Chapter 3

Theoretical Background II (Pro-drop in Acquisition)

3.0 Introduction

The goals of this chapter are (1) to introduce the major issues involved in acquisition research, particularly as these relate to the question of pro-drop, (2) to review relevant second language acquisition empirical studies that have been conducted on the pro-drop parameter, and (3) to review recent attempts to explain second language developmental issues related to pro-drop from an Optimality perspective.

Section 3.1 begins with a discussion of how second language acquisition studies have mirrored developments in linguistic theory more generally. This section will suggest that evolving developments in linguistic theory have been applied to specific, longstanding problems in linguistics, such as pro-drop, making these types of problems barometers of success for a theory. Section 3.2 considers the key acquisitional issue that any grammatical theory must address: the logical problem of language acquisition (Chomsky 1981). Section 3.3 then discusses two concerns related to the logical problem of language acquisition: the subset principle and learnability. The issue of learnability is central to any developmental account, and this dissertation is based on a particular learning algorithm emerging from research in learnability: the Robust Interpretative Parsing/Constraint Demotion Algorithm of Tesar and Smolensky (2000), for which this dissertation will offer further supporting evidence. Section 3.4 takes up two remaining critical assumptions for the argument developed in this dissertation: the accessibility of

Universal Grammar for L2 learners and the role of first language (L1) transfer in second language learning. Section 3.5 then reviews specific studies that have considered these issues through empirical tests of pro-drop from a parameter-setting perspective. These will underscore how the prevailing linguistic theory at the time of the research was empirically tested by developmental accounts. The need for a reconsideration of the pro-drop parameter from a developmental perspective, the focus of this dissertation, will become clear. Section 3.6 reviews the only SLA accounts related to pro-drop from an OT perspective to date: Park (2000) and LaFond, Hayes, and Bhatt (2001). As we will see, Park's account was limited to null arguments in child second language acquisition, and LaFond, Hayes, and Bhatt dealt only with null subjects, one piece of the pro-drop phenomenon. Thus, this chapter sets the stage for remaining questions and a clear direction for further study. These questions and hypotheses, as well as the research design they suggest, are discussed in the next chapter.

3.1 Linguistic theory and SLA

This section situates the questions of SLA and the research agenda of this dissertation within the broader domain of linguistic theorizing. In the modern era, as Braidi (1999) and Mitchell and Myles (1998) have observed, L2-acquisition studies have mirrored developments in linguistic theory more generally. Early (pre-SLA) studies in the 1950s and 1960s were influenced by behavioral psychology and structural linguistics. The influence of Bloomfieldian views of language acquisition as habit formation were applied to language teaching, culminating in the development of the Audio-Lingual

Method, a method that involved drilling language patterns until proper language 'habits' were formed.[1]

Researchers such as Fries (1945) and Lado (1957) claimed that by analyzing contrasts between native and target languages, problem areas for language learners might be predicted. Lado (1957:2) additionally claimed that learners 'transfer' the forms and meanings of their native language to the target language. Subsequent research supported claims for the existence of transfer, but overall, the 'Contrastive-Analysis' approach was shown to be less predictive than hoped.

Corder (1967) argued on the basis of Chomsky's competence/performance distinction that learner errors provide a window into the learner's linguistic knowledge. Like Chomsky, Corder sought to carefully distinguish unsystematic performance problems from the systematic errors that were a natural outcome of the system the learner was using. This approach treated errors not as the result of bad habits, but as clear markers of developmental stages of the language learner. Corder's application of linguistic theory to questions of second language acquisition began a practice that has continued to the present day.

Both Contrastive-Analysis and Error Analysis shared a common assumption that learner grammars are deficient, transitional systems. This view was challenged by a third approach to L2 acquisition that viewed learner grammars as systematic grammars in their own right. Selinker (1972:214) describes these learner grammars as 'a separate linguistic system based on the observable output which results from a learner's attempted production of a TL [target language] norm'. Selinker brought the concept of 'interlanguage' to the fore of SLA studies and highlighted a number of processes that are

involved in the language acquisition puzzle: language transfer, overgeneralization, transfer of training, and L2 learning and communication strategies.

During the two decades following Selinker (1972), SLA researchers increasingly worked within the generative program and moved away from earlier ties to language pedagogy. Whereas earlier there had been an implicit assumption that SLA theories would provide teachers with recipes for successful practice, the radical shift away from practical requirements resulted in SLA developing as an autonomous field of inquiry with distinctive theoretical orientations, methodologies, and goals. As this field separated from language teaching, it began to explore various issues: To what extent does L2 learning mirror that of L1 learning? What role does the first language play in the acquisition of a second? Is access to Universal Grammar still available for the second language learner? What role do psychological, social, and environmental factors play in acquisition? To what degree are L2 learning and processing similar to other more general cognitive processes? Numerous questions have been raised; few definitive answers have been forthcoming.

Much of the early work in SLA attempted to apply research that had been done in L1 acquisition to an L2 context. For example, Brown's (1973) morpheme study inspired investigations into developmental orders for the same grammatical morphemes among L2 learners. Dulay and Burt (1974) focused on child second language acquisition and found very similar acquisitional patterns in the L2 acquisition of English as for L1 acquisition. Bailey, Madden and Krashen (1974) extended this work to adult second language acquisition and found a great number of similarities to child L1 acquisition — those

acquiring English either as a first language or as a second language learn grammatical morphemes in a relatively set order, regardless of instruction.

Current work in SLA continues to explore important interfaces between L1 and L2 acquisition. For example, Brown's (2000) research on Japanese, Korean, and Chinese learners of English as a second language suggests that in infant speech perception there is a direct link between the development of a feature geometry and the decline of perceptual capabilities. Brown posits that the acquisition of phonological structure imposes specific boundaries so that later input filters L2 allophones so that they are perceived in learners' L1 phonemic categories. For example, Chinese speakers in this study differed from both Japanese and Korean speakers in their abilities to discriminate /l/ and /r/. According to Brown, this is expected because the presence of the [coronal] feature in Chinese permits a distinction on this dimension, while the absence of this feature in Korean and Japanese 'funnels' the acoustic signal for these two sounds into a single perceptual category (2000:40).

In another study, Young-Scholten and Archibald (2000), proceed from findings regarding the L1 acquisition of consonant clusters in certain quantity-sensitive Germanic languages to consider L2 interaction between segmental features and syllable structure. Young-Scholten and Archibald found that what is transferred from an L1 is not only the canonical CV structure, but also the complex interaction of the segmental inventory that determines the feature geometry of a segment, influencing what sequences of segments are allowed in the developing L2. Young-Scholten and Archibald also looked at sonority distancing and found that there is greater difficulty involved in acquiring L2 consonant clusters when contrasts in that L2 require an adjustment of the L1 feature inventory.

From the earliest SLA studies (Fries 1945, Lado 1957, Corder 1967, Brown 1973, Dulay and Burt 1974, etc.) to the present (Brown 2000, Young-Scholten and Archibald 2000, Park 2000, LaFond, Hayes, and Bhatt 2001, etc.), SLA studies have attempted to draw insights from L1 theorizing and, together with the specific insights that L2 research brings, draw conclusions about human processes of language acquisition. Each stage of the development of linguistic theory has seen a reinterpretation of earlier results in light of the current state of the theory, applied first to specific, persistent problems in linguistics (e.g. pro-drop), and then to the issue of second language acquisition.

Hence, the linguistic advance of a parameter-setting model (Chomsky 1980) came first, followed by various applications of this model to the pro-drop question (e.g. Rizzi 1982, Jaeggli 1982), and only later it was extended to questions of second language development (White 1985, Lakshmanan 1986, et al.). It was not until linguistic theorizing proposed that discourse issues must be admitted into an understanding of crosslinguistic differences (e.g. Vallduvi 1992, Erteschik-Shir 1993), that these concerns became incorporated into SLA research on pro-drop (e.g. Liceras and Díaz 1995, Pérez-Leroux and Glass 1997). The advent of Minimalism (Chomsky 1995) yielded another new theory of grammar that was applied first to pro-drop (Speas 1994, Radford 1997) and only later was followed by the predictable extension to an SLA context (Park 2000).

