Non-universality in Optimality Theory: An Information-based Model

Hye-Sun Cho
(Seoul National University)

Cho, Hye-Sun. 2003. Non-universality in Optimality Theory: An Information-based Model. *SNU Working Papers in English Language and Linguistics 2,* 171-187. This paper focuses on non-universality in standard OT. Non-universality in OT can be considered in two aspects: language-specific constraints and item-specific constraint ranking. Contrary to the basic assumption of OT, universality of constraints, there are language-specific non-universal constraint. In addition, applicability of the trisyllabic laxing rule implicates that different morphemes require different constraint ranking. Russell's morpheme-constraint model proposes abstract signature of morphemes and gatekeeper which controls different rankings for different morphemic items. This paper expands his idea and proposes a special component, termed Supervisor (SUP). Since operation of SUP is activated based on the information contained in input, information-contained input becomes critical in my proposal. My proposal thus proffers a solution for non-universality issue of standard OT. **(Seoul National University)**

Keywords: Optimality Theory, non-universality, specificity, morpheme, input, ranking, trisyllabic laxing

1. Introduction

This thesis investigates the innate problems that the standard Optimality Theory (OT: Prince & Smolensky 1993) cannot avoid due to its universalistic approach to phonological phenomena. Constraint universality is one of the basic assumptions of OT that causes controversial problems of specific constraints. In this paper, the meaning of specificity may cover the following cases: specific constraints for specific languages and specific constraint ranking for specific items.

Contrary to the assumption that all constraints are universal, some constraints are active only in a small number of languages. It is hard to see those constraints are also universal. Recently, researchers develop more and more constraints that are hard to see as universal constraints. We should consider how standard OT could include such language-specific, non-universal, specific constraints in its system.

In addition to language-specific constraints, we often see constraints ranked differently in one language for different morphological items. Let us assume that there is a constraint hierarchy which most morphemes follow. However, some other group of morphemes follows another hierarchy. Assuming the first group of morphemes is regular, the second group of morphemes may be seen as irregular. In this case, specificity means the idea of irregularity against regularity in a language.

This thesis attempts to give clarification about how OT may work for specific cases based on information of the input. It does not intend to deny OT entirely, but to give more clear description of the process. Admitting advantages of OT, the purpose of this thesis lies in adjusting the OT framework to be compatible with such idiosyncratic part as well as language-universal aspect.

Let us now briefly examine the architecture of OT. The OT architecture, introduced by Prince and Smolensky (1993), consists of a pair of functions, Generator (Gen) and Evaluator (Eval):

(1) The structure of OT grammar
$\mathrm{Gen}(\mathrm{in}_i)$ = {cand1, cand2}
Eval({cand1, cand2}) = out_{real}

Gen generates a candidate set from an input, which is in turn evaluated in EVAL by ranked constraints. Constraints are hierarchically ranked: Violation on higher-ranked constraints is fatal while violation of lower-ranked ones is tolerable. Candidate that has the least important violation will be selected as an optimal output form.

(2)

/Input/	Constraint 1	Constraint 2
☞ Candidate a		*
Candidate b	*!	

In the above tableau (2), both candidate (a) and (b) incur one violation of the constraint in question. However, candidate (a) is selected as an optimal output because it satisfies the higher-ranked constraint. Candidate (b) is suboptimal since violation of the higher-ranked constraint is fatal despite of satisfaction of the lower-ranked constraint.

One of the achievements of OT is that OT seeks to find out functional unity of phonological rules, unlike the previous rule-based phonology. Rule-based theory has separate rules to explain different phonological phenomena while Optimality Theory shows functional unity of the separate rules. Functional unity is expressed by constraints including phonological markedness constraints that are often phonetically-grounded. This is one of the advantages of optimality theory. By integrating the ultimate goal of separate rules into a constraint, we can have a new insight to analyze various phonological phenomena, and investigate the structure how the universal grammar is built. Archangeli points out the theoretical advantages of OT as opposed to rule-based phonology as listed in (3).

(3) **Advantages of OT** (Archangeli 1997:27)

1. It defines a clear and limited role of constraints.
   a. Each constraint is universal.
   b. Constraints are ranked in EVAL.
2. It eliminates the rule component entirely. Different constraint rankings in EVAL express language variability.
3. It focuses research directly on language universals. Each constraint is universal.
4. It resolves the ɔroblem. Universals dont play the same role in every language.

