GRADIENT PHONOTACTICS AND THE COMPLEXITY HYPOTHESIS

ABSTRACT: Lexical items can be more or less well-formed depending on the phoneme combinations they contain. This phenomenon is called gradient phonotactics. We propose an approach to gradient phonotactics based on Optimality Theory (Prince and Smolensky 1993/2004). At the heart of the proposal is the Complexity Hypothesis that attributes the relative well-formedness of a lexical item to its relative grammatical complexity measured in terms of ranking information: the more complex the lexical item, the less well-formed it is. The theory orders linguistic structures in an implicational hierarchy that reflects their relative well-formedness. Some implications are universal; others depend on language-specific rankings. The Complexity Hypothesis is supported by phonotactic data from Muna (Austronesian) as recently analyzed by Coetzee and Pater (2008).

1. INTRODUCTION

PHONOTACTICS is the study of permissible and impermissible phoneme combinations in a language. It has often been noted that phonotactic principles appear to be GRADIENT: lexical items can be more or less well-formed depending on the phoneme combinations they contain.

The gradience of phonotactics emerges in at least two ways. First, some types of lexical items are statistically overrepresented, others statistically underrepresented, depending on their phonotactic structure. For example, in Arabic, there is a well-known dissimilatory constraint against homorganic consonants in adjacent positions within the verbal root (Frisch, Pierrehumbert, and Broe 2004; Greenberg 1950; McCarthy 1988, 1994; Coetzee and Pater 2008; Pierrehumbert 1993): the more similar the consonants are, the less commonly they co-occur in actual lexical items. For similar gradient generalizations in other languages, see e.g. Berkley 1994a, 1994b, 2000; Coleman and Pierrehumbert 1997; Hammond 2004; Hay, Pierrehumbert, and Beckman 2004 (English); Pater and Coetzee 2005; Coetzee and Pater 2006, 2008 (Muna). Second, it has been observed that novel words ("wug words") show gradient acceptability that depends on their phonotactic structure. Thus, speakers of English judge nonsense words like *stin* to be rather good, *smy* to be less good, and *bzarshk* to be rather bad (Albright 2006). For similar effects, see Bailey and Hahn 2001; Coleman and Pierrehumbert 1997; Frisch, Large, and Pisoni 2000; Frisch and Zawaydeh 2001; Greenberg and Jenkins 1964; Ohala and Ohala 1986; Vitevitch, Luce, Charles-Luce, and Kemmerer 1997, among others.

Gradient phonotactic generalizations are a challenge for phonological theory. How should such generalizations be explained? There are two main possibilities. The first possibility is a GRAMMATICAL EXPLANATION: phonological grammars are formalized in such a way as to predict the relative likelihoods of segment combinations in terms of their relative markedness, perhaps stated over natural classes. Some segment combinations would be so ill-formed as to end up being absolutely ungrammatical, while others would be more or less grammatical along different dimensions of markedness. The second possibility is a LEXICAL EXPLANATION: gradient judgments arise by consulting the lexicon. On this view, a novel word would derive support from existing words depending on the number of its lexical neighbors, defined in terms of e.g. string edit distance

(Kruskal 1983), which may be weighted by lexical token frequency, similarity, etc. An extreme version of such a model would deny the phonological grammar any role in gradient well-formedness judgments. The speakers would simply consult the existing lexical items in judging the well-formedness of novel words, not abstract markedness relations stated over combinations of natural classes. It is likely that a successful explanation of gradient phonotactics will ultimately involve both grammatical and lexical factors. The best approach seems to be to develop explicit theories of both types and try to figure out what kind of division of labor is empirically justified (Coetzee 2008).

In a recent paper, Coetzee and Pater (2008) propose a grammatical theory of gradient phonotactics stated in terms of weighted constraints in the sense of Harmonic Grammar (HG) (see e.g. Smolensky and Legendre 2006). They note that "there seems not yet to be a satisfactory account of gradient phonotactic acceptability available within [Optimality Theory, OT]" (Prince and Smolensky 1993/2004). This paper proposes such an account. The hypothesis is that the relative well-formedness of a phonotactic structure depends on its grammatical complexity in the following sense: the more ranking information a phonotactic structure requires in order to surface faithfully, the less well-formed it is. We call this the COMPLEXITY HYPOTHESIS:

(1) The Complexity Hypothesis: The probability of an <input, output> mapping is inversely correlated with its grammatical complexity.

Coetzee and Pater (2008) focus on the question of learnability, showing that a HG weighting can be learned that reflects the quantitative patterns in the data. The present paper focuses on the question how phonological theory constrains possible quantitative patterns. We first show that an OT grammar orders linguistic structures in an implicational hierarchy that reflects their grammatical complexity. The Complexity Hypothesis then makes the following prediction: if a more complex structure *P* is well-formed, then a less complex structure *Q* is at least as well-formed. Some of these implications are universal and thus do not need to be learned; others depend on language-specific rankings. Working out such predictions is important because they reveal what the theory admits and what it excludes, both universally and under given language-particular conditions. Finally, we show that the Complexity Hypothesis is supported by phonotactic data from Muna (Austronesian) as recently analyzed by Coetzee and Pater (2008).

The present paper draws upon two lines of recent research. Our proposal is deeply indebted to Coetzee and Pater's work on gradient phonotactics (Pater and Coetzee 2005, Coetzee and Pater 2006, 2008): we follow their analysis as closely as possible, but develop it in a very different direction, proposing an optimality-theoretic reinterpretation of their results. The proposal draws much of its inspiration from Prince's work on Elementary Ranking Conditions (Prince 2002a, 2002b, 2006, 2007; Brasoveanu and Prince 2005) which provides formal tools for determining the necessary and sufficient conditions for phonological mappings. This paper provides a concrete illustration of the usefulness of these notions in empirical work.

The paper is structured as follows. Section 2 introduces the proposal and illustrates it based on a simple example from Arabic. Section 3 contains a more detailed discussion of the gradient phonotactics of Muna based on the work of Coetzee and Pater (2006, 2008). Section 4 concludes the paper.

2. THE PROPOSAL

2.1 *The Arabic example*

In Arabic, roots are normally composed of three consonants, e.g. *ktb* 'write'. There is a well-known dissimilatory constraint against homorganic consonants in adjacent positions within the verbal root (Frisch, Pierrehumbert, and Broe 2004; Greenberg 1950; McCarthy 1988, 1994; Pater and Coetzee 2005, Coetzee and Pater 2008; Pierrehumbert 1993). For example, root morphemes with adjacent labial consonants (**fbm, *bfk, *kbm*) are ill-formed (McCarthy 1988:88). However, it is also well known that the pattern shows gradience. Frisch, Pierrehumbert and Broe (2004) argue that the strength of the dissimilatory effect is a gradient function of the similarity of the consonants in the pair: the more similar the consonants, the less frequently they co-occur in actual lexical items.

Frisch, Pierrehumbert, and Broe (2004) study the gradient phonotactics of Arabic based on 2,674 verbal roots from a dictionary of standard Arabic (Cowan 1979). They describe these patterns in terms of OBSERVED/EXPECTED (O/E) VALUES. The O/E value is the ratio of the observed number of occurring consonant pairs (O) to the number that would be expected if the consonants combined at random (E). An O/E value greater than 1 indicates that there are more observed combinations than expected, i.e. the combination is favored. An O/E value smaller than 1 indicates that there are fewer observed combinations than expected, i.e. the combination is disfavored.[1] O/E values for pairs of adjacent consonants in Arabic verbal roots are shown in (2). We have omitted gutturals.

(2) Observed/Expected values in Arabic (Frisch, Pierrehumbert, and Broe 2004:186): labials = b,f,m; dorsals = k, g, q; coronal sonorants = l, r, n; coronal fricatives = θ, ð, s, z, sˤ, zˤ, ʃ; coronal plosives = t, d, tˤ, dˤ

	labial	dorsal	coronal sonorant	coronal fricative	coronal plosive
labial	0.00				
dorsal	1.15	0.02			
coronal sonorant	1.18	1.48	0.06		
coronal fricative	1.31	1.16	1.21	0.04	
coronal plosive	1.37	0.80	1.23	0.52	0.14

[1] The O/E value of the pair [t-d] can be calculated as follows: the probability of [t] is the number of pairs with [t] as the first member divided by the number of all pairs; the probability of [d] is the number of pairs with [d] as the second member divided by the number of all pairs; the expected value of [t-d] is the probability of [t] as the first member multiplied by the probability of [d] as the second member multiplied by the number of all pairs; the O/E value is the observed number of [t-d] pairs divided by the expected number of [t-d] pairs.

The O/E values for pairs of homorganic consonants are near zero (shaded cells), but within coronals there is gradience (bottom right hand box). The central observations are stated in (3) (Frisch, Pierrehumbert, and Broe 2004:187):

(3) The quantitative patterning of adjacent coronals
  (a) If the coronals are both sonorants, fricatives, or plosives, O/E is low;
  (b) If the coronals are fricative + plosive, O/E is higher;
  (c) If the coronals are sonorant + fricative or plosive, O/E is high.

The challenge is to derive such quantitative patterns from the grammar. In the next section, we will show how the coronal pattern can be derived in Optimality Theory from the interaction of ranked and violable constraints.

2.2 *T-orders*

We start by adopting five phonological constraints from Pater and Coetzee 2005 (see also McCarthy 1988, 1994 and Padgett 1995): one faithfulness constraint that requires the input and the output to be identical and four markedness constraints against identical feature specifications on adjacent segments where adjacency is defined on the consonantal root. The markedness constraints are special cases of the Obligatory Contour Principle (OCP) that bans adjacent identical elements (Goldsmith 1976, Leben 1973).

(4) Constraints (Pater and Coetzee 2005)

FAITH	Input and output are identical.
OCP-COR	No adjacent coronals (e.g. /t-n/)
OCP-COR[−son]	No adjacent coronal obstruents (e.g. /t-s/)
OCP-COR[+son]	No adjacent coronal sonorants (e.g. /l-n/)
OCP-COR[−son][αcont]	No adjacent coronal obstruents agreeing in continuancy (e.g. /t-d/)

The violations assigned by these constraints are illustrated in (5). We consider four inputs: plosive-plosive (/t-d/), plosive-fricative (/t-s/), sonorant-sonorant (/l-n/), and plosive-sonorant (/t-n/). Each input has two output candidates: the faithful candidate and the unfaithful candidate OTHER. We will write input-output mappings between angled brackets, e.g. <t-d, t-d> is the faithful mapping, <t-d, OTHER> is the unfaithful mapping. A phonotactic sequence is well-formed if the faithful mapping wins, else it is ill-formed. Which mapping wins depends on the language-specific constraint ranking which is not specified in (5).

(5) Violation table

		FAITH	OCP-COR [−son, αcont]	OCP-COR [+son]	OCP-COR [−son]	OCP-COR
(a) t-d	t-d		*		*	*
	OTHER	*				
(b) t-s	t-s				*	*
	OTHER	*				
(c) l-n	l-n			*		*
	OTHER	*				
(d) t-n	t-n					*
	OTHER	*				

The crucial observation is that some mappings entail others. For example, if the sequence [l-n] is well-formed in a language, the sequence [t-n] will also be well-formed in the same language. This can be verified by examining the violation table: the mapping <l-n, l-n> requires that FAITH dominate both OCP-COR[+son] and OCP-COR, whereas the mapping <t-n, t-n> only requires that FAITH dominate OCP-COR. If the former holds, the latter must also hold, hence the well-formedness of [l-n] entails the well-formedness of [t-n]. The rankings required for each faithful mapping are shown in (6).