Given this pattern, it is natural to expect that the evolution of Optimality Theory (Prince and Smolensky 1993) should see the extension of this grammatical theory to the persistent question of pro-drop (e.g. Grimshaw and Samek-Lodovici 1995, Speas 1997) and then to the issue of pro-drop within SLA. But this final step of the progression has not yet been made for adult second language acquisition of syntax, with the exception of

LaFond, Hayes, and Bhatt (2001), and this paper addresses only the acquisition of null subjects in SLA. As is clear, pro-drop has been viewed by linguistic researchers as one of those persistent problems that a linguistic theory must be able to address, making it something of a barometer of the successfulness of a linguistic theory. This further confirms the need for a developmental account of pro-drop from an OT perspective.

Creating such an account of pro-drop may also address the need for an analysis that can explain important interfaces between discourse and syntax. Early SLA research followed the lead of the Chomskyan 'revolution' by considering how second language learners acquire parameters thought to be responsible for their competence. But as Chafe (1994) suggests, the initial Chomskyan approach was less 'revolutionary' than commonly imagined. According to Chomsky's early formulation (1957:13), language is

> ...a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements...each language has a finite number of phonemes (or letters in its alphabet) and each sentence is representable as a finite sequence of these phonemes (or letters), though there are infinitely many sentences.

Although this approach was quite different than behaviorism, it still uncritically assumed that language was observable only via its forms, and that context specific meanings (i.e. pragmatics, discourse, social setting) were of little interest to any scientist attempting to explain core human competencies regarding language.

Recent developments in linguistic theory during the last decade suggest something far more revolutionary—that we need to recognize that language acquisition involves the complex functioning of a multifaceted, mental system. This means that a theory of grammar must, at least potentially, be able to both address interactions (and competitions) between various levels of language knowledge (e.g. conflicts between syntax and discourse) and to integrate questions of language learning and learnability into

a broader field of linguistic theorizing. Some theories of language (e.g. OT) are better prepared to meet this challenge than others.

3.2 The logical problem of language acquisition

Chomsky (1986) put forward a set of questions that he believed formed the essential framework for further inquiry into the phenomenon of language. They are questions that still are critical to linguistic theory and language acquisition today:

1. What is the system of knowledge?
2. How does this system of knowledge arise in the mind/brain?
3. How is this knowledge put to use in speech (or writing)?
4. What physical mechanisms serve as the material basis for this system of knowledge and this use of language?

Chomsky's second question relates to what has been labeled the 'logical problem of language acquisition' (Hornstein and Lightfoot 1981). A clear formulation of this problem is found in Crain and Thornton (1998:283):

> Many aspects of grammatical knowledge are represented as constraints, that is as sanctions against linguistic analysis of one kind or another. Constraints are negative statements. It is safe to assume that not all children, perhaps no children, encounter evidence pertaining to constraints. The pertinent evidence would be information about which linguistic expressions and meanings are prohibited in the target language. It follows from the absence of such negative evidence in children's experience that knowledge encoded by constraints is not learned from experience. If not, then this aspect of linguistic competence must be innately specified, as part of Universal Grammar.

This problem is also referred to as the 'poverty-of-stimulus' argument or 'Plato's Problem,' referring to a passage in *The Meno* where Plato records how Socrates led an uneducated child to discover theorems of geometry by only asking questions. Plato pondered how the child found the truths without being given any information,[2]

concluding that this was evidence of an earlier existence, knowledge in the child's mind that was reawakened when stimulated by questions. Chomsky argues (as Leibniz did) that Plato's conclusion was essentially correct if by 'preexistence' we understand instead that certain aspects of our knowledge are innate, part of our genetic/biological endowment.[3]

The theorems of geometry are really quite basic compared to what children must learn about language. While children must be exposed to input before language acquisition takes place, once input is present, they acquire very complex grammatical principles (structure dependency, subjacency, binding, etc.) in a brief period of time, without explicit instruction and on the basis of input that is impoverished, imprecise, and variable. For example, Lakshmanan (1994:3) shows that children consistently choose 3.1b, which applies a computationally complex structure-dependent rule, over 3.1c, which uses a computationally simple structure-independent rule.

(3.1.) a. *The book which is on the table is dull*.

b. *Is the book which is on the table dull?*

c. **Is the book which on the table is dull?*

Somehow, children know that what moves to the front of interrogatives of this type is the verb in the matrix clause, rather than simply moving the first verb in a linear ordering. Children also, without instruction, learn that 'them' in 3.2a can be coreferential with 'the men' in 3.2a, and also that it could refer to some other contextually salient persons, but that in 3.2b 'them' permits only the latter interpretation.

(3.2.) a. *I wonder who the men*$_i$ *expected to see them*$_i$.

b. **The men*$_i$ *expected to see them*$_i$.

How do children know these fine points related to the syntactic ordering of elements, the interpretation of pronouns, and semantics?

Fromkin (2000:472), notes that, in Italian, participle constructions can contain both atelic verbs, as in 3.3a and telic verbs, as in 3.3b.[4]

(3.3.) a. *Gianni ha ballato* (*per un ora*)
'John has danced (for an hour).'

b. *Gianni è caduto* (**per un ora*)
'John fell (*for an hour).'

Without explicit instruction, 2-year-old Italian children restrict their use of the past participle construction to telic predicates such as *cadere* 'fall', *rompere* 'break', *arrivare*, 'arrive'. They do not use this past participle construction with atelic predicates such as *volare* 'fly', and *ballare* 'dance'. How did Italian children acquire the telic/atelic distinction, and why do they undergeneralize at this stage of their acquisition?

In the area of morphology, ever since Brown's (1973) study on the order of acquisition for 14 grammatical morphemes it has been repeatedly demonstrated that children's acquisition follows clear developmental patterns. Mitchell and Myles (1998:27) point out that, regardless of the language being learned, children begin to use negation around the same age and mark the negative in similar ways, by first attaching some marker to the outside of the sentence and only later moving this marker inside the sentence. Since adult grammars differ cross-linguistically in the way negation is handled, why do children learning various languages follow similar developmental paths?

Given all these types of evidence from various areas of children's grammars, we are left with the question of how children come to know these subtle distinctions and nuances of language without being taught. Answers pointing to mimicry of parents, the hearing of simplified input, or explicit correction are all easily repudiated. Arguing that children do not produce certain agrammatical forms because they have not heard them before does not explain why children will produce forms such as 'Joey goed there with me.' The types of developmental errors children make on their path to adult language acquisition are predictable, and are little affected by any overt attempts to instruct them.

This, then, is the logical problem of language acquisition, a problem that generative frameworks specifically attempt to address. For Chomsky, the solution to Plato's problem is simply that the properties of the mind/brain include certain principles of the language faculty — UG (or a Language Acquisition Device, a linguistic 'black box') — that, given a rich enough linguistic environment, is able to determine the value of certain parameters and provide interpretation of linguistic expressions, even those that a child learning a given language has never encountered.

Cook (1991) outlines four stages to the poverty-of-the-stimulus argument. Given a particular aspect of linguistics knowledge, for example pro-drop, the argument would be stated as follows (adapted from Cook 1991:83-84):

Step A:	Native speakers of a language have specific intuitions about certain aspects of the pro-drop.
Step B:	These aspects of grammar could not have been acquired from the language input typically available to children.
Step C:	These aspects of grammar have not been taught.
Step D:	Therefore, these aspects must be built into the mind.

As the stages show, linguistic theory begins by looking at the knowledge of the native speaker. Since the linguistic input the child receives is insufficient to teach this principle, other sources must be sought. Cook and Newson write, 'The poverty-of-stimulus argument is fundamentally simple; whenever you find something that the adult knows which the child cannot acquire, it must already be within the child' (1996:85).