Despite these advantages over rule-based theory, researchers have pointed out problems challenging OT. Archangeli argues that OT eliminates the rule component entirely. However, as we will see in the coming chapters, some linguistic phenomena still need parochial rules such as r-intrusion rule of Eastern Massachusetts English (McCarthy 1993). Not all of the phonological phenomena can be caught within the seemingly well-organized net of OT framework. Blevins (1999: 228) also maintains, OT explains the core of universal grammar, but not the language-particular phonological rules that exist at the periphery. OT often fails to explain some language-particular and morpheme-specific phenomena within its standard framework. It is those exceptional cases that this thesis focuses on.

Some constraints developed to handle such cases cannot always be seen as universal, contrary to the basic assumption of OT, universality of constraints. If constraints are freely developed for each different language without any basis of universality, as argued in (3), the basic assumption of constraints universality will be threatened, and constraints will become no other than a different name of rules. McMahon (2000:9) also indicates the challenging aspects that undermine OT's basic structure as follows:

(4) **Challenges to OT** (McMahon 2000:9)

1. The interests, abilities and limitations of speakers, hearers and learners may lead to particular language-specific developments, and not all of these are reconcilable with a deeply universalist model like OT.
2. Current works have developed hundreds of constraints and mechanisms. However, they are not truly universal or explanatory unless the constraints themselves are controlled and restricted.
3. OT is in danger of failing to cope with the language-specific part.

It is unreasonable in the first place to integrate some obviously exceptional cases in one theoretical framework. It is natural to define a Universal Grammar (UG) on the one hand, and to deal with those peripheral phenomena on the other hand. The key point is not to forcefully include the exceptionality into universality, but to find out the effective way to handle those exceptionalities. Once OT can deal with exceptional cases properly, it will become more error-free and accurate system. This is not to deny OT as a model of Universal Grammar, but to extend it to handle the clearly exceptional cases. This thesis hopes to contribute OT in this light.

The organization of this thesis is as follows: In chapter 2, I discuss the problematic cases of specificity. Chapter 3 illustrates several OT models proposed in previous work, which are concerned with the problems of language-specificity. In chapter 4, I propose a Supervisor as a solution to those problems.

2. Discussion: Specificity

By Specificity, I mean an aspect of OT grammar that shows non-universal constraints and item-specific constraint set, as I have mentioned in chapter 1. Specificity significantly weakens one of the fundamental notions of OT that all constraints are universal. OT is clumsy in dealing with language-specific aspects of language even though successful with language-universal parts. Specificity in OT can be summarized as follows:

(5) **Specificity**
  a. Language-specificity : Non-universal constraints
  b. Item-specificity : Different rankings for different items

As we will see in this section, different hierarchy may be needed for different morphemes, and parochial constraints or even rules may be necessary for a specific language. Let us now examine the two aspects of specificity, starting with language-specific constraints.

2.1 Language-Specificity

The strongest interpretation of constraint universality is that all constraints are part of Universal Grammar. However, the universality should be given more flexible interpretation since all constraints are not equally active in all languages as Kager argues. Contrary to the assumption that all constraints are universal, some constraints are active only in a small number of languages. It is hard to see those constraints are also universal. Strong interpretation of constraint universality is shaking from the root as more and more empirical researches give examples militating against this fundamental notion. For example, look at the nominative truncation in Lardil (Prince & Smolensky 1993:100-101).

(6)

	Underlying Stem	Nominative	
a.	C Loss from stem		
	ŋ uk	ŋ u	'story'
	wuŋ unuŋ	wuŋ unu	'queen-fish'
b.	V Loss from stem		
	yiliyili	yiliyil	'oyster sp'
	mayar^{x}a	mayar^x	'rainbow'

In the example (6a), the final consonant is unparsed in nominative form. They propose the constraint Coda-Condition which means that a coda consonant can have only coronal place or else no place specification of its own at all. However, the final vowels are also lost in (6b), for which the constraint Free-V is proposed.

(7) **Free-V** : Word-final vowels must not be parsed (in the nominative)

However, Prince & Smolensky admits that it would not be one of the canonical examples of universal markedness constraints. Later in their discussion, they also argue that all the constraints they are proposing are universal and that only Free-V involves a language-particular idiosyncrasy. However, there are much more constraints that are language-specific than they expected. Look at the Nootka example in the following.

