Optimality Theory and Hierarchy Construction

GIOVANA FERREIRA GONÇALVES BONILHA
CARMEN LÚCIA BARRETO MATZENAUER

Abstract

Arising from the application of Optimality Theory (OT) to issues of language acquisition, the term "stratum" may represent (i) a single constraint, and therefore the strict ranking proposed by standard OT, as shown especially in research using data of the target language, or (ii) a group of constraints which do not present a fixed ranking among themselves. This kind of stratum has been often interpreted as a mere grouping of constraints without any relation of dominance, the optimum output being that which violates the smallest number of constraints in the stratum. However, the way the theoretical model works allows for a different interpretation, based on the concept of "floating" ranking, which may be responsible for variable outputs. Based on the phonological acquisition of Brazilian Portuguese, this paper discusses the several interpretations which may be attributed to the notion of "stratum" in the construction of constraints hierarchy and establishes connections with phonological variation in adult speech.

In Optimality Theory the notion of constraints hierarchy and, consequently, of dominance are essential to the model, the relevance of these notions arising from the possibility of violating constraints. Violability is thus important because it affects one of the basal points of the formal model – the constraints – which, along with Gen and Eval, constitute Universal Grammar (UG). It is exactly due to the fact that constraints can be violated that every language in the world can have its specific ordering of universal constraints. Being violability one of the properties that characterize OT, in spite of its necessarily minimal nature, according to McCarthy & Prince (1993:05), the idea of dominance must be fundamental for the theory. It is based on minimal violation, that is, on the violation of constraints that are lower in the ranking of a specific language that an output is considered optimal among all the candidates provided by Gen.

The process of acquiring a language, according to OT, implies the acquisition of the ranking of constraints which characterizes it. Following the

 ISSN 1645-4537

learning algorithm proposed by Tesar & Smolensky (2000), the child starts from an initial stage in which the constraints of markedness dominate those of faithfulness – this hierarchy being responsible for the choice of outputs with unmarked structures and segments. Linguistic development occurs by the demotion of markedness constraints, which generates different grammars until the acquisition of the target system.

Since the constraints demotion process can motivate the formation of strata made of either a group of constraints or just one constraint, one needs to question whether the stratum – especially when composed by a group of constraints – should be interpreted as (i) just one constraint and, therefore, subjected to the strict ordering proposed by standard OT, or (ii) a group of constraints without any fixed ranking among them.

Analyses of how different languages function have interpreted the stratum as a single constraint, that is, as a group of constraints which do not present a relation of dominance among them, the choice of optimal output being determined by the total number of violations of the constraints that form it. As McCarthy (2002:205) puts it, the ranking returned by the constraint demotion learning algorithm is a stratified partial order: constraints are grouped into blocks, called strata; strata are ranked relative to other strata; but constraints within a stratum are nonconflicting and therefore unrankable and unranked.

An example of this interpretation can be seen in (1), in which candidate (a) is chosen as optimal because it violates only once one of the three constraints that share the stratum – C3, C4, C5.

(1) Tableau 1

/input/	C1	C2	C3	C4	C5
☞ a) cand1			*		
b) cand2				*	*
c) cand3		*!			

In this paper, however, another possible reading is suggested: a stratum that shares constraints can reveal, in fact, the possibility that its constituting constraints present a “floating” ordering, very much like what was suggested by Antilla (1995) in terms of adult variable data. With this new interpretation attributed to a “complex stratum” – one that places constraints into groups – the tableau shown in (1) can have a new reading, as shown in (2).

(2) Tableau 2

/input/	C1	C2	C3	C4	C5
☞ a) cand1			*		
☞ b) cand2				*	*
c) cand3		*!			