Although the logical problem of language acquisition initially arose from a first language perspective, SLA researchers have argued that it applies to second language contexts as well (White 1985, 1989, Flynn 1987, Cook 1996). As in L1 acquisition, L2 learners acquire the grammatical properties of the target language in the face of insufficiently rich and insufficiently precise input. For example, Spanish learners of English as a second language must learn that English is not a pro-drop language, despite the fact that the input they receive in informal conversation, email, etc., often uses subjectless sentences.[5]

The logical problem of language acquisition has led second language researchers to ask specific questions related to the relationship between L1 and L2 acquisition. They have queried to what extent the L2 learner has access to UG, how the L1 interacts with the acquisition of an L2, and how we might account for the variation in ultimate attainment of a second language. For Chomsky, the differences between L1 and L2 acquisition are greater than the similarities. It is true that while first language acquisition happens naturally and with little apparent effort, learning a second language for adults requires much more effort with uncertain results. Nevertheless, for most of the world, learning second languages is a normal human activity, and Plato's problem is no less

important in second language contexts than for first language. We must somehow account for how a second language learner acquires constructions that could not otherwise be reasonably acquired on the basis of the input alone. This implicates some role for UG also in second language learning. As Flynn and O'Neill (1988:2) claim, SLA legitimately interfaces with linguistic theory in its 'attempt to examine the extent to which a theory of UG could be useful in explaining the L2 acquisition process...' and in its desire to identify 'evidence for the role of UG in L2 acquisition' and the domains in which this evidence may be found.

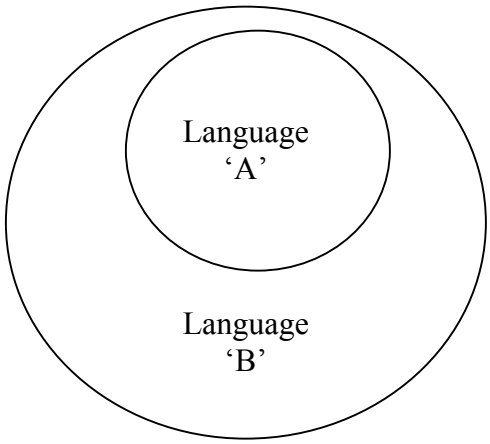
3.3 The Subset Principle and learnability

Since learning a grammar is difficult (at least for adult second language learners), and the input available to language learners is not sufficient to achieve this learning, the logical problem of language acquisition raises two additional important issues: the subset principle and the issue of learnability. These issues involve questions concerning the restriction of the learning space of possible grammars and the search method learners must use to arrive at the target grammar.

One attempt at a comprehensive theory of learnability for L1 acquisition is found in Pinker (1996), who argues that children acquire language 'by exploiting rich formal and substantive constraints on the types of rules that languages may have' (1996:358). Pinker claims that children entertain a small subset of the possible hypotheses consistent with the input they receive, with certain triggering conditions sparking reevaluation of the hypotheses. This reference to the Subset Principle (Angluin 1978, Berwick 1985, Manzini and Wexler 1987) is important also for L2 acquisition.

The subset-superset relationship is one that has played an important role in parameter-setting accounts, which hold that children choose the parameter with the fewest possible assumptions, based on the language input received, shown in Figure 3.1.

Figure 3.1 Subset relationship



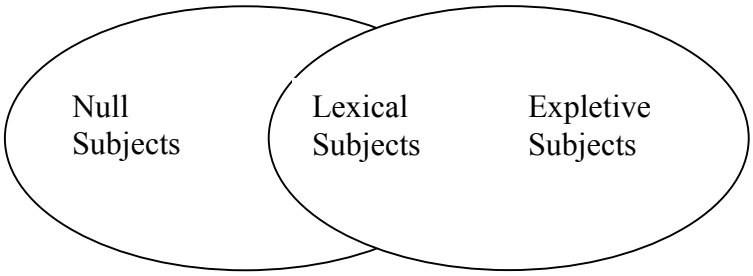
In Figure 3.1, 'Language A' is in a subset relationship to 'Language B'. This means that all possible sentences in 'Language A' are also possible in 'Language B', but the reverse is not true. The assumption, then, is that children must begin with the most restrictive grammar (enlarging it as the data permits), because if they were to start at the larger grammar, they would never receive enough evidence to restrict it to Language A, since the sentences found in the smaller grammar are also found in the larger grammar.

The subset principle has led some researchers (Berwick 1985, Phinney 1987, et al.) to assume that non-pro-drop languages are subsets of pro-drop languages. This permits predictions to be made regarding the difficulties involved in taking a particular learning path. For example, speakers of English (the subset language) would only require positive Spanish input (e.g. a sentence with a null subject) to begin to realize that they must adopt a larger sphere of language possibilities than their L1 requires. However,

learners of Spanish are in a more difficult situation. Positive input from English will yield overt subject pronouns, but these pronouns are also allowed in Spanish. Spanish speakers, being in a superset relationship to English, will require negative evidence that their pro-drop sentences are not grammatical in English. The requirement for additional negative evidence should make the learning process more difficult.

While some researchers believe the subset principle applies to the pro-drop relationship between Spanish and English, others have challenged this on several grounds (e.g. Hyams 1986, Wexler and Manzini 1987, MacLaughlin 1995). First, the subset condition is not fully met. For example, both English and Spanish use overt lexical subjects, but only English uses expletive subjects and only Spanish uses null referential subjects. This leads to an intersecting relationship between two languages such as Spanish and English, illustrated in Figure 3.2.

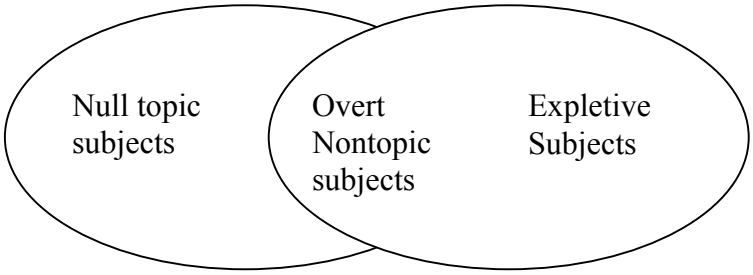
Figure 3.2 Intersecting relationship of subjects in Spanish and English



Second, even the expanded diagram in Figure 3.2 does not take into account that learning Spanish involves more than a general acceptance of null subjects. As Galván (1998) points out, learners of Spanish must learn both that null subjects are possible, and that certain discourse factors constrain their use. Any inclusion of discourse conditions would require Figure 3.2 to be amended so that nontopic null subjects would fall in the

intersection between the two languages, but null topic subjects would fall only on the Spanish side (Figure 3.3):

Figure 3.3 Discourse conditions in Spanish and English



Finally, although the Subset Principle provides one way to restrict the hypothesis space for learner grammars, it does not necessarily hold that this greater restrictiveness will automatically translate into a more learnable grammar. For learnability purposes, the manner of search is more important than the total size of the search area. Consequently, the choice of search method becomes critical in a developmental account of L2 learning.

The learning algorithm used in this dissertation, that of Tesar and Smolensky (2000), is able to arrive at target grammars without reference to the Subset Principle. It does so even assuming no variance in the set of possible inputs to the grammars of all languages; when the grammar of a given language is supplied with this set of universal inputs, the grammatical inventory of that language is defined as the output forms that emerge as a result of the operation of the grammar. This concept is called the 'richness of the base' (Prince and Smolensky 1993).

Although the learning algorithm of Tesar and Smolensky is not obligated to assume a particular initial state (i.e. it arrives at the target hierarchy regardless of the initial hierarchy assumed), an initial state could be hypothesized. Tesar and Smolensky

report on a suggestion made to them in personal communication with Prince (1993) that the initial hierarchy of L1 learners may be such that faithfulness constraints are lower ranked than markedness constraints, with the result that structural constraints are demoted below faithfulness constraints only 'in response to the appearance of marked forms in observed overt structures' (2000:76).

