Andries W. Coetzee
University of Massachusetts Amherst, November 2002

Between-Language Frequency Effects in Phonological Theory

1. General explanation of problem

In the typological literature there has always been an acute awareness of the fact that many features of human language are not randomly distributed within the set of possible languages. Some of these non-random distributions should undoubtedly be ascribed to socio-historical factors. However, even if such factors are taken into consideration many linguistic features with non-random distributional patterns still remain. In this paper it is assumed that one of the central aims of linguistic theory, should be to account for all significant generalizations about human language. In so far as these non-random distributional patterns are significant, their explanation therefore falls within the scope of formal linguistic theory.

This is a departure from most of the recent linguistic theories that aim to explain simply what is possible and what not (Hale & Rice 2000). However, there are many examples, especially from the typological literature, that claim that between-language frequency effects should also be explained by formal linguistic theory – see for instance Croft (1990), Comrie (1989), Saporta (1966), Greenberg (1966) and countless other texts on typology and general introductions to linguistic theory.

In this paper focus will be on accounting for between-language frequency effects in phonological theory. There is no mechanism in modern phonological theory that can account for between-language frequency effects – not in rule-based theories, rule-and-constraint-based theories, or pure constraint-based theories. This paper argues that Optimality Theory (Prince & Smolensky 1993) naturally lends itself to be extended in a way that will empower the theory to make predictions about between-language frequency distributions. To this effect, two proposals for the extension of OT are made in this paper:

(i) It is argued that the OT-model that Anttila (1995, 1997) developed to account for within-language variation and frequency effects, can be extended straightforwardly to make predictions about between-language frequency effects.

(ii) It is argued that, in order to account adequately for many non-random cross-linguistic frequency distributions, a new family of constraints (preference constraints) should be added to the grammar. Preference constraints state what the preferred ranking between pairs of other constraints is. By favoring certain rankings over others, preference constraints introduce a ranking bias into the system. They enable the theory to account for non-random patterns of between-language frequency distributions.

Although this paper focuses on phonological theory, the general issues discussed here are equally true of other components of the grammar also. The theory developed here for phonology should therefore ultimately be applied to fields like syntax also.

This paper is presented in the following sections: In §2 a few examples of the type of non-random distributional patterns that will be the focus of this paper are presented. In §3 Anttila's (1995, 1997) OT-model is extended so that it can also account for between-language frequency distributions. This section also contains an example of an application of the extended OT-model to the non-random distribution of roundness in vowel inventories. The failure of the theory to adequately account for cross-linguistic distribution of this feature, prompts the introduction of preference constraints in section §4. This section forms the central part of this paper. The next section (§5) discusses some of the theoretical implications of introducing preference constraints into the grammar. Section §6 considers a few alternatives to preference constraints, and concludes that preference constraints are to be preferred to the alternatives. Lastly, §7 summarizes and evaluates the proposals made in the paper.

2. Non-random between-language distribution and phonological theory

There is no formal mechanism in phonological theory to account for between language frequency distribution. In a rule-based theory like SPE (Chomsky & Halle 1968) it is at best possible to predict that a more natural rule (expressible with fewer features) should be encountered more frequently in the world's languages. However, this is a relative frequency prediction – more frequently than some other rule. Predictions about absolute frequencies are impossible in such a theory.

In rule-and-constraint-based theories (Paradis 1988, Rice 1987, Yip 1988) frequency predictions became even more difficult. In these theories, well-formedness conditions functioned as universal filters or constraints to which output patterns had to comply. However, these conditions were typically inviolable and therefore intended to express patterns that are universally observed. Variation between languages depended primarily on the way in which a specific language repaired a form that didn't conform to the universal well-formedness conditions, and there never was a fully developed theory to explain how the choice between different repair strategies was to be made.

OT, as a purely constraint-based theory of phonology, places strong limits on the possible variation between languages (via factorial typology). Different sound patterns and phonological processes are the result of different rankings of a set of universal constraints. But there is no formal way in classic OT to make predictions about the frequency with which certain patterns or processes will be encountered. Stated in terms of constraint rankings – there is no way to explain why certain rankings between the constraints are observed more frequently than others.