(6) The rankings required for each faithful mapping

(a) <t-d, t-d>:
FAITH >> OCP-COR ∧
FAITH >> OCP-COR[−son] ∧
FAITH >> OCP-COR[−son, αcont]

(b) <t-s, t-s>:
FAITH >> OCP-COR ∧
FAITH >> OCP-COR[−son]

(c) <l-n, l-n>:
FAITH >> OCP-COR ∧
FAITH >> OCP-COR[+son]

(d) <t-n, t-n>
FAITH >> OCP-COR

Similar entailments hold among unfaithful mappings. For example, if the sequence [t-n] is ill-formed in a language, the sequence [l-n] will also be ill-formed in the same language. This can be verified as follows: the mapping <t-n, OTHER> requires that OCP-COR dominate FAITH, whereas the mapping <l-n, OTHER> only requires that OCP-COR dominate FAITH or that OCP-COR[+son] dominate FAITH. If the former holds, the latter must also hold, hence the ill-formedness of [t-n] entails the ill-formedness of [l-n]. The rankings required for each unfaithful mapping are shown in (7).

(7) The rankings required for each unfaithful mapping

(a) <t-d, OTHER>:
OCP-COR >> FAITH ∨
OCP-COR[−son] >> FAITH ∨
OCP-COR[−son, αcont] >> FAITH

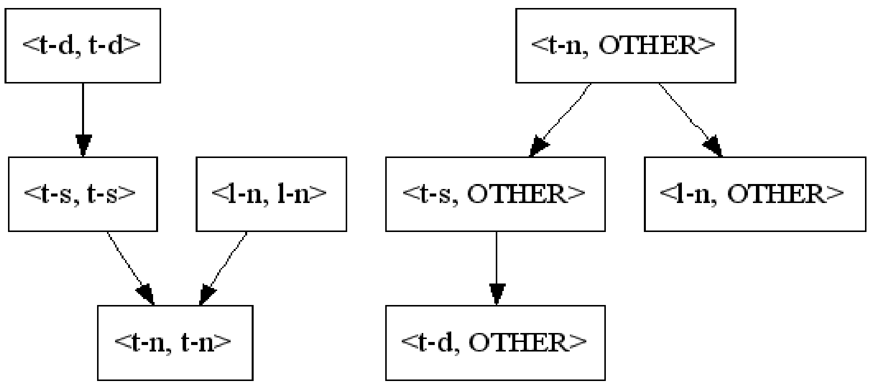
(b) <t-s, OTHER>:
OCP-COR >> FAITH ∨
OCP-COR[−son] >> FAITH

(c) <l-n, OTHER>:
OCP-COR >> FAITH ∨
OCP-COR[+son] >> FAITH

(d) <t-n, OTHER>:
OCP-COR >> FAITH

The entailments among phonological mappings constitute a partial ordering that can be visualized as a directed graph. The graph in (8) shows all the entailments hidden in the grammar in (5). Transitive arrows have been removed for perspicuity.

(8) Entailments among phonological mappings as a directed graph



Another way of discovering the graph in (8) is to examine the languages predicted by the grammar. The grammar in (5) has five constraints which can be ranked in 5! = 120 possible ways. Using OTSOFT (Hayes, Tesar, and Zuraw 2003) we can verify that these rankings yield 7 distinct languages. This is the FACTORIAL TYPOLOGY shown in (9).

(9) The factorial typology

	#1	#2	#3	#4	#5	#6	#7
/t-d/:	t-d	t-d	OTHER	OTHER	OTHER	OTHER	OTHER
/t-s/:	t-s	t-s	t-s	t-s	OTHER	OTHER	OTHER
/t-n/:	t-n	t-n	t-n	t-n	t-n	t-n	OTHER
/l-n/:	l-n	OTHER	l-n	OTHER	l-n	OTHER	OTHER

The factorial typology reveals several interesting asymmetries. For example, there are two languages where /t-d/ surfaces faithfully (#1, #2). These languages are a subset of the four languages where /t-s/ surfaces faithfully (#1, #2, #3, #4). This generalization can be stated as an IMPLICATIONAL UNIVERSAL that relates two <input, output> mappings:

(10) An implicational universal
<t-d, t-d> → <t-s, t-s> If /t-d/ is realized faithfully, so is /t-s/.

This implicational universal is true of every language in the factorial typology. In other words, it is true no matter how the constraints are ranked. If we systematically work through the factorial typology we discover seven more implicational universals. The resulting eight implicational universals are summarized in (11). Taken together, they describe the graph in (8).

(11) Eight implicational universals

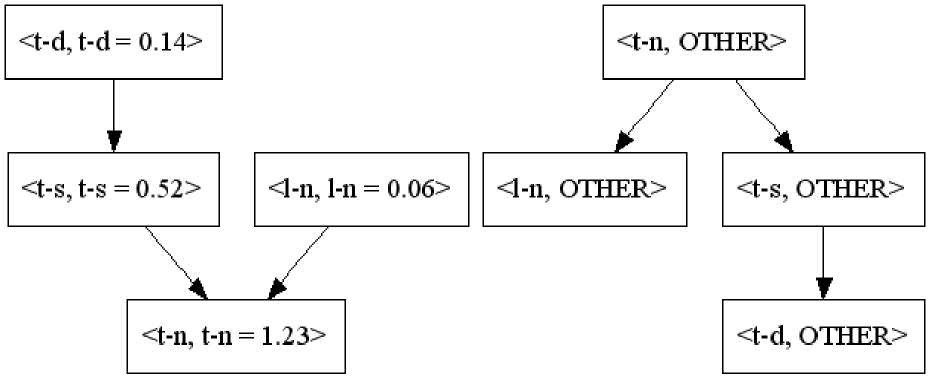
(a) <t-d, t-d> → <t-s, t-s>
(b) <t-d, t-d> → <t-n, t-n>
(c) <t-s, t-s> → <t-n, t-n>
(d) <l-n, l-n> → <t-n, t-n>
(e) <t-s, OTHER> → <t-d, OTHER>
(f) <t-n, OTHER> → <t-d, OTHER>
(g) <t-n, OTHER> → <t-s, OTHER>
(h) <t-n, OTHER> → <l-n, OTHER>

Let us now summarize what has been done. We have identified a structure hidden in an optimality-theoretic constraint set that describes entailments among phonological mappings. We have arrived at this structure in two ways: INTENSIONALLY in terms of constraint rankings and EXTENSIONALLY in terms of the predicted language typology. The structure describes the relative grammatical complexity of phonological mappings in the following sense: the higher in the graph the mapping resides, the more ranking information is needed to describe it grammatically and the fewer languages instantiate it typologically. We call this structure a TYPOLOGICAL ORDER, or T-ORDER.[2]

How are T-orders relevant for gradient phonotactics? In (12), we have annotated the T-order with the O/E values reported by Frisch, Pierrehumbert, and Broe (2004) for Arabic. The annotated graph shows that the O/E values grow in the direction of the arrows. This means that the T-order is reflected quantitatively in the Arabic lexicon: phonological mappings that are grammatically complex and typologically marked are also quantitatively underrepresented.

[2] A reviewer asks whether the same entailments would follow if the candidate OTHER were “broken up” into more specific candidates. In Optimality Theory, a candidate is defined by its constraint violation profile (Samek-Lodovici and Prince 2005). OTHER is the name we have given to the candidate that violates FAITH, but none of the OCP-constraints. Assuming that constraints are binary, 5 constraints define $2^5 = 32$ possible candidates. With the exception of the faithful candidate and OTHER, all these candidates are conceptually impossible, such as a consonant pair that violates OCP-COR[−son], but not OCP-COR, or a consonant pair that violates both OCP-COR[−son] and OCP-COR[+son]. Introducing more candidates such as [p-t] or [t-m] would involve positing additional constraints. Whether the same entailments would still follow or not depends on these constraints. For example, adding an ad hoc constraint *T-N violated by [t-n] would remove all entailments of the form X → <t-n, t-n> which would result in a different T-order. This illustrates the fact that typological predictions can be subverted by ill-chosen additional constraints.

(12) The T-order and Arabic O/E values



Based on this initial discovery, we now state our hypothesis:

(13) The Complexity Hypothesis (preliminary version): The O/E value of a phonotactic sequence is inversely correlated with its grammatical complexity.

Why should the Complexity Hypothesis hold true?[3] A language's vocabulary is largely a matter of history and the result of extragrammatical events such as language contact. The systematic patterning of O/E values in the lexicon therefore comes as a surprise. One possible explanation is that over time, all else being equal, a language's lexical stock will favor less complex items over more complex items (Kiparsky 1995, 2008). This is consistent with the observation that in many languages stems form prosodically natural classes (e.g. monosyllables or disyllables) and affixes may be restricted to natural classes (e.g. the Germanic dentals for inflection). Under the present proposal, a lexical item that requires more ranking information is at a disadvantage in comparison to a lexical item that requires less ranking information. It is this strive for simplicity that we see reflected quantitatively in the Arabic lexicon: roots that require less ranking information are favored; roots that require more ranking information are disfavored.

2.3 *T-orders and language-specific rankings*

The Complexity Hypothesis makes specific predictions about the universality and language-particularity of phonotactic patterns. For example, the above grammar predicts that the ordering of O/E values [t-d] ≤ [t-s] ≤ [t-n] and [l-n] ≤ [t-n] should be universal. If we assume that constraints are innate, it follows that a newborn knows this ordering before hearing the first word. In contrast, nothing is predicted about the relative ordering of O/E values between [l-n] and [t-d] and between [l-n] and [t-s]. These orderings depend on the language-particular constraint ranking. They must thus be learned from the data which means that they should be able to vary from language to language.

[3] I thank Michael Kenstowicz and an anonymous reviewer for raising these larger issues and for suggesting ways to think about them.

Consider adding two Arabic-specific rankings into the grammar (cf. Pater and Coetzee 2005). The rankings are shown in (14); the resulting typology is shown in (15); and the corresponding T-order is shown in (16)-(17).

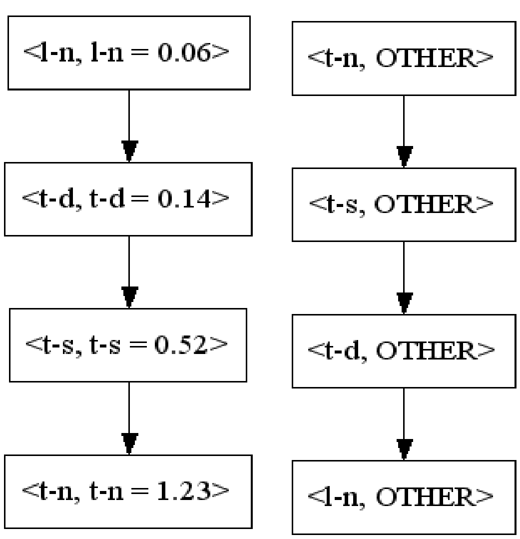
(14) Language-specific rankings for Arabic
- (a) OCP-COR[+son] >> OCP-COR[−son, αcont]
- (b) OCP-COR[+son] >> OCP-COR[−son]

(15) Typology

	#1	#2	#3	#4	#5
/t-n/:	t-n	t-n	t-n	t-n	OTHER
/t-s/:	t-s	t-s	t-s	OTHER	OTHER
/t-d/:	t-d	t-d	OTHER	OTHER	OTHER
/l-n/:	l-n	OTHER	OTHER	OTHER	OTHER

(16) T-order as pairs of <input, output> pairs
- (a) <t-d, t-d> → <t-s, t-s> universal
- (b) <t-d, t-d> → <t-n, t-n> universal
- (c) <t-s, t-s> → <t-n, t-n> universal
- (d) <l-n, l-n> → <t-n, t-n> universal
- (e) <l-n, l-n> → <t-d, t-d> language-specific
- (f) <l-n, l-n> → <t-s, t-s> language-specific
- (g) <t-s, OTHER> → <t-d, OTHER> universal
- (h) <t-n, OTHER> → <t-d, OTHER> universal
- (i) <t-n, OTHER> → <t-s, OTHER> universal
- (j) <t-n, OTHER> → <l-n, OTHER> universal
- (k) <t-s, OTHER> → <l-n, OTHER> language-specific
- (l) <t-d, OTHER> → <l-n, OTHER> language-specific

(17) T-order as a directed graph



The rankings in (14) add four new entailments into the T-order ((16e,f,k,l)), turning the universal partial order into a language-specific total order. The resulting T-order corresponds perfectly to the ordering of the empirical O/E values ((17)).