Tesar and Smolensky (2000) argue, however, that it is not the Subset Principle that is critical to the learnability of a grammar, but rather the particular learning mechanism that is employed. They focus on a particular problem in language learning — how learners, who often receive overtly ambiguous language data, are faced with a serious paradox: they cannot determine a grammar's hidden structure until they have constructed a grammar based upon their interpretation of the overt forms they hear, but they cannot construct a grammar without some analysis of the hidden structure. To address this paradox, Tesar and Smolensky have proposed a learning procedure where learners' first guesses at a structural analysis are used to improve their grammar, and this improved grammar is then used to improve the analysis. In other words, through successive approximation, learners acquire progressively better interpretations and a progressively better grammar simultaneously.

Tesar and Smolensky (2000) look to OT for the core principles that inform this learning strategy, and in their proposed model, Robust Interpretive Parsing / Constraint Demotion Algorithm (RIP/CDA), and they provide evidence for the accuracy and computational efficiency of their proposed model through a series of computer simulations and by a set of formal proofs. Their central claim is that OT provides the learning mechanism (RIP/CDA) through which the interdependence of grammars and

structural descriptions is overcome, allowing the learner both to assign structure and to learn grammar at the same time. The learning problem is decomposed into several parts—deducing hidden structure in language data, using the data to improve the existing model, assigning an improved hidden structure to the original overt data, and once again learning the grammar (using a 'robust' parser). This divides the problem into one of parsing and grammar learning.

The centerpiece of Tesar and Smolensky's theory is 'Constraint Demotion,' the notion that constraints violated by grammatical structural descriptions must be demoted, in the total ranking of constraints, below constraints violated by competing (ungrammatical) structural descriptions. The CDA operates in the following manner: When the hearer attempts to create a grammar based on the output of the speaker, Robust Interpretative Parsing first computes an input for the speakers' productions. The hearer's grammar then compares constraint violations of the target optimal output (winner) to those of the current grammatical system (loser). The algorithm cancels out the constraints violated by both winner and loser candidates as illustrated in 3.4:

(3.4.) <input A>

Candidate X (loser): Constraint 2, Constraint 3
Candidate Y (winner): Constraint 1, Constraint 3

Candidate X (loser): Constraint 2, ~~Constraint 3~~
Candidate Y (winner): Constraint 1, ~~Constraint 3~~

The hearer's grammar then demotes the constraints violated by only the winner below those violated by the loser, resulting in a grammar that will produce the intended outputs with fewer violations. This process of constraint demotion proceeds recursively until there are no more mismatches between the perceived output and the grammatical system.

The algorithm that Tesar and Smolensky propose accomplishes more than traditional parameter-setting models. It reveals not only a workable way to arrive at a desired end state, but also in that it provides a clear description of the developmental steps expected as individual constraints are demoted in the total hierarchy. But does parameter setting provide a more learnable model, since it is usually thought to contain only binary choices for a limited number of parameters? Does not the OT model require that more interactions are involved, resulting in a far larger number of possible grammars and, thus, a less learnable system?

Tesar and Smolensky answer these concerns by demonstrating that, although the total number of possible rankings in an OT system may be quite high with even a limited number of constraints, the restrictiveness of the structure OT places on the grammar permits learners to efficiently arrive at a target grammar in a reasonable number of learning steps. Parameter-setting models, even with their more limited set of possible grammars, do not improve upon the learnability of a language, if an uninformed learning method is used — an exhaustive search of all possibilities:

> Comfort from the finiteness of the space of possible grammars is tenuous indeed. For a grammatical theory...might be well structured, permitting *informed* search that converges quickly to the correct grammar—even though uninformed exhaustive search is infeasible...a well-structured theory admitting an infinity of grammars could well be feasibly learnable, while a poorly constructed theory admitting a finite but very large number of possible grammars might not.
>
> (Tesar and Smolensky 2000:2-3)

Tesar and Smolensky remark that a parameter-setting approach with n parameters admits at least 2^n grammars (assuming only binary parameters). This means that the search space increases exponentially with each proposed parameter, and the exhaustive search method quickly becomes unfeasible.

Tesar and Smolensky are not the first to observe this problem, and there are other proposals to deal with issues of learnability within a parameter-setting framework. For example, Dresher and Kaye (1990) proposed a 'cue learning' approach that could be applied specifically to learning metrical stress. But the use of a particular algorithm for each component of grammar is less preferable than an approach that can account for multiple components of a linguistic theory.

Other proposals, such as the Triggering Learning Algorithm (Gibson and Wexler 1994) or modifications of it (e.g. Niyogi and Berwick 1996), may be applied to a more general class of components, but as Tesar and Smolensky note, these algorithms are minimally informed by grammatical theory. Gibson and Wexler's algorithm randomly flips parameters to arrive at an analyzable input, and in Niyogi and Berwick's version, the randomly flipped parameters do not even need to directly result in analyzability. Consequently, these approaches are simply generic search algorithms that could be employed for any parameterized system; they make no use of grammatical theory or its unique properties. In contrast, learning in Tesar and Smolensky's algorithm is derived solely from general grammatical structure and informed by a specific theory of grammar. The particular strengths of their algorithm make it the choice for this dissertation's analysis of Spanish L2 data.

3.4 Theory of access to UG and L1 transfer

The analysis provided in this dissertation will critically assume both that learners of an L2 have access to Universal Grammar (through the OT learning mechanism), and that the initial hierarchical ranking of constraints for these learners is the ranking of their

L1 (i.e. 'transfer'). Neither of these positions has unanimous support among SLA researchers; therefore, this section briefly discusses these ideas.

Although generative linguistic research has proceeded from the assumption that humans have innate access to Universal Grammar (Chomsky 1965, 1975, 1986) for their acquisition of their first languages, access to UG for second language acquisition has been more controversial, and differing positions have been taken as to the extent to which learners have access to and use UG as they learn a second language. In terms of access, three broad positions emerge: first, the position that, unlike learning of the L1, no access to UG is available for adult L2 learners (Lenneberg 1967, Clahsen and Muysken 1986, Bley-Vroman 1989); second, the position that learners have access to UG, but only partially, not in the same direct and unmediated way that there is access to UG for the acquisition of the first language (Schachter 1989, Strozer 1992, Vainikka and Young-Scholten 1991, Bhatt and Hancin-Bhatt 1997); finally, the position that L2 learners have full and direct access to UG, making this one way in which L1 and L2 acquisition are similar (Epstein, Flynn and Martohardjono 1996, Lakshmanan 1993, Schwartz and Sprouse 1996).

Section 3.1 stated that the early Contrastive Analysis and Error-Analysis theories were challenged by the interlanguage approach of Selinker (1972), who viewed learner grammars in terms of a complex interaction between first language influences (transfer) and innate language learning processes (access to UG). This did not involve a complete disregard of first language influences, and, in fact, Gass and Selinker argued for such influences. They claimed that the 'overwhelming evidence that language transfer is indeed a real and central phenomenon that must be considered in any full account of the

second language acquisition process', and they further argued that SLA involves hypothesis testing using UG, L1 data, L2 data, and the knowledge of interlanguages (1983:7).

Nevertheless, debate concerning the notion of transfer continues. For example, Martohardjono & Flynn (1995) analyzed control structures in English, Japanese, Chinese and Spanish and concluded with the strong claim that L2 learners ignore their L1 syntax and rely instead solely upon principles of UG. In contrast, White (1989) argues that despite the learner's use of UG, the L1 does indeed play a significant role in L2 acquisition. White claims that L1 parameter settings are part of the second language learners' interlanguage, and that they influence attempts to both understand and produce the target language.

White (1989:48-9) outlined five logical possibilities for the relationship between transfer of the L1 and the accessibility of UG for L2 acquisition (Figure 3.4):

Figure 3.4 Possible relationships between access to UG and L1 transfer

a. UG is accessible and functions as it does in L1 acquisition.
b. UG is accessible, but learners initially transfer the settings of the L1.
c. UG is accessible, but only via the settings of the L1.
d. UG is accessible, but does not function identically as in L1 acquisition.
e. UG is inaccessible.