If all sound patterns and processes were randomly distributed between the languages of the world, the inability of phonological theory to predict frequency distributions would have been a non-issue. Then the distribution could have been ascribed to chance. However, there are many examples of non-randomly distributed patterns in the world's languages. A particularly rich source for such patterns is the UPSID database of phonemic inventories (Maddieson & Precoda 1992). In the compilation of this database, Maddieson and his co-workers have gone to great lengths to prevent genetic relationships between the languages in the database and also socio-historical factors to introduce biases in the data (cf. Maddieson 1984: 5-7). It is therefore accepted that the non-random patterns observed in this database are truly non-random patterns. There are also many other examples from especially the typological literature of non-random cross-linguistic distributional patterns. Below is a non-exhaustive list of examples. This list is presented here to motivate the claim that these non-random patterns of distribution are widely attested, and that phonological theory will gain much in terms of explanatory power if it can account for patterns like these. (Unless stated otherwise, the following patterns were identified in the UPSID database.)

(i) Both front rounded and back unrounded vowels are possible sounds. However, not one of these two groups of sounds is widely attested. Front rounded vowels are attested in roughly 10% of the world's languages, and back unrounded vowels in about 25%. There is also strong dependency relation between these two groups of sounds. Only 3% out of the languages with front rounded vowels, have these sounds to the exclusion of back unrounded vowels.[1] (For more on this pattern see §4.)

(ii) The distribution of aspiration on stops is very similar to that of roundness on vowels. Both voiced and voiceless aspirated stops are underrepresented in phonemic inventories, but voiced aspirated stops are more restricted in their distribution than voiceless aspirated stops.[2]

(iii) Metathesis is a possible repair for unacceptable syllable structure, but it is encountered much less frequently than either epenthesis or deletion (Hume 1998 & to appear, Ultan 1978, Hock 1985). (See §4.1 for more.)

(iv) Both leftward and rightward spreading of nasality is possible. However, rightward spreading is much more common than leftward spreading (Walker 1998: 65).[3]

(v) In general, if a language allows nasal consonants into its inventory, the coronal nasal /n/ will be among the nasal consonants. However, there are a few languages that are exceptions and that have /m/ as their only nasal consonant (Ferguson 1966).[4]

---

1 English is an example of a language with neither front rounded, nor back unrounded vowels. Farsi falls in the group of languages with back unrounded vowels, but no front rounded vowels (Thackston 1993:xv-xvi). German is an example of a language with both front rounded and back unrounded vowels (Maddieson 1984:265). Albanian has a front rounded vowel, but no back unrounded vowel (Maddieson 1984 and Newmark 1957).

2 Most languages, like English, has no phonemic aspiration. There are several languages with aspiration only on voiceless stops – Norhern Sotho for instance (Poulos & Louwrens 1994:420-421). Languages like Hindi has phonemic aspiration both on voiced and voiceless stops (Maddieson 1984:270). It sometimes claimed that Javanese has voiced aspirated stops but no voiceless aspirated stops in its inventory. For more on the Javanese case, see footnote 19 below.

3 Languages with rightward spreading are plentiful. Walker mentions among others Sundanese, Dayak, Madurese, Warao, and Tuyaca. There are fewer languages with leftward spreading. Walkers mentions *inter alia* the following as examples: Yoruba, Isoko, Hindi, and English. (Walker 1998:67-79.)

4 Ferguson (1966:56) mentions Winnebago and Yoruba as possible examples of languages with phonemic /m/, but no phonemic /n/.

(vi) Most languages do not allow any syllabic consonants. It is also generally true that a language that allows a consonant of certain sonorancy to be syllabic, will also allow all consonants of higher sonorancy to be syllabic. However, there are exceptions to this generalization. There is a small group of languages that allow non-sonorant consonants into syllabic peak position, but do not tolerate syllabic sonorant consonants. Out of a sample of 182 languages Bell (1978) found about 5.5% of this type.[5]

Phonological theory can account for the fact that all of these patterns are possible in human language. But phonological theory has no formal way for accounting for the relative or absolute frequency with which certain patterns or processes are found. This paper claims that an adequate theory of phonology should be able to do this. A solution to this problem is proposed within the constraint-based theory of phonology, OT.

