) to connect *$ø_C$ with the Animacy Constraints (17) and the Definiteness Constraints (18), and (ii) to find a way to neutralize *$ø_C$ in the relevant cases.

In order to connect *$ø_C$ with (17) and (18), Aissen proposes to apply the operation of Local Constraint Conjunction (Smolensky 1995). According to this, by conjoining two constraints C_1 and C_2 in a domain D, we obtain the constraint C_1 & C_2, which is violated if both C_1 and C_2 are violated. The local constraint conjunction between *$ø_C$ and the constraints (17) and (18) results in a set of subject and object-oriented constraints that include *$ø_C$ in each point of the scale. Here, I will represent only the object-oriented constraints (that is, the result of conjoining *$ø_C$ with (17b) and (18b), respectively), since they are the relevant ones for DOM—keep in mind, however, that there are languages that

exhibit Differential Subject Marking (DSM), although they are far less common (Aissen 2003:472-474).[8]

(20) a. LOCAL CONJUNCTION OF *øC WITH THE ANIMACY CONSTRAINTS
*Object/Human & *øC **>>**
*Object/Animate & *øC **>>**
*Object/Inanimate & *øC

b. LOCAL CONJUNCTION OF *øC WITH THE DEFINITENESS CONSTRAINTS
*Object/Pronoun & *øC **>>**
*Object/Name & *øC **>>**
*Object/Definite & *øC **>>**
*Object/ Specific & *øC **>>**
*Object/ Non Specific & *øC

[Aissen 2003: 448]

This ranking, however, enforces case marking in all objects. We need a way to neutralize *øC. Aissen proposes to do that by posing an economy constraint to penalize overt case marking:

(21) *STRUCC: Penalizes a value for the morphological category CASE.
[Aissen 2003: 448]

Now *STRUCC can be inserted in any point in the ranking established by (20), and this will obtain the different cutting points that we see in DOM languages. Just to mention an example, in Hebrew, where pronouns, names and definite objects are morphologically marked, *STRUCC will be inserted between the corresponding constraints in (20b), rendering this ranking in the relevant point:

(22) …*Object/Definite & *øC **>>** *STRUCC >> *Object/ Specific & *øC…

Obviously, this system is not yet ready for Spanish DOM, because it can only apply to languages where just one dimension (Animacy or Definiteness) is used for DOM constructions. As discussed, Spanish uses both.

To make the system suitable for two dimensional DOM languages, Aissen proposes to conjoin the Animacy Constraints (17) and the Definiteness Constraint (18). This will give us a partial ranking, in the sense that there will be no way to establish a ranking between some combinations of the

[8] Of course, a question arises immediately regarding why Differential Subject Marking is so rare. Notice that nothing prevents a local conjunction with the subject-oriented constraints. This will also have acquisitional implications. I will come back to this.

constraints. The highest combined constraint will be the combination of the two highest individual constraints, that is:

(23) *Object/Human & Object/Pronoun

But for the second combined highest constraint we have two possibilities: the combination of each one of the individual highest constraints with the second highest individual constraint in the other dimension. There is no way to establish a ranking between these two combined constraints:

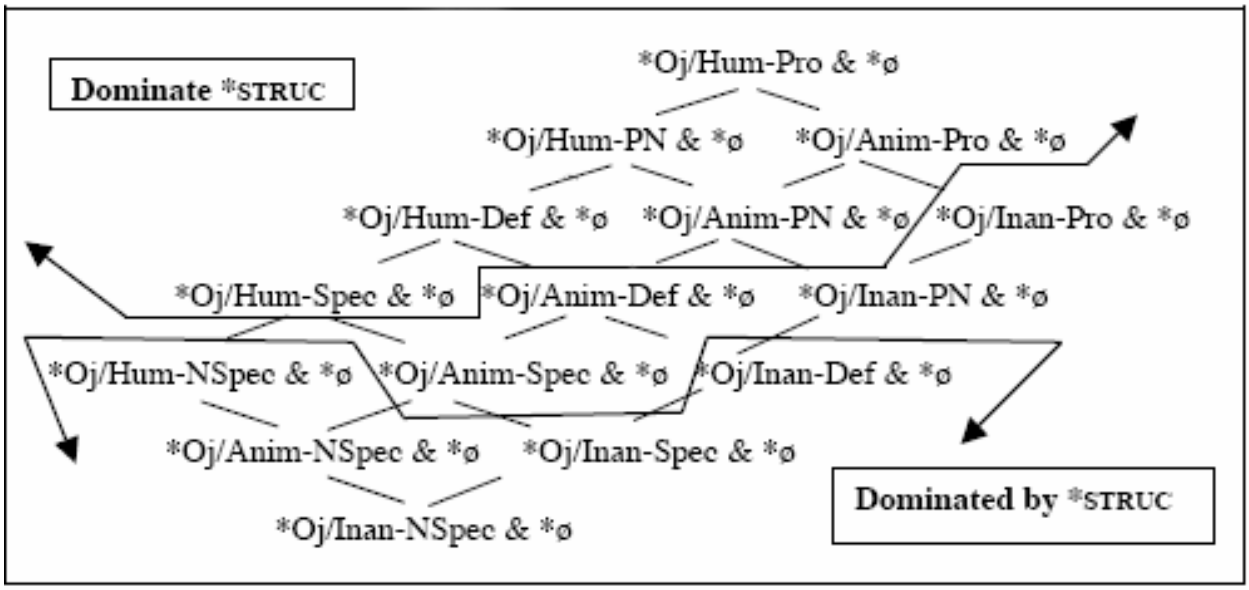
(24) *Object/Human & Object/Name
*Object/Pronoun & Object/Animate

As should be apparent, this situation extends to the combinations of lower constraints as well, configuring a partially ranked hierarchy. Although some sets of constraints are ordered, inside those sets the constraints cannot be ordered. As in the one dimensional cases, these partially ranked constraints need to be conjoined with *øC, obtaining constraints of this form:

(25) [*Object/Human & Object/Pronoun] & *ØC

This configures a huge set of possibilities for a cutting point, that is, the point where *STRUCC can be inserted. An additional complication arises: according to Aissen 2003, in some constructions DOM is optional, then, *STRUCC can be reranked at will, at least in some portion of the ranking. This is her proposed hierarchy for Modern Spanish:

(26) DOM in Modern Spanish[9]



[Adapted from Aissen 2003: 472]

According to (26), there are three situations that are relevant for the application of DOM in Spanish. First, the marking is mandatory in the cases where the constraints always dominate *STRUC$_C$; in these cases, the object has one of these characteristics:

(27) MANDATORY MARKING
- a. Human Pronoun
- b. Human Proper Noun
- c. Animate Pronoun (non-human)
- d. Human Definite
- e. Animate Proper Noun (non-human)
- f. Human Specific (non-definite)

Second, the marking should be optional when *STRUC$_C$ and the constraints can be reranked; in these cases, the object can be:

(28) OPTIONAL MARKING
- a. Inanimate Pronoun
- b. Animate Definite (non-human)
- c. Inanimate Proper Noun
- d. Animate Specific (non-human, non-definite)

[9] In this figure, a constraint like:
i. [*Object/Human & Object/Pronoun] & *ø$_C$
is abbreviated in this way:
ii. *Oj/Human-Pro & *ø$_C$

Finally, the marking should be impossible if *STRUC$_C$ always dominates the constraints; in these cases, the object can be:

(29) MANDATORY NON-MARKING
a. Human Non-Specific
b. Inanimate Definite
c. Animate Non-Specific
d. Inanimate Specific (non-definite)
e. Inanimate Non-Specific

As mentioned in the previous section, this explanation makes some wrong empirical predictions. For instance, the Human Non-Specific quantifier ***nadie*** ("nobody") should appear without the preposition A, but this is actually obligatory (here I repeat (6) as (30)):

(30) No mató *(a) nadie
NEG killed PREP nobody
He didn't kill anybody

Although this system is not without problems, both empirical and theoretical (see Carnie and Jelinek 2003, Haspelmath 2004, Carnie 2004, 2005), it is strong enough to deserve the considerable attention it has obtained, since it does predict the behavior of DOM in the vast majority of the cases. [10] I will put aside these problems, and in the next section I will concentrate in the acquisitional predictions that this system makes with respect to the acquisition of DOM in Spanish.[11]

---

[10] Aissen acknowledges that (26) "oversimplifies in some respects" (2003: 471). This situation is not limited to Spanish DOM. The situation in Hindi DOM is similarly problematic, as Aissen also recognizes: "The facts about animates [in Hindi] have been less discussed in the literature, so the treatment [of Hindi DOM] is somewhat arbitrary" (2003: 467). Nevertheless, she thinks that in Hindi and Spanish the facts "are consistent with the basic hypothesis: if overt marking is possible with direct objects with property α, then it is possible with direct objects with property β, where β dominates α [in (26)]" (2003: 468). In this paper, I will grant this, and put these problems aside.