These possibilities are exemplified in Table 3.1 for a parameter-setting model of pro-drop where the pro-drop parameter is given two settings [+pd] or [-pd] (Adapted from Braidi 1999:62-64):

Table 3.1 Parameter model of acquisition of pro-drop

Condition: UG is accessible and functions as it does in L1 acquisition.				
L2 French Input [-pd]	→	L1 Spanish UG Default [+pd]	→	Set to [-pd]
L2 French Input [-pd]	→	L1 English UG Default [+pd]	→	Set to [-pd]
Condition: UG is accessible, but learners initially transfer the settings of the L1.				
L2 French Input [-pd]	→	L1 Spanish [+pd] UG	→ L1 transfer [+pd]	Reset parameter to [-pd]
L2 French Input [-pd]	→	L1 English [-pd] UG	→ L1 transfer [-pd]	Resetting unnecessary
Condition: UG is accessible, but only via the settings of the L1.				
L2 French Input [-pd]	→	L1 Spanish [+pd]	→ L1 transfer [+pd]	Cannot reset without UG
L2 French Input [-pd]	→	L1 English [-pd]	→ L1 transfer [-pd]	Resetting unnecessary
Condition: UG is accessible, but does not function identically as in L1 acquisition.				
L2 French Input [-pd]	→	UG & other components	→	Variable success in acquisition of pro-drop due to other components
L2 French Input [-pd]	→	UG & other components	→	
Condition: UG is inaccessible.				
L2 French Input [-pd]	→	L1 Spanish & general learning mechanisms	→	Variable success due to reliance on general learning mechanisms
L2 French Input [-pd]	→	L1 English & general learning mechanisms	→	

Table 3.1 assumes a default setting of [+pro-drop] for a pro-drop parameter. How the learner grammar will handle L2 input depends both on the transfer/access possibilities and on the relationship between the default setting of the parameter, the setting of the parameter in the L1, and the setting of the parameter of the L2. The parameter may require 'setting', 'resetting', 'maintaining', 'transferring', or some other treatment.

White (1989) harmonizes with Schwartz and Sprouse (1996) who hold, in contrast to Martohardjono and Flynn (1995), that all L1 parameters initially transfer and that it is only due to the failure of the L1 grammar to adequately represent the facts of the L2 that learners must restructure their grammar with the options available to them from UG.

This position, referred to as Full Transfer/Full Access, was supported in Schwartz and Sprouse's study by spontaneous production data of an adult native speaker of Turkish who was learning German. Turkish and German are very dissimilar in surface syntax, and as this position would predict, the learner was quickly forced to UG-constrained restructuring of his syntactic system. Selinker and Lakshmanan (1994) provide numerous additional examples from Dutch, Czech, Spanish, French, Hebrew, and Hindi, to show that transfer of L1 syntax to the L2 not only occurs, but also prolongs the restructuring, particularly when there are multiple effects at work.

Although questions related to transfer and access have regularly been posed within the parameter-setting framework, the same questions are applicable to the OT approach used in this study. From an OT perspective, although there are no parameters to transfer, what may transfer are the constraint rankings of the L1. In terms of access to UG, from an OT perspective this could mean either access to default constraint rankings supplied by UG or, more broadly, to the UG instantiated learning mechanism, represented by the operation of the OT grammar. Since there is no a priori reason to assume a default ranking of constraints, and since the successful operation of the RIP/CDA learning algorithm does not depend on an initial default ranking, the analysis provided in this dissertation will assume that while learners have full access to UG, UG is defined as the OT grammar learning mechanism together with the universal set of constraints.

3.5 Pro-drop in SLA

This section reviews the findings and methodologies of several SLA studies testing aspects of the pro-drop. For most of the early studies (e.g. White 1985, Emberson 1986, Hilles 1986, Phinney 1987, and Liceras 1988, 1989), this involved establishing the facts of pro-drop, submitting theoretical parametric accounts to empirical testing, and deciding on the default option for the proposed 'pro-drop parameter.' Very few SLA studies related to pro-drop have moved beyond the parameter-setting model; these few will be covered in Section 3.6. This section concludes with a short statement of the set of facts that SLA research on pro-drop has revealed, as well as remaining questions these studies leave unanswered.

The concerns of several early SLA researchers revolved around extending the L1 acquisitional work of Hyams (1983, 1986) to an L2 context, with some of the Hyams' hypotheses either carried into, or challenged in, these L2 studies. For example, whereas Hyams (1983) argued that pro-drop was the unmarked setting and that, if the target language was not pro-drop, a switch of parameter settings was required, White (1985) argued that the unmarked setting is [-pro-drop].

White (1985) argued that learners of a language at variance with their L1 in regards to pro-drop do not immediately or easily reverse this parameter when they begin to acquire the L2; rather, they initially transfer the setting of their L1 into the L2. White was also concerned about the types of input needed to trigger the change.

White's study tested the acquisition of English, specifically related to the three characteristics most closely associated with pro-drop: null subjects, subject inversion, and *that-trace* effects. Her study involved 73 adult L2 learners of English at McGill

University in Montreal, Canada. Of these 73 learners, 54 were native speakers of Spanish and 19 were native speakers of French.[6] White used a grammaticality judgment task where subjects were given 31 English sentences, some of which were well-formed English sentences and others that were not. The ill-formed sentences contained null subjects, inverted subjects, and violations of *that-trace*. White hypothesized that Spanish speakers would be likely to reject an English sentence such as (3.5) because the Spanish equivalent in (3.6) requires the presence of the complementizer *que*.

(3.5.) *Who do you believe will be the next president?*

(3.6.) *¿Quién crees* **(que) será* *el próximo presidente?*
Who believe-2sg that be-3sg-fut the next president

In contrast, White thought that French speakers would be less likely to make pro-drop related errors in their assessment of English sentences, because French and English share the characteristic of both being non-pro-drop languages.

White's results showed that Spanish speakers did indeed have more difficulty than the French speakers in correctly judging grammaticality when null subjects were used, although both groups also had difficulties judging sentences with expletive subjects ('*It* seems that Fred is unhappy'). Both the French and Spanish groups performed well in regards to subject inversion and poorly in regards to *that-trace*. Although White concludes that pro-drop is a parameter with a set of related consequences, the results of her study clearly show that the various pieces of pro-drop are not simultaneously acquired.

Emberson (1986) argued against access to UG in second language acquisition. Emberson focused on acquisitional differences between 'core' (i.e. pro-drop) and 'non-

core' (i.e. the present progressive tense) aspects of Spanish. The study involved 41 English speakers in a fourth-semester Spanish class at the University of Texas in Austin. Emberson reports on correlations of student scores on eight grammaticality judgment tests.

Emberson concluded that UG is not active in second language acquisition, based both on a lack of correlation between aspects of pro-drop and on better performance on the progressive tense (a non-core aspect of the grammar) than on pro-drop (a core aspect of Spanish grammar). This, of course, critically assumes that a core/non-core distinction exists for the Spanish progressive and pro-drop, an assumption not shared elsewhere in the pro-drop literature. Emberson's results do, however, indicate an important place for transfer, since L1 English learners of Spanish did better on the present progressive (a feature that Spanish and English share) than they did on pro-drop related tasks (where differences exist between Spanish and English). Rather than demonstrate that UG is no longer active in second language acquisition, Emberson simply provides more evidence of a legitimate role for transfer in SLA.