More generally, adding rankings into the grammar can add entailments into the T-order, but not remove any existing entailments. Consider the possible relations between two <input, output> mappings in a factorial typology. There are three distinct cases:

(18) Relations between <input, output> mappings

(a) Neither mapping entails the other:

	#1	#2	#3	#4	#5	#6	#7
/t-s/:	t-s	t-s	t-s	t-s	OTHER	OTHER	OTHER
/l-n/:	l-n	OTHER	l-n	OTHER	l-n	OTHER	OTHER

(b) One mapping entails the other:

	#1	#2	#3	#4	#5	#6	#7
/t-d/:	t-d	t-d	OTHER	OTHER	OTHER	OTHER	OTHER
/t-s/:	t-s	t-s	t-s	t-s	OTHER	OTHER	OTHER

(c) Each mapping entails the other:

	#1	#2	#3	#4	#5	#6	#7
/t-n/:	t-n	t-n	t-n	t-n	t-n	t-n	OTHER
/s-n/:	s-n	s-n	s-n	s-n	s-n	s-n	OTHER

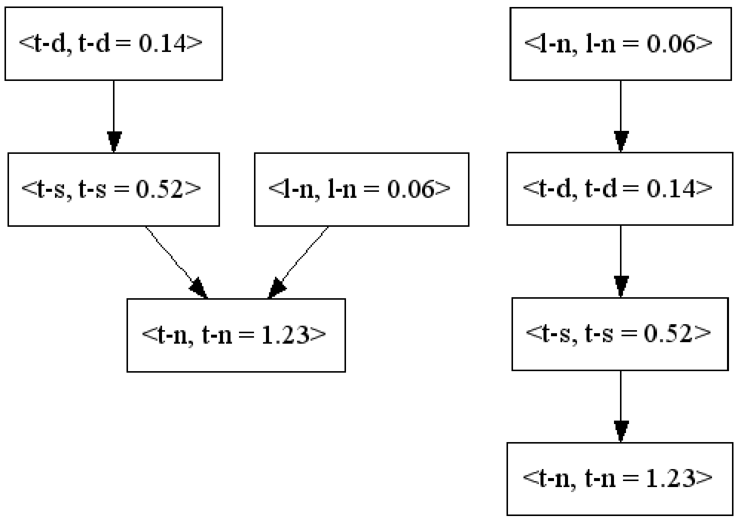
Adding rankings will either eliminate columns or have no effect. Clearly, adding rankings can never add new columns into the factorial typology because new rankings restrict the set of languages rather than expand it. Now, consider the possible effects of ranking in each case. In (18a), we start out with no entailments. Eliminating #5 would add the entailment [l-n] → [t-s]; eliminating #2 and #4 would add the reverse entailment [t-s] → [l-n]. In (18b), we start out with the entailment [t-d] → [t-s]. Removing #3 and #4 would add the reverse entailment [t-s] → [t-d], resulting in entailments both ways, but it is not possible to remove the existing entailment. This is because the implicational universal holds true of every column in the factorial typology and we can only remove columns. Finally, in (18c) we start out with entailments both ways and neither adding nor subtracting entailments is possible. We conclude that adding rankings can only add entailments into the T-order, but never remove any existing entailments.

2.4 *Evaluating T-orders*

It is useful to have some way of determining how well a T-order matches the data. In this paper, we have chosen to compare two T-orders: the actual T-order derived from the grammar and the ideal T-order that correctly captures the empirically observed frequency ordering among the predicted <input, output> mappings. We illustrate this using the Arabic grammar with no rankings: (19) uses the graph notation; (20) uses the pair notation. The unfaithful mappings are omitted because we only have O/E values for the faithful mappings.

(19) Actual and ideal T-orders as directed graphs

(a) The actual T-order (b) The ideal T-order



(20) Actual and ideal T-orders as pairs of pairs

(a)	<t-d, t-d> → <t-s, t-s>	actual, ideal
(b)	<t-d, t-d> → <t-n, t-n>	actual, ideal
(c)	<t-s, t-s> → <t-n, t-n>	actual, ideal
(d)	<l-n, l-n> → <t-n, t-n>	actual, ideal
(e)	<l-n, l-n> → <t-d, t-d>	ideal
(f)	<l-n, l-n> → <t-s, t-s>	ideal

In this case, the actual and ideal T-orders are not identical. In order to determine how well the two match, we calculate two evaluation measures commonly used in information retrieval: PRECISION and RECALL (see e.g. Manning and Schütze 1999: 267-271).

(21) (a) PRECISION is the ratio of the number of pairs that are in both T-orders to the number of pairs that are in the actual T-order. This number indicates how many of the predicted quantitative relationships are observed. In our example, precision = 4/4 = 100%.

(b) RECALL is the ratio of the number of pairs that are in both T-orders to the number of pairs that are in the ideal T-order. This number indicates how many of the observed quantitative relationships are predicted. In our example, recall = 4/6 = 67%.

What do these numbers tell us about the quality of the linguistic analysis? Assume that the T-order is based on universal constraints, but no language-specific rankings. Perfect precision (= 100%) means that all the predicted quantitative relationships are observed in the data. In other words, the universals have stood the test of one particular language. In this sense, high precision is an indicator of DESCRIPTIVE SUCCESS. Perfect recall (= 100%)

means that all the quantitative relationships observed in the data are predicted. This describes the unlikely event that all the quantitative relationships in the data are universal. However, the higher the recall value, the fewer language-specific stipulations (= rankings, learning) will be needed to account for the residual facts. In our example, 67% of the quantitative relationships are universal; the rest have to be learned from the Arabic data. In this sense, high recall is an indicator of EXPLANATORY SUCCESS.[4] The obvious next step is to develop an evaluation metric that is sensitive to the magnitude of the difference between two nodes. This task is left for future work.

2.5 *Finding T-orders*

Given a grammar, how can we find the T-order? We have seen two methods. The direct method involves figuring out the rankings required for each <input, output> mapping and the entailments among these rankings. This approach has been developed by Prince (2002a, 2002b, 2006, 2007) and Brasoveanu and Prince (2005) who formalize the problem in terms of ELEMENTARY RANKING CONDITIONS (ERCs) and provide a calculus for working with ERCs. The indirect method involves figuring out the implicational universals among the <input, output> mappings in a factorial typology. As Prince (2006, 31) notes, both perspectives are valuable and one may be more helpful than the other depending on the circumstances. For small grammar fragments, the T-order is easy to work out under either method with paper and pencil. For grammars of realistic size this quickly becomes a tedious exercise and the only viable option is to use a computer.

All the T-orders in this paper have been produced by T-ORDER GENERATOR (Anttila and Andrus 2006), a free open-source Python program for computing and visualizing T-orders.[5] The current development version (November 2007) allows the user to compute T-orders either directly from constraint violation patterns or indirectly from factorial typologies. Under the direct method, the program reads Microsoft Excel files written in the traditional tableau format and computes T-orders using ERCs which remain invisible to the user. The ERC algorithm is described in the README file that accompanies the software. Under the indirect method, the program reads factorial typology files produced by OTSoft and computes T-orders using the following algorithm:

(22) Finding T-orders from factorial typologies

- For all <input, output> pairs in the factorial typology, construct all the directed edges consisting of a start pair and an end pair, with different inputs.
- For each edge <$pair_0$, $pair_1$>, look through all the output patterns in the factorial typology. If for some output pattern, $pair_0$ appears but $pair_1$ does not, discard the edge. If $pair_1$ appears whenever $pair_0$ appears, keep the edge.

---

[4] Two specific remarks about computing recall are in order. First, <input, output> pairs that stand in an entailment relationship always have distinct inputs. Pairs with identical inputs, e.g. </kɑst me/, [kɑst me]> vs. </kɑst me/, [kɑs me]> 'cost me' (variable *t*-deletion) never stand in an entailment relationship. Since such pairs never appear in actual T-orders, they do not appear in ideal T-orders either and thus do not figure in the computation of recall. Second, we assume that two pairs with exactly the same empirical number (e.g. O/E-value) entail each other in the ideal T-order. Such entailments do figure in the computation of recall.

[5] The program can be downloaded from http://www.stanford.edu/~anttila/research/software.html.

2.6 *Summary*

A set of optimality-theoretic constraints defines a partial order that describes the entailments among optimal <input, output> mappings. We called this structure a T-order. A T-order describes the relative grammatical complexity of <input, output> mappings in the following sense: the higher in the T-order the mapping resides, the more ranking information is needed to describe it grammatically and the fewer languages instantiate it typologically. T-orders are not a new theoretical device, but a consequence of standard Optimality Theory: every optimality-theoretic grammar has a T-order that just needs to be spelled out. Our substantive proposal is the Complexity Hypothesis which attributes the relative well-formedness of a phonotactic sequence to its grammatical complexity. The Complexity Hypothesis is an empirical claim that can be true or false. In Section 4, we will see that the Complexity Hypothesis is supported by phonotactic data from Muna (Austronesian) as analyzed by Coetzee and Pater (2006, 2008).

T-orders are in no way limited to phonotactics: they simply order <input, output> mappings in terms of grammatical complexity and they are implicitly present in all domains of linguistics that involve typological and quantitative generalizations. In phonotactics the mappings are faithful e.g. </t-d/, [t-d]>. In phonological alternations the mappings are unfaithful, e.g. </kɑst me/, [kɑs me]> 'cost me' (*t*-deletion). In syntax and semantics the mappings relate semantic inputs to syntactic outputs. An area of particular interest is language variation; see e.g. Anttila 2007 and Anttila, Fong, Benus, and Nycz to appear (phonological variation) and Anttila 2008 (syntactic variation). A more general statement of the Complexity Hypothesis that is neutral with respect to the nature of the <input, output> mapping is given in (23):

(23) <u>The Complexity Hypothesis (final version)</u>: The probability of an <input, output> mapping is inversely correlated with its grammatical complexity.

3. COETZEE AND PATER

In a series of important papers, Coetzee and Pater have proposed two theories of gradient phonotactics and tested them on Arabic and Muna. The earlier proposal (Pater and Coetzee 2005, Coetzee and Pater 2006) is framed in terms of Optimality Theory. The later proposal (Coetzee and Pater 2008) rejects strict ranking in favor of weighting. In this section, we briefly outline Coetzee and Pater's proposals and compare them to our proposal.

We will continue with the familiar Arabic example. Recall the phonotactic patterns among Arabic coronals:

(24) Adjacent coronals in Arabic
- (a) If the coronals are both sonorants, fricatives, or plosives, O/E is low;
- (b) If the coronals are fricative + plosive, O/E is higher;
- (c) If the coronals are sonorant + fricative or plosive, O/E is high.

Pater and Coetzee's (2005) optimality-theoretic analysis proposes that markedness (here: OCP) constraints are ranked in an order that reflects the relative O/E values of the consonant pairs. They posit the following ranking that reflects the Arabic O/E values:

(25) The ranking among OCP-COR constraints in Arabic (Pater and Coetzee 2005:6)

{OCP-COR[−son][αcont], OCP-COR[+son]} >> OCP-COR[−son] >> OCP-COR

violated by (24a), low O/E	violated by (24b), intermediate O/E	violated by (24c), high O/E

Gradience is captured in this model by positing lexically indexed faithfulness constraints. The ordering of O/E values is derived by assuming a stage of evaluation where each lexical item is submitted to the grammar with each lexical indexation and assigned a well-formedness score proportional to the number of evaluations that make the faithful candidate optimal. This is illustrated for the input /t-d/ in (26).