[11] For alternative analysis of Spanish DOM constructions see Torrego 1998, which could be considered a standard explanation for the phenomenon in a derivational syntax. According to Torrego 1998, an accusative a-DP (a Direct Object marked with A) undergoes overt raising to the Specifier of *v*P. She assumes that *v* has a D feature that attracts the marked DO, and that the dative preposition A has also a D feature to check with *v*. Furthermore, she proposes that marked DOs have an inherent case (on top of the structural one). See Pesetsky and Torrego 2004 and Rodríguez-Mondoñedo 2006a, for other accounts related with Torrego 1998. For different views see Isenberg 1968, Luján 1978, Demonte 1987, Zubizarreta 1994, Pensado 1995a,b, Brugé and Brugger 1996, Leonetti 2003, 2004 among many others. It is fair to say that every analysis on the market resorts to some idealization of the data (assuming some variation of (5), for instance), to provide a systematic explanation. In addition, for analysis of phenomena related with hierarchies (not necessarily DOM) in a derivational syntax see Isaak 2000, Carnie and Jelinek 2003, Béjar and Rezac 2004, Carnie 2004, Haspelmath 2004, McGinnis 2005, among others.

3. Acquisitional Predictions

3.1 Learnability in OT

A system that depends on the (re)ranking of constraints to express the difference among languages, as OT does, in principle, predicts a potentially huge number of languages. For a N number of constraints, the potential number of languages is N!, which means that a child has N! possible grammars to match the input. If, for instance, in a given domain there are six constraints, this will give the child 720 possible rankings. Given the number of constraints involved in DOM, in principle, we face a huge number of possibilities. Assuming an error-driven learning algorithm (Wexler and Culicover 1980), as is standard in OT learnability theories (Tesar and Smolensky 1998, 2000), the amount of input necessary to come out with the target ranking is potentially huge. As should be apparent, this predicts that the children will make chaotic choices before s/he figures out the right grammar (something that it is indeed very unlikely). In the OT framework, however, there are some means to greatly reduce the number of hypotheses a child could entertain.

First, it is normally assumed that constraints derived from harmonic alignments are in a universal ranking,[12] that is, that, different from simpler constraints, they cannot be rearranged to produce linguistic variation (Aissen 2003: 442). In other words, children cannot rerank (17) and (18). With respect to DOM, this assumption trims down the potential number a rerankings that a child could entertain. In fact, the system outlined here leaves children with only one possibility of reranking: different cutting points in the hierarchies, that is, different positions for *STRUC$_C$.

As mentioned, the key to obtaining cross-linguistic differences is the insertion of *STRUC$_C$ at different points with respect to the rest of the constraints. This still predicts a great amount of variability among languages. Figure (26) alone predicts 16 different types of two-dimensional DOM languages.[13] As Aissen (2003: 461) acknowledges, there is no evidence that all these points are actually used for languages, but there is not sufficient evidence for the contrary either, given the limited research on most DOM languages.[14]

[12] Although this fact could be derived from the universal ranking of the primitive hierarchies (animacy and specificity), it is not clear why these hierarchies (that have a simple form) cannot be source of linguistic variation. In other words, these hierarchies have a status different from the normal constraints.

[13] This number could increase, given that, as we saw in previous section, *STRUC$_C$ can be placed in the middle of a constraint. This number could also raise if we add optionality to the picture; as we announced the issue of optionality has been put aside for the purposes of this paper.

[14] In this sense, Spanish is exceptional, since the distributional patterns of DOM have been studied for a long time, at least in the Hispanic tradition (see Pensado 1995b for a survey).

In addition, since the Animacy Scale and the Definiteness Scale do not need to be conjoined, the insertion of *STRUC$_C$ in these scales predicts four types of DOM languages with respect to the Animacy Constraints (20a) and six types with respect to the Definiteness Constraints (20b). Interestingly, all these languages are attested, according to Aissen 2003.

There is an additional source for cross-linguistic variation, however. In some languages, *STRUC$_C$ must be ranked "somewhere within the basic category, rather than at its boundary" (Aissen 2003:457). For instance, in Hindi (also a two-dimensional DOM language, like Spanish), an object can get a marker with some animals (elephants, lions), but not with others (peacocks, mice). Obviously, this multiplies the number of possible languages that the system can generate.

The possibilities for cross-linguistic variation can be directly translated to the possibilities children have before acquiring their language. This still leaves a large number of hypotheses (that is, alternative rankings, different positions for *STRUC$_C$) that children could entertain before coming up with the target ranking. Potentially, this could lead to several mistakes (with respect to the adult language).

In the OT framework, however, there is another way to reduce the potential number of errors (that is, wrong hypotheses) that children could make. A central piece of the OT learning algorithm is the notion of Constraint Demotion (CD). Before showing how it works, let me explain some of the notions necessary for CD to function.

According to Tesar and Smolensky (2000: 37), the OT learning algorithm does not work with the whole range of constraints, but with the space of stratified hierarchies, which has this form:

(31) Stratified domination hierarchies
{C1, C2…C3} >> {C4, C5…C6} >> {C7, C8…C9}
[Tesar and Smolensky 2000: 37]

In this figure, {C1, C2…C3} is the first stratum, and its constraints are not ranked with respect to each other, but each of them ranks higher than the constraints in the second stratum {C4, C5…C6}, which in turn rank higher than the constraints in the third stratum {C7, C8…C9}, and so on. This is the space where CD works, by demoting a constraint from a higher stratum to a lower one.

Additionally, we should remember that in the OT framework, the basic notion is that a grammar matches a given *input* to a particular structural description (the *output*). There are, however, several *candidates* that compete to match the input. Each candidate is confronted with all the constraints and they receive a

mark (in the form of an asterisk "*") each time they violate each constraint. Since the constraints are ranked, so are the marks. The higher in ranking the constraint violated is, the less harmonic (that is, the more marked) the candidate is. The *winner* is the more harmonic candidate, that is, the candidate whose marks are outranked by the marks of each of the *losers*—to achieve this, in a loser-winner pair, it is not necessary that *all* loser-marks outrank *all* winner-marks; it is enough that *one* loser-mark outranks *all* the winner-marks:

(32) loser-marks >> winner-marks
[Tesar and Smolensky 2000: 38]

This gives a specific form to the task of the learner in an OT learnability theory: given a number of loser/winner pairs, the learner must rank the constraints in such a way that at least one of the marks of each loser outranks all the marks of the winner. The tool to achieve this task, according to Tesar and Smolensky 1998, 2000 is CD:

(33) Principle of Constraint Demotion (CD)
For any constraint C assessing an uncancelled winner mark, if C is not dominated by a constraint assessing an uncancelled loser mark, demote C to immediately below the highest ranking constraint assessing an uncancelled loser mark.[15] [Tesar and Smolensky 2000: 36-34]

Let me explain this definition through an example. In order for CD to work, the learner must start with an initial stratified hierarchy (something like (31)). No special initial hierarchy needs to be assumed (Tesar and Smolensky 2000: 46), as long we assume CD. For illustration, let's suppose an initial hierarchy with one stratum that contains four constraints:

(34) {C1, C2, C3, C4}

Remember that inside a stratum the constraints are not ranked, so in (34)and (35) we have an unordered set of constraints. Now let's assume that we have a pair of competing structural descriptions A and B with these marks with respect to the constraints in (34):[16]

[15] For the sake of brevity, I will ignore here the notion of "(un)cancelled mark," which refers to marks that are common (or not common) to a pair of candidates. The algorithm cancels all marks for a given constraint if both members of a pair are marked with respect to such constraint. All surviving marks are uncancelled. See Tesar and Smolensky (2000: 25-28 and 35-36) for discussion about how this affects the learning algorithm.