(8)

a.	Dorsals become labialised after round vowels		
	K→ ^{w} / o____	ʔ k^{w}i:ɬ	'making it'
		cf. ki:ɬ	'making'
b.	Syllable-final labiodorsals delabialize		
	K^w→ / ____.	ɬ l.š x	'to take pity on'
		cf. ɬ k^{w}iqnak	'pitiful'

To explain the above examples, McCarthy proposes that two constraints must be active in Nootka. Rounding prohibits plain dorsals after rounded vowels, while Unrounding rules out round dorsals syllable-finally.

(9) Constraints for Nootka
ROUNDING: *oK
UNROUNDING: $*K^{w}]_{\sigma}$

McMahon (2003: 19-24) argues that constraints such as ROUNDING and UNROUNDING in Nootka as well as Free-V in Lardil are all language-specific and rule-like. She maintains that apparently language-specific statements like this are presented as constraints, and it is becoming a common strategy of OT researchers. Maintaining that non-universal constraints do exist in current OT researches, she claims that the problem is to implicitly posit language-specific constraint while explicitly maintaining all constraints universal. In this sense, I propose that OT should clarify how to deal with such non-universal constraints.

2.2 Item-Specificity

This section shows how constraints are reranked based on different group of morphemes. Let us consider two rules: Trisyllabic laxing rule (TSL) and Flapping rule. TSL (10) makes a vowel lax when the vowel is followed by two syllables that are not stressed as in divine-divinity. Flapping rule (36) alters the intervocalic consonant with a flap sound when it follows a stressed vowel. Flapping rule always applies in the context of $V_1\_V_2$ (V_1 is accented) without exception as in (11a), (11b), and (11c).

(10) TSL (Trisyllabic Laxing Rule)
$\bar{V} \rightarrow \breve{V}$/ __ V_1CV_2 (where V_1 is not stressed.)

(11) Flapping rule: t→ / V_1___V_2 (V_1 is accented)
a. Morpheme-internally: a[D]om

b. Morpheme-boundary: mee[D]-ing
c. Phrasal level: wha[D] is wrong?

According to Kenstowicz (1994: 195), there is a critical difference between TSL and Flapping rule. While the latter is applied without exception, TSL is applied differently for each morpheme. The rule regularly applies when the suffixes [-ify], [-ual], [-ize], and [-ous] are attached to the stem as in (12).

(12) TSL-regular morphemes

[-ify]	vī	vĭ ify
	clē	clă ify
[-ual]	rī	rĭ ual
	grā e	gră -ual
[-ize]	tȳrant	tўrann-ize
	pē al	pĕnal-ize
[-ous]	tȳrant	tўrann-ous
	fā le	fă ul-ous

However, some morphemes do not undergo TSL. As in the following instances, the environment of ___CVCV is not sufficient to describe the application of TSL. Look at the following exceptions.

(13) TSL-irregular morphemes
a. Non-derived words: nī htingale, stē edore, ī ory
b. Derived-but-non-applicable: brā ery, mī htily, pī ting

The ___CVCV string of (13a) does not shorten the precedent vowel. With these examples, we can assume that application of TSL is confined to derived environment, in other words, when the stems are suffixed. However, we will soon recognize that this assumption is wrong because there are instances that TSL does NOT apply for derived words, as in (13b).

Let us see how it might be explained in Optimality Theoretic account. TSL shortens the long vowel before a string of two light syllables. In other words, it prevents a sequence of $\bar{V}\sigma\sigma$. Therefore

I propose the constraint *V̄σσ. The constraints for TSL are listed in (14):

(14) Constraints for TSL

a. ***V̄σσ**: Long vowel must not be followed by two syllables without accent.
b. **FtBin:** Feet are binary under moraic or syllabic analysis.
c. **NonFin**: No prosodic head is final in PrWd.
d. **Align-R**: The right edge of a Grammatical Word coincides with the right edge of a syllable.
e. **Parse-σ**: Syllables are parsed by feet.
f. **Max(μ** : Moras of input must be preserved in the output.

The hierarchy among these constraints must be as follows:

(15) Constraint ranking for TSL
*V̄σσ≫ FtBin, NonFin ≫ Align-R ≫ arse-σ Max(μ

(16)

/divi:n+ity/	*V̄σσ	FtBin	NonFin	Align-R	Parse-σ	Max(μ
a. dɪ i:(nɪ ɪ	*!				*!*	
b. (dɪ i:n)ɪ ɪ	*!			*	**	
c. (dɪ i)(nɪ ɪ			*			*
d. dɪ vi:.nɪ tɪ	*!			*	**	
e.☞ ɪ vɪ ɪ tɪ				*	**	*
f. (dɪ ɪ nɪ ɪ				*!*	**	*

The hierarchy reaches at the correct output form for divinity. The suffixes in (12), for which TSL regularly applies, follow the hierarchy as the following tableaux.