The study of Hilles (1986) offered another test of Hyams (1983) regarding the triggers involved in resetting the pro-drop parameter to [-pro-drop]. Hyams had argued that the presence of expletives reveal to the learner the need for parameter resetting. Hilles (1986) pursues the same line of reasoning as Hyams (1983) but departs from the earlier hypothesis by posing that the presence of modals, rather than expletives, supply the triggering data. Hilles attempted to determine when lexical subject pronouns first surface through the use of a longitudinal study of a single subject, Jorge. Jorge was a 12-year old native speaker of Spanish who had not formally studied English. From a

mixture of elicited and spontaneous data collection methods (none rigorously defined), Hilles gathered information regarding the number of instances of missing subjects. She also attempted to count the instances where pro-drop could have occurred (if the text had been in Spanish) but did not occur in English. Hilles then developed a formula by which she divided the total number of null subject instantiations by the sum of that number added to the number of instances where null subjects would have been expected to occur in Spanish.[7] Interpreting the data through the use of this formula, Hilles concluded that Jorge's use of null subjects was declining over time.

The remaining question for Hilles concerned what triggered the parameter change. Hilles noticed that the use of null subjects diminished as lexical material in the AUX category emerged, a finding consistent with the observation Hyman (1986) had made for L1 acquisition. For this reason, Hilles concluded that modals and expletives both may work together as triggers for parameter setting and, in the case of L2 acquisition, resetting.

Hilles' account has several obvious limitations. While longitudinal data may be very useful, the use of single subject makes generalizabilty of Hilles' conclusions difficult. Furthermore, there is a great deal of subjectivity in taking English utterances, reconstructing how they 'might' have been said in Spanish, and then drawing conclusions based on that reconstruction. Finally, a raw count of missing subjects is not sensitive to discourse conditions in Spanish (or conversational English!) where subjects may have been properly deleted.

Phinney (1987) also began with Hyams' (1983) hypothesis and attempted to empirically test it in a classroom environment through a contrastive study of L1 English

speakers learning Spanish and L1 speakers of Spanish learning English. Learners wrote compositions, which were then analyzed for the presence or absence of subject pronouns and the use of agreement morphology. Both groups performed well in regards to verbal agreement, but there were significant differences regarding the use of lexical and null subjects. L1 Spanish speakers omitted both expletive and referential pronouns in their L2 Spanish compositions, but L1 English speakers correctly omitted both types of pronouns. On the basis of this production data, Phinney (1987) concluded that the pro-drop parameter is reset easily and early for English learners of Spanish, but with greater difficulty for Spanish learners of English. Phinney then related her results to a theory of markedness, claiming that [+pro-drop] is the unmarked setting of the pro-drop parameter, making [-pro-drop] a marked setting. For Phinney (1987) this result meant that it was harder to change a parameter from an unmarked setting to a marked setting than it was to change from a marked setting to an unmarked setting.

Several questions remain unanswered in Phinney's account. First, the two test groups composed their compositions under differing conditions—for one group the exercise was a test; for the other, the exercise was simply a class activity. It is uncertain what role the differing conditions may have played on the results. Second, as Phinney also notes, most of the forms were written in the first person, which may have skewed results. Third, Phinney's results showed that the L1 Spanish speakers omitted subjects more frequently when there were errors in verb agreement, leaving open the question of the source of the errors. Finally, Phinney does not indicate whether subject pronouns were overt or covert in nontopic environments, or in those discourse situations where overt use of the pronoun is required also in Spanish.

Liceras (1988) provided additional production and grammaticality judgment tests of pro-drop and its properties. The 1988 study involved French and English learners of L2 Spanish. In this study, Liceras found no persuasive evidence that gaining a command of the Spanish inflectional system coincided with the various properties associated with pro-drop (e.g. inversion, optional or obligatory subject pronouns, or *that-trace* effects). She further found that even advanced learners 'did not have native-like competence with respect to stylistic conventions that govern the use of inversion in Spanish' (1989:115). Liceras (1988) concluded that different triggers may be required for different effects, with inversion possibly being triggered by the need in Spanish to have the preposition *a* before direct objects that refer to people, and *that-trace* possibly being triggered by the realization that empty complementizers are not permitted in Spanish relativization.

Liceras (1989) involved a grammaticality judgment of 17 items administered to 30 L1 English and 32 L1 French learners of Spanish from four levels of Spanish proficiency. The 17 items were manipulated to reveal information null subjects, overt expletives, inversion, and *that-trace*. The learners in this study were asked to make judgments, correct sentences with errors, and translate sentences into their L1.

From the results of these tasks, Liceras drew a number of conclusions: First, there was further confirmation of her 1988 finding that learners do not interpret *that-trace* sequences correctly, regardless of whether the complementizer is present or not, and that learners had more difficulties with inversion than with null subjects. Second, no lexical expletives in Spanish were accepted by learners. Most importantly, Liceras posited an implicational hierarchy regarding the order in which at least some of the properties associated with pro-drop are acquired. Liceras proposed that null subjects must be

acquired before inversion, and inversion before *that-trace*, at the same time granting that construction complexity and the structural properties of Spanish also play a role in the acquisition. Liceras noted that, if the Spanish setting for *that-trace* has been acquired, then inversion and null subjects have also been acquired, but the presence of inversion does not imply that acquisition of *that-trace* has taken place. The hierarchy of acquisition that emerges is: null subjects > inversion > *that-trace*.

This hierarchy would decompose the pro-drop parameter into different grammatical elements, the acquisitional order of which could then be empirically tested. If indeed the various grammatical phenomena that have been associated with pro-drop are acquired at different times, it would be reasonable to assume that there is no single parameter at all, but rather a progressive restructuring of a grammar in such a way that each new restructuring is evidenced by new grammatical effects. Liceras does not go so far as making this claim, but this insight is a key point of this dissertation.

In another study, Liceras (1989) combined grammaticality judgments with a translation task, providing a fuller insight into the language competence of the L2 learner; however, one complication of this study was that sentences often involved more than one feature, leaving some uncertainty as to what feature the reader was judging as grammatical or ungrammatical. This is a drawback to grammaticality judgment tests in which learners give only a binary response to the grammaticality of an item. In Liceras' study, this problem is mitigated by including a correction and translation, but a clearer approach might be to provide subjects with clear choices varying only in the presence or absence of a targeted feature. An additional drawback to the Liceras (1989) study is the number of conditions that are spread out over a small number of test items. This resulted

in some properties of pro-drop being tested in only a couple of items. An improved methodology would propose more items per condition, and more total items.

Few SLA studies have provided true longitudinal data of the development of non-native Spanish speakers. Hilles (1986) attempted this but, as we saw, the study was limited to a single subject. Liceras (et al., 1997) represented another attempt to obtain longitudinal data from 5 high school and 11 university students acquiring L2 Spanish. Data collection took place through three recorded, 30-minute interviews eliciting spontaneous speech. The interviews occurred after students' formal exposure to Spanish reached 50 hours, 65 hours, and 80 hours.

Liceras (et al.) evaluated the recorded data for the total number of sentences, the use of pro, the use of personal pronouns, and instances of incorrect morphology. The goal of this study was to evaluate several current hypotheses in grammatical theory as they relate to the status of null subjects and clitic pronouns: the minimal sentence hypothesis (Radford 1990), the short sentence hypothesis (Meisel and Müller 1992), the full sentence hypothesis (Hyams 1994, Rizzi 1994), and the VP-internal hypothesis (Zagona 1982, Koopman and Sportiche 1991). The details of each of their findings regarding each of these hypotheses are not significant here, and the results of their study were somewhat inconclusive,[8] but Liceras (et al., 1997) is important for this dissertation in two regards: first, it provides another example of SLA research providing empirical tests of current grammatical theory; second, Liceras (et al., 1997:128) conclude that '...further research is necessary to define the various stages in the development of the non-native pronominal system', and they suggest that a comprehensive account of non-native competence requires not only longitudinal data but also '...specific tests given at

different stages of the development of the non-native grammar' (1997:128). This dissertation seeks to make progress precisely in these areas.

Liceras and Díaz (1999) represents another attempt to account for the distribution of null subjects and pronominal subjects in L2 Spanish using advances in grammatical theory. Liceras and Díaz use proposals by Rizzi (1994) and Hyams (1994) to account for pro, and they conclude that as non-native learners construct their grammars, they use default licensing that allows null pronouns, provided the null pronouns can be identified.