3. Between-language frequency predictions in Optimality Theory

3.1 Extending Anttila's model to cross-linguistic variation

In OT, cross-linguistic variation follows from factorial typology. All possible rerankings of the constraints in the universal constraint set should result in possible human languages. Also, all possible human languages should follow from this. Through factorial typology OT is able to account for much of the typological variation in the languages of the world. However, it cannot account for non-random distribution of patterns and processes within the set of possible languages.

Anttila's (1995, 1997) model of an OT grammar has partially addressed this problem. Anttila has extended the role that factorial typology plays in OT to include also predictions about frequency.[6] He proposes that intra-language variation results from crucial non-ranking between conflicting constraints. Every time the grammar evaluates an input-output pair, one total ranking of the unranked constraints is randomly chosen. Variation is then the result of choosing different rankings of the unranked constraints.

But Anttila also makes frequency predictions. Even in cases of free variation, it often happens that some variant is found more frequently than another. Anttila assumes that each of the possible rankings between the unranked constraints has an equal likelihood of being chosen. Here it is important that nominally distinct rankings between the constraints can converge on the same output. From this it then follows that a variant that is chosen as optimal in a greater portion of the possible rankings between the unranked constraints, is predicted to occur more frequently.

---

5 Bell (1978:158) mentions as examples with only obstruent syllabics Acoma, Caro, Sinhalese and Chipaya among other.

6 See also Anttila and Cho (1998), Cho (1998) and Zubritskaya (1995, 1998).

This model of within-language variation can easily be extended to between-language variation. Factorial typology will then dictate not only what the possible patterns and processes are, but also with what frequency different patterns and processes are found.

(1) Frequency predictions with regard to inter-language variation

If **n** is the number of possible rankings resulting in a certain pattern or process, and **t** is the total number of possible rankings, then the frequency with which this specific pattern or process will be encountered cross-linguistically is predicted to be **n/t**.[7]

Without adding anything to the architecture of the grammar, it is therefore possible to enable an OT grammar to make predictions about cross-linguistic frequency distribution. It simply requires adding an additional function to constraint reranking (factorial typology), a device already available in theory.

An account can now be given of not only the possible and impossible, but also of the more and less frequent. This places an additional burden on explanation in OT – it should also account for significant patterns of frequency distribution.

3.2 The cross-linguistic distribution on front rounded and back unrounded vowels

By giving factorial typology and extended role, an OT grammar can now also make predictions about the frequency with which some pattern should be encountered. The next important question is then whether an OT grammar is able to adequately account for non-random patterns of distribution. To answer this question, the cross-linguistic distribution of vowel roundness in phonemic inventories is investigated in more detail in this section. Below the distribution of roundness in vowel inventories in the UPSID database is presented first. Then the predictions that an OT grammar (extended as explained above in §3.1) makes about this distribution are explained. Comparison between the actual distributional patterns and the predictions made by OT, will show that OT cannot adequately account for non-random patterns of distribution. This will prompt the introduction of preference constraints in section §4.

Three generalizations can be made about the distribution of roundness on vowels: (i) There is a strong tendency for phonemic inventories not to include front rounded or back unrounded vowels. (ii) There are significantly fewer languages with front rounded, than with back unrounded vowels. (iii) The presence of front rounded vowels also depends on the presence of back unrounded vowels. The table below summarizes the co-occurrence patterns of back unrounded and front rounded vowels in the UPSID database.

---

[7] This frequency is independent from extra-grammatical influences. Factors such as genetic relationship between languages, language contact, historical prominence, etc., can cause deviations from the predicted frequency.