In (2), candidates (a) and (b) are potential optimal outputs, that is, the choice between them will depend on the ranking that the floating constraints present in the stratum. If one considers this "floating possibility" of the constraints in the stratum, in a given production attempt, for example, C3 can be dominating C4 and C5, and this ranking will determine the selection of candidate 2 as optimal form, and in a different production attempt, C4 and C5 can be dominating C3, with candidate 1 being chosen as optimal form.

This proposal seems to corroborate the variation found in phonological acquisition data. When inquired about the pertinence of this new proposal, McCarthy said he believed that if the child still does not know the correct ordering that a specific number of constraints have in her language, she will get an ordering in one production and another one in a different production at random.

It must be pointed out that the term "floating" used here should not be mistaken for the one proposed by Reynolds (1994) for adult variation data. For Reynolds only one constraint or a delimited group of them – within a stratum with strict domination – can change positions in the hierarchy. Observe (3).

(3) Figure 1

{C1}>>{C2}>> { ......C4...... C3>>C5 }

According to (3), only C4 can change positions in the stratum, as the dominance relation between C3 and C5 must be maintained. Reynolds (1994), when analyzing adult data based on a total ranking of the constraints, does not emphasize the strata that share constraints. However, if one considers that probably even in adult grammar not all the constraints present a dominance relation, it can be inferred that in Reynolds' proposal strata that share constraints can still be read as choosing the candidate that violates the smallest number of constraints.

In order to demonstrate the validity of the reading of constraints that share a stratum shown in (2), we will take as a basis phonological acquisition data and the learning algorithm proposed by Tesar & Smolensky, since in the acquisition process the use of the algorithm may result in the formation of the strata that share constraints. According to Tesar & Smolensky (2000), the demotion process in language acquisition should always be minimal, that is, a constraint is demoted as high as possible in the hierarchy, even if this implies its positioning in a stratum already occupied by a constraint. Let us observe in (4), (5) and (6) an example of the formation of a stratum which shares constraints, taken from a study of the acquisition of falling diphthongs in Brazilian Portuguese (Bonilha 2000).

(4)
Hierarchy H1
{NotComplex (nucleus), NoCoda}>>{ Max, Dep}>>{Onset}

Starting from Hierarchy H1 – which allows for the production of CV and V syllables in Portuguese –, the use of the learning algorithm establishes that the constraint Not complex (nucleus) must be demoted below faithfulness constraints so that a diphthong can be produced by the learner. The demotion of Not complex (nucleus) is made minimally, that is, there is no creation of a new stratum, for the analysis of suboptimal/optimal pairs according to (5) does not determine that Not Complex (nucleus) be dominated by Onset. According to the workings of the algorithm, the constraint violated by the optimal candidate – Not Complex (nucleus) – must be demoted below the constraint violated by the suboptimal candidate – Max I/O and Dep I/O –, if we consider the pairs *b<a* and *c<a*, respectively. The constraints NotComplex (nucleus) and Onset can, therefore, share the stratum.

(5) Constraints violated by suboptimal/optimal pairs *pa.pa* < *pa.paj* ‘daddy’ e *pa.pa.pi* < *pa.paj*

Loser < winner		Loser-marks	Winner-marks
b<a	pa.pa < pa.paj	MAX I/O	Not Complex (nucleus)
c<a	pa.pa.pi < pa.paj	DEP I/O	Not Complex (nucleus)

The ordering in (6) with NotComplex (nucleus) sharing a stratum allows for the emergence of syllables CVV in Portuguese.

(6)
{NoCoda}>>{ Max, Dep}>>{Onset, NotComplex (nucleus)}

To work with acquisition data demands an appropriate reading of the strata that share constraints, since differently from the analyses of the target form – in which most strata are formed by a single constraint –, the researcher will constantly have to face this kind of stratum as a result of the process of hierarchy construction in the target language on the part of the child.

Grijzenhout & Joppen (2000), in a study about the early stages of syllable acquisition in German, interpret the stratum that shares constraints according to the example shown in (1), in which constraint violations are counted, as can be observed in (7).