(26) Evaluating /t-d/ (Pater and Coetzee 2005)

		FAITH-L2	OCP-COR [−son, αcont]	OCP-COR [+son]	**FAITH-L1**	OCP-COR [−son]	**FAITH**	OCP-COR
(a) t-d	☞t-d		*			*		*
	OTHER	*!						
(b) t-d	t-d		*!			*		*
	☞OTHER				*			
(c) t-d	t-d		*!			*		*
	☞OTHER						*	

The input /t-d/ is indexed to FAITH-L2 in (26a), FAITH-L1 in (26b), and FAITH in (26c). The faithful candidate [t-d] only wins in (26a), i.e. under 1/3 of the indexations. This well-formedness score is interpreted relatively: the higher the score, the more acceptable the lexical item. The well-formedness scores for other inputs are derived in an analogous manner. The results of evaluating the four sample inputs are shown in (27).

(27) Well-formedness scores for four faithful mappings (Pater and Coetzee 2005)

	MAPPING	SCORE	O/E
(a)	/l-n/ → [l-n]	1/3	0.06
(b)	/t-d/ → [t-d]	1/3	0.14
(c)	/t-s/ → [t-s]	2/3	0.52
(d)	/t-n/ → [t-n]	3/3	1.23

In sum, Pater and Coetzee's (2005) theory has three parts:

(28) An outline of Pater and Coetzee's (2005) theory
- (a) Markedness constraints are ranked according to the frequency with which they are violated in the lexicon
- (b) Interspersed among the markedness constraints are lexically indexed faithfulness constraints.
- (c) The relative well-formedness of a lexical item is determined by submitting it to the grammar with each lexical indexation.

Our reanalysis builds directly on Pater and Coetzee's (2005) analysis. The analyses share the same constraints and the strict ranking assumption. However, the reanalysis takes Pater and Coetzee's analysis one step further by recognizing the existence of the T-order and the limits it imposes on O/E values. T-orders explicitly show which orderings among the O/E values are ranking-dependent, hence language-specific, hence something that must be learned, and which orderings are ranking-independent, hence universal, hence something known a priori. This is relevant for testing the proposed constraints: the prediction is that cross-linguistic variation should be limited to the ranking-dependent orderings. Finally, the reanalysis assumes no lexically indexed constraints and hence no stage where lexical items are evaluated with different lexical indexations. Instead, the well-formedness of a lexical item follows directly from its grammatical complexity as defined by the T-order.

The new analysis proposed by Coetzee and Pater (2008) is different in two major ways. First, the constraints are not the same. The differences relevant for the Arabic example are summarized in (29).

(29)		PATER AND COETZEE 2005	COETZEE AND PATER 2008
	(a)	OCP-COR [−son, αcont]	OCP-COR [αstricture]
	(b)	OCP-COR[+son], OCP-COR[−son]	OCP-COR[αsonorant]
	(c)	--	OCP-COR[αvoice]

Overall, the new constraints are less fine-grained. The constraint OCP-COR[αstricture] assigns a violation mark to pairs of stops, pairs of fricatives, and pairs of liquids. The constraint OCP-COR[αsonorant] targets both [+son] and [−son]. The new constraint OCP-COR[αvoice] plays a minor role in Arabic, but is relevant for our example.

Second, Coetzee and Pater (2008) adopt Harmonic Grammar, a predecessor of Optimality Theory where constraints are not strictly ranked, but numerically weighted (Legendre, Miyata and Smolensky 1990, Smolensky and Legendre 2006; see Keller 2006 for the closely related Linear Optimality Theory). The weights for Arabic given in Coetzee and Pater (2008, 36) are shown in (30). These weights were learned by the Linear OT learning algorithm implemented in Praat (Boersma and Weenink 2007).

(30) Arabic in Harmonic Grammar

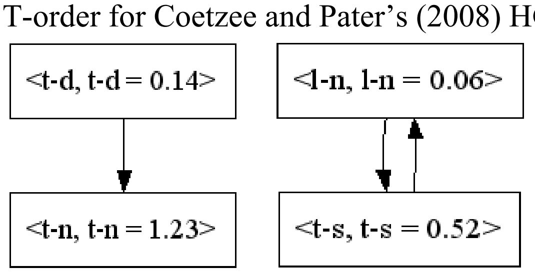
		227.25	65.74	69.17	45.61	0	*H*	*A*	*O/E*
		ID-PLACE	OCP-COR [αson]	OCP-COR [αstric]	OCP-COR [αvoice]	OCP-COR			
t-d	t-d		–1	–1		–1	–134.91	92.34	0.14
	OTHER	–1					–227.25	–92.34	
t-s	t-s		–1		–1	–1	–111.35	115.9	0.52
	OTHER	–1					–227.25	–115.9	
l-n	l-n		–1		–1	–1	–111.35	115.9	0.06
	OTHER	–1					–227.25	–115.9	
t-n	t-n			–1		–1	–69.17	158.08	1.23
	OTHER	–1					–227.25	–158.08	

In Harmonic Grammar, the relative strengths of the constraints are expressed as numerical weights. In tableau (30), the weights are non-negative real numbers and appear in the top row. Constraint violations are negative integer scores and appear in the cells. The candidate's harmony (column *H*) is calculated by multiplying each score by its weight and by summing the weighted constraint scores.

In gradient phonotactics, the crucial question is how to compare candidates across inputs. For example, we would want to predict that [t-d] is less well-formed than [t-n]. As Coetzee and Pater (2008) point out, it is not possible to compare raw harmony scores across inputs because the harmony of a winner is relative to its competitors and a winner in one tableau may have a lower harmony than a loser in another tableau (Boersma 2004; Legendre, Sorace, and Smolensky 2006). Coetzee and Pater (2008) solve the problem by relativizing a candidate's acceptability to its competitors: a candidate's acceptability (column *A*) is obtained by subtracting from its harmony value the harmony value of its most harmonic competitor. Thus, if *x* is a candidate and *y* its most harmonic competitor, the acceptability of *x* is $A(x) = H(x) - H(y)$. Coetzee and Pater (2008) evaluate the success of their analysis by calculating the strength of the correlation between the acceptability values and the O/E values.

We can identify the limits of the grammar fragment in (30) by finding its T-order. We took the constraints in (30), computed the typology of languages predicted under Harmonic Grammar with the aid of OT-Help (Becker, Pater, and Potts 2007), and used this typology to find the T-order. The resulting graph is shown in (31), annotated with the familiar O/E values. Note that the mappings <t-s, t-s> and <l-n, l-n> form a cycle: each mapping implies the other.

(31) T-order for Coetzee and Pater's (2008) HG analysis of Arabic coronals



The graph in (31) shows the three implicational universals that hold true irrespective of weighting. First, [t-d] (*A* = 92.34) should be universally less acceptable than [t-n] (*A* = 158.08) because the former incurs a superset of the violations of the latter. Second, [t-s] and [l-n] should be equally acceptable (*A* = 115.9) because no constraint in (30) distinguishes between them. The T-order computed under Optimality Theory is identical to that computed under Harmonic Grammar.

We can now compare the T-orders for the old and the new constraints:

(32)

PATER AND COETZEE 2005	COETZEE AND PATER 2008
<t-d, t-d = 0.14> → <t-n, t-n = 1.23>	<t-d, t-d = 0.14> → <t-n, t-n = 1.23>
<l-n, l-n = 0.06> → <t-n, t-n = 1.23>	<l-n, l-n = 0.06> → <t-s, t-s = 0.52>
<t-d, t-d = 0.14> → <t-s, t-s = 0.52>	<t-s, t-s = 0.52> → <l-n, l-n = 0.06>
<t-s, t-s = 0.52> → <t-n, t-n = 1.23>	

Which constraints are better? The new analysis has one precision problem: the ordering [t-s] ≤ [l-n] is predicted to be universal, but this is falsified by both Arabic and Muna. The problem arises because no constraint differentiates between the two consonant pairs. Note that the old analysis had no problem here because OCP-COR[αson] was split into two independent constraints: OCP-COR[+son] and OCP-COR[−son]. There are also differences in recall. The old analysis predicts that the three pairwise orderings [l-n] ≤ [t-n], [t-d] ≤ [t-s], and [t-s] ≤ [t-n] should hold universally, whereas the new analysis predicts that all three should be language-specific and a matter of learning. In reality, [t-d] ≤ [t-s] is contradicted by Muna, supporting the new analysis. On the other hand, the other two predicted orderings do hold up in both languages, which can be taken as evidence for the old analysis. Finally, the new analysis predicts that the ordering [l-n] ≤ [t-s] should be universal, whereas the old analysis predicts it to be language-specific and a matter of learning. Since this prediction holds up in both languages, we have some measure of support for the new analysis. Naturally, these predictions ought to be tested against many more languages before we can decide which set of constraints better approximates the cross-linguistic phonotactic facts.

Several conclusions emerge from this brief review. First, we have seen that an OT account of gradient phonotactics is possible without assuming lexically indexed constraints or weighted constraints. The alternative we have proposed is that the O/E values reflect the grammatical complexity of phoneme sequences defined in terms of constraint rankings. Second, we have seen that T-orders can be computed for any theory that predicts typological generalizations. We illustrated this briefly for Coetzee and Pater's (2006) Harmonic Grammar analysis of Arabic. Another illustration can be found in a recent analysis of English rhyme structure by vander Wyk and McClelland (2007) who derive partial orderings identical to T-orders from a set of weighted constraints. Third, we have seen that minor changes in the constraints can result in different predictions about what is universal and what is language-particular in gradient phonotactic patterns. Deriving such predictions is an important part of phonological analysis. Here T-orders turn out to be a useful grammar debugging tool: a precision error signals a problem that cannot be remedied by ranking or weighting. Instead, new constraints will be needed.

4. Gradient phonotactics in Muna

In this section, we test the Complexity Hypothesis on gradient phonotactic data from the Austronesian language Muna (van den Berg 1989) based on the work of Coetzee and Pater (2006, 2008). Their data consist of 7,892 adjacent consonant pairs from 5,854 roots of shapes (V)CVCV and (V)CVCVCV listed in an electronic version of van den Berg and Sidu's (1996) Muna dictionary. Coetzee and Pater (2006) describe the data in great detail, down to the level of individual segment pairs, providing an excellent testing ground for theories of gradient phonotactics. We will see that Coetzee and Pater's linguistic analysis is of high quality and stands up to scrutiny, some residual problems notwithstanding. The purpose of this section is to show how Coetzee and Pater's analysis can be recast in terms of Optimality Theory without assuming either lexically specific constraints or weighted constraints. In all other respects, we have tried to stay as close to Coetzee and Pater's analysis as possible. In cases where they present several analyses, we have chosen the alternative that yields the best empirical fit.

4.1 *The phonotactics of obstruents and nasals*

Coetzee and Pater (2006) start by accounting for the phonotactics of Muna obstruents and nasals. Liquids are brought into the discussion later. We follow this presentational order here.[6] The analysis focuses on those segment types that can be compared across places of articulation. These segments are listed in (33).