(2) Back unrounded and front rounded vowels in the UPSID database[8]

Front rouded?	Back unrounded?	Number	Percentage
No	No	324	72
No	Yes	80	18
Yes	Yes	32	7
Yes	No	15	3
	Total	451	

Table (2) shows that, as is well known, back vowels tend to be rounded, and front vowels to be unrounded. There are perceptual reasons for this. Back and front vowels of the same height differ in that back vowels have a lower F_2. Lip rounding in back vowels increases the length of the cavity in front of the tongue constriction, pushing F_2 down further.[9] This reinforces the acoustic difference of back vowels from front vowels. In front vowels lip rounding also lowers F_2. In order to maximize the difference between front and back vowels, it is therefore better not to round the lips in front vowels. There is an obvious perceptual advantage to have contrasting phonemes as different as possible acoustically (Kirchner 1998, Flemming 1995, Crothers, 1978). This accounts for the general avoidance of both back unrounded and front rounded vowels.

However, table (2) also shows that front rounded vowels are more underrepresented than back unrounded vowels. One possible reason is that the formation of front rounded vowels is more costly than that of back unrounded vowels. In front rounded vowels, at least two active articulators are involved - the tongue and the lips. In the back unrounded vowels only the tongue is active – under the assumption that [unrounded] and [spread lips] are different features. Both front rounded and back unrounded vowels are marked on perceptual grounds, but front unrounded vowels are also marked on articulatory grounds.[10]

Roundness on vowels can be regulated through the markedness constraints against the front rounded and back unrounded vowels, *FRRD and *BKUNRD, and the faithfulness constraint requiring faithful parsing of the feature [round], IDENT(rnd). In OT there are two ways in which these constraints can interact. They can be allowed to rerank freely, or the ranking between the markedness constraints can be fixed: *FRRD >> * BKUNRD. Fixed rankings are typically used in OT to capture implicational relationships – the presence of front rounded vowels implies the presence of back unrounded vowels.

---

8 A χ^2-test on these numbers yields a value of 53.01 and a p-value of 1.8×10^{-11}. The pattern of distribution is significantly different from chance.

9 See Stevens (1998: 290-4) for a discussion of the interaction of lip rounding and F_2-values.

10 See Kirchner (1998) and Flemming (1995) for a formal incorporation of articulatory cost into OT.

Since the rankings that are possible in each of these two ways of thinking about the interaction between the constraints are different, they will result in different predictions about the frequency with which output patterns will be observed. The rankings possible with each of these options and the output patterns associated with each ranking are given below.

(3) Free reranking – 6 possible rankings

Ranking	**Front round?**	**Back unround?**
*FRRD >> *BKUNRD >> IDENT(rnd)	No	No
*BKUNRD >> *FRRD >> IDENT(rnd)	No	No
*BKUNRD >> IDENT(rnd) >> *FRRD	Yes	No
*FRRD >> IDENT(rnd) >> *BKUNRD	No	Yes
IDENT(rnd) >> *FRRD >> *BKUNRD	Yes	Yes
IDENT(rnd) >> *BKUNRD >> *FRRD	Yes	Yes
Total	**3**	**3**

(4) Universally fixed ranking between the markedness constraints: *FRRD >> *BKUNRD

Ranking	**Front round?**	**Back unround?**
*FRRD >> *BKUNRD >> IDENT(rnd)	No	No
*FRRD >> IDENT(rnd) >> *BKUNRD	No	Yes
IDENT(rnd) >> *FRRD >> *BKUNRD	Yes	Yes
Total	**1**	**2**

Using the metric to make cross-linguistic frequency predictions that was introduced in (1), the predictions that each of these two options make can be computed. Consider the example where the constraints are allowed to rerank freely: with two out of the six possible rankings, neither front rounded nor back unrounded vowels are attested. It is therefore predicted that 2/6 = 33% of languages should have this pattern. There are also two rankings that will result in inventories with both front rounded and back unrounded vowels. It is therefore also predicted that 33% of languages should show this pattern. There is one ranking that results in a inventory with front rounded but no back unrounded vowels, and also one that results in an inventory with back unrounded but no front rounded vowels. It is therefore predicted that 1/6 = 16.5% of languages should fall in each of these two classes. By similar reasoning the predictions made a model with fixed rankings can also be computed. In the table below, these predictions are compared to the actually observed patterns in UPSID.