[16] Notice that this is not the typical OT table (where the winner is picked by choosing the structural description that violates the lower constraints), but just a table with unordered constraints and how the different structural description A and B behave with respect to the constraints. The winner will be the structural description that matches the data.

(35)

	C1	C2	C3	C4
A		*		
B			*	

If A were the winner, CD must demote C2 to the second stratum, in fact, it must create the second stratum:

(36) {C1, C3, C4} >> {C2}

On the other hand, if B were the winner, CD needs to demote C3 to create the second stratum:

(37) {C1, C2, C4} >> {C3}

The process continues examining winner/loser pairs until a total ranking is achieved. Since the learning algorithm does not work with the space of total rankings, this greatly reduces the amount of data necessary to reach the target grammar. Given a number of constraints N, the maximum amount of data necessary to achieve total ranking is N(N-1)/2 (Tesar and Smolensky 2000: 44), which represents a huge reduction from the possible N! of total rankings.[17]

As discussed, the data necessary to achieve the total ranking represents the maximum number of mistakes a child can make before reaching the total ranking. In the next subsection, I will discuss how the two means to reduce the number of possible mistakes that we have presented here (innate hierarchies and CD) combine together to present a possible picture of the acquisition of DOM in Spanish from an OT framework. We will see that, even if the number of possible mistakes is reduced to the minimum, the system still predicts that the child will undergo a state where his or her hypothetical grammar induces him or her to make some mistakes.

3.2 OT Learnability and DOM

The system discussed for DOM so far gives us, generally speaking, the following possibilities:

[17] The proof for this assertion is the following: "[Given that the initial hierarchy] H_0 dominates every hierarchy, […] therefore [it] must dominate H_L. The greatest distance will be when H_L is a totally ranked hierarchy. The furthest constraint from the top stratum will be the one in the bottom stratum, which has offset (N-1). The next lowest constraint has offset (N-2), and so forth. Thus, the h-distance will be: (N-1) + (N-2) +…+1 + 0, which is precisely N(N-1)/2" (slightly simplified from Tesar and Smolensky 2000: 96)

(38) i. The language does not mark any object
ii. The language marks some objects in the Animacy Scale
iii. The language marks some objects in the Definiteness Scale
iv. The language marks some objects in the two-dimensional hierarchy
v. The language marks all objects

To get their particular languages, children have these tools with respect to DOM:[18]

(39) SCALES AND CONSTRAINTS
i. Relational Scale (13)
ii. Animacy Scale (14a)
iii. Definiteness Scale (14b)
iv. Star Zero (19)
v. *STRUC$_C$ (21)

(40) OPERATIONS
i. Harmonic Alignment
ii. Local Conjunction

(41) LEARNING ALGORITHM: Constraint Demotion (CD)

We must assume that, in addiction to CD, the package in (39)-(40) is innate. What is different across language is how the elements in (39)-(40) interact. We have already seen how (38ii-iv) are obtained: in order to get cross-linguistic differences among the objects that are marked, *STRUC$_C$ moves across the other constraints, which were formed by Harmonic Alignment and Local Conjunction.

It is not clear, however, how (38i) and (38v) are obtained, that is, it is not clear how to get languages that mark all objects and languages that do not mark any object. The question is whether in these cases the language makes use of Harmonic Alignment and Local Conjunction. Since *STRUC$_C$ can be placed at the top or at the bottom of the constraints, these operations could still be used, and we could get (38i) and (38v) by moving *STRUC$_C$ to the edges of the ranking (in fact, this is what Aissen seems to have in mind). If this is possible, then (38i) and (38v) are easy to obtain. It is enough to put *STRUC$_C$ at the top (or the bottom) of one ranking; however, this necessarily implies that we need to put *STRUC$_C$ also at the top (or the bottom) of the other one, otherwise, a contradiction will arise. Imagine that we put *STRUC$_C$ at the top of the Animacy Constraints: we are saying that no object should be marked; if in addition we put *STRUC$_C$ at the bottom of the Definiteness Constraints, we would say that

[18] Additionally, we should add the capacity to morphologically mark a given object. For Spanish, this means that children should be able to use A. Snyder and Lillo-Martin 2001 show that A is an oblique-case marker and that it is used very early, which is consistent with the findings in this papers. When children are able to use Spanish DOM, they are already using A.

all objects must be marked, contradicting the previous result. The same situation arises if we try to use Star Zero to cancel just one dimension; if we do not conjoin Star Zero with the Animacy Constraints, we will be saying that no object should be marked, which will contradict us if we try to conjoin Star Zero with the Definiteness Constraints. This is not a problem to obtain (38i) and (38v), that is, we can always obtain languages where all objects are marked or languages where no object is marked—we can do that by moving *STRUC$_C$ to the same edge in both dimensions or by avoiding to conjoin Star Zero in both dimensions. However, this has an important consequence for one dimensional DOM languages. In these cases, where *STRUC$_C$ must run across one of the rankings (precisely the one in which Star Zero must be present), the other cannot be formed; otherwise, the same kind of contradictions will arise. That is, we cannot put *STRUC$_C$ either at the end or at the bottom of the other ranking;[19] nor can we avoid conjoining Star Zero in the other ranking. In other words, we cannot use *STRUC$_C$ or Star Zero to cancel one dimension.

Given the fact that, for one-dimensional DOM languages, we must avoid the formation of the other rankings, it is more economical to assume that in non-DOM languages no ranking is formed in any dimension. To be more precise, Harmonic Alignment and Local Conjunction with respect to (39) happen only in DOM languages—but, of course, they are still possibilities of Universal Grammar (UG), and could be used by non-DOM languages in domains other than (39).

Given this, what children need to decide is the following:[20]

(42) i. Whether they must harmonize the Relational Scale with the Animacy Scale and/or the Definiteness Scales (obtaining Animacy Constraints and/or Definiteness Constraints)
ii. Whether they must use the Object tier (for DOM) or the Subject tier (for DSM) of the Constraints
iii. Whether they must conjoin the Animacy Constraints with the Definiteness Constraints
iv. Whether they must conjoin Star Zero with the Animacy Constraints or with the Definiteness Constraints (or with the conjunction of them if the decision in (42iii) comes up positive)
v. In which position they must place *STRUC$_C$ in the ranking resulting from the decision in (42iv)

[19] Notice that it is irrelevant where we assume *STRUC$_C$ to start in the hierarchy. The point here is not how to move *STRUC$_C$ across one hierarchy, but if we can use *STRUC$_C$ to cancel one dimension.

[20] It is worth noticing that (42i-iv) look more like parameters than like constraints. It is not impossible, of course, to imagine a constraint-based approach to deal with these options. I will leave this issue open.