(7) Tableau 3

/a:p /	C-Place	V-Place	Onset	*Struc
a) a:	*		*	*!*
b) a:p			*	**!*
☞ c) ∅	*	*		
d) pa:				***!

Independently of discussing here the appropriateness of the constraints used by the authors in their analyses, it must be observed that the third candidate – zero production – is chosen because it presents two violations, as against the four violations presented by candidates (a) and (b). The optimal candidate here is, therefore, chosen because of the total number of violations, in the same way that happened in (1).

Going back to the idea that one of the central elements of Optimality theory is exactly the ordering of constraints, that is, the candidates are chosen as ideal outputs because they only violate constraints that are dominated by others – the total number of violations should be considered only when two candidates tie in relation to a constraint, as shown in (8); if this does not occur, the standard choice of the optimal form is established by the ordering of the constraints.

(8) Tableau 4

/input/	C1	C2	C3
a) cand1		*	**!
☞ b) cand2		*	*
c) cand3	*!		

In the *Tableau* in (8), candidates (a) and (b) were tied in terms of the violations of C1 and C2, and in this case the total violation count was necessary to establish the optimal output.

The theory, however, privileges the dominance relation between constraints and not the total number of violations, a principle which the present study does not intend to deny. Our proposal is motivated by the idea that, when faced with the partial rankings that emerge from the acquisition data – when the child is building its hierarchy –, the researcher needs a reading that does not affect the principles of the theory.

What is being defended in the current research, therefore, is a reading of the stratum that shares constraints that is coherent both with the essence of the theory and with the variation presented by the learner in the gradual process of language acquisition.

According to this proposal – that the constraints that share a stratum can have the position altered within this domain reflecting dominance relation –, in (7) only the second candidate could not be chosen as optimal form because, considering the potential rankings, three candidates could be chosen at a specific production phase. Let us now observe the tableaux in (9), considering that the constraints that share a stratum can potentially change positions within the stratum.

(9) a. Tableau 5

/a:p /	Onset	*Struc	C-Place	V-Place
a:	*!	**	*	
a:p	*!	***		
☞ ∅			*	*
pa:		*!**		

b. Tableau 6

/a:p /	V-Place	C-Place	Onset	*Struc
a:		*!	*	**
a:p			*!	***
∅	*!	*		
☞pa:				***

c. Tableau 7

/a:p /	V-Place	*Struc	C-Place	Onset
☞a:		**	*	*
a:p		***!		*
∅	*!		*	
pa:		***!		

As we can see in (9a), if the ordering at a specific production stage is Onset>>*Struc>>C-Place>>V-Place[1], the optimal candidate will be the third one, without any form being produced for the target [a:p] 'monkey'; in (9b), the potential ordering change V-Place>>C-Place>>Onset>>*Struc allows for the emergence of the last candidate as optimal form; in (9c), with the ordering V-Place>>*Struc>>C-Place>>Onset, the first candidate is chosen as ideal output.

Tesar (2000) points out the difficulty of working with stratified hierarchies when using the learning algorithm, despite the fact that the algorithm is paradoxically responsible for the construction of the strata that share constraints.

According to the author, it is necessary to establish a reading for this kind of stratum in order to allow for the relative harmony of a pair of candidates. The reading used is the one referred in (1), with the total number of violations made by the candidates; in this case, if two candidates present the same number of violations, the next stratum is the one that will be considered to define the ideal output.

[1] This dominance exists as a possibility once constraints are floating; for this reason, the lines used in the tableau are dotted.

Even though Tesar assumes this position, he points out that this kind of reading does not always work for a learner trying to reach the target hierarchy for a specific production. Another possible reading would be the one in which the candidates would be tied, that is, two candidates would be considered optimal because one constraint would choose candidate (a) and another one would choose candidate (b), as seen in (10).