(5) Comparison between UPSID, free reranking, and universally fixed ranking

Front round?	**Back unround?**	**Actual**	**Free reranking**	**Fixed ranking**
No	No	72	33	33
No	Yes	25	16.5	33
Yes	Yes	7	33	33
Yes	No	3	16.5	0

The free reranking option does not account for the observed pattern of distribution at all. Under an option with free reranking of the constraints, a uniform distribution of front rounded and back unrounded vowels is predicted. Half of all languages are predicted to have front rounded vowels, and similarly for back unrounded vowels. The general avoidance of both of these classes of sounds is not captured. The dependence of front rounded in back unrounded vowels is also not captured at all.

The option with the universally fixed ranking between the markedness constraints does considerably better. It predicts that only 33% of all languages should have front rounded vowels – this is still far from the actually observed 10%, but the tendency to avoid these sounds is predicted. However, this approach also predicts that back unrounded vowels should be observed in 66% of all languages, which is far of the actually observed 32%. The fixed ranking option does capture the dependency of front rounded vowels on back unrounded vowels. However, it expresses this as an absolute fact, while it is only a strong tendency. It is predicted that no language with front rounded vowels but no back unrounded vowels should exist, while 3% of the languages in UPSID are classified into this group.

Even though OT can now make predictions about the frequency with which certain patterns should be observed, its predictions are not very accurate. Uniform distribution can easily be accounted for by allowing free reranking between constraints. However, as soon non-uniform patterns of distribution are observed the wrong predictions are made. There are two basic problems with the predictions: (i) If something is allowed as possible through factorial typology, it is very hard to restrict its occurrence considerably, while it is often the fact that some possible pattern is attested only very rarely. (ii) Universally fixed rankings account better for implicational universals (dependency relationships between patterns). However, whenever such a dependency relationship is only a strong tendency and not an absolute universal, a fixed ranking will exclude certain scarce but possible patterns from existing.

In §2 several examples were given of non-random distributional patterns and also of strong universal tendencies that are not absolute universals. The vowel roundness example is therefore not an isolated occurrence. It is indicative of a problem that will be faced with each of the examples discussed in §2, and with all other patterns like that. In the next section a new family of constraints, preference constraints, are introduced to address this problem.

4. Accounting for the non-random cross-linguistic frequency distributions

There are four possible patterns of co-occurrence of front rounded and back unrounded vowels. Based on the numbers in table (2) these four groups can be ordered from the most frequently to the least frequency attested. These four groups are ordered in this way in the table below. The constraint rankings that will result in each of these output patterns are also indicated in the table.

(6) Distribution patterns of vowel roundness in phonemic inventories

	Front rounded	**Back unrounded**	**Ranking**
Most frequent	No	No	*FRRD, *BKUNRD >> IDENT(rnd)
	No	Yes	*FRRD >> IDENT(rnd) >> *BKUNRD
	Yes	Yes	IDENT(rnd) >> *FRRD, *BKUNRD
Least often	Yes	No	*BKUNRD >> IDENT(rnd) >> *FRRD

From this table, it is possible to make the following inferences: (i) The most frequently attested pattern results only when both markedness constraints outrank the faithfulness constraint. (ii) There are more output patterns that follow from a ranking *FRRD >> *BKUNRD, than from the ranking *BKUNRD >> *FRRD. These two observations can be stated as ranking tendencies – IDENT(rnd) tends to be ranked in the lowest position, and *FRRD tends to outrank *BKUNRD.

In order to account successfully for the observed distributional pattern of font rounded and back unrounded vowels, it is necessary to formally incorporate ranking tendencies such as these into the theory. To achieve this a new family of constraints, preference constraints, are introduced into the theory. Preference constraints state the preferred ranking between pairs of constraints. Rankings that conform to these preferences are favored by these constraints. In this way a statistical bias is introduced into the predictions about frequency.

