The Second Language Acquisition in Optimality Theory: Starting Point and Its Result

Incheol Choi
(Kyungpook National University)

Choi, Incheol. 2006. The Second Language Acquisition in Optimality Theory: Starting Point and Its Result. *The Linguistic Association of Korea Journal, 14*(2), 35-58. This paper aims to develop a second language acquisition model on the basis of Optimality Theory. Following the principle of *richness of the base*, the proposed model suggests that the second language acquisition is a process of acquiring the ranking of a target language. Specifically, in this model, the learning is facilitated by Constraint Demotion which was originally suggested for the learning mechanism in the first language acquisition (Tesar and Smolensky 1996). However, contrary to the first acquisition, the learning begins with either learners' mother tongue or the same starting point as their first language aquisition. This discrepancy, together with other distinctive factors, plays a significant role in the explanation of language transfer, asymmetric learning between subset and superset languages, and markedness phenomena that characterize the second language acquisition.

Key Words: second language acquisition, optimality theory, Constraint Demotion, markedness, subset theory, acquisition model

1. Introduction

In spite of several controversies, Universal Grammar (UG) has been an important theoretical and empirical foundation of the study of second language (L2) acquisition over the last two decades. Particularly, with the development of the principles and parameters framework and its succeeding theories, it has been an important issue in researches in L2 acquisition.

However, the results of such investigations have been at most vague or controversial. Such dissatisfaction may be ascribed to the fundamental differences of L2 learning from the first language (L1) learning. For example, L2 learners, in contrast to L1 learners, hardly show uniformity in their eventual achievement[1]. Furthermore, their learning of L2 may not lead them into truly native-like competence. Another main source of the difficulty comes from the uncertainty of UG itself. Although various works and empirical evidence of UG have been dedicated to the research of L2, its specific role and contents in L2 acquisition are still in a fog. The former obstacle may be hard to get over and still disturb our correct understanding of the L2 acquisition process. However, it has become possible for us to define UG more explicitly. The emergence and development of Optimality Theory (OT) during last decade have enabled us to pursue such a task.

According to the principle of *richness of the base* in OT (Prince and Smolensky 1993), the source of all systematic cross-linguistic variation is constraint reranking. That is, the set of constraints and the *Inputs* are identical across languages since they are assumed to belong to UG. In this theoretical perspective, L1 acquisition is defined as a process of acquiring the language particular ranking and the lexicon (Smolensky 1996). According to Tesar and Smolensky (1996), this process involves Constraint Demotion whose main function is to reorder the raking of constraints. On the basis of OT learnability theory, this article provides a second language (L2) acquisition model which is aimed to show the process of L2 acquisition more explicitly and specifically.

The appropriate L2 acquisition model has to explain many factors that are not observed in the L1 acquisition process. Among these, the most prominent factors include the transfer of rule systems from L1 into L2, the systematic patterns among L2 learners at each stage of learning, and the eventual incompleteness for the majority of L2 learners (Towell & Hawkins 1994). The model suggested in this paper rises to the challenge of an appropriate explanation of such issues. Specifically, I propose that properties of the starting point in L2 learning play a significant role and affect the

1) However, I do not deny that there exist rough patterns in the process and the result of the L2 acquisition.

whole procedure of L2 learning. The starting point is mainly determined by the properties of L2 data and the complete L1 ranking. If the L1 ranking matches the L2 data, language transfers are expected. If L1 ranking does not match the L2 data, the L2 learners do not persist in their L1 rankings for L2 learning. In this case I suggest that the starting point of the L2 learning becomes the *initial state* from which the L1 acquisition starts. The *initial state* reflects the ranking in which structural constraints dominate faithfulness constraints (Smolensky, 1996). The L2 acquisition model suggested in this paper explains why language transfer applies to a whole parameter rather than a separate phenomenon, why unmarked properties are easier to learn than marked properties, and finally how the subset and superset distinction of language input plays a significant role in the L2 acquisition process (Berwick 1985).

2. Acquisition Model in OT

In this section, I will first provide a brief description of the Constraint Demotion and L1 initial state suggested in OT based learnability theory. In section 2.3, I will suggest that the Constraint Demotion operates on the L2 acquisition process while setting up a distinct initial state for L2 learners that would partially explain the difference between the L1 and L2 acquisition processes.

2.1. Constraint Demotion

In OT, cross-linguistic variation is due to the different rankings of the universal constraints. According to Tesar and Smolensky (1996), language learning starts with the universal components of any OT grammar: the set of possible *Inputs*, the function *Gen* generating the output candidates for any possible Input and the constraints *Con* on well-formedness. These innate universal components, together with other linguistics and biological abilities, make it possible for language learners to establish a raking of the constraints in a particular language. According to Tesar and Smolensky (1996), the main method that sets up the particular ranking is 'Constraint

Demotion'.