(10) Tableau 8

	C1	C2	C3	C4
☞Cand (a)		**		
☞Cand (b)			***	

The example in (10) shows that the total number of violations does not choose the optimal candidate because, in this case, only candidate (a) would be the chosen output. However, the author's proposal (2000:26) is limited, since it has hypothetical examples that consider only two constraints in the same stratum without mentioning the stratum formed by many constraints as in (7): two candidates have conflicting outcomes on a stratum if one of the constraints of the stratum prefers one candidate, while another of the constraints in the stratum prefers the other candidate.

The question presented by Tesar opens up the possibility for the present proposal based on the learning algorithm, which, considering the total of constraints, has not been successful in dealing with strata which share constraints. The new model proposed here is able to explain the variation shown in the data of the same child in the same phase of phonological development. The linguistic production of Bruno (1:2,10 (years: months, days)) is an example: in the same data gathering, it presents the variable form [papu] ~ [pako] for the input */pato/* ‘duck’. The explanation for the choice of different outputs as variable forms can result from the different ordering attributed to constraints that share the same stratum and which can “float” in it, as shown in (11).

(11) Tableau 9

/pato/	*[coronal]	Ident I/O (coronal)	Ident I/O (labial)	Ident I/O (dorsal)	Max I/O	Onset
a) pa.tu	*!					
☞b) pa.pu		*	*			
☞c) pa.ku		*		*		
☞d) pa.u					*	*

In the period of one hour of recording, the subject evidences the possibility of producing a CV target by using the repair strategy CV→CV[2], with substitution of the segment positioned in onset (*[papu] ~ [paku]*). If the choice of the optimal candidate did not consider a dominance relation among the constraints that share the complex stratum – {Ident I/O (coronal), Ident I/O (labial), Ident I/O (dorsal), Max I/O} –, candidate (d) for input /pato/ would emerge, as it violates just Max constraints, while candidates (b) and (c) violate two constraints. However, if it is postulated that the constraints that share a stratum have the potentiality of changing positions in the hierarchy, producing a dominance relation, candidates (b) and (c) can be chosen as optimal, according to (12).

(12) a. Tableau 10

/pato/	*[coronal]	Max I/O	Ident I/O (labial)	Ident I/O (coronal)	Ident I/O (dorsal)	Onset
a) pa.tu	*!					
b) pa.pu			*!	*		
☞ c) pa.ku				*	*	
d) pa.u		*!				*

b. Tableau 11

/pato/	*[coronal]	Max I/O	Ident I/O (dorsal)	Ident I/O (coronal)	Ident I/O (labial)	Onset
a) pa.tu	*!					
☞ b) pa.pu				*	*	
c) pa.ku			*!	*		
d) pa.u		*!				*

It is interesting to note that if we considered that Max I/O was ranked higher than Ident constraints, the candidate (d) in (12) would not have been selected as optimal form and the example in (12) would not be an evidence of a "floating ranking". Therefore, it is necessary to mention that the presence of Max I/O in the same stratum of the Ident family constraints is justified by the fact that the subject has produced, in the same data gathering, the output [a.o] for the input /alo/ 'hallow'. See tableau in (13).

(13) Tableau 12

/alo/	*[coronal]	Ident I/O (labial)	Ident I/O (dorsal)	Ident I/O (coronal)	Max I/O	Onset
a) a.lo	*!					
b) a.ko			*!	*		
☞c) a.o					*	*

[2] CV → CV is considered a repair strategy when there is an exchange of segments.

If Max I/O were positioned higher than the mentioned complex stratum, the output [a.o] effectively produced by the child would not be chosen.

It should be observed that the tableau in (11) introduces the constraints Ident I/O(coronal), Ident I/O(labial) and Ident I/O(dorsal) that are interpreted in the literature (McCarthy & Prince (1995:226)), as shown in (14).

(14)
Let α be a segment en S_1 and β be any correspondent of α in S_2.
If α is [γF], then β is [γF].