Liceras and Díaz' study looked at two different production data: (1) half-hour interviews in which beginning learners, 12 year olds (n=5) and university students (n=6), answered questions and were asked to create stories regarding characters and actions depicted in a comic strip, and (2) narratives produced by advanced-intermediate, non-native speakers (n=15) in which speakers were asked to tell a story based on one of their favorite films. The first languages of the subjects in the second task were Chinese, English, French, German, and Japanese. Three native speakers of Spanish were also asked to perform both of the tasks. This study was ambitious in its breadth, with subjects of differing ages, L1s, proficiency levels, and even tasks; however, given the relatively small number of subjects, this amount of diversity may be more of a weakness than a strength.

Liceras and Díaz found that even in early interlanguage learners produce null subjects both in matrix and subordinate clauses. Using the theoretical insights of Rizzi (1994) and Hyams (1994), they give an analysis where identification via discourse of f-features overrides identification through subject pronouns. Liceras and Díaz believed this to be a natural result for two reasons: (1) there are a large number of null subjects in the

Spanish input (saliency), and (2) learning the subjective, objective, and oblique inventories of Spanish pronouns is difficult (avoidance).

The type of production data provided here nicely complements a comprehension task, but it should be noted that Liceras and Díaz's study is representative of many other studies reviewed in this section in that it does not demonstrate how a parameter-setting framework accounts for developmental stages reflecting the interaction of discoursal and syntactic requirements. The binary nature of a parameter-setting framework does not easily account for stages in developmental route, and even Liceras and Díaz grant that interlanguages do not appear to be instantiations of coherent parametric options. Interlanguage evidence does not suggest that, for example, English second language learners of Spanish may simply 'turn on' a null subject parameter at some point. Rather, the evidence reveals just what a parameter-setting model should not: learners do not move directly from a 'non-pro-drop' stage to a 'pro-drop' stage; instead, there are clear developmental stages in which more than a single binary parameter is involved

With the exception of a handful of studies (notably, Liceras 1988, 1989, Liceras and Díaz 1995, Pérez-Leroux and Glass 1997, 1999, Pérez-Leroux, et al. 1999), most L2 research on pro-drop has ignored semantic and discoursal interfaces with syntax in learners' grammars. Many earlier studies operated with the premise that null-subjects are optional in languages such as Spanish and Italian, or simply described the conditions under which pro-drop is permissible. More recent L2 research (e.g. Pérez-Leroux et al. 1999, LaFond, Hayes, and Bhatt 2001) argues that such 'optionality' evaporates when concerns for information structure are included in the equation. Second language

learners face the challenge of learning that subjects are obligatorily absent in some contexts and obligatorily present in other contexts.

Pérez-Leroux, et al. (1999), demonstrate that interfaces between discourse and syntax in pro-drop languages control the distribution of null and overt pronouns. In Spanish, a null subject is permissible only when the antecedent is a topic; it is prohibited when the antecedent has nontopic status, as shown in 3.7 and 3.8 (Adapted from Pérez-Leroux, et al.1999): [9]

(3.7.) a. *¿Quién canta?*
'Who sings?'

b. *Ella*/*Ø *canta*.
'She sings.'

(3.8.) a. *¿Qué canta Cecilia?*
'What does Cecilia sing?'

b. Ø/??*Ella canta boleros*.
'She sings boleros.'

In 3.7, *ella* has not yet been activated in the discourse, prohibiting a null pronoun, but in 3.8, e*lla* is noticeably worse than the null pronoun, because the discourse contains a recoverable 3sg feminine reference.

These examples also suggest that issues of focus may be involved in the distribution of null subjects. The answer to 3.7a must be focused, and when this is the case, null subjects are prohibited. Constraints related to focus play a significant role in the analysis used in this dissertation. As Grimshaw and Samek-Lodovici (1995) noted, subjects that undergo what was previously believed to be 'free'inversion are not free at all; these subjects are actually focused in the discourse. Thus, concerns for information

structure relations may place conditions on inversion. If focus, like topic, interfaces with syntax in significant ways to determine the acceptability of sentences, then it, too can not be ignored in an account of L2 development.

Although this section's review of the literature may seem to illustrate that there is little agreement on how the L2 acquisition of pro-drop proceeds, a certain set of facts regarding the acquisition of null subjects does emerge. First, despite the fact that null subjects appear early in L2 learners from a non-pro-drop language who are beginning to acquire a pro-drop language (Phinney 1987, Liceras and Díaz 1995), these learners initially overgenerate overt pronouns. This was noted early on by Fleming (1977), and confirmed by Liceras (1988), and Al-Kasey and Pérez-Leroux (1998). Second, despite their own overgeneration of overt pronouns in production tasks, L2 learners can sometimes detect the ungrammaticality of overgenerated overt pronouns, as shown in Liceras (1988). Third, there appears to be an implicational hierarchy regarding the order in which at least some of the properties associated with pro-drop are acquired (Liceras 1989). Finally, as Al-Kasey and Pérez-Leroux (1998) note, L2 learner errors are systematic, patterned, and related to the influence of the L1. If learners' errors are systematic and patterned and there is an implicational hierarchy to the properties of pro-drop, then learners should follow a predictable developmental path as they acquire the pro-drop. The details of such a path have yet to be explicitly described. One of the contributions this dissertation makes to SLA research, in addition to an analysis offered within a particular theoretical framework (Chapter 6), is a clearer statement of the facts related to the path taken by L2 learners regarding pro-drop (Chapter 5).

3.6 SLA accounts of pro-drop from an OT perspective

This section completes the selected review of SLA literature by discussing two accounts related to pro-drop from an OT perspective: Park (2000) and LaFond, Hayes, and Bhatt (2001). Both of these studies begin to fill a gap between linguistic theorizing and its application to second language acquisition, but both of these studies have weaknesses that this dissertation attempts to address. Park's account is limited to null arguments in child second language acquisition, and as Park herself admits (2000:38) child L2 acquisition may have more in common with L1 acquisition than adult L2 acquisition. LaFond, Hayes, and Bhatt's account deals more directly with pro-drop from an adult SLA perspective, but this study is restricted to a consideration of null subjects, leaving further pieces of the pro-drop puzzle to later research. This dissertation supplies the further development of ideas prompted by this study.

For her study of null arguments (both subject and object) and wh-questions, Park (2000) used production data (collected by the National Center for Bilingual Research) from six Korean children learning English as a second language in a bilingual education school program. This data was part of a corpus of natural and elicited-interview speech. Park's goal was to examine whether two current theories, MP and OT, 'could account for language acquisition phenomena in real time' (2000:226), a goal that she concluded neither theory in its current state is prepared to handle.

According to Park, the feature-checking model of MP does not account for the distribution of null arguments. To make an MP account work, Park proposed that Korean has agreement features similar to Spanish-type languages, but that these features show up very late in children's grammars (past the age of six to eight years!), manifested by the

honorific marker *si* on the verb. Park concedes that such a claim involves quite a stretch for MP theory.

In regards to OT, Park used the constraints found in Speas (1997) and discussed in Chapter 2 of this dissertation. Once again, Park assumed a very late development of agreement features, perhaps as the result of inactivity of constraints related to agreement and phi-features. Park also argued that DROPTOPIC is a floating constraint, variably ranked along the constraint hierarchy. She notes that if this is the case, the placement of DROPTOPIC is more useful for explaining the distribution of null subjects in differing registers than in language development. Regarding this possibility, Park writes (2000:231):

> This implies that OT may be more of a grammar of pragmatics than a grammar of syntax. Those who consider that pragmatics is not a part of linguistic competence may want to argue that OT is a grammar of linguistic performance.

Park reports that she considers pragmatics to be a module of linguistic competence, so she would not herself draw the conclusion that OT is a performance grammar.