In Tesar and Smolensky's acquisition model, language learning is the process of finding the ranking which makes the grammatical structure more harmonic than the other competing ones. For example, when a language learner receives a positive data *p* of the target language, he will consider *p* grammatical, a winner. Then every other parse *p'* of the same *Input* becomes the losing competitor. If a property of the learner's current grammar differs from the corresponding property of the target language, it becomes the loser in the target language. It means that the learner's current ranking is not harmonic with the target language. To adjust the learner's current grammar to the harmonic one of the target language, the learner needs to establish a new ranking in which the constraints violated by the loser out-rank those violated by the winner of the target language. Constraint Demotion accomplishes this reranking task by demoting the constraints violated by the winner so that they are dominated by the constraints violated by the loser. According to Tesar and Smolensky (1996), this Constraint Demotion is the main method with which L1 learners have access to their native language.

2.2. Initial Ranking

One of the basic assumptions in L1 acquisition is that learners' *initial states* are the same across languages. OT defines all of the universal components such as *Input* and constraints as innate. If such linguistic abilities are already built in language learners, the question of *initial state* is reduced to what is the initial ranking of the universal constraints.

> The same linguistic structures which are marked in the sense of later-emerging in child language tend also to be marked in adult languages. Using Optimality Theory, and developing a proposal of Prince & Smolensky (1993), this generalization can be reduced to the requirement that the *initial state* H0 possess the property that structural markedness constraints outrank faithfulness constraints

(Smolensky 1996, p. 1).

Structural constraints force early child outputs to hold unmarked structural properties. In early stages of language learning, the structural constraints dominate the faithfulness constraints to prevent faithfulness to *Inputs* from forcing the early child outputs to have unmarked structures. During the acquisition process, children change their constraint ranking to that of adults via Constraint Demotion.

Smolensky (1996) suggests an example that explains the *initial state* of language learning. Suppose a language contains only morphemes constructed from .CV. syllables. That is, every syllable consists of an obligatory onset consonant and a nucleus vowel. If the language does not have any segment deletion or insertion, it will have no disparity between deep and surface forms. Unmarked syllable structures would result from the unmarked ranking where the structural constraints such as ONSET and NO-CODA dominate the faithfulness constraints requiring input-output identity.[2] However, in this language the correct outputs do not violate faithfulness constraints or structural constraints because there are no disparities between deep and surface forms. Therefore the lexical items in this language are optimal under all constraints rankings because the correct outputs violate no constraints. This language, however, shows a strong regularity, in that it consists of only the unmarked syllable structures, CV. According to *Richness of the base*, this regularity results from the grammar alone, assuming no limitations on Inputs. Thus, to produce just .CV. syllables with unrestricted Inputs, e.g. /CVCVC/, this language should have the ranking that the syllable structure constraints outrank faithfulness constraints. If all constraints are located in one stratum without any ranking, the learners in this language can never set up the ranking. The reason is that Constraint Demotion cannot work because all constraints satisfy the unmarked syllable structures. Then, to set up the unmarked ranking, the *initial state* should already have the property such that structural constraints dominate

2) The constraint ONSET requires that syllables have onset consonants and the constraint NO-CODA blocks syllables from having coda consonants.

faithfulness constraints.

This idea, suggested by Smolensky (1996), explains the *initial state*. The structural constraints force outputs to be unmarked. The unmarked structures reflect the *initial state*. That is, structural constraints dominate faithfulness constraints in the *initial state*.

2.3. L2 acquisition model

Although it is controversial whether UG plays a role in L2 acquisition, I adopt the Full Access Hypothesis in which UG guides the process of L2 learning (Epstein et al, 1996). Given the hypothesis, like L1 acquisition, L2 language learners have to find a new ranking in order to acquire the target language. I suggest that this task is also achieved by the Constraint Demotion. That is, Constraint Demotion is the basic learning mechanism in L2 acquisition process as well as in L1 acquisition, although its application may have to involve various additional factors that are not found in its L1 application.

The learning task in OT based L2 acquisition theory can be formalized as in (1).

(1) Learning Problems in L2 acquisition

Given: The Universal components of any OT grammar
The set of possible *Inputs*
The function Gen generating the candidate outputs for any possible Input
The constraints Con on well formedness
L1 language particular OT grammar:
a ranking (or set of rankings) of the constraints in Con
L2 learning data in the form of full structural descriptions of grammatical form

Find: A language-particular OT grammar consistent with all the L2 data: a ranking (or set of rankings) of the constraints in Con

Like L1 learners, L2 learners are already given the universal components defined in OT grammar when they start learning. However, in contrast to L1 learners, L2 learners already have a fully established L1 ranking. The existence of the complete ranking of the L1 gives rise to the question of what is an L2 learner's starting point. Specifically, the question is whether L2 learners start from the same *initial stage* as L1 learners. The positive answer will explain the pervasiveness of the language transfer in the L2 acquisition process. Conversely, it may not explain why the language transfer cannot fully cover the early stage of the L2 acquisition. For instance, according to White (1986) and Liceras (1989), English speakers who learn Spanish drop subjects as frequently as Spanish native speakers do. If the complete L1 ranking were what L2 learners start from, those subjects would not be dropped even when they are evident from contexts. This is because English requires an overt subject in sentences. On the other hand, if it is assumed that the starting point of L2 learning is the same as the initial state of L1 acquisition, then it must be explained again why language transfer not only occurs, but is also pervasive in L2 acquisition.