According to McCarthy, this definition of Ident was based on the binary features originated in classic generative phonology. From this theoretical position, the formulation of Ident referred to in (14) can explain the proposed analyses. However, when considering privative features, it becomes necessary to reread Ident according to the tableau in (11), in order to contemplate not only the movement input $\rightarrow$ output, but also that of output $\rightarrow$ input, since this kind of feature will or will not be present in the segment – whether it be from input or from output. In the present paper, therefore, a new definition for Ident is suggested, as seen in (15).

(15)
Let α be a segment en S_1 and β be any correspondent of α in S_2.
If α is [γF], then β is [γF]; if β is [γF], then α is [γF].

The new definition suggested in (15) explains why candidate [papu] violates not only the constraint Ident I/O(coronal), but also Ident I/O(labial) and the candidate [paku] violates both Ident I/O(coronal) and Ident I/O(dorsal).

Another example of variation that is often seen in phonological acquisition data refers to the use of the segments /s/ and /ʃ/ for the palatal fricative of Portuguese. This variation can be explained by the choice of two possible outputs as a consequence of the fact that two markedness constraints operating in the language (*Coronal/-anterior and *Coronal/+anterior) still share the same stratum and thus dominate faithfulness constraints. The dominance alternating between one markedness constraint or another, as shown in (16a) and (16b), is responsible for the varying forms present in the children's speech. This grammar allows for the choice of the two outputs independently of the type of coronal fricative that appears in the input (Matzenauer 2001).

(16) a. Tableau 13

/ʃave/	*COR/–ANT	*COR/+ANT	IDENT-IO(ant)
☞ a) savi		*	*
b) ʃavi	*!		

chave ′key’

b. Tableau 14

/ʃave/	*COR/+ANT	*COR/–ANT	IDENT-IO(ant)
a) savi	*!		*
☞ b) ʃavi		*	

Actually, the way floating constraints work can be observed in the process of phonological stabilization of different pertinent features in the language system. In the process of acquisition of Brazilian Portuguese, variant features are frequently produced by the child in a single target segment of the same data gathering, as the examples below (in 17) illustrate.

(17)

a) Target segment [v] (Lara – 2:0)

	adult form	child form
livro 'book'	[‘livru]	[´iʃu]
vela 'candle'	[‘vɛla]	[‘dɛla]
vamos 'we go'	[‘vãmus]	[‘vãmu]

b) Target segment [s] (Maria – 2:0)

	adult form	child form
céu 'sky'	[‘sɛw]	[´tɛw]
esse 'this'	[´esi]	[´eʃi]
massa 'paste'	[´masa]	[´masa]

c) Target segment [g] (Vitor – 2:1)

	adult form	child form
gato 'cat'	[‘gatu]	[‘katu]
gatinho 'kitty'	[ga’tʃĩɲu]	[ga’tʃĩɲu]
garfo 'fork'	[´garfu]	[´dafu]

d) Target segment [k] (Paulo – 2:4)

	adult form	child form
cabelo 'hair'	[ka’belu]	[ta’belu]
carro 'car'	[‘kaRu]	[‘kaRu]
cachorro 'dog'	[ka’ʃoRu]	[a’soRu]

By means of the “floating” ranking, it can be possible to predict forms emerging from variation in the phonological acquisition – especially those that minimally violate constraints that share the same stratum; the ordering presented by these constrains in a specific moment of the linguistic production will define the chosen output.

The “floating” ranking can also appropriately explain variation in adult speech. As an example of this possibility, let us take the analysis proposed by

Hora (2002) for the variable production of the coronal fricative in coda as [s] ~ [h] ~ [∅], which considers the functioning of "floating" constraints having the above mentioned proposal by Reynolds (1994) as a basis. Considering this variation of the Brazilian Portuguese in the light of the reading model of complex stratum proposed here, one may begin to understand it as a consequence of constraints that share the same stratum and that, depending on the ranking they present, will be responsible for the variant that will be used. The example is in (18)[3].