(17) [ify] Suffixation

/vayl+ifay/	*V̄σσ	FtBin	NonFin	Align-R	Parse-σ	Max(μ
a. (vay.lı fay	*!			*	*	
b. ☞ ı1 ı fay				*	*	*

However, the exceptions mentioned in (13a) and (13b) do not have the same results as in the following tableaux. "☜ ıeans incorrect output. Tableau (18) shows that the current hierarchy chooses the wrong candidate in case of non-derived words of (13a):

(18)

/ayvorı	*V̄σσ	FtBin	NonFin	Align-R	Parse-σ	Max(μ
a. (ay.vo).rı	*!			*	*	
b. ☜(ı 'o).rı				*	*	*

Tableau (19) shows that the current hierarchy chooses the wrong candidate in case of bravery. It is one of the examples in which TSL does not apply though they are derived words, which are listed in (13b).

(19)

/breyvə ı	*V̄σσ	FtBin	NonFin	Align-R	Parse-σ	Max(μ
a. (brey.və rı	*!			*	*	
b. ☜(bre.və rı				*	*	*

As we have seen so far, the constraint ranking (19) does not work for the words with such suffixes as *ly*, *-ing*, and *-ery*. From the discussion in this section, we can conclude that different morphemes may require a different constraint set.

3. Previous OT models

3.1 Prince & Smolensky (1993): Standard OT

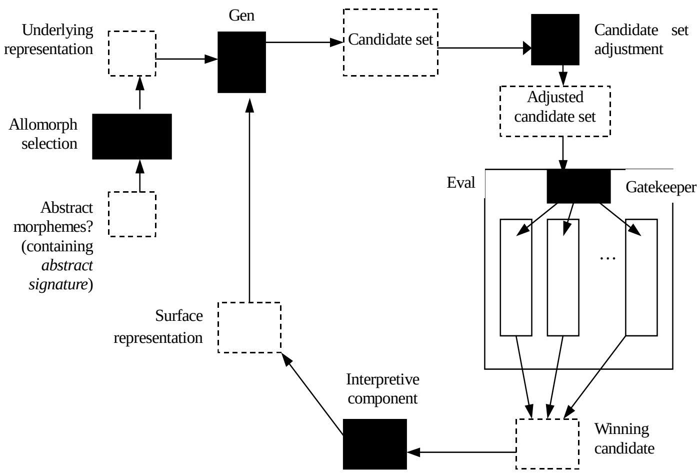
The Harmonic Parallelism of Prince & Smolensky (1993) is the standard OT model that provides the standard process introduced in chapter 1. Input goes into Gen to generate a candidate set, which is in turn submitted to Eva(luation). By Eval, the optimal output is selected.

(20)

3.2 Russell (1995): Morpheme-as-constraint approach

The following diagram (21) is Russells (1995) model. He addresses how to deal with the phonological information associated with morphemes, maintaining that morphemes are constraints that contain phonological information. The model consists of several other components that do not exist in the standard OT, such as Allomorph Selection, Candidate Set Adjustment, Gatekeeper, and Interpretive component.

(21)



Let us look at how each component works one by one. Allomorph Selection component chooses the appropriate morpheme to submit Gen based on some uniquely identifying index or diacritic feature, which is the collection of morphological information, called abstract signature of the morph. The most basic type of information comes in the form of notions such as stem, root, or suffix.

The output of Allomorph Selection is referred to as Underlying Representation (UR) in the standard OT. The UR goes into Candidate Set Adjustment and Candidate Set Adjustment produces the adjusted candidate set. Russell shows as an example possibility of using language-particular rule such as *r*-insertion rule to restrict candidate generation. However, greater part of the operation of Candidate Set Adjustment remains unexplained and requires further studies.

The surviving candidates from the processing of Candidate Set Adjustment are put through Gatekeeper before being evaluated. Gatekeeper is pre-processing step contained in Eval, and selects hierarchy for idiosyncratic morphemes. Assuming a single language have many different constraint hierarchies, the choice of which hierarchy should be used is sensitive to the lexical demands of individual morphs or sets of morphs.