At a glance, the discussion above may lead us to further confusion since it implies that neither the universal initial state nor learners' L1 rankings can be the sole answer to the question about L2 learners' starting points. However, what we have to deny is the uniqueness and absoluteness of the starting point of the L2 acquisition. What I suggest here is that the staring point of L2 learning involves various possibilities. This flexibility is not only a characteristic property of L2 learning but also a tenet of OT. That is, OT can well embrace this property by its nature. Specifically, I suggest that the initial state and the learners' L1 ranking compete with each other to be the starting point of L2 learning. When L2 learners encounter L2 data, they will realize that the aspects of the L2 data are either identical to or different from their L1 parameters. For example, when a Korean learns Japanese, he will find that the word order of Japanese is identical or similar to Korean. Then, little learning may take place with respect to Japanese word order structure. That is, the Korean learner may adopt the word order of Korean to acquire thatof Japanese. On the other hand, suppose an English speaker learns Japanese. He will easily realize that the word order of Japanese is

different from that of English. In this case, the starting point of learning is not the ranking of English because the English speaker knows that the L1 ranking cannot produce the correct outputs of Japanese word order structure. Thus the English speaker will go back to the initial state of L1 instead of persisting in the complete L1 ranking.

Evidence for this assumption can be shown in the acquisition of the null subject parameter. When English speakers learn the null subject parameter of Spanish, they easily acquire the null subject property. If learners start their learning from the English ranking, we cannot explain why they easily acquire the null subject property. According to Liceras (1989), null subject parameter reflects the unmarked property with respect to the subject realization of sentences. That is, the null subject parameter is near initial state. This unmarked property makes the null subject parameter in Spanish easy for English speakers to acquire. Therefore, it may be persuasive to assume that L2 learners start their learning from the initial state, when the L1 ranking does not match L2 data.

In sum, the starting points of the L2 learning are not the same across languages in contrast to L1 initial state. Whether the starting point of L2 learning becomes the learners' L1 rankings or the initial statedepends on how L2 learners analyze the L2 data. It is well known fact that L2 acquisition differs in degree of success and uniformity of the resulting systems (Bley-Vroman 1989). The non uniformity of the starting point of L2 learning partially explains the problem of L2 acquisition.

Now the model suggested in (2) explains how L2 learners set up the ranking of the target language.

(2) L2 Acquisition Model
    Starting Point for L2 Learning:
        Language Particular L1 ranking
        Otherwise, Language Universal initial ranking:
            Structural Constraints >> Faithfulness Constraints
    How to Learn L2 ranking: Constraint Demotion

3. Phenomena in L2 acquisition

In this section, I will show how the L2 acquisition model proposed in the previous section can apply to various phenomena in L2 acquisition reported in literature. First, in section 3.1, I will show that the proposed model well explains the language transfer phenomenon with respect to the head direction parameter. Second, in section 3.2, I will show how the markedness theory which explains asymmetric acquisition between marked properties and unmarked properties can be reinterpreted in the proposed model. In addition, I will show that this alternative explanation can avoid some problems that the previous markedness theory may have to confront. Finally, in section 3.3, I will show that the proposed model does not need any additional mechanism to explain the phenomena that used to be explained by the Subset Principle.

3.1. Language Transfer

Flynn (1989) shows that a match or a mismatch in the head direction of the L1 and the L2 determine the acquisition pattern of L2 relative clause constructions. In her study, two groups of adult L2 learners of English - Spanish and Japanese speakers with equivalent ESL proficiency levels - were tested. Her study shows that errors made by the Spanish and Japanese speakers differ qualitatively. Japanese learners show more difficulty with head direction than Spanish learners. Thus Flynn (1989) concluded: "If principles involve parameter, L2 learners at early stages recognize a match or a mismatch in the values of the parameters between the L1 and the L2. When the L1 and the L2 mismatch in the parameter, acquisition is disrupted as learners must assign a new value to the parameter in acquisition."

One of the questions that arises from this result is whether the language learners transfer some part of properties of their L1 relative clauses to the target language or the whole head direction parameter. Unfortunately Flynn's experiment cannot give full answer to this question. Spanish has a SVO word order like English thus it shows head initial property in VP. Spanish also shows the head initial property in relative clauses, in that the

head noun precedes the relative clause. Because both properties are identical to English, we cannot determine whether the result of Flynn (1989) is from each separate transfer of VP structure and relative clauses or from the whole transfer of the head direction parameter. Evidence that supports the latter possibility can be drawn from Japanese informants.