(18) a. Tableau 15

Candidates	*PARSE/ Fricativa	PARSE-RN	PARSE-PN
a) mes.mo	*!		
☞ b) meh.mo			*
c) me<s>.mo		*!	

mesmo 'same'

b. Tableau 16

Candidates	*PARSE/ Fricativa	PARSE-PN	PARSE-RN
a) mes.mo	*!		
b) meh.mo		*!	
☞c) me<s>.mo			*

c. Tableau 17

Candidates	PARSE-PN	PARSE-RN	*PARSE/ Fricativa
☞a) mes.mo			*
b) meh.mo	*!		
c) me<s>.mo		*!	

What is proposed here, therefore, is that even constraints grouped in a same stratum present a dominance relation among them, but with a particularity: the dominance among these constraints can be "floating". The possibility of "floating" constraints would be a property of complex strata, and it is this property that responds for the variable forms present in the language acquisition process, and also for variation in adult speech.

The present proposal introduces the possibility of generating different output forms from a single input in a single grammar, which must predict a constraint

---

[3] (Hora 2002) makes use of the constraints PARSECoda-RN (the root node of a segment in coda is linked by the mora) and PARSECoda-PN (the root node of a segment in coda is linked by RN).

hierarchy with two kinds of dominance relation – strict and “floating” –, the domain of the “floating” constraint being delimited by the complex strata, that is, those strata that group constraints and that are established by the demotions caused by the learning algorithm in the process of language acquisition. In this sense there seems to be an intrinsic relation between the variation in the acquisition process and the variation in adult speech.

Acknowledgements

We would like to thank John McCarthy and Gisela Collischonn for their comments and suggestions. We remain, however, fully responsible for any problem that may have remained.

References

Antilla, A. (1995) *Deriving Variation from Grammar: a Study of Finnish Genitives*. (available on Rutgers Optimality Archive).

Bonilha, Giovana F.G. (2000) *Aquisição dos ditongos orais decrescentes*: uma análise à luz da Teoria da Otimidade. Master Dissertation. Pelotas: UCPEL.

Grijzenhout, J. & S. Joppen (2000) *First Steps in the Acquisition of German Phonology: A Case Study.* (available on Rutgers Optimality Archive).

Hora, D. (2002) Teoria fonológica e variação: a fricativas coronal /s/. *Letras de Hoje* **37** (1), 199-219.

Matzenauer-Hernandorena, C. L. (1990) *Aquisição da Fonologia do Português: estabelecimento de padrões com base em traços distintivos*. PhD Dissertation. Porto Alegre: PUCRS.

Matzenauer-Hernandorena, C. L. (2001) On the acquisition of fricatives in Brazilian Portuguese. In *Proceedings of the GALA'2001 Conference on Language Acquisition* (J. Costa & M. J. Freitas, editors). Lisboa: APL.

McCarthy, J. and A. Prince (1993) *Prosodic Morphology*. New Brunswick: Rutgers University Center for Cognitive Science.

McCarthy, J. and A. Prince (1995) Faithfulness and Identity in Prosodic Morphology. In *The Prosody Morphology Interface* (R. Kager & Zonneveld, editors). Cambridge: Cambridge University Press.

Reynolds, W. T. (1994) *Variation and Optimality.* PhD Dissertation. Pennsylvania: University of Pennsylvania.

Tesar, B. (2000) *Using Inconsistency Detection to Overcome Structural Ambiguity in Language Learning* (available on Rutgers Optimality Archive).

Tesar, B. and P. Smolensky (2000) *Learnability in Optimality Theory.* Massachusetts: MIT Press.

Giovana Bonilha
Universidade Católica de Pelotas
Brasil
gfgb@terra.com.br

Carmén Lúcia Matzenauer
Universidade Católica de Pelotas
Brasil
carmenluc@terra.com.br