(33) Muna obstruents and nasals

labials	/p, b, f, m/
coronals	/t, d, s, n/
dorsals	/k, g, ʁ, ŋ/

The following segment pairs are compared:

(34) Segment pairs

	LAB	COR	DOR
voiced stop + voiceless stop	b-p	d-t	g-k
nasal + voiced stop	m-b	n-d	ŋ-g
nasal + voiceless stop	m-p	n-t	ŋ-k
nasal + fricative	m-f	n-s	ŋ-ʁ
fricative + voiced stop	f-b	s-d	ʁ-g
fricative + voiceless stop	f-p	s-t	ʁ-k

Coetzee and Pater (2008) observe that agreement in the features [αsonorant], [αstricture], and [αvoice] contributes to the OCP effect. They use linear regression to show that each

[6] Following Coetzee and Pater 2006, we will not discuss prenasals. These are analyzed in Coetzee and Pater 2008.

feature independently contributes to the lowering of the O/E value of the segment pair. This motivates positing the OCP-constraints in (35):

(35) Constraints for Muna obstruents and nasals

OCP-DOR[αson, βstric]	OCP-LAB[αson, βstric]	
OCP-DOR[αvoice]	OCP-LAB[αvoice]	OCP-COR[αvoice]
OCP-DOR[αstric]	OCP-LAB[αstric]	OCP-COR[αstric]
OCP-DOR[αson]	OCP-LAB[αson]	OCP-COR[αson]
OCP-DOR	OCP-LAB	OCP-COR

The constraints in (35) are identical to those in Coetzee and Pater 2006, except that we have followed Coetzee and Pater (2008) in replacing the feature [αcont] with the feature [αstric]. The difference will matter in fricative-liquid pairs: the sequences [s-l] and [s-r] both violate OCP-COR[αcont] because both segments are [+continuant], but they do not violate OCP-COR[αstric] because the segments differ in stricture (fricative vs. liquid). This modification results in a better empirical fit.

The constraints are interpreted as follows. A general OCP-constraint (e.g. OCP-LAB) is violated if and only if adjacent consonants share place of articulation. A specific OCP-constraint (e.g. OCP-LAB[αcont]) is violated if and only if adjacent consonants share place of articulation and the value of the specified feature. This is illustrated in (36).

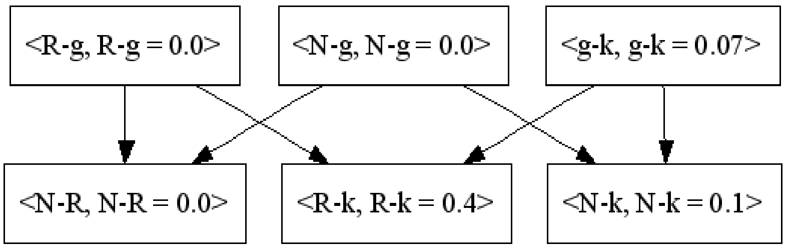
(36) Sample constraint violations

		FAITH	OCP-LAB [αstric]	OCP-LAB [αson]	OCP-LAB
(a) m-p	m-p		*		*
	OTHER	*			
(b) f-b	f-b			*	*
	OTHER	*			
(c) m-f	m-f				*
	OTHER	*			

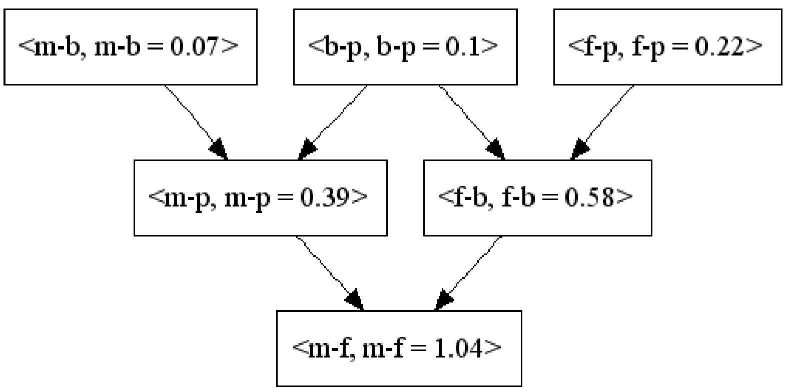
We start by computing the T-order based on the eighteen inputs in (34) and the fourteen constraints in (35), without assuming any rankings. The resulting T-order consists of three disjoint graphs, one for each place of articulation. These graphs are shown in (37). The nodes have been annotated by the Muna O/E values reported in Coetzee and Pater 2008. Notice two typographical substitutions: N = ŋ, R = ʁ.

(37) T-order with no rankings: precision = 0.958, recall = 0.146

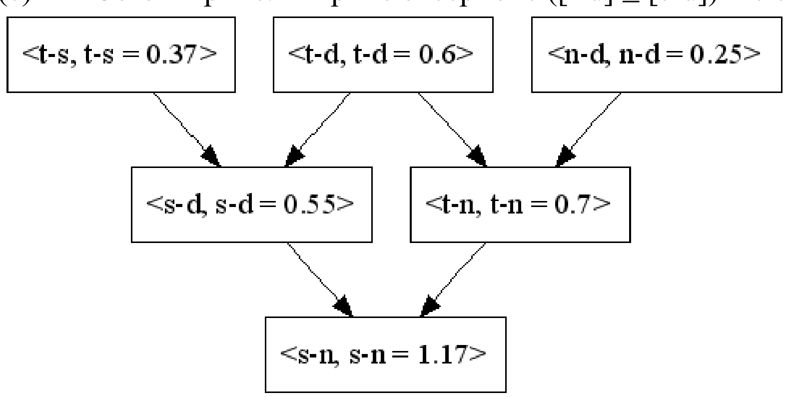
(a) Dorsal pairs. All pairs are correctly ordered.



(b) Labial pairs. All pairs are correctly ordered.



(c) Coronal pairs. All pairs except one ([t-d] ≤ [s-d]) are correctly ordered.



The predicted implicational universals hold up in the dictionary very well. The precision value (precision = 0.958) suffers from one minor error: [s-d] is incorrectly predicted to have a higher O/E value than [t-d]. The source of the problem is diagnosed by the violation tableau in (38): [t-d] incurs a superset of the violations of [s-d] due to the constraint OCP-COR[αstric]. This means that the rankings required for the faithful mapping <t-d, t-d> entail those required for the faithful mapping <s-d, s-d>, which means that the former dominates the latter in the T-order.

(38) Comparing the violation patterns of </s-d/, [s-d]> and </t-d/, [t-d]>

		FAITH	OCP-COR[αstric]	OCP-COR[αson]	OCP-COR
s-d	s-d (O/E = 0.55)			*	*
	OTHER	*			
t-d	t-d (O/E = 0.60)		*	*	*
	OTHER	*			

The low recall value (recall = 0.144) indicates that there are many quantitative relationships in the data that are not captured by the universal constraints themselves. This suggests that language-specific rankings are needed as well.

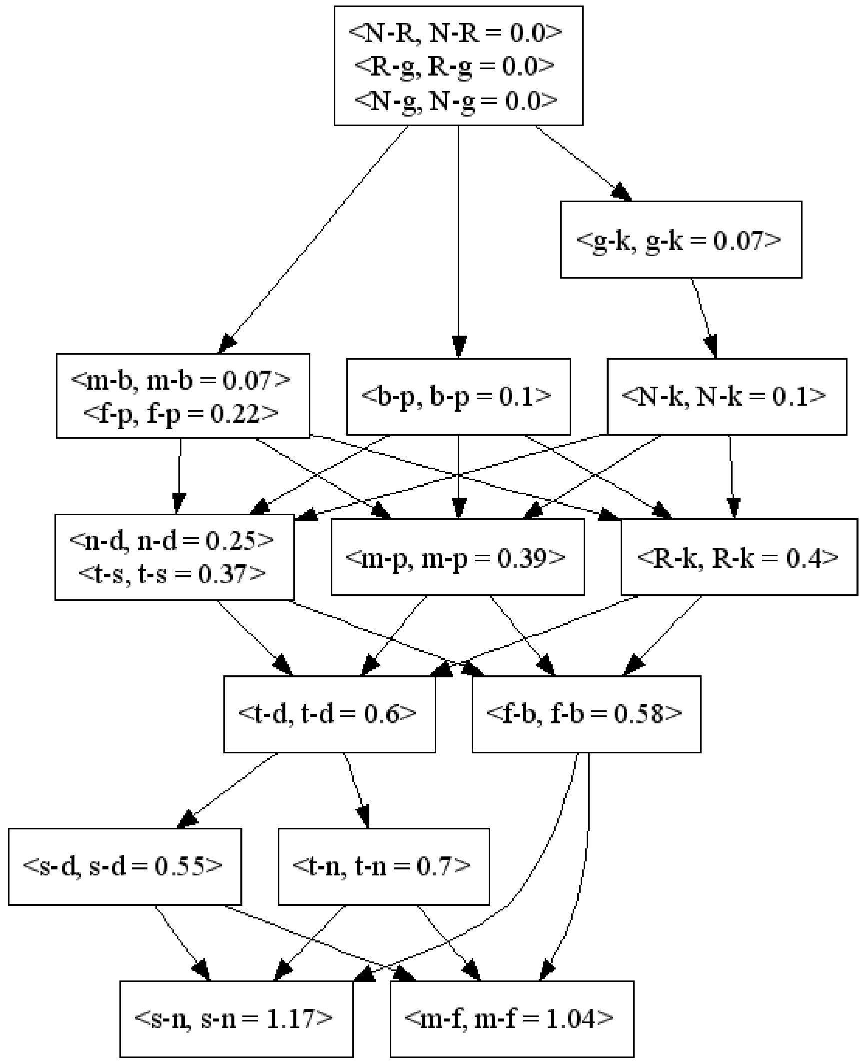
In addition to the constraints in (35), Coetzee and Pater (2006) posit rankings specific to Muna. They establish five internally unranked strata using the following heuristic: the smaller the O/E value of a consonant pair, the higher-ranked the markedness constraint targeting that pair. The ranking for Muna is shown in (39). The five strata correspond to the five groups of O/E values listed in the right hand column. The O/E values range from unattested (e.g. [ŋ-g], O/E = 0) to overrepresented (e.g. [n-s], O/E = 1.17).

(39) The constraint ranking for Muna

	TARGETED PAIRS	O/E VALUES
Stratum 1:		
OCP-DOR[αvoice]	ŋ-g, ʁ-g, ŋ-ʁ	0, 0, 0
Stratum 2:		
OCP-DOR[αson, βstric]	g-k	0.07
OCP-DOR[αstric]	ŋ-k	0.10
OCP-LAB[αson, βstric]	b-p	0.10
OCP-LAB[αvoice]	f-p, m-b	0.22, 0.07
Stratum 3:		
OCP-DOR[αson]	ʁ-k	0.40
OCP-DOR		
OCP-LAB[αstric]	m-p	0.39
OCP-COR[αvoice]	n-d, s-t	0.25, 0.37
Stratum 4:		
OCP-LAB[αson]	f-b	0.58
OCP-COR[αson]	s-d, d-t	0.55, 0.60
OCP-COR[αstric]	d-t, n-t	0.60, 0.70
Stratum 5:		
OCP-LAB	m-f	1.04
OCP-COR	n-s	1.17

We now add these rankings into the grammar and recompute the T-order. The result is shown in (40). The diagram is simplified by collapsing cycles into boxes.

(40) The T-order for Muna obstruents and nasals with the rankings posited by Coetzee and Pater (2006): precision = 0.979, recall = 0.867



This graph shows that Coetzee and Pater's (2006) constraints and rankings predict the ordering of the O/E values very well. The precision value is now nearly perfect (precision = 0.979) and the recall value improves dramatically (recall = 0.867). There are three minor problems on the precision side. The first problem is familiar: [t-d] (O/E = 0.60) and [s-d] (O/E = 0.55) come out in the wrong order. This problem is inherited from the constraints and for this reason cannot be fixed by adding rankings. The remaining two problems originate from the rankings: [m-b] and [f-p] (O/E = 0.07, 0.22) are predicted to have identical O/E values; so are [n-d] and [t-s] (O/E = 0.25, 0.37). The problem lies in the high-ranked constraints OCP-LAB[αvoice] and OCP-COR[αvoice], respectively. As shown in (41) and (42), these two constraints render irrelevant all the lower-ranked constraints that would potentially distinguish between these sequences. As we will see shortly, the problem with [n-d] and [t-s] in (42) will disappear with the introduction of more constraints on coronals.