Japanese is a head final language thus, heads are located in the final position of phrases, as shown in (3)

(3) a. [PP nihon-ni] (prepositional phrase)
Japan-in
"in Japan"
b. Watashi-wa [VP nihonjin desu]. (verb phrase)
I-topic Japanese am
"I am Japanese"

This head direction is also kept in adverbial adjunct clauses. For example, the preposed left branching of the adverbial clause in (4a) correlates with the head final parameter whereas the postposed right branching in (4b) correlates with the head first parameter.

(4) a. PREPOSED [When the actor finished the book, [the woman called the professor]
b. POSTPOSED [The worker called the owner [when the engineer finished the plans]] (Epstein et al, 1996)

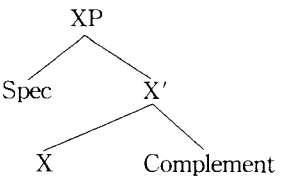
In Japanese only (4a) is allowed. According to Epstein et al. (1996), at an early stage, Japanese adult L2 learners of English do not find the preposed sentence structures significantly easier either in imitating or in comprehension. Later in acquisition, the Japanese speakers show a significant preference for postposed sentence structures.

Although (4b) is a more canonical structure, both are allowed in English. If transfer applies to each separate phenomenon rather than a whole parameter, we can predict that (4a) would be more preferred by Japanese speakers because only the structure in (4a) is allowed in Japanese. The

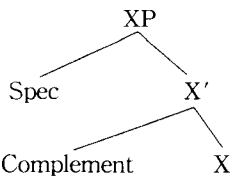
contradictory result, however, shows that language transfer does not apply to each separate phenomenon. When Japanese learners encounter English sentences, they come to realize that English and Japanese differ in head direction. Thus, from an early stage they do not prefer (4a), even though it not only reflects the head final structure, but is also allowed in English data. At the later stage of language acquisition, the Japanese speaker comes to prefer (4b), which reflects the head initial parameter. This is possible because language transfer does not apply to separate phenomena, but to a whole parameter. That is, when Japanese learners acquire adjunct adverbial clauses, they have already acquired the head direction parameter from the basic word order of sentences.

In OT the result above can be interpreted as transfer or non-transfer of language particular ranking. Now, it is possible to show how the Constraint Demotion mechanism operates on the acquisition of the head direction parameter. With respect to phrase structures, the relative order among specifier, complement and head can be illustrated by the tree structures in (5).

(5) a. English, Spanish



b. Japanese



The head parameter illustrated in (5) can be reinterpreted by the universal OT principles. I suggest that the different head directions in (5) can be earned by the different rakings of the universal constraints. The universal constraints are given in (6).

(6) Spec-First: align specifier with the left edge of a phrase
Head-First: align head with the left edge of a phrase
Head-final: align head with the right edge of a phrase

When Spanish learners first confront English declarative sentences, they come to find that English word order is SVO like their mother tongue. The ranking of the constraints that produces SVO word order is illustrated in (7).

(7) English declarative sentence with respect to head direction

/study (John, syntax)	Spec-First	Head-First	Head-Final
[$_{spec}$John] [$_{head}$ [$_{head}$ study] [$_{com}$ syntax]] ☜		*	*
[$_{spec}$John] [$_{head}$ [$_{com}$ syntax][$_{head}$ study]]		**!	
[$_{head}$ [$_{head}$ study] [$_{com}$ syntax]] [$_{spec}$John]	*!		

The ranking in (7) is also relevant to relative clauses. In English the head noun always precedes the relative clause. This is the result of a ranking in which Head-First dominates Head-Final. Spanish speakers learning English do not need to set up a new ranking because their L1 ranking can apply to the L2 data. What Spanish learners need to do is realize that their L1 ranking can apply to the L2 data. On the other hand, in the same situation, Japanese learners will find that their L1 ranking never produces the correct output that matches English sentences because Japanese has SOV word order[3]. (8) illustrates the ranking that produces the SOV word order in Japanese.

(8) Japanese declarative sentence with respect to head direction

/study (John, syntax)	Spec-First	Head-Final	Head-First
[$_{spec}$John] [$_{head}$ [$_{head}$ study] [$_{com}$ syntax]]		*!	*
[$_{spec}$John] [$_{head}$ [$_{com}$ syntax][$_{head}$ study]] ☜			**
[$_{head}$ [$_{head}$ study] [$_{com}$ syntax]] [$_{spec}$John]	*!	**	

3) This realization and following learning process may be seriously influenced by individual learning circumstances, ability and motivation. These factors may result in the nonuniformity in L2 learners' achievement. The model proposed here does not cover this problem.