As mentioned above, the value of Park's account for adult SLA may be limited. Her bold proposal regarding agreement in Korean may spark further study for L1 studies of language development in Korean, but it admittedly falls short of providing a developmental account that is useful for adult L2 acquisition.

LaFond, Hayes, and Bhatt (2001) more directly addressed the issue of learning null subjects in a second language from an OT perspective. This study investigated the L2 acquisition of Spanish and Italian by L1 English learners at three universities: the University of South Carolina, the Pennsylvania State University, and the University of Arizona. A total of 210 subjects, divided into five proficiency levels for Spanish and four

proficiency levels for Italian, were given a written grammaticality judgment that investigated the use of null subjects in conversational dialogues.

Unlike most previous studies, these dialogues established a clear discourse context, and learners were asked to make a choice between two responses that best completed the dialogue. The responses represented syntactic minimal pairs, differing only in their presence or absence of a null subject. For some dialogues, the discourse context required a null subject; in others an overt subject was preferred. Error rates for topic and nontopic subjects were computed, and the results of this study revealed that very early learners initially undergenerate null subjects. At the next level of proficiency, learners use null subjects quite generally, without discriminating discourse context, resulting in overgeneration of null subjects. The appropriate subject retention in nontopic contexts was achieved only in advanced stages of acquisition. These results revealed a developmental path that LaFond, Hayes, and Bhatt interpreted from an OT perspective. To do this, they begin by assuming the three constraints in 3.4 (from Grimshaw & Samek-Lodovici, 1995):

Table 3.2 Constraints implicated in the initial surfacing of null topics

PARSE	Parse input constituents (Failed by unparsed elements in the input)
SUBJECT (SUBJ)	The highest A-specifier in an extended projection must be filled (Failed by clauses without a subject in the canonical position)
DROPTOPIC (DROPT)	Leave arguments coreferent with the topic structurally unrealized (Failed by overt constituents which are coreferential with the topic)

Following the logic of OT, differences between English and Spanish regarding null subjects are the result in variance in the rankings of these constraints. The English

ranking, PARSE » SUBJECT » DROPTOPIC, rejects candidates with null arguments Tableau 3.1:

Tableau 3.1 English grammatical system

input: <*leave* (x), x = topic, x = *they*, y = *at eight*> tense = past	PARSE	SUBJ	DROPT
a. ø *left at eight*	*!	*	
b. (null)	*!*		
c. K *They left at eight*			*
d. *Left they at eight*		*!	*

In contrast, the Spanish ranking, DROPTOPIC » PARSE » SUBJECT, selects candidates with null subjects as optimal when referencing a topic subject.

Tableau 3.2 Spanish grammatical system

input: <*salir* (x), x = topic, x = *ellos*, y = *las ocho*> tense = past	DROPT	PARSE	SUBJ
a. K ø *Salieron a las ocho*		*	*
b. (null)		**!	
c. *Ellos salieron a las ocho*	*!		
d. *Salieron ellos a las ocho*	*!		*

LaFond, Hayes, and Bhatt discuss the developmental moves from the L1 English to the L2 Spanish or Italian. L1 English learners of Spanish initially transfer their L1 constraint ranking to the L2 data. Therefore, their grammars select a choice that is inappropriate for the L2. Following the Constraint Demotion Algorithm of Tesar and Smolensky (2000), the path of acquisition involves pairs of constraints being matched against each other, with the higher-ranked constraint (winner) being moved (demoted) lower than the loser-violated constraint. The initial demotion is shown in 3.3:

Tableau 3.3 Initial constraint demotion: SUBJECT below DROPTOPIC

	SUBJ	DROPT
K a. [IP *ellos* [VP ...]		*
b. [IP ø [VP ...]	*!	

The resulting interlanguage hierarchy in Tableau 3.4 still does not converge on the L2 ranking; its choice is suboptimal:

Tableau 3.4 Interlanguage 1: PARSE » DROP TOPIC » SUBJECT

input: subject=topic	PARSE	DROPT	SUBJ
K a. [IP *ellos*[VP ...]		*	
b. [IP ø [VP ...]	*!		*

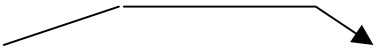
Therefore, the learning algorithm applies once again and continues to apply as long as there is a mismatch between loser and winner marks (Tableau 3.5):

Tableau 3.5 Mark-data pair

loser/winner pairs	marks' (loser)	marks' (winner)
[IP ø [VP ...] < [IP *ellos* [VP ...]	DROPT	PARSE

At the next stage, PARSE is demoted below DROPTOPIC:

Tableau 3.6 Constraint demotion: PARSE below DROPTOPIC



input: subject=topic	PARSE	DROPT
K a. [IP *ellos* [VP ...]		*
b. [IP ø [VP ...]	*!	

Once this demotion occurs, the resulting hierarchy finally chooses the target optimal candidate, Tableau 3.7b:

Tableau 3.7 Interlanguage 2 (Target): DROPTOPIC » PARSE » SUBJECT

input: subject=topic	DROPT	PARSE	SUBJ
a. [IP *ellos* [VP...]	*!		
K b. [IP ø [VP ...]		*	*

The ranking in Tableau 3.7 converges on the ranking of Spanish, so the learning algorithm halts. The final restructured constraint ranking correctly reflects the observation that topics will be dropped; this ranking will necessarily retain nontopics.

This analysis given by LaFond, Hayes, and Bhatt accounts for the acquisition of null subjects, but does not yet explain how other properties associated with pro-drop languages fit into the picture. In the next chapter, the specific research questions and hypotheses that grow out of this development will suggest that inversion and *that-trace* effects may also be characterized in similar fashion as the null subject analysis of LaFond, Hayes, and Bhatt (2001). The null subject analysis of LaFond, Hayes, and Bhatt (2001) will also be given further refinement in light of the new data this dissertation provides. In so doing, this dissertation will provide a developmental account that provides a clear contribution to SLA research — it will offer an explanation of why certain pro-drop effects surface before others and why neither discoursal nor syntactic explanations can solely explain the path that learners take in their acquisition of Spanish.

Notes

[1] The Audio-Lingual Method was the name given to the pronunciation and pattern drills used by the Army Specialized Training Program (ASTP) during World War II and based on the structural approach described by Fries (1948).

[2] This 'thought experiment' provides evidence of another principle — that questions themselves indoctrinate and direct us to look at the world in specific ways.

[3] In the context of second language acquisition, it might be interesting to pursue the flip side of Plato's problem, what Chomsky has called 'Orwell's problem,' i.e. how is it that human beings know so little given the amount of information to which they have access. Taken together, they lead us to ask why, given input that is so rich in some ways and so impoverished in others, humans acquire what they do.

[4] 'Atelic' verbs are those that refer cumulatively, without a specific endpoint. In contrast, 'telic' verbs are delimited by a clear point at which the action of the verb is ends.

[5] As in messages such as 'Guess what? ∅ went to work today and Mary wasn't there, so ∅ went downstairs and ∅ asked Sue if I could use her parking space...'

[6] The lower number of French speakers was the result of White treating this group as a control group, since French (like English) is a non-pro-drop language. For those operating within a parameter-setting model, this assumption may be uncontroversial, but in a OT account such as that proposed in this dissertation, differences between French and English constraint hierarchies may challenge the grouping of French and English together in this manner.

[7] X / (Y+X), where x=the number of instances of null subjects and y=the instances where null subjects would be predicted to have occurred in Spanish, but did not appear in English. Hilles did not include null expletives in this count because she thought possible triggers must be counted separately.

[8] For example, Liceras et al. (1997:128) could not determine whether non-native productions of pronouns had the same value as native production of clitics because they could not determine 'whether they are grammatical items or morphological parts of the verb'.

[9] Liceras and Díaz (1995) hold that L2 learners' null subjects are *not* necessarily instances of topic-drop. They argue that the learner's production of null subjects is the result of local restructuring options (not reparameterization). As Pérez-Leroux and Glass (1997) note, if Liceras and Díaz are right, the overt/null alternation of subject pronouns may involve different aspects behaving independently.