(41) OCP-LAB[αvoice] is ranked too high

		FAITH	OCP-LAB [αvoice]	OCP-LAB [αstric]	OCP-LAB [αson]	OCP-LAB
m-b	m-b (O/E = 0.07)		*	*		*
	OTHER	*				
f-p	f-p (O/E = 0.22)		*		*	*
	OTHER	*				

(42) OCP-COR[αvoice] is ranked too high

		FAITH	OCP-COR [αvoice]	OCP-COR [αstric]	OCP-COR [αson]	OCP-COR
n-d	n-d (O/E = 0.25)		*	*		*
	OTHER	*				
t-s	t-s (O/E = 0.37)		*		*	*
	OTHER	*				

These residual problems aside, the coverage of the analysis is impressive, showing that Coetzee and Pater's constraints and rankings are on the right track. The fact that the relative ordering of O/E values in the Muna lexicon can be largely derived from simple and general phonological constraints ranked in a language-specific manner lends strong support to the view that gradient phonotactics has a grammatical basis instead of just being a by-product of lexical frequencies. A grammatical theory of gradient phonotactics predicts that only some quantitative patterns should be attested. In this section, we have seen that these predictions are largely borne out in Muna. The situation is very different in a theory with no grammatical constraints over possible lexicons. In such a theory, it would appear that any quantitative patterns should be possible.

We conclude with a caveat. A limitation inherent in the dictionary methodology is that the number of observations is often very small precisely because of the OCP-effect. An extreme example is the difference between /ʁ-g/ (O/E = 0/9 = 0) and /ŋ-ʁ/ (O/E = 0/8 = 0) where adding one word of the former type would contradict the predicted universal ordering. Another example is the difference between /g-k/ (O/E = 2/30 = 0.07) and /ŋ-k/ (O/E = 3/30 = 0.1) where adding two words of type /g-k/ or subtracting two words of type /ŋ-k/ would have the same effect. In other cases, the difference is more robust. For example, in the case of /s-d/ (O/E = 21/38 = 0.55) vs. /s-n/ (O/E = 50/43 = 1.17) one would need 24 more /s-d/-words to contradict the prediction. Psycholinguistic experimentation would be an alternative way to explore the relative grammaticality of these consonantal sequences. In any case, the theory derives clear and fine-grained predictions that can be straightforwardly tested once experimental data become available.

4.2 *Adding in the liquids*

After covering the phonotactics of obstruents and nasals, Coetzee and Pater (2006) extend the analysis to pairs that contain the liquids /l,r/. This results in 9 additional pairs:

(43) Segment pairs containing liquids

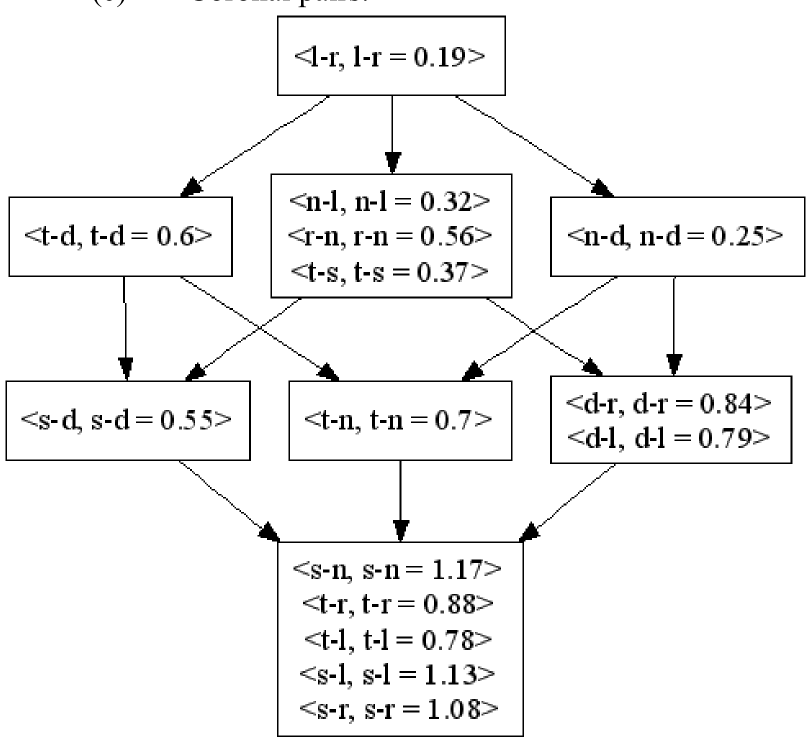
	/l/	/r/	/l+r/
voiceless stop	t-l	t-r	l-r
voiceless fricative	s-l	s-r	
voiced stop	d-l	d-r	
nasal	n-l	r-n	

Coetzee and Pater (2006) introduce the two additional constraints shown in (44). The first targets the pairs [n-d] and [l-r], where the segments agree in both stricture and voice. The second targets [t-s], [l-r], [n-l], and [r-n], where the segments agree in both sonorancy and voice.

(44) Two additional constraints for liquids
(a) OCP-COR[αstric, βvoice]
(b) OCP-COR[αson, βvoice]

We start by recomputing the T-order, again first without rankings. The result is shown in (45). Only the coronal graph is shown. The labial and dorsal graphs remain the same as before.

(45) T-order with no rankings: precision = 0.845, recall = 0.275
(a) Dorsal pairs. All pairs are correctly ordered. (= (37))
(b) Labial pairs. All pairs are correctly ordered. (= (37))
(c) Coronal pairs:



We now compare the success of the augmented grammar (obstruents, nasals, liquids) to the success of the original grammar (obstruents, nasals). The grammar with liquids has a higher recall, but a lower precision than the grammar without liquids.

(46) Precision and recall before and after liquids, with no rankings assumed

	INPUTS/CONSTRAINTS	RANKINGS	PRECISION	RECALL
(a)	Obstruents, nasals	none	0.958	0.146
(b)	Obstruents, nasals, liquids	none	0.845	0.275

Why do we get the drop in precision? The T-order turns out to contain 18 incorrect entailments. 14 of them arise because the grammar is not fine-grained enough to distinguish among candidates with distinct O/E values. The violation patterns in (47) reveal three groups of indistinguishable consonant pairs: all the pairs within each group ((47a), (47b), (47c)) violate exactly the same constraints. The observed differences among the O/E values thus cannot be derived from the present grammar. Either more constraints will be needed or the differences must have an extragrammatical explanation.

(47) Indistinguishable pairs

	FAITH	OCP-COR [αson, βvce]	OCP-COR [αson]	OCP-COR [αvce]	OCP-COR
(a) s-n, s-l, s-r, t-r, t-l					*
OTHER	*				
(b) n-l, r-n, t-s		*	*	*	*
OTHER	*				
(c) d-r, d-l				*	*
OTHER	*				

(48) Indistinguishable pairs and actual O/E values

(a)	s-n	1.17	(b)	r-n	0.56	(c)	d-r	0.84
	s-l	1.13		t-s	0.37		d-l	0.79
	s-r	1.08		n-l	0.32			
	t-r	0.88						
	t-l	0.78						

The remaining four incorrect entailments involve cases where the grammar predicts an asymmetry, but the O/E values pattern in the opposite way. As shown in (49), in all these cases the numerical differences are small. The violation patterns in (50) show that the constraints responsible for these asymmetries are OCP-COR[αstric], OCP-COR[αvce], and OCP-COR[αson, βvce]. The incorrect asymmetry in (50c) is familiar as it already occurred in the original T-order that did not include liquids.

(49) Predicted asymmetries and actual O/E values

(a)	d-r, d-l (O/E = 0.84, 0.79)	≤	t-l (O/E = 0.78)
(b)	r-n (O/E = 0.56)	≤	s-d (O/E = 0.55)
(c)	t-d (O/E = 0.60)	≤	s-d (O/E = 0.55)

(50) Incorrect asymmetries

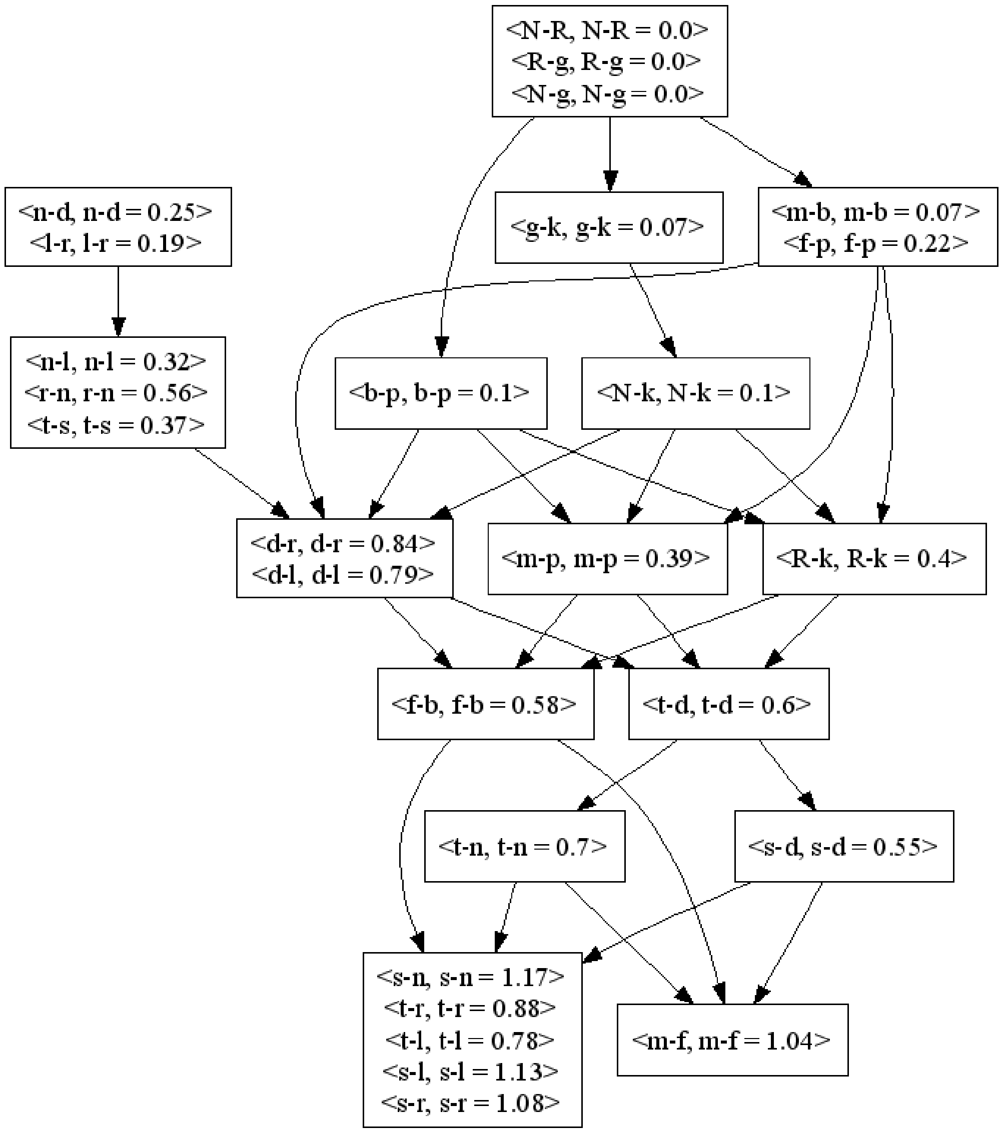
	FAITH	OCP-COR [αstric]	OCP-COR [αson, βvce]	OCP-COR [αson]	OCP-COR [αvce]	OCP-COR
(a1) d-r, d-l					*	*
OTHER	*					
(a2) t-l						*
OTHER	*					
(b1) r-n			*	*	*	*
OTHER	*					
(b2) s-d				*		*
OTHER	*					
(c1) s-d				*		*
OTHER	*					
(c2) t-d		*		*		*
OTHER	*					

Finally, Coetzee and Pater (2006) rank the new constraints as shown in (51). They do not explicitly indicate where the new constraints belong in the existing stratum ordering. This of course is not a flaw, but simply means that the ranking remains partial. For this reason we use the labels “Stratum X” and “Stratum Y”. The complete grammar with all the constraints and rankings yields the T-order in (52).