As illustrated in (8), the Japanese L1 ranking will not produce L2 English data. Thus, Japanese learners will try another way to set up the L2 English word order parameter. One possibility of learning English word order ranking is to cancel all L1 rankings with respect to word order. Since there is no conflict between structural constraints and faithfulness constraints with respect to head direction, all constraints will be located in one stratum. After that, as in L1 acquisition, Constraint Demotion will adjust the initial ranking to the target ranking. As a result, the starting point of Japanese learners for learning English word order will be the same as the initial state of children. The initial state for English word order illustrated in (9) exhibits no fixed ranking of the principles.

(9) Initial state for Head direction parameter:
{Spec-First, Head-First, Head-Final}

When Japanese learners encounter English data, they will realize that the winner in the data of the target language has SVO word order. Then the loser will be the other candidates such as SOV and VOS. (10) illustrates the loser and winner pair with respect to the relative order of subject and VP.

(10) Mark-data pair

loser < winner pairs	marks (loser)	marks (winner)
[$_{head}$ study syntax] [$_{spec}$ John] < [$_{spec}$ John] [$_{head}$ study syntax]	*Spec-First	*Head-First

According to Constraint Demotion, constraints of loser marks should out-rank the constraints of winner marks. Thus, Constraint Demotion forces Head-First to be demoted below Spec-First. The resulting hierarchy will be that in (11).

(11) {Spec-First, Head-Final} >>{Head-First}

Now another loser/winner pair is to be selected. Namely, learners need to set up the relative order between the head and its complements. This task

can be accomplished by analyzing VP internal structures of English sentences as shown in (12).

(12) Mark-data pair

loser < winner pairs	marks (loser)	marks (winner)
[$_{com}$ syntax] [$_{head}$ study] < [$_{head}$ study] [$_{com}$ syntax]	*Head-First	*Head-Final

Constraint Demotion again demotes Head-Final violated by the winner below Head-First violated by the loser. Thus the language learner can reach the totally ranked stage as shown in (13).

(13) { Spec-First} >> {Head-First} >> {Head-Final}

The fundamental difference between Japanese and Spanish speakers is that only Japanese speakers need to reset the head direction parameter for acquiring English head direction. Thus when both group have a similar level of fluency in English, Spanish speakers have less difficulty with English head direction than Japanese speakers. As we previously observed in adverbial adjunct clause test, Japanese speakers do not transfer their L1 property after they set up the ranking in which Head-First dominates Head-Final. This implies that L2 learners do not transfer each separate fact to the target language but they transfer the ranking to the target language.

3.2. Markedness theory

Liceras (1989) indicates that resetting the pro-drop parameter of Spanish L1 to the non pro-drop parameter in English L2 is more difficult than the reverse situation. The same observations are found in White (1985, 1986) and Phinney (1987). Based on these observations, Liceras and Phinney argue that the pro-drop parameter is an unmarked linguistic property.

Various proposals and ideas have been suggested regarding markedness phenomena, particularly in the principles and parameters theory of UG. For

example, Chomsky (1981, 1986) suggests that language learners learnthe core grammar on the basis of the principles and parameters of UG. Peripheral aspects are exceptions to the settings of core grammar or idiosyncratic features of the language. Thus, the marked peripheral aspects should be acquired later than the unmarked core grammar. As explained in section 2.2, in my approach, such properties can be explained by the initial unmarked ranking in which structural constraints dominate faithfulness constraints.

Regarding the pro-drop phenomenon, Grimshaw & Samek-Lodovici (1998) say that null subjects are possible only if they are licensed by an antecedent which is interpreted as a discourse topic. On the basis of this generalization, Grimshaw & Samek-Lodovici (1998) propose that the pro-drop phenomenon follows from the constraint DROPTOPIC requiring arguments whose antecedent is a topic to be unexpressed. Thus they conclude, "The cross-linguistic distribution of null subjects follows from the ranking of DROPTOPIC relative to the constraints SUBJECT and PARSE, which favor those competitors which express the thematic subjects overtly." The relevant constraints are given in (14).

(14) DROPTOPIC: Leave arguments coreferent with the topic structurally unrealized.
PARSE: Parse Input constituents.
SUBJECT: The highest A-specifier in an extended projection must be filled.

According to Grimshaw & Samek-Lodovici (1998), Inputspecifies the abstract information regarding a lexical head and a mapping of its argument structure onto another lexical head, including information about which phrases are foci and which are coreferent with the topic. Thus, if the subject of verb sing is coreferent with the topic, the Input will be that in (15) (Grimshaw & Samek-Lodovici, 1998).