As should be apparent, (42) still leaves a lot of room for possible mistakes that children can make in the process of acquiring DOM. We can grant that the child will easily find out that he or she is in a DOM or a DSM language; so (42i-ii) are not really a problem, that is, for instance, that we don't expect that the children could entertain the possibility of being DOM when their language is not. A single case of marked object or subject will suffice to establish the relevant tier.[21]

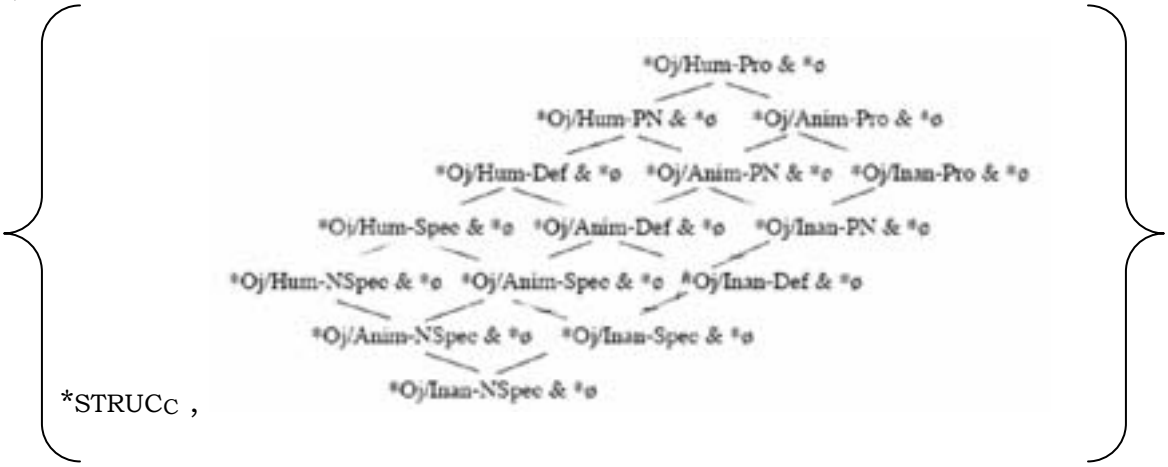
This is more difficult to say about (42iii-iv), that allow the children to decide which dimensions are relevant. Here it is important to stress that conjoining is an operation that works in a local domain, that is, the system does not allow conjunction between unrelated constraints (which would make the possibility of mistakes very huge); in accordance to this, we can grant that the children would entertain only the conjunction of the Animacy and the Definiteness constraint, not any arbitrary constraint. However, a marked object is not by itself evidence that the constraints need to be conjoined; the child needs a set of different marked objects to conclude this. Potentially, this situation could lead to overgeneralizations. For instance, if the child hears the sentence *María vio al perrito* (Mary saw the doggy), this could mean three different things: (i) that animate objects are always marked, or (ii) that definite objects are always marked, or (iii) that objects that are both animate and definite are marked. If, in addition to the previous sentence the child hears *María vio un perrito* (Mary saw a doggy), he or she could disregard (i) but not (ii) or (iii). In order to do so, the child needs to hear, in addition to the previous sentences, something like *María vio la mesa* (Mary saw the table). Notice that these data will help to establish only the type of conjunctions to be made; they do not yet allow us to establish the cutting point for DOM. In addition, it is worth noticing that the choice between one-dimensional DOM or two-dimensional DOM necessarily requires memory of previous data. At the very least, this means that the learning algorithm is not purely data-driven, but it needs some specific assumptions for the process of acquiring a language.

Even if the conjunction is made easily, (42v) implies a major risk: to decide exactly which objects must be marked. Given that the final position of *STRUC$_C$ is essentially unrestricted (remember that it can be even at the middle of a constraint, splitting the number of objects under a single constraint), this predicts a number of potential mistakes.

[21] An interesting question arises here with respect to the possibility that a language could be both DSM and DOM. It is not clear if such languages exist. Although the system outlined here does not exclude this possibility, it presents a serious obstacle for the existence of these languages (and DSM languages in general). The system requires an overt mark of case that could be dropped in certain contexts. Since subject agreement markers are not considered case markers, this leaves aside DOM languages where there is also a split pattern of subject agreement—like, for instance, Persian, a DOM language where Animacy is also relevant to subject agreement (Sedighi 2005). A framework where case and agreement are two sides of the same coin could extend the number of DSM languages. See also section 5.

Remember, however, that we have some means to restrict the number of potential mistakes, as explained in the previous subsection. The first way was to assume that the constraints formed by Harmonic Alignment are universally ranked, that is, that the learning algorithm, that is CD, cannot modify the ranking. Let me draw a general picture of how CD could work for Spanish DOM. I am going to assume that the combined constraint-hierarchy of Animacy and Definiteness has been already established (but see footnote 17), and the only remaining problem is the position of *STRUCC. If so, the initial stratified hierarchy will have this stratum:

(43) Initial Hierarchy



According to (43), the initial stratum contains a subset of constraints that are already ranked, so portions of this hierarchy will have to be demoted to the following strata until the total ranking is achieved (the total ranking was given in(26)), preserving the initial ranking. Another possibility is that (43) does not really contain a ranked hierarchy, but that when the constraints are demoted, they must be ranked in such a way that respects the ranking created by Harmonic Alignment and the Conjunctions. In any case, (43) predicts that the child will demote *STRUCC if, as first input, he or she hears a marked object; this implies that *STRUCC is at this moment at the bottom of the ranking (in the second stratum). Conversely, if, as first input, the child hears an unmarked object, he or she will demote the whole ranked hierarchy, leaving *STRUCC at the top (in the first stratum). Successive demotions will achieve the final ranking (26).

Notice that in this case each demotion creates a possible grammar, that is, a possible cutting point for *STRUCC. This is important because nothing forbids the child to start using the grammar s/he creates each time s/he demotes a constraint. Since not all steps of demoting will match the Spanish ranking, this predicts some possible mistakes. It must be stressed that, given the

combination of CD with a universal ranking, the amount of data necessary to achieve the total ranking is considerably less. There are 16 constraints in (43), 15 in the Constraint Hierarchy and *STRUC$_C$. If CD had to apply to order these independently, 120 data points would be required (see footnote 16). However, the 15 in the Hierarchy have already been ranked using other procedures (Harmonic Alignment and Local Conjunction). This leaves only the ranking of *STRUC$_C$ with respect to the others, requiring only a maximum of 16 data points.[22]

It is important to notice that CD does not require any special initial hierarchy (Tesar and Smolensky 2000: 46-47). So, in principle, CD could begin by placing *STRUC$_C$ at the bottom of the ranked hierarchy. If this is the case, the system predicts errors of commission (that is, misuse of A). Suppose the child places *STRUC$_C$ just over the constraint [*Oj/Anim-NSpec & *Ø], which could happen if the child hears an unmarked object that is animate and non specific, provided that *STRUC$_C$ is at the bottom in the previous step. At this moment, this will imply that all other objects over this constraint must be marked, which is not true of Spanish. On the other hand, if CD places *STRUC$_C$ first at the top, this predicts errors of omission (that is, lack of A). Suppose the child places *STRUC$_C$ right under the constraint [*Oj/Hum-Pro & *Ø], this will imply that all other objects are unmarked, which again is not true of Spanish.

This means that even if we stipulate the initial position of *STRUC$_C$ some errors are still predicted. Suppose that we stipulate the *STRUC$_C$ begins at the top of the hierarchy, and that it moves down to its final position (which is different across languages). The first problem that arises has to do with the fact that, as we discussed, in order to build the two dimensional hierarchy the child needs to know that his/her language is DOM. It would be strange that s/he builds the hierarchy just to cancel its effects (by putting *STRUC$_C$ at the top). Notice that this will imply that the process of learning DOM is not purely data-driven (as we expect in the OT framework), but it needs assumptions that are specific for the process of acquisition. But even if we grant this, accepting that the initial position of *STRUC$_C$ is given, we still predict errors. If *STRUC$_C$ begins at the top and moves down, it could move to a point where it is still too high, predicting that some objects that should be marked in the adult language will be without the preposition (that is, it predicts errors of omission, as mentioned).

Even more, the system is not free of errors of commission without further assumptions; if the child hears an unmarked object s/he could move *STRUC$_C$ to a position that is too low, allowing the marking of some intermediate objects—for instance, the child could hear *Vio un escritorio* “Saw a desk” and move *STRUC$_C$ just over the constraint with Inanimate and Non Specific Objects, which will allow her to incorrectly mark *el escritorio* “the desk”; to avoid this,

[22] See, however, footnote 13.

we must assume that children do not move *STRUC$_C$ when confronted with unmarked objects.