(51) The four-stratum constraint ranking for constraints on liquids

	TARGETED PAIRS	O/E VALUES
Stratum X:		
OCP-COR[αcont, βvoice]	n-d, l-r	0.25, 0.19
Stratum Y:		
OCP-COR[αson, βvoice]	n-l, r-n, t-s	0.32, 0.56, 0.37
Stratum 4:		
OCP-COR[αson]	s-d, t-d	0.55, 0.60
Stratum 5:		
OCP-COR	t-l, t-n, d-l, d-r, t-r, s-l, s-r, s-n	0.78, 0.70, 0.79, 0.84 0.88, 1.13, 1.08, 1.17

(52) T-order with all rankings: precision = 0.906, recall = 0.758



Assuming Coetzee and Pater's (2006) rankings, we now compare the success of the complete grammar to the success of the grammar for obstruents and nasals only. The inclusion of liquids turns out to lower both precision and recall:

(53) Precision and recall before and after liquids, with rankings

	Inputs/Constraints	Rankings	Precision	Recall
(a)	Obstruents, nasals	C&P 2006	0.979	0.867
(b)	Obstruents, nasals, liquids	C&P 2006	0.906	0.758

We have already identified the problems that arise from the constraints. How about the problems that arise from the rankings? We can identify them by taking the T-order with rankings ((52)), the T-order without rankings ((45)), and by computing their difference. Of the resulting 182 entailments that originate from the rankings 10 predict an asymmetry opposite to the observed O/E values. Let us consider each of these entailments in turn.

First, the high ranking of OCP-COR[αvoice] results in the eight problematic entailments in (54): the O/E values of [d-r] and [d-l] are higher than predicted. Illustrative tableaux are shown in (55) and (56).

(54) Predictions and O/E values (8 entailments):

(a) d-r, d-l (O/E = 0.84, 0.79) ≤ f-b (O/E = 0.58)

(b) d-r, d-l (O/E = 0.84, 0.79) ≤ t-n, t-d, s-d (O/E = 0.7, 0.6, 0.55)

(55) Problem: OCP-COR[αvoice] dominates OCP-LAB[αson] and OCP-LAB

		OCP-COR [αvce]	FAITH	OCP-LAB [αson]	OCP-LAB	OCP-COR
d-r, d-l	d-r, d-l	*!				*
	☞ OTHER		*			
f-b	☞ f-b			*	*	
	OTHER		*!			

(56) Problem: OCP-COR[αvoice] dominates OCP-COR[αstric] and OCP-COR[αson]

		OCP-COR [αvce]	FAITH	OCP-COR [αstric]	OCP-COR [αson]	OCP-COR
d-r, d-l	d-r, d-l	*!				*
	☞ OTHER		*			
t-d, t-n, s-d	☞ t-d, t-n, s-d			*	*	*
	OTHER		*!			

Second, the high ranking of two voice-related constraints predicts identical O/E values for the consonant sequences in (57). The problem is that OCP-COR[αstric, βvoice] and OCP-LAB[αvoice] render irrelevant all the lower-ranked constraints that could potentially distinguish between these sequences. The problem in (57b) is familiar as it was already part of the grammar without liquids. Illustrative tableaux are shown in (58) and (59).

(57) Predictions and O/E values (2 entailments):

(a) n-d (O/E = 0.25) = l-r (O/E = 0.19)

(b) m-b (O/E = 0.07) = f-p (O/E = 0.22)

(58) OCP-COR[αstric, βvoice] dominates OCP-COR[αson, βvoice] and OCP-COR[αson]

		OCP-COR [αstric, βvce]	FAITH	OCP-COR [αson, βvce]	OCP-COR [αvce]	OCP-COR [αstric]	OCP-COR [αson]	OCP-COR
n-d	n-d	*!			*	*		*
	☞ OTHER		*					
l-r	l-r	*!		*	*	*	*	*
	☞ OTHER		*					

(59) OCP-LAB[αvoice] dominates OCP-LAB[αstric] and OCP-LAB[αson]

		OCP-LAB [αvoice]	FAITH	OCP-LAB [αstric]	OCP-LAB [αson]	OCP-LAB
m-b	m-b	*!		*		*
	☞ OTHER		*			
f-p	f-p	*!			*	*
	☞ OTHER		*			

We conclude by an interim summary. We took Coetzee and Pater's (2006) optimality-theoretic constraints and rankings for Muna and recast their analysis in terms of the Complexity Hypothesis that assumes neither lexically indexed constraints nor weighted constraints. The Complexity Hypothesis attributes the relative well-formedness of a lexical item to its relative grammatical complexity measured in terms of ranking information: more complex items should have lower O/E values; less complex items should have higher O/E values. We saw how the theory orders linguistic structures in an implicational hierarchy that is predicted to correlate with their relative well-formedness: if a more complex structure *P* is well-formed, then a less complex structure *Q* should be at least as well-formed. Some of these implications are predicted to hold for all languages, no matter how the constraints are ranked; other implications are predicted to hold given certain language-specific rankings. We then showed concretely how the entire implicational structure can be computed. Finally, we found that the Complexity Hypothesis is remarkably consistent with the Muna facts as analyzed by Coetzee and Pater (2006, 2008), some residual problems notwithstanding.

4.3 *Alternative analyses*

There are several ways one can try to improve the above analyses of Muna phonotactics. One possibility is to try different constraints. In our discussion of Muna, we assumed the constraints and rankings proposed by Coetzee and Pater (2006), except that we followed Coetzee and Pater (2008) in replacing the feature [αcont] with the feature [αstric] which results in better precision and recall. How would the result change if we adopted all of Coetzee and Pater's (2008) constraints? It turns out that the universal T-order would remain exactly the same. In particular, all the precision errors would carry over to the new analysis. This suggests that adopting the new constraints would not result in any major improvements.

Another possibility is to try different rankings. We saw that several minor precision errors arose from the high ranking of constraints related to [voice]. The obvious fix is to keep these constraints in the grammar, but leave them unranked. It turns out that this indeed helps in terms of precision, but hurts in terms of recall. If we leave the constraints OCP-LAB[αvoice], OCP-COR[αvoice], and OCP-COR[αstric, βvoice] unranked, we gain in precision (precision = 0.931, up from 0.906), but lose in recall (recall = 0.680, down from 0.758). This illustrates the trade-off between precision and recall: it is easy to achieve high precision at the cost of low recall and vice versa. Put differently, it is easy to avoid incorrect predictions by not making predictions at all and by making predictions one runs the risk of making some incorrect ones as well. The challenge is to modify the rankings in ways that improve precision without simultaneously compromising recall.

Yet another possibility is to experiment with different sources of implicational universals (Kiparsky 1994, De Lacy 2002, McCarthy 2002:19-22, Pater and Coetzee 2006:25). One possibility is to assume UNIVERSAL RANKINGS. In this case, by ranking OCP-Place constraints for labials and dorsals universally higher than the corresponding OCP-Place constraints for coronals we would guarantee that OCP-Place effects are weakest in coronals. This is illustrated in (60).

(60) OCP-Place constraints in universal ranking

		OCP-LAB	OCP-DOR	OCP-COR
b-p	☞ b-p	*		
	OTHER			
g-k	☞ g-k		*	
	OTHER			
t-d	☞ t-d			*
	OTHER			

Another possibility is to assume that the OCP-Place constraints stand in a STRINGENCY RELATION. In such a grammar, there are no constraints that refer to the unmarked coronal place. Instead, there are general constraints that register OCP-Place violations irrespective of place. As a consequence, OCP-Place violations for labials and dorsals would be a proper superset of OCP-Place violations for coronals and the implicational universals would arise from the constraints themselves. This is illustrated in (61).

(61) Place-OCP constraints in a stringency relation

		OCP-LAB	OCP-DOR	OCP-PLACE
b-p	☞ b-p	*		*
	OTHER			
g-k	☞ g-k		*	*
	OTHER			
t-d	☞ t-d			*
	OTHER			

Which of these alternatives works better for Muna? The results for four different grammars are given in (62) in order of descending precision and recall:

(62) Comparing four alternative grammars (all segment pairs included)

	CONSTRAINTS	RANKINGS	PRECISION	RECALL
(a)	original	C&P 2006	0.906	0.758
(b)	original	universal	0.895	0.5
(c)	stringent	none	0.895	0.5
(d)	original	none	0.845	0.275

The grammar with Coetzee and Pater's constraints and rankings ((62a)) is clearly the best of the four in terms of both precision and recall. The grammar with only the universal rankings ((62b)) is almost identical in terms of precision, but clearly worse in terms of recall, showing that Coetzee and Pater's (2006) language-specific rankings do real analytical work. We then compared universal rankings and stringent constraints ((62c)): this was done by replacing the OCP-COR constraints by OCP-PLACE constraints and by assuming that each violation of an OCP-LAB or OCP-DOR constraint also violates the respective OCP-PLACE constraint, if one exists. This grammar turned out to yield exactly the same result as the previous one, giving us no reason to prefer one formulation over the other. Finally, the lowest precision and the lowest recall resulted from the grammar with Coetzee and Pater's constraints but none of their rankings ((62d)).

5. CONCLUSION

Why do gradient phonotactic patterns exist? Some phoneme combinations are more marked than others. How is this reflected in the grammar? Marked phoneme combinations are grammatically more complex than unmarked phoneme combinations. The goal of this paper has been to make these intuitions explicit. Building on Coetzee and Pater's recent work on gradient phonotactics (Coetzee and Pater 2006, 2008) and Prince's work on ranking entailments (Prince 2002a, 2002b, 2006, 2007, Brasoveanu and Prince 2005), we have proposed an optimality-theoretic account of gradient phonotactics that assumes neither lexically indexed constraints nor weighted constraints. At the heart of the proposal is the Complexity Hypothesis that attributes the relative well-formedness of a lexical item to its relative grammatical complexity measured in terms of ranking information: the more complex the lexical item, the less well-formed it is. The theory derives an implicational hierarchy that orders linguistic structures in terms of their complexity and (by hypothesis) their relative well-formedness: if a more complex structure *P* is well-formed, then a less complex structure *Q* should be at least as well-formed. Some of these implications are predicted to hold for all languages, no matter how the constraints are ranked; other implications depend on language-specific rankings. We showed that the Complexity Hypothesis is supported by phonotactic evidence from Muna (Austronesian) as recently analyzed by Coetzee and Pater (2006, 2008), some residual problems notwithstanding.

Coetzee and Pater (2008) focus on the important question of learnability, showing that a HG weighting can be learned that reflects the gradient phonotactic patterns in the data. The present paper focuses on a different question: how does the theory constrain possible phonotactic patterns? More specifically, what does the theory admit and what does it exclude? What quantitative asymmetries must hold in all languages, no matter how the constraints are ranked (or weighted)? What quantitative asymmetries must hold given particular language-specific rankings (or weightings)? As we have seen, these questions are easy to answer in OT. It will be interesting to see how they can be answered in other frameworks, in particular HG.