(15) <sing (x), x=topic, x=John>

With respect to the Input in (15), pro-drop languages such as Spanish and Italian always satisfy DROPTOPIC. It means that thesubject coreferent with the topic does not appear in sentences in these languages. Contrary to pro-drop languages, non pro-drop languages such as English and French satisfy PARSE by sacrificing DROPTOPIC. The two tableaux in (16) show the rankings that result in the pro-drop and non pro-drop parameters, respectively.

(16) a. DROPTOPIC >> PARSE (Spanish, Italian)
Input: <sing (x), x=topic, x=he>

Candidate	DROPTOPIC	PARSE	SUBJECT
a. has sung ☜		*	*
b. he has sung	*!		

b. PARSE >> DROPTOPIC (English, French)
Input: <sing (x), x=topic, x=he>

Candidate	PARSE	DROPTOPIC	SUBJECT
a. has sung	*!		*
b. he has sung ☜		*	

According to Smolensky (1996), all systematic cross-linguistic variation is the result of reranking of universal constraints. Thus the generalization by Liceras (1989), White (1985, 1986) and Phinney (1987) can be reinterpreted in terms of reranking of constraints. That is, English L1 speakers easily acquire the new ranking in which DROPTOPIC dominates PARSE whereas Spanish speakers have some difficulty in acquiring the ranking PARSE >> DROPTOPIC. The reason for this asymmetric acquisition process is that only the ranking DROPTOPIC >> PARSE is close to the initial ranking inwhich structural constraints dominate faithfulness constraints. As stated in (14), DROPTOPIC is a restriction on the appearance of the topic related thematic argument. On the other hand, PARSE is a prohibition of the

omission of an element specified in Input on the surface expression. Consequently, DROPTOPC is considered as a structural constraint and PARSE as a faithfulness constraint. Now we can draw the initial state whose property is basically the same as that of the pro-drop languages.

(17) *initial state*: DROPTOPIC >> PARSE

When non pro-drop language speakers encounter the data from pro-drop languages, thev realize that their L1 ranking does not allow pro-drop. Then they adopt the initial state in (17), which is the same as the target state for the L2. Thus, non pro-drop L1 speakers do not need any learning to set the L2ranking producing null subject. However, when pro-drop language speakers confront the data from non pro-drop languages, neither their L1 ranking nor initial state guarantees the successful acquisition of the non-pro drop parameter. What is needed for pro-drop L1 speakers to learn a non pro-drop L2 parameter is to set a ranking in which PARSE outranks DROPTOPIC. This learning can be carried out by the Constraint Demotion. The difference between adopting the ranking of the initial state and learning a new ranking via the Constraint Demotion explains the asymmetric result observed by Liceras (1989), White (1985, 1986) and Phinney (1987).

One of the potential controversies arising from the explanation above would be whether the initial sate in (17) can be justified. Specifically, if the ranking in (17) is a conjecture, the explanation above will also be only a possibility. Of course, someone may appeal to the fact that subject omission is a pervasive phenomenon for children. However, such facts may be due to a simple processing limitation (Bloom 1990, Pinker 1984 and Valian 1991). The actual evidence for the proposed initial state in (17) comes from Hyams & Wexler (1993). They observe the proportions of lexical and pronominal subjects used by two children, Adam and Eve, during the developmental period. In a state in which Adam is no longer omitting subjects, the proportion of lexical to pronominal subjects is 1:2. At the same stagefor Eve, the proportion is 1:8. However, in a much younger null subject period, Adam shows a 3:1 ratio in favor of lexical subjects and Eve a 1:3 ratio in favor of pronominal. This experiment shows that younger Adam and Eve are using

more lexical subjects than pronominal subjects compared to older Adam and Eve. Here, the question is why the relative scarcity of pronouns in the younger Adam and Eve's speech takes place. According to Hyams &Wexler, the reason is that null subjects are used in most of the contexts where non-null subject languages use pronouns. Thus they argue that there exists a trade off between null subjects and pronouns if Adam and Eve are developing from a null subject stage to non null subject stage. Similar observations can be found in Wang et al (1992). Their experiments on English speaking children show that the chance of omitting a pronoun is greater than the chance of omitting a lexical subject. Another result of their experiment is that a younger child's chance of omitting a pronoun is greater than an older child's. This result matches with our prediction that DROPTOPIC dominates PARSE in the initial ranking. Pronoun subjects usually refer to the topic or the antecedent in the previous sentences. Then DROPTOPIC will restrict the appearance of pronoun rather than a lexical subject. It means that the initial stage that we have suggested in (17) has independent justification.

So far I have shown that the ranking producing the null subject has the property of the initial state. This explanation based on OT covers the phenomena that have been explained by the markedness hypothesis. Furthermore, this alternative explanation can avoid certain criticisms that may be inevitable in other frameworks. For example, Towell & Hawkins (1994) criticizes the basic idea of markeness hypothesis:

This strong implication that unmarked parameter values will be easier to set than marked ones has been found to raise a problem, however. Studies of children learning first language which have set both unmarked and marked parameter values have found that they do not necessarily develop the unmarked value first. For example, a study by French (1985:137) of the acquisition of pied-piping and preposition stranding in English by 3-5 year olds found no significant difference in the comprehension of stranding across the three age groups, suggesting that stranding is not late to emerge.