Even if we enrich the learning algorithm with these assumptions, there are more sources of variation. As discussed, *STRUC$_C$ can be reranked, at least with respect to certain constraints, to create instances of optional DOM. Assuming that reranking is a UG option, children need to figure out which constraints allow reranking with respect to *STRUC$_C$. Even if we further assume that the child is conservative, in particular, that s/he waits until encountering the relevant input, the reranking is not hierarchically constrained, in the sense that nothing in the system prevents that *STRUC$_C$ from being reranked with respect to a low constraint but not with respect to a higher one—and in fact this possibility is allowed in (26). Then, in this case, there are no entailments to be derived from an Object in the input with respect to other possible Objects in different places in the hierarchy. Even constraints in the same level can behave differently with respect to the reranking, as (26) also shows. Notice that it is possible that *STRUC$_C$ could be reranked with respect to all constraints, producing a language like Korean, where the accusative marker is truly optional with all kinds of Objects (Kang 2005). So, there is room for some more or less chaotic optionality, which could create both errors of commission and omission.[23] The fact that optionality is allowed also implies that the system needs to have some memory of previous data, which also means that the learning algorithm is not purely data-driven.

In conclusion, Aissen's 2003 explanation for DOM languages, together with standard assumptions for learning in an OT framework, makes a very distinct prediction with respect to the possibilities of DOM acquisition: the child will have to entertain several hypotheses with respect to the right cutting point for DOM, which will lead to detectable errors. In the next section we will find out that such a prediction is not borne out with respect to the acquisition of Spanish DOM.

4. Case-Studies

We will study this phenomenon by reviewing the data from the CHILDES database, which has the record of six native Spanish children: LINAZA, MONTES (Montes 1987, 1992), ORNAT (López Ornat 1994), ROMERO, SERRA, and VILA. This is a longitudinal study (up to the age of three years).

[23] Aissen (2003:464) realizes the potential risk of being too permissive in this respect, and suggests that Stochastic OT maybe could provide means to restrict the reranking. Of course, it could also create new acquisitional challenges. Lee 2002 presents an analysis of Korean in Stochastic OT. Here, we are concerned with classical OT only.

I have selected Objects where the marker must be present or where the marker must be absent. I have disregarded Objects that Aissen's theory predicts to be optional. Since we are looking for possible mistakes that the child can make, optionally marked objects do not allow us to conclude anything (they are always right). The objects that Aissen's Theory predicts to be optional are the following (here I repeat (28) as (44)):

(44) OPTIONAL MARKING
a. Inanimate Pronoun
b. Animate Definite (non-human)
c. Inanimate Proper Noun
d. Animate Specific (non-human, non-definite)

Consequently, I have excluded all object-types in (44) from the counting. In addition, I have only counted in situ objects, that is, objects that are in the frame [V O]. Topicalized objects or wh-objects are not included. I also have excluded the object if the verb is not present (e.g. in answers). I must point out, however, that no mistake has been found associated with these types of object.

All cases studies allow the same conclusion: there are not significant errors in child performance. This is exactly the opposite that we could expect from the system outlined before.

4.1 ORNAT: María

María is a Spanish child, from Madrid, Spain. She was studied by Susana López Ornat (López Ornat 1994) from 1988 to 1991. María belongs to a middle-class professional family. She was videotaped at home during bath, play or feeding interactions with her parents.

In the period observed (from 1;7 to 3;00) María produced a total of 441 Objects. The first Object marked with A was at 1;09.

(45) Table 1: MARÍA
Fisher Exact Test $p < .001$

		A-required		
		No	**Yes**	
A-present	**Yes**	2	24	26
	No	413	2	415
		415	26	441

Only 26 objects were marked with the preposition A; of these, only two are a mistake. 415 objects were not marked with A; in this case, two of them were wrongly unmarked. This makes a total of four mistakes. It is important to considerer that it is not clear to what extent these are mistakes, since there could be an alternative explanation. Although it is possible that they are not mistakes, I will count them as such. Let me explain, however, why they could

be considered mistakes at all (which would mean that Maria makes no mistake at all). I will comment on each mistake.

The first omission is made at 2;03:

(46) LACK OF A
---
*CHI voy a bu(s)ca(r) los gigantes [% en el libro]
I'm going to look for the giants [% in the book]

Here the DP [los gigantes] is human and definite, then, it should have A. However, María's utterance apparently does not have A as a marker. However, notice that the María actually pronounces the verb "buscar" as "buca"; the researcher reconstructed the missing parts (as it is obvious from the parentheses). So this leaves open the possibility that the A could have been reduced. María is pronouncing "voy a buca los gigantes," she could be saying "voy a bu(r)ca(r) (a) los gigantes," with the A. This is a speculation, of course, so I will count this as an error.

The second omission is made at 2;07:

(47) LACK OF A
---
*CHI: Ahí se va miranda Papá
There he goes looking at Dad

Here [Papá] is specific and animate, so it should have A. However, it is possible that it has one. The gerundive "miranda" has an incorrect "a" as ending (it should be "mirando"). So it is possible to assume that this is the marker A that has taken the place of the final vowel of the gerundive.

The two errors of commission are made at 2;05:

(48) MISUSE OF A
---
*CHI: y cojo a tambo(r)
and I get PREP drum

Here [tambor] is not animate, so it should not have A. However, it is evident from the context that "tambor" (drum) refers to a bear doll that has a drum. As a matter of fact, María mentions this in the very next phrase and in this case she uses A correctly:

(49)
---
*CHI: y cojo a tambo(r) y cuando cojo al osito Wily
and I get PREP drum and when I get the little Willy Bear

In fact, some moments later in the same transcript, she mentions Willy Bear as "tambor" again (which is the next mistake):

(50) MISUSE OF A
---
*CHI:
viene la policía a saca(r) a la, al tambor, al osito Wily
the police comes to catch PREP the, PREP +the drum, PREP +the little Willy Bear
---

So it is possible that all of these mistakes are not really mistakes, but since this conclusion is not guarantied, I will consider them as real mistakes.

A Fisher Exact Test over (45) gives us a $p < .001$, which means that this contingency is not random, but the result of a good use of the rule. Maria has an adult-like grammar with respect to DOM. The errors, even in the event that they are errors, are negligible.

4.2 MONTES: Koki

Koki is a Mexican child. She belongs to a middle-class family. The mother is a native speaker of Spanish, but the father is native speaker of English, and he was learning Spanish at the time of the record. Before starting the record, Koki and her family lived in Poland (the first six months), Argentina (the second six months) and USA (two months). At the time of the record she was living in Mexico. In the period observed (from 1;7 to 2;11) Koki produced a total of 334 Objects. The first Object marked with A was at 2;01.

(51) Table 2: KOKI
Fisher Exact Test $p < .001$

		A-required		
		No	**Yes**	
A-present	**Yes**	3	12	15
	No	316	4	320
		319	16	335

Only 15 objects were marked with the preposition A; of these, only three can be considered a mistake. 320 objects were not marked with A; in this case, four of them were wrongly unmarked. This makes a total of 7 mistakes. As in the

previous case, it is not clear to what extent these are mistakes, since there could be an alternative explanation. Again, although it is possible that they are not mistakes, I will count them as such. I will also explain in this case if they could be considered mistakes at all (leaving Koki with no mistake at all).

The first two possible errors of comission were made at 1;11 (notice that this is before the appearance of the first good case of A):

(52) MISUSE OF A
---
*CHI:
Pseudo-phonetic transcription, according to the transcriber (adapted):
amelo al lapi'sito ea ibuj… e ibujar a una o

Written form:
Dámelo al lapicito, voy a dibuj… voy a dibujar a una o
Give-me-CL A+the little pencil, I-am going to dr… I-am going to draw A an o
Give the little pencil, I'm going to draw one "o"

This looks like a real mistake, since both objects are [-animate]. We cannot suggest that the first instance is mispronunciation of the article *el* (*the*), since this is not likely with respect to the second instance, because the article there is *una* (*a*, feminine-plural). However, Koki utters this sentence right after her mother uses A with an [-animate] object:

(53)
---
*MOT:
dónde lo dejó a (e)l lapicito ?
Where CL left PREP+(the) little pencil
Where did he leave the little pencil?