Finally, Interpretive component makes some adjustment on the output thus produced. For example, unparsed material undergoes stray deletion, or epenthetic segments receive default specifications. However, it is also undefined how powerful this component may be.

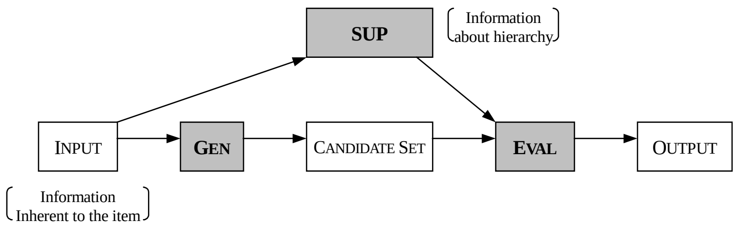
Russells model makes use of information or specification of morphemes in input level, which makes Gatekeeper select appropriate hierarchy. Through whole process, from input (Abstract morphemes) to Surface Representation, each step is controlled by each appropriate component. Contrary to Russells model, the standard OT does not make explicit how the specification of a morpheme can influence the constraint reranking.

4. My Proposal: Information-based model

4.1 Definition

Instead of employing various components as Russell proposes, I intend to connect abstract signature of the morphemes with Gatekeeper in a component like (22). It replaces the previously proposed components, termed Supervisor (SUP). Look at the diagram below for further discussion about the operation of SUP.

(22)



SUP controls constraint hierarchy based on the information of the input. It may vacuously operate for universal and regular phenomena that have no marked information in input. To function like this, SUP has necessary information about the constraint ranking of specific items. Based on the information of the input, SUP gives EVAL instruction to rerank constraint hierarchy just as Russells Gatekeeper, or control candidate generation to be properly restricted.

4.2 Recognition of Information-contained input

Since SUP gives instructions based on specific information of items, the information contained in the input becomes critical in my proposal. The information may include which language the lexical word comes from, or which type of suffixes the input has. The information may play a crucial role in determining the output.

In the model that I am proposing, it is inevitable that the role of the input becomes critical since the activity of SUP sets off according to the information of the input. There are significant instances showing language and morpheme specificity. Such cases are hard to explain without reference to the individual input.

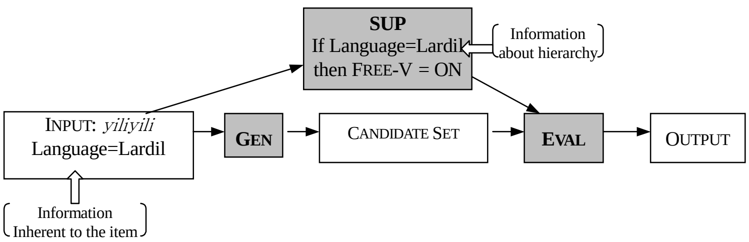
One might argue that depending on input is giving up the core principle of OT that is output-based. My model does not refuse output-based system. OT is an output-based system, yet we can refer to input when we cannot see the phonological motivation at the output level, such as non-surface-true opacity. The input holds the key for such cases.

4.3 Function

4.3.1 Language-specificity

As I have mentioned in 2.2, Lardil has a specific constraint Free-V. Input contains information that is inherent to the item (Langauge=Lardil), and SUP has the hierarchy information based on the information of the input (If Language is Lardil, then turn Turning on a constraint means that the constraint is active and significantly high-ranked in the language. Turning off a constraint means that the constraint is inactive and lowest-ranked in the language. on the constraint Free-V.)

(23)



4.3.2 Morpheme-specificity

SUP has the information about specific morphemes and respective rankings. The same constraint ranking that is used for TSL-applicable words does not work for the words with such suffixes as *-ly*, *-ing*, and *-ery*. They should be explained under different constraint set. Let us see what happens if we do not consider the highest-ranked constraint *V̄σσ.