Although the above mentioned criticism may be correct, it cannot apply to

the unmarkedness definition based on OT. In the principles and parameters theory, the unmarkedness is defined in terms of parameters. That is, when one parameter is more inclusive than the other, the former is the unmarked one. For example, English allows both pied piping and preposition stranding, whereas French allows only pied piping. The pied piping can then be assumed as an inclusive value. However, such a markedness relation observed between pied-piping and preposition stranding is not maintained in the OT markedness definition. As suggested in Smolensky (1996), the unmarked structure in OT is the result of the ranking that structural constraints dominate faithfulness constraints. As we have alreadydiscussed, null subject parameter is one of the examples where the structural constraint DROPTOPIC dominates the faithfulness constraint PARSE. However, the difference between pied-piping and preposition stranding may not be drawn by the conflict between structural constraint and faithfulness constraint, but it is explained by the output structural differences. Consequently, the proposed acquisition model is not undermined by the criticism in Towell & Hawkins (1994).

In sum, if an L2 learner confronts a group of L2 data that differ from the property of their L1 grammar, they will not transfer their L1 ranking to the L2 grammar, but they will start their learning of L2 grammar from the initial state. The initial statehas a ranking producing the unmarked outputs in which structural constraints dominate faithfulness constraints.

3.3. Subset Principle

Another important theoretical apparatus in the research of the L2 acquisition is the subset principle, which is said to reveal children's learning mechanisms. The subset principle suggests that learners do not initially assume the over-inclusive grammar which requires negative evidence for disconfirmation (Berwick 1985, Wexler and Manzini 1987). This is because language acquisition must be done solely on the basis of positive evidence. Thesubset principle divides languages by subset and superset. For example, the CV syllable continuum discussed in section 2.2 violates no structural constraints such as No-Coda and Onset. Therefore, it is more likely to be

the initial property. Because all languages include CV syllable in their syllable structure inventory, the initial syllable structure becomes a subset of all other languages. Now let's look at the relationship between Japanese and English syllable structure. Both are supersets of the initial syllable structure. However, Japanese syllable structure becomes a subset of English syllable structure. The first reason is that Japanese does not allow complex onset and complex coda whereas English allows both of them. In other words, Japanese observes the structural constraint, COMPLEX, that prohibits the occurrence of the complex onsets. The second reason is that Japanese does not allow syllabic consonants whereas English allows syllabic consonants such as *l,r* and *m*.

At a glance, now we are in a contradictory situation. This is because Japanese syllable structure is more unmarked but less inclusive than English. Now the markedness theory makes us predict that English L1 speakers acquire Japanese syllable structure easier than Japanese L1 speakers acquire English syllable structure. On the other hand, the subset principle makes the opposite prediction because resetting superset to subset is difficult. This problem, however, does not occur in the OT based learning model proposed here. Surely Japanese has a more unmarked syllable structure than English in that in its syllable structure the structural constraints such as NO CODA and COMPLEX dominate some faithfulness constraints. On the other hand, to explain the English syllable structure, some faithfulness constraints should dominate NO CODA and COMPLEX. Suppose English L1 speakers are exposed to Japanese syllable structuresto learn. Will it be possible that they find any motivation for giving up their L1 ranking. This question is appropriate because English L1 speakers learning L2 Japanese can produce all of the Japanese syllable structures with their L1 ranking and will have no positive evidence that will reveal that their ranking is wrong. Consequently, although the Japanese syllable structure is close to the initial state and easy to learn, the English L1 speakers will not go back to the initial state.

The discussion above shows that the markendess theory does not wholly explain the L2 acquisition process. That is, when a language has a superset value of the target language, its speakers do not have to come back to the

initial ranking to acquire the target language. This is because their L1 ranking can be compatible with the data of the L2 regardless of whether it reflects the actual ranking of the target language. In this case, the L2 learners may eventually fail to set up the exact ranking of the target language.

4. Conclusion

The basic idea of OT acquisition theory is that language learning is a process of acquiring a language specific ranking of the universal constraints. I have shown that this idea can solve various problems that may arise fromL2 acquisition process. Specifically, OT language learning is facilitated by Constraint Demotion. Contrary to L1 acquisition, L2 acquisition starts from both initial stateand L1 complete ranking. When L1 ranking can produce the L2 data, language transfers are observed. As illustrated by head direction parameter, language transfer is a ranking transfer. The ranking transfer is not always successful. One example is the subset problem that has been explained in 3.3. For instance, English speakers do not set the ranking of the L2 Japanese syllable structure because English ranking is compatible with the Japanese syllable structure. However when L1 parameters significantly differ from L2 parameters, the learners do not start the learning from their L1 ranking but they go back to the initial statein which no ranking of the universal constraints exists or structural constraints dominate faithfulness constraints. The fact that null subject parameter is easier to acquire than obligatory subject parameter supports the idea. Thus in this OT based learning model the markedness principle and the subset principle are not in conflict. Instead, they explain the properties of L2 acquisition in cooperation.