As mentioned in the introduction, this is possible in the adult language, with objects that are doubled by a clitic (CL). So, actually, Koki is trying to target an adult-like rule. I will count this as an error, though.

The last error of commission is made at 2;11

(54) MISUSE OF A
---
*CHI: quiero quitar a(l) techo
I-want to remove PREP +the roof
---

(54) looks like a real mistake. It is a clear non animate specific object (it is similar to (52)). However, we must point out that the kid only pronounces [a], and the transcriber reconstruct [al]. The adult-like form should be [el] in this case. It is worth noticing that [a] is sometimes, at least in earlier stages, a variant of [el], for this child—but , again, the second mistake in (52) weakens this speculation. As in the other cases, I consider this an error.

The first error or omission is made at 2;01:

(55) LACK OF A
---
*CHI: yo veo (l)a(s) niñas to[24] (l)a ventana
and I see (the) girls through the window

As we can see, the transcriber reconstructs the expression “a” with “(l)a(s)”, that is, s/he interprets this “a” as a part of the Spanish article “las” (which is the plural feminine form of *the*). It is not wrong to proceed in this way, since this is something Koki does several times, in fact, the very next article in the same sentence is “(l)a”, with the reduced form. However, if the phonetic form of the article is now [a], then this opens the possibility of a reduction with the preposition A. Additionally, this sentence is actually a repetition of a sentence the mother has just stated; this sentence contains a mistake (also pointed out by the transcriber, so it is not an error of transcribing), which switches the preposition A and the verb:

(56)
---
*MOT: y yo a veo las niñas por la ventana .
and I PREP see the girls through the window

So we can speculate that the kid is just mimicking the relevant part of the mother’s sentence, or, even more interestingly, that she is trying to correct the mother. We will count this as an error, though, as explained.

The next error or omission is made at 2;06.

(57) LACK OF A
---
*CHI: quiero la niña que está diciendo a su mamá
I-want the girl that is saying to her mother
---

[24] The expression “to” is an unrelated mistake; the kid target is the Spanish preposition “por”.

As the transcriber notices (by marking it with a "[?]"), this is very odd sentence. Not only it is completely out of context with respect to the rest of the conversation (the transcriber points out that it may refer to something in a tape they listened to before, presumably Koki's tape), but it is not clear what she is referring to. One possibility is to interpret [la niña que está diciendo a su mamá] simply as "the tape", that is, as an inanimate object.[25] After a few other interventions, Koki repeats her request (although incompletely), which I count as the next omission:

(58) LACK OF A
---
*CHI: quiero la niña que que…
I-want the girl that that…

The same considerations apply here. As in the other cases, I will count these two cases as errors. The last omission is:

(59) LACK OF A
---
*CHI: todavía no le pongas los niños
yet not it put the kids
Don't put the kids yet

When stating (59), Koki is building a small toy-house, and she is putting some plastic pieces as "kids" (and also parents), so the objects she refers to are actually non-animate (actually, (59) could be acceptable in the adult language under this interpretation).

Despite counting these mistakes, the Fisher Exact also gives us a $p < .001$, which again means that this contingency is not random, but the result of a good use of the rule. Koki has an adult-like grammar with respect to DOM. The errors are negligible.

4.3 LINAZA: Juan

Juan is a boy from Spain. The files mainly contain recordings made since the boy was two years old (although there is one file from 0;9 years). In the period

[25] So Koki is saying that she wants to see the tape that contains the girl that is talking to her mother. As a matter of fact, if an adult were to state this sentence with this interpretation, s/he would not use A.

observed (from 0;09 to 2;09), Juan produced a total of 46 Objects. The first Object marked with A was at 2;04.

(60) Table 3: JUAN
Fisher Exact Test $p < .001$

		A-required		
		No	**Yes**	
A-present	**Yes**	2	8	10
	No	36	0	36
		38	8	46

Juan produces only two mistakes, both at 2;04:

(61) LACK OF A
---
*CHI: sí tú tú pega Chacha
yes you you hit Chacha
---
(62) LACK OF A
---
*CHI: puja puja Chacha
push push Chacha

In both cases, there is a proper noun that does not receive the preposition. As in similar cases with other children, we can also speculate that, since the verb ends in [a], the child is suppressing the A for phonetic reasons. In fact, right after stating (61), the child corrects himself, emphasizing the A:

(63)
---
*CHI: tú pega a Chacha
you hit PREP Chacha

As in the other cases, we count both of these omissions as mistakes.

In this case too, the Fisher Exact also gives us a $p < .001$, that is this contingency is not random. Juan has an adult-like grammar with respect to DOM. The errors are negligible.

4.4 VILA: Emilio

Emilio is a Spanish-speaking boy in Catalunya. In the period observed (from 0;11 to 2;11) Emilio produced 169 objects. There are four mistakes in total (one is an omission; three more are misuses of A). Although we do get a significant

result (p < .046), the proportion of errors is considerably bigger than in the other cases. In fact, there are only two objects marked with A, and one of them is a mistake.

(64) Table 4: EMILIO
Fisher Exact Test p < .046

		A-required		
		No	**Yes**	
A-present	**Yes**	1	1	2
	No	164	3	167
		165	4	169

There might be, however, a very good reason for this behavior. Emilio lives in a Catalan-speaking community. In Catalan, no object receives special marking. It has been observed already that Spanish-speaking children in bilingual communities (when one of the languages does not have DOM) drop A occasionally (Silva-Corvalán 1994, Luján 1996, Montrul 2004 present data with respect to Spanish-speaking children in English-speaking communities). I want to suggest that this accounts for the lack of marked objects in Emilio's data. This conclusion is reinforced by some mistakes Emilio makes in a type of construction that is related with DOM, namely, the existential construction (for a discussion regarding DOM and Existentials in Spanish see Rodríguez-Mondoñedo 2006a). For instance, at the age of 2;11, Emilio states:

(65)
---
*CHI: sí que hay Nacho
yes that there-is Nacho
It is true that there is Nacho

This sentence is ungrammatical in Spanish.[26] Different from English, the existential verb *hay* (there-is) does not accept proper nouns as internal nominals, not even under list-readings, as the intended reading of (65) seems to be (Suñer 1982, Leonetti 2005, Rodríguez-Mondoñedo 2006b). Instead of *hay*, Spanish must use *estar* in cases like this:

(66) Está/*hay Nacho
is/there-is Nacho
There is Nacho

[26] (65) is not the only instance of [+animate] and [+specific] object under existential *haber* in Emilio's records. There are three other instances. I'm not counting these instances neither as mistakes nor as good objects. As it is obvious, they do not belong to the Table 4 (literally speaking, they are cases of NO-A-present and NO-A-required, that is, "good" objects, which is rather artificial, given that they are not Spanish adult-like language).

In Catalan, however, the use of proper nouns with the existential verb *haver* is grammatical in these cases (Rigau 1994, 1997, Brucart 2002, Leonetti 2005):

(67) Hi ha en Joan
there-is the John
There is John
[Leonetti 2005: (4)]

This suggests that we are dealing here with a case of interference from Catalan. It is worth stressing, however, that even if we accept Emilio's mistakes as legitimate, we still have significance ($p < .046$).

For the sake of completeness, let me offer Emilio's mistakes with respect to DOM:

(68) LACK OF A (at 2;04)
---
*CHI: no, tú dibuja el nene
no you draw the baby
---

(69) LACK OF A (at 2;07)
---
*CHI: quiero ver el Simón
I want to see the Simon
I want to see Simon
---

(70) LACK OF A (at 2;07)
---
*CHI: todos están matando ella
everybody is killing she
Everybody is killing her
---

(71) MISUSE OF A (at 2;08)
---
*CHI: pónmela a esta canción
put-me-CL PREP this song
Put this song.
---

(68)-(70) are clear cases of lack of A. It is worth noticing that in (69) there is an additional mistake: the proper noun has an article, which is not grammatical in Spanish, but it is possible in Catalan. This confirms the interference from Catalan. (71) contains a use of A with a non animate object; however, it is an

instance of clitic doubling, which, as discussed, sometimes licenses A in the adult language. We will count all of these as errors, as in the other cases.