We conclude by pointing out an empirical complication that will ultimately have to be addressed by any theory of lexical regularities: phonotactic constraints do not hold equally for all lexical items. This suggests that the phonotactic grammar of a language is unlikely to consist of a single constraint ranking or a single set of weighted constraints. The best-known case is that of expressives. For example, Finnish generally avoids syllables with identical onset and coda consonants, especially word-initially, especially if the consonants are sonorants (Karlsson 1982, 129). Thus, syllables like *lVl*, *rVr*, *mVm*, and *nVn* are systematically absent, but they are found in descriptive and hypocoristic words, and they are in productive phonostylistic use in expressives, e.g. *mimmi* ‘girl’, *mömmö* ‘drug’, *nynny* ‘wimp’, *löllö* ‘paunch’, *lälläri* ‘softie’. This is not a language-specific quirk: in a recent study of Dutch and Kambera (Austronesian), Klamer (2002) discusses similar patterns and proposes that semantically complex items favour structurally complex forms, and vice versa. Moreover, there are reasons to believe that this phenomenon is not limited to expressives: nouns, verbs, and adjectives are often phonologically different in systematic ways (see e.g. Anttila 2002, 2006; Smith 1998). Ultimately, individual lexical items may differ from each other, suggesting that phonotactic patterns may need to be learned lexical item by lexical item. This is entirely compatible with the present proposal: lexical items may subscribe to different phonological constraint rankings; these rankings may differ in complexity; and differences in complexity are reflected in phonotactic probabilities, and perhaps even in semantics.

ACKNOWLEDGEMENTS

Parts of this work were presented at the Indiana University Phonology Fest (June 2006), Stanford Phonology Workshop (October 2006), and the 2007 LSA Linguistic Institute at Stanford University (July 2007). I thank Adam Albright, Lev Blumenfeld, Adrian Brasoveanu, Vivienne Fong, Michael Kenstowicz, Paul Kiparsky, Joe Pater, Alan Prince, Nathan Sanders, Paul Smolensky, and two anonymous reviewers for helpful comments and/or discussion. In particular, I would like to thank John McCarthy and Marc van Oostendorp for pointing out the connection between T-orders and Prince’s Elementary Ranking Conditions and Curtis Andrus for programming T-Order Generator and for many inspiring discussions. I am solely responsible for any errors.

REFERENCES

Albright, Adam. 2006. ‘Gradient phonotactic effects: lexical? grammatical? both? neither?’ Paper presented at the 80th Annual Meeting of the Linguistic Society of America, Albuquerque, NM.

Anttila, Arto. 2002. ‘Morphologically Conditioned Phonological Alternations’, *Natural Language and Linguistic Theory* 20, 1-42.

Anttila, Arto. 2007. ‘Word stress in Finnish’, Paper presented at the 81st Annual Meeting of the Linguistic Society of America, Anaheim, CA. Handout available at http://www.stanford.edu/~anttila/research/papers.

Anttila, Arto. 2008. 'Phonological constraints on constituent ordering', in Charles B. Chang and Hannah Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, Cascadilla Proceedings Project, Somerville, MA.

Anttila, Arto and Curtis Andrus. 2006. 'T-Order Generator', software package, Stanford University, http://www.stanford.edu/~anttila/research/software.html.

Anttila, Arto, Vivienne Fong, Stefan Benus, and Jennifer Nycz. to appear. 'Variation and opacity in Singapore English consonant clusters', *Phonology* (accepted with minor revisions).

Bailey, Todd M. and Ulrike Hahn. 2001. 'Determinants of wordlikeness: Phonotactics or lexical neighborhoods?' *Journal of Memory and Language* 44, 568-591.

Becker, Michael, Joe Pater and Christopher Potts. 2007. 'OT-Help 1.1', software package, University of Massachusetts, Amherst, http://people.umass.edu/mbe/software.html.

van den Berg, René. 1989. *A Grammar of the Muna Language*, Foris, Dordrecht.

van den Berg, René and La Ode Sidu. 1996. *Muna-English Dictionary*, KITLV Press, the Netherlands.

Berkley, Deborah Milam. 1994a. 'Variability and Obligatory Contour Principle Effects', in Beals et al., (eds.), *CLS 30: Proceedings of the 30th Annual Meeting of the Chicago Linguistic Society, Volume 2: The Parasession on Variation and Linguistic Theory,* Chicago Linguistic Society, Chicago. pp. 1-12.

Berkley, Deborah Milam. 1994b. 'The OCP and gradient data', *Studies in the Linguistic Sciences* 24, 59-72.

Berkley, Deborah Milam. 2000. *Gradient Obligatory Contour Principle Effects*, Ph.D. dissertation, Northwestern University.

Boersma, Paul. 2004. 'A stochastic OT account of paralinguistic tasks such as grammaticality and prototypicality judgments', ROA-648.

Boersma, Paul & David Weenink. 2007. PRAAT, a system for doing phonetics by computer. Institute of Phonetic Sciences of the University of Amsterdam, http://www.praat.org/.

Brasoveanu, Adrian and Alan Prince. 2005. 'Ranking and Necessity, Part I: The Fusional Reduction Algorithm', ROA-794.

Coetzee, Andries W. 2008. 'Grammaticality and ungrammaticality in phonology', *Language* 84. Pre-print version available as ROA-945.

Coetzee, Andries W. and Joe Pater. 2006. 'Lexically ranked OCP-Place constraints in Muna', Ms., University of Michigan and University of Massachusetts, Amherst, ROA-842.

Coetzee, Andries W. and Joe Pater. 2008. 'Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic', Ms., University of Michigan and University of Massachusetts, Amherst. To appear in *Natural Language and Linguistic Theory*. Also available as ROA-950.

Coleman, John and Janet Pierrehumbert. 1997. 'Stochastic phonological grammars and acceptability', in *Computational Phonology. Third Meeting of the ACL Special Interest Group in Computational Phonology*, Association for Computational Linguistics, Somerset. pp. 49-56.

Cowan, J. 1979. *Hans Wehr: A Dictionary of Modern Written Arabic*, Otto Harrasowitz, Wiesbaden, Germany.

De Lacy, Paul. 2002. *The Formal Expression of Markedness*. Ph.D. dissertation, University of Massachusetts, Amherst, ROA-542.

Frisch, Stefan A., Nathan R. Large, and David B. Pisoni. 2000. 'Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords', *Journal of Memory and Language* 42, 481-496.

Frisch, Stefan and Bushra Zawaydeh. 2001. 'The Psychological Reality of OCP-Place in Arabic', *Language* 77, 91-106.

Frisch, Stefan A., Janet B. Pierrehumbert, and Michael B. Broe. 2004. 'Similarity avoidance and the OCP', *Natural Language and Linguistic Theory* 22(1), 179-228.

Goldsmith, John. 1976. *Autosegmental phonology*, Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Greenberg, Joseph. 1950. 'The patterning of root morphemes in Semitic', *Word* 6, 162-181.

Greenberg, J. H. and J. J. Jenkins. 1964. 'Studies in the psychological correlates of the sound system of American English', *Word* 20, 157-177.

Hammond, Michael. 2004. 'Gradience, phonotactics, and the lexicon in English phonology', *International Journal of English Studies*, 4(2), 1-24.

Hay, Jennifer, Janet Pierrehumbert, and Mary Beckman. 2004. 'Speech Perception, Well-Formedness, and the Statistics of the Lexicon', in J. Local, R. Ogden, and R. Temple, (eds.), *Papers in Laboratory Phonology VI*, Cambridge University Press, Cambridge UK, pp. 58-74. Also available at http://www.ling.canterbury.ac.nz/people/hay.shtml.

Hayes, Bruce, Bruce Tesar, and Kie Zuraw. 2003. 'OTSoft 2.1', software package, http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Karlsson, Fred. 1982. *Suomen kielen äänne- ja muotorakenne* [*The Phonological and Morphological Structure of Finnish*], Werner Söderström Osakeyhtiö, Helsinki.

Keller, Frank. 2006. 'Linear Optimality Theory as a Model of Gradience in Grammar', in Gisbert Fanselow, Caroline Féry, Ralph Vogel, and Matthias Schlesewsky, (eds.), *Gradience in Grammar: Generative Perspectives*. Oxford University Press, Oxford, pp. 270-287.

Kiparsky, Paul. 1994. 'Remarks on Markedness', Handout from TREND 2 (Trilateral Phonology Weekend), January 22, 1994, University of California, Santa Cruz, http://www.stanford.edu/~kiparsky/Papers/trend.pdf.

Kiparsky, Paul. 1995. 'The phonological basis of sound change', in John A. Goldsmith, (ed.), *The Handbook of Phonological Theory*, Blackwell Publishers, Cambridge, Massachusetts, pp. 640-670.

Kiparsky, Paul. 2008. 'Dvandvas, blocking, and the associative: the bumpy ride from phrase to word', Ms., Stanford University, http://www.stanford.edu/~kiparsky/Papers/sanskritcocompounds.pdf.

Klamer, Marian. 2002. 'Semantically motivated lexical patterns: A study of Dutch and Kambera expressives', *Language* 78(2), 258-286.

Kruskal, J.B.1983. 'An overview of sequence comparison: Time warps, string edits, and macromolecules', *SIAM Review* 25(2), 201-237.

Leben, William R. 1973. *Suprasegmental phonology*, Ph.D. dissertation, Massachusetts Institute of Technology.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Harmonic Grammar – A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations, in *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, Erlbaum, Hillsdale, NJ, pp. 388-395.

Legendre, Géraldine, Antonella Sorace, and Paul Smolensky. 2006. 'The Optimality Theory—Harmonic Grammar Connection', in Paul Smolensky and Géraldine Legendre, (eds.), *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*, *Vol. 2*. MIT Press, Cambridge, Massachusetts, pp. 339-402.

Manning, Christopher D. and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*, The MIT Press, Cambridge, Massachusetts.

McCarthy, John. 1988. 'Feature geometry and dependency: A review', *Phonetica* 45, 84-108.

McCarthy, John. 1994. 'The phonetics and phonology of Semitic pharyngeals', in Patricia Keating, (ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology* 3, Cambridge University Press, Cambridge, pp. 191-233.

McCarthy, John. 2002. *A Thematic Guide to Optimality Theory*, Cambridge University Press, Cambridge.

Ohala, J. and M. Ohala. 1986. 'Testing hypotheses regarding the psychological manifestation of morpheme structure constraints', in *Experimental Psychology*, Academic Press, Orlando, FL, pp. 239-252.

Padgett, Jaye. 1995. *Stricture in Feature Geometry*. Dissertations in Linguistics, Center for the Study of Language and Information, Stanford, CA.

Pater, Joe and Andries Coetzee. 2005. 'Lexically Specific Constraints: Gradience, Learnability, and Perception', *Proceedings of the Phonology-Morphology Circle of Korea,* Seoul, Korea, pp. 85-119.

Pierrehumbert, Janet. 1993. 'Dissimilarity in the Arabic verbal roots', in A. Schafer, (ed.), *NELS 23: Proceedings of the North East Linguistic Society,* GLSA, University of Massachusetts, Amherst, pp. 367-381.

Prince, Alan. 2002a. 'Entailed Ranking Arguments', ROA-500.

Prince, Alan. 2002b. 'Arguing Optimality', ROA-562.

Prince, Alan. 2006. 'Implication & Impossibility in Grammatical Systems: What it is & How to find it', ROA-880.

Prince, Alan. 2007. 'The pursuit of theory', in Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, Cambridge University Press, Cambridge. pp. 33-60.

Prince, Alan and Paul Smolensky 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*, Blackwell Publishing, Malden, Massachusetts.

Raffelsiefen, Renate. 1999. 'Phonological constraints on English word formation' in Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1998*, pp. 225-287.

Samek-Lodovici, Vieri and Alan Prince. 2005. 'Fundamental Properties of Harmonic Bounding', ROA-785.

Smith, Jennifer L. 1998. 'Noun faithfulness: On the privileged status of nouns in phonology', ROA-242.

Smolensky, Paul and Géraldine Legendre. 2006. *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*, MIT Press, Cambridge, Mass.

Vitevitch, M. S., P. A. Luce, J. Charles-Luce, and D. Kemmerer. 1997. 'Phonotactics and syllable stress: Implications for the processing of spoken nonsense words', *Language and Speech* 40, 47-62.

vander Wyk, Brent C. and James L. McClelland. 2007. 'Graded Constraints on English Word Forms', Poster presented at the *Workshop on Variation, Gradience and Frequency in Phonology*, Stanford University, July 7, 2007.