The main benefit of this OT based L2 acquisition model is that it shows the specific interaction process of the universal components of language and L2 learning process. That is, it can show the explicit process in which learners' established ranking is reset to that of the target language. The second benefit to be considered is that it can solve some problems that the previous approaches may suffer from. As we have observed, the proposed

model reveals when the markedness theory plays a role or the language transfer takes place. Finally, it explains how the L2 acquisition is different from the L1 acquisition by suggesting the different initial state such as the established L1 ranking or adopting the learning mechanism such as the subset principle.

References

Berwick, R. (1985). *The acquisition of syntactic knowledge.* Cambridge, Mass: MIT Press.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? *Linguistic perspectives on second language acquisition.* (pp. 41-68). Cambridge University Press.

Bresnan, J. (1998). The Emergence of the Unmarked Pronoun II. Unpublished manuscript.

Chomsky, N. (1981). *Lectures on government and binding: the pisa lectures.* Dordrecht, The Netherlands: Foris.

Chomsky, N. (1986). *Knowledge of Language: its nature, origin, and use.* New York: Praeger.

Epstein, S., Flynn, S., & Matohardjono, G. (1996). Second language acquisition: theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences 19,* 677-758.

Flynn, S. (1996). A parameter-setting approach to second language acquisition. *Handbook of second language acquisition* (pp. 121-158). Academic Press.

Flynn, S. (1989). The role of the head-initial/head-final parameter in the acquisition of English relative clauses by adult Spanish and Japanese speakers. *Linguistic perspectives on second language acquisition* (pp. 89-108). Cambridge University Press.

Flynn, S. (1987). *A parameter-setting model of L2 acquisition: experimental studies in anaphora.* Dordrecht, The Netherlands: Reidel.

Gass, S. (1996). Second language acquisition and linguistic theory: the

role of language transfer. *Handbook of second language acquisition* (pp. 317-348). Academic Press.

Grimshaw, J. & Samek-Lodovici, V. (1998). Optimality theory and human sentence processing. *Is the best good enough?* (pp. 193-220), Mass: MIT Press.

Hyams, N. & Wexler, K. (1993). On the grammatical Basis of null subject in child language. *Linguistic Inquiry 24*(3), 421-459.

Liceras, J. M. (1989). On some properties of the "pro-drop" parameter: looking for missing subjects in non-native Spanish. *Linguistic perspectives on second language acquisition* (pp. 109-133). Cambridge University Press.

Phinney, M. (1987). The pro-drop parameter in second language acquisition. In T. Roeper and E. Williams (Eds), *Parameter setting* (pp. 221-238). Dordrecht, The Netherlands: Reidel.

Prince, A. & Smolensky, P. (1993). Optimality theory: constraint interaction in generative grammar. Technical Report, Rutgers University, New Brunswick, Nj, and University of Colorado, Boulder. Rutgers Optimality Archive.

Smolensky, P. (1996). The initial state and 'richness of the base' in Optimality theory. Technical Report, Cognitive Science Department, Johns Hopkins University. Rutgers Optimality Archive.

Tesar, B. & Smolensky, P. (1996). Learnability in Optimality theory (long version). Technical Report, Cognitive Science Department, Johns Hopkins University. Rutgers Optimality Archive.

Towell, R. & Hawkins, R. (1994). *Approaches to second language acquisition*. Clevedon: Multilingual Matters.

Wang, Q. & Lillo-Martin, D & Best, C & Levitt, A. (1992). Null-Subjects and objects acquisition of Chinese. *Language Acquisition, 2*, 221-254.

Wexler, K. and Manzini. M. (1987). Parameters and learnability in binding theory. T. Roeper and E. Williams (Eds), *Parameter Setting* (pp. 41-76). Dordrecht. The Netherlands: Reidel.

White, L. (1985). Is there a logical problem of second language acquisition? *TESL Canada Journal, 2*(2), 29-41.

White, L. (1986). Implications of parametric variation for adult second language acquisition: an investigation of the pro-drop parameter. In V. Cook, (Ed). *Experimental approaches to second language acquisition* (pp. 55-72). London: Pergamon Press.

Incheol Choi
Department of English Education
Kyungpook National University
1370 Sankyuk-dong, Buk-gu
Daegu 702-701, Korea
Phone: 053-950-5831
Email: incheol@knu.ac.kr

Received: 30 Mar, 2006
Revised: 2 Jun, 2006
Accepted: 19 Jun, 2006