Beside the fact that Emilio does make mistakes, the Fisher Exact also gives us a $p < .046$, that is, a significant correlation, which implies that Emilio too has an adult-like grammar with respect to DOM.

4.5 ROMERO (Alfonso) and SERRA/SOLE (Eduardo)

I have analyzed the records of two other children: Alfonso (from the Romero database) and Eduardo (from the Serra/Solé database). No mistakes were found in those recordings. It is necessary to point out, however, that the records in these cases are considerably smaller than in the others. Consequently, the number of objects is smaller too. In both cases, there is no instance of DOM objects. Alfonso presents two unmarked objects (by the age of 2;00) and Eduardo has only one unmarked object (by the age of 3;01).

5. Conclusions and suggestions for further research

A constraint-based approach to DOM, such as the OT system proposed by Aissen 2003, even under the learnability assumptions of the OT framework (such as those made by Tesar and Smolensky 1998, 2000), leads to the conclusion that the child will entertain a number of grammars different from the target grammar, before acquiring the final ranking of constraints. This predicts that children will make mistakes in their performance. The data examined here, from Spanish-speaking children in the CHILDES database, does not support this expectation. It is very clear that children master Spanish DOM virtually without mistake. This raises serious doubts regarding the capacity of the OT framework to explain the process of acquisition. I also have shown that the process of DOM acquisition under the OT framework cannot be entirely data-driven, but some special assumptions (e.g. memory of previous data) are needed to account for the acquisition.

With respect to DOM, the remaining question is, of course, what could be the reason why young children master this phenomenon so well, which is especially surprising if we take in consideration, as discussed in the first section, that the explicit rules for DOM are yet to be properly stated. Although the scope of this paper does not allow me to fully discuss an alternative, I would like to sketch a line of reasoning whose final details I will leave for future research.

One important problem of Aissen's proposal is that in her system DOM languages are somehow special with respect to the others. As we discussed in section 4.2, it is more economical to assume that in languages without DOM all the operations needed in DOM languages do not take place. This is consistent

with a narrow view of DOM, according to which it is a phenomenon that takes place in a subset of languages. However, all the machinery needed for DOM has to be considered universal. It is a surprise that this machinery has to be suspended by Universal Grammar in non DOM languages.

Here I want to propose that DOM is a consequence of a broader phenomenon, namely, the fact that Case and Agreement are two sides of the same coin. This is an old but controversial idea, and I won't even attempt to justify it in this paper (see Chomsky 2000, 2001a,b for some discussion from a theoretical perspective). Let me assume it. Let me further assume that universally there are two different positions that license the Direct Object (see Torrego 1998, Rodríguez-Mondoñedo 2006a for some discussion with respect to Spanish DOM, see also Carnie 2005 for a more general implementation).[27] If this is correct, this will mean that all languages have an object split; what is different is the types of object that are licensed in each position and the ways that languages use to mark this licensing. For Aissen 2003, DOM is limited to languages where one type of objects is marked and other objects are unmarked. This excludes, for instance, Finnish, where there are two morphological markers for the objects, but still a semantic difference between the objects correlates with the markers. It also excludes languages like German where all objects receive the same morphological marker, but there is a semantic split between the objects that correlates with their position in the sentence (a phenomenon usually called object-shift). And of course, it excludes English, where all objects are morphologically unmarked. I would like to propose a system where all these possibilities are instances of the same phenomenon.

With respect to the marker, the picture that emerges from these considerations is the following:

(72)

	DO Position 1	DO Position 2
Finnish	A	B
German	A	A
Spanish, Turkish	A	Ø
English	Ø	Ø

In other words, there are two universal positions for the Direct Object. Finnish marks each of them with a different marker, German with the same one, Spanish and Turkish only mark one position, and English does not mark any of them. Since we are assuming that objects must agree with the head that

[27] I will leave open the possibility that these positions could be both base-generated, or that one of them could be the result of a Last Resort operation to license the Case of an object when the other position is somehow defective or incomplete—see Rodríguez-Mondoñedo (2004, 2006a) for some suggestions for the latter with respect to Spanish DOM.

licenses their case, we can conclude that this head is responsible for the semantic split that we see behind DOM; that is, the difference in the cutting point for DOM, the exact types of object that are marked or unmarked, depends exclusively on the lexical properties of the licensing head. For instance, in Spanish, the head for DO Position 1 licenses only specific and animate objects; in Turkish, the same head licenses only specific objects. If this is correct, it means that the syntax for all DOM languages is basically the same. Notice that this greatly improves the child's ability to acquire DOM, since it implies that actually all languages are DOM; the only thing that the child must learn is how to mark these positions.

An additional advantage of this perspective is that we can propose that the low number of languages with DSM is just apparent. Although it is true that there are few languages that mark one type of subjects and leave the other subjects unmarked, there are an important number of languages that have split pattern of agreement with the subject. Under the assumption that Case and Agreement are two sides of the same coin, these languages will be DSM too. In fact, these suggestion will allow us to consider quirky subjects as an instance of DSM; which is not allowed in Aissen's system, since languages with quirky subjects (for instance Icelandic) have a morphological marker for quirky subjects (usually the dative marker) and a different morphological marker for the other subjects (nominative).[28]

I am aware that these speculations raise a number of questions that I am not addressing here. An important one is related with the way the split is achieved in DOM languages. I would like to suggest that, since this split is semantic, we should bring the acquisition of semantics into the picture. This will give us a complete picture of DOM. As mentioned, I leave the implementation of these suggestions for future research.

[28] It is interesting to notice that the Dative marker is also the most used marker in DOM languages (Bossong 1991). In fact, Spanish has also quirky subjects (with verbs like *gustar* "to like"), which, if these considerations are correct, that it is also a DSM language. See also footnote 19.

References

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 213:435-448.

Béjar, Susana, and Rezac, Milan. 2004. Cyclic Agree. Ms.

Bossong, G. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in der neuiranischen Sprachen*. Tübingen: Narr.

Bossong, G. 1991. Differential object marking in Romance and beyond. In *New Analyses in Romance Linguistics, Selected Papers from the XVIII Linguistic Symposium on Romance Languages 1988*, eds. D. Wanner and D. Kibbee, 143-170. Amsterdam: Benjamins.

Brucart, José Maria. 2002. Els determinants. In *Gramàtica del català contemporani*, ed. Jaume Sola. Barcelona: Empuries.

Brugé, Laura, and Brugger, Gerhard. 1996. On the Accusative a in Spanish. *Probus* 8:1-51.

Carnie, Andrew, and Jelinek, Eloise. 2003. Argument hierarchies and the mapping principle. In *Formal Approaches to Function in Grammar. In honor of Eloise Jelinek*, eds. Andrew Carnie, Heidi Harley and MaryAnn Willie, 265–296.

Carnie, Andrew. 2004. Some Remarks on Markedness Hierarchies: A Reply to Aissen 1999 and 2003. Ms., *University of Arizona*.

Carnie, Andrew. 2005. A Phase-geometric approach to multiple marking systems. Martha McGinnis, Norvin Richards (eds.) Perspectives on Phases. MIT Working Papers in Linguistics.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, eds. Roger Martin, David Michaels and Juan Uriagereka, 89-155. Cambridge, Mass: MIT Press.

Chomsky, Noam. 2001a. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 1- 52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001b. Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics* 20.

Croft, William. 1988. Agreement vs. Case Marking and Direct Objects. In *Agreement in Natural Language: Approaches, Theories, Descriptions*, eds. M. Barlow and C. Fergunson, 159-179. Stanford: CSLI.

Demonte, Violeta. 1987. C-command, prepositions, and predications. *Linguistic Inquiry* 18:147-157.