(24)

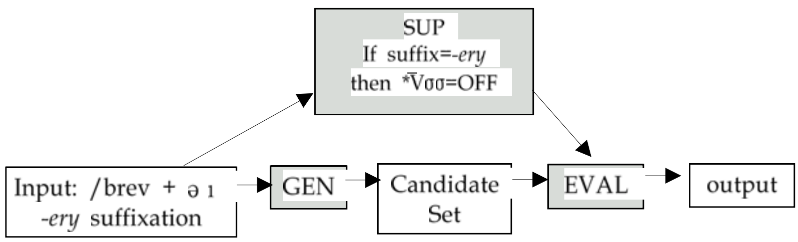
/breyvə ı	*V̄σσ	FtBin	NonFin	Align-R	Parse-σ	Max(μ
a. (brey.və rı	*!			*	*	
b. ☞(bre.və rı				*	*	*

Since brav-ery is not affected by the highest constraint *V̄σσ, candidate (a) is chosen as the optimal output because it satisfies Max(μ Therefore, the highest-ranked constraint *V̄σσ should not be active for the words suffixed with *ery*, *-ily*, and *-ing*. If the input is suffixed with one of the suffixes, SUP gives instruction to CON to make the highest-ranked constraint *V̄σσ inactive. The following tableau (25) includes the operation of SUP. The diagram (26) shows the function of SUP.

(25)

/breyvə ı	*V̄σσ	FtBin	NonFin	Align-R	Parse-σ	Max(μ
a. (brey.və rı	*!			*	*	
b. ☞(bre.və rı				*	*	*
SUP	OFF					

(26)



5. Conclusion

This thesis investigates the innate problems that the standard Optimality Theory. In this paper, the meaning of specificity covers both specific constraints for specific languages and specific constraint ranking for specific items. Contrary to the assumption that all constraints are universal, some constraints are active only in a small number of languages, such as the constraint Free-V in Lardil and constraint for Nootka (De)labialization. In addition to language-specific constraints, we often see constraints ranked differently in one language for different morphological items. Unlike Flapping rule which always applies in the context necessary to the rule, TSL (Trisyllabic laxing rule) shows irregular application based on types of morphemes.

To specify morpheme-specific hierarchies, Russells model proposes additional components, among which Gatekeeper adjust constraint hierarchies based on the morphemes abstract signature. Based on his model, I try to integrate abstract signature of the morphemes and the component Gatekeeper into one component, SUP. SUP controls constraint hierarchy based on the information of the input. It not only reranks constraints based on input morphemes, but also brings specific constraints for each specific language. Since SUP gives instructions based on specific information of items, the information contained in the input becomes critical in my proposal. I termed my proposal as information-based model because input has inherent information of the input items, and SUP also consists of information about which hierarchy EVAL may use.

References

Archangeli, Diana (1998) Optimality Theory. Blackwell: Oxford.

Archangeli, Diana and Keiichiro suzuki (1997) he Yokuts Challenge In I. Roca (ed.), Derivations and Constraints in Phonology. Clarendon Press: Oxford.

Kager, Rene (1999) Optimality Theory. Cambridge University Press: NY.

Kenstowicz, Michael (1994) Phonology in Generative Grammar. Blackwell: Massachusetts.

Lee, Jae-Young (1996) Some Aspects of English Phonology: An Optimality Theoretic Approach. Ph.D. dissertation. University of Illinois.

_____________ (2002) Suffix Classes and Stress Patterns in English The Journal of English Language and Literature 48 (1). 215-241.

The English Language and Literature Association of Korea.

McCarthy, John J., and Alan S. Prince (1993) Prosodic Morphology I: Constraint Interaction and Satisfaction. Ms., University of Massachusetts, Amherst, and Rutgers University.

McMahon, April (2000) Change, Chance, and Optimality. Oxford University Press: New York.

Mohanan, K. P. (1982) Lexical Phonology. Indiana University Linguistics Club: Bloomington.

Myers, James (1999) Lexical Phonology and the Lexicon Studies in Phonetics, Phonology and Morphonology 25. 74-170. The Phonology-Morphology Circle of Korea.

Paradis, Carole and Darlene Lacharité (1993) Introduction: The Emergence of Constraints in Generative Phonology and a Comparison of Three Current Constraint-Based Models Canadian Journal of Linguistics 38 (2). 127-153.

Prince, Alan and Paul Smolensky (1993) Optimality Theory: Constraint Interaction in Generative Grammar. Rutgers University: Piscataway.

Russell, Kevin (1995) Morphemes and Candidates in Optimality Theory. University of Manitoba.

Smolensky, Paul (1993) Harmony, markedness, and phonological activity. Handout to talk presented at Rutgers Optimality Workshop 1, 23 October 1993, New Brunswick, N.J. [ROA-87, http://ruccs.rutgers.edu/roa.html]

Hye-Sun Cho
rubylily@knou.ac.kr