Farkas, Donka. 2002. Specificity Distinctions. *Journal of Semantics* 19:213-243.

Farkas, Donka, and Heusinger, Klaus von. 2003. Referential stability and differential case marking in Romanian. Ms.

Haspelmat, Martin. 2004. Against markedness (and what to replace it with). Ms., *Max-Planck-Institut für evolutionäre Anthropologie*.

Heusinger, Klaus von. 2002. Specificity and definiteness in sentence and discourse structure. *Journal of Semantics* 19:245-274.

Heusinger, Klaus von, and Kaiser, Georg A. 2003. Animacy, Specificity, and Definiteness in Spanish. In *Proceedings of the Workshop Semantic and Syntactic Aspects of Specificity in Romance Languages*, eds. Klaus von Heusinger and Georg A. Kaiser, 41-65. Fachbereich Sprachwissenschaf: Universität Konstanz.

Isaak, André. 2000. Split Case Marking and Prominence Relations, University of Massachusetts, Amherst. Ph.D. Dissertation.

Isenberg, Horst. 1968. *Das Direkte Objekt im Spanischen*: Studia Grammatica 9. Berlin: Akademie-Verlag.

Kang, Bosook. 2005. Acquisition of language particular properties under impoverished input, University of Connecticut: Ph.D. Dissertation.

King, Larry D. 1984. The Semantics of Direct Object "A" in Spanish. *Hispania* 67:397-403.

Kliffer, Michael. 1982. Personal a, Kinesis and Individuation. In *Papers from the XII Linguistic Symposium on Romance Languages*, ed. P. et al Baldi, 195-216. Amsterdam: John Benjamins.

Laca, Brenda. 1995. 'Sobre el uso del acusativo preposicional en español', In *El complemento directo preposicional*, ed. Carmen Pensado, 11-59. Madrid: Visor.

Lee, Hanjung. 2002. Referential Accessibility and Stylistic Variation in OT: A Corpus Study. Paper presented at *38th Meeting of the Chicago Linguistic Society*.

Leonetti, Manuel. 2003. Specificity and Object Marking: the Case of Spanish a. In *Proceedings of the Workshop "Semantic and Syntactic Aspects of Specificity in Romance Languages".* eds. Klaus von Heusinger and Georg A. Kaiser, 67-101. Fachbereich Sprachwissenschaft: Universität Konstanz.

Leonetti, Manuel. 2004. Specificity and Differential Object Marking in Spanish. *Catalan Journal of Linguistics* 3:75-114.

Leonetti, Manuel. 2005. Definiteness Effect and the Role of the Coda in Existential Constructions. Ms

López Ornat, Susana. 1994. *La adquisición de la lengua Española*. Madrid: Siglo XXI.

Luján, Marta. 1978. Direct object nouns and the preposition "a" in Spanish. *Texas Linguistic Forum* 10:30-52.

Luján, Marta. 1996. Clitic-doubling and the acquisition of agreement in Spanish. In *Perspectives on Spanish linguistics*, eds. Javier Gutiérrez-Rexach and L. Silva Villar, 119–138. Los Angeles: UCLA, Department of Linguistics.

MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum.

Martin, Juan. 1999. The Syntax and Semantics of Spanish Accusative a. In *Advances in Hispanic Linguistics: Papers from the 2nd Hispanic Linguistics Symposium*, eds. Javier Gutiérrez-Rexach and Fernando Martínez-Gil, 469-485: Cascadilla Press.

McGinnis, Martha. 2005. On markedness asymmetries in person and number. *Language* 81:699-718.

Montes, Rosa G. 1992. *Achieving understanding: Repair mechanisms in mother–child conversations*. Unpublished doctoral dissertation, Georgetown University.

Montes, Rosa. 1987. *Secuencias de clarificación en conversaciones con niños (Morphe 3-4)*. Universidad Autónoma de Puebla.

Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition* 7:125–142.

Pensado, Carmen ed. 1995a. *El complemento directo preposicional*. Madrid: Visor.

Pensado, Carmen. 1995b. El complemento directo preposicional: estado de la cuestión y bibliografía comentada. In *El complemento directo preposicional*, ed. Carmen Pensado, 11-59. Madrid: Visor.

Pesetsky, David, and Torrego, Esther. 2004. Tense, Case, and the Nature of Syntactic Categories. In *The Syntax of Time*, eds. Jacqueline Guéron and Jacqueline Lecarme, 495-539. Cambridge, MA.: MIT Press.

Prince, Allan and Smolensky, Paul. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report 2, Center for Cognitive Science, Rutgers University. To be published by MIT Press, Cambridge, MA.

Rigau, Gemma. 1994. Catalan Presentational Sentences and the properties of AGR. In *Paths Towards Universal Grammar. Studies in Honor of Richard S. Kayne.*, eds. Guglielmo Cinque, Jan Koster, J. Y. Pollock, Luigi Rizzi and Rafaella Zanuttini, 343-360. Washington D.C.: Georgetown University Press.

Rigau, Gemma. 1997. Locative Sentences and Related Constructions in Catalan: ésser/haver Alternation. In *Theoretical Issues at the Morphology-Syntax Interface*, eds. Maria Uribe-Etxebarria and Amaya Mendikoetxea, 395-421. Bilbao/Donostia: Universidad del País Vasco.

Rodríguez-Mondoñedo, Miguel. 2004. Differential Object Marking in Double Object Constructions. Ms. University of Connecticut

Rodríguez-Mondoñedo, Miguel. 2006a. Spanish Existentials and other Accusative Constructions. In *Minimalist Essays*, ed. Cedric Boeckx. Amsterdam: John Benjamins. 326–394

Rodríguez-Mondoñedo, Miguel. 2006b. A [person] restriction on the Definiteness Effect in Spanish. Paper presented at NELS 37. University of Illinois at Urbana-Champaign.

Sánchez, Liliana. 2003. *Quechua-Spanish bilingualism: interference and convergence in functional categories*. Amsterdam: Johns Benjamins.

Sedighi, Anousha. 2005. Subject-predicate Agreement Restrictions in Persian, University of Ottawa: PhD Dissertation.

Silva-Corvalán, Carmen. 1994. *Language contact and change*. Oxford: Oxford University Press.

Smolensky, P. 1995. On the internal Structure of the constraint component Con of UG. ROA-86-000. http://roa.rutgers.edu

Snyder, William, and Lillo-Martin, Diane. 2001. On the acquisition of Spanish goal PP's. Paper presented at *4th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, University of Illinois at Urbana-Champaign

Suñer, Margarita. 1982. *Syntax and Semantics of Presentational Sentence-Types*. Washington: Georgetown University Press.

Tesar, Bruce, and Smolensky, Paul. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29:229-268.

Tesar, Bruce, and Smolensky, Paul. 2000. *Learnability in Optimality Theory*. Cambridge: MIT Press.

Torrego, Esther. 1998. *The dependencies of objects*: Linguistic inquiry monographs 34. Cambridge, Mass.: MIT Press.

Torrego, Esther. 1999. El complemento directo preposicional. In *Gramática Descriptiva de la Lengua Española*, eds. Ignacio Bosque and Violeta Demonte, 1779-1806. Madrid: Espasa.

Weissenrieder, Maureen. 1985. Exceptional uses of the Accusative a. *Hispania* 68:393-398.

Weissenrieder, Maureen. 1990. Variable Uses of the Direct-Object Marker a. *Hispania* 73:223-231.

Wexler, Ken, and Culicover, Peter. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Zagona, Karen T. 2002. *The Syntax of Spanish*. Cambridge: Cambridge University Press.

Zubizarreta, Ana María. 1994. Word Order in Spanish and the Nature of Nominative Case. In *Beyond Principles and Parameters. Essays in memory of O. Jaeggli*, eds. K. Johnson and I. Roberts, 223-250. Dordrecht: Kluwer Academic Press.

02/2006

Miguel Rodríguez-Mondoñedo
University of Connecticut
miguelmrm AT gmail.com