In section §4.1 below, preference constraints are introduced. The logic behind them and how they actually work are explained. In §4.2 the predictions that an OT grammar with preference constraints makes about the distribution of front rounded and back unrounded vowels are discussed, and compared to the predictions of an OT grammar without these constraints.

4.1 Introducing Preference Constraints

Preference constraints require pairs of other constraints to be ranked in a specific order. They therefore evaluate the ranking between constraints. However, OT is a theory in which candidate output forms are evaluated for their harmony relative to some constraint hierarchy. The ranking between the constraints, the hierarchy itself, cannot be evaluated in an OT grammar. The preference constraints are therefore formulated in such a way that they evaluate constraint rankings only indirectly through evaluating candidates. Preference constraints penalize candidates that benefit from non-preferred rankings. This is equivalent in effect, even if not in principle, to penalizing the non-preferred rankings, and therefore to favoring the preferred rankings. Preference constraints are defined as follows:

(7) General schema of preference constraint: C_1 prefers to outrank C_2

[C_1 >> C_2]
In a grammar with the ranking C_2 >> C_1, assign every candidate violating C_1 one violation mark.

This constraint states that the preferred ranking between the constraints C_1 and C_2, is C_1 >> C_2. When this preferred ranking is observed, candidates that violate C_1 will violate the higher ranking of the two constraints. They will therefore be less harmonic than candidates that violate C_2. Under the non-preferred ranking, C_2 >> C_1, this situation is reversed. Now candidates that violate C_1 will be more harmonic than candidates violating C_2. Candidates that violate C_1 therefore benefit from the non-preferred ranking. The preference constraint [C_1>>C_2] penalizes C_1-violators in grammars with the non-preferred ranking.

By penalizing the candidate that benefits from the non-preferred ranking, the preference constraint penalizes the non-preferred ranking, and therefore indirectly favors the preferred ranking. In this way it also introduces a statistical bias in the output patterns that the theory will predict. To see how this works, consider a constraint set that contains only the constraints C_1 and C_2. Suppose also that there are only two candidates to consider, namely C_1-violators [*C_1], and C_2-violators [*C_2]. There are two rankings possible between the constraints. Under one of them [*C_1] will be optimal, and under the other [*C_2].

(8) A grammar with only two constraints

a. C_1 >> C_2

	C_1	C_2
[*C_1]	*!	
☞ [*C_2]		*

b. C_2 >> C_1

	C_2	C_1
☞ [*C_1]		*
[*C_2]	*!	

Suppose that C_1 preferably outranks C_2. This also implies that the candidate [*C_2] should be optimal more than [*C_1]. This preference will be expressed by a preference constraint [C_1>>C_2]. When the preferred ranking, C_1 >> C_2, is observed, the default output will be [*C_2] – see (8a) above. Since the preference constraint does not assign any violations in these circumstances, it also has no influence on the predicted outputs. All grammars with the preferred ranking, will have the default output [*C_2]. However, when the non-preferred ranking, C_2 >> C_1, is observed, the default output is [*C_1] – see (8b) above. However, this output will not always be chosen. In these cases, the preference constraint will assign an additional violation to [*C_1]. If this preference constraint outranks C_2, then [*C_2] will be chosen as optimal candidate in spite of the ranking C_2 >> C_1. Therefore, in these grammars only in those instances where the preference constraint ranks below C_2, will the default output be chosen.

This is illustrated in the tableaux below. Tableau (9a) represents grammars where the preferred ranking is observed. Tableaux (9b) and (9c) represent grammars with the non-preferred ranking. In (9b) the preference constraint is ranked low enough that it will have no influence on the output. In (9c) it outranks the C_2, and therefore forces the choice of the non-default output.

(9) Preference constraint in action

a. [C_1 >> C_2] obeyed, and therefore its ranking has no influence on the output

	C_1	C_2	[C_1 >> C_2]
[*C_1]	*!		
☞[*C_2]		*	

b. [C_1 >> C_2] violated and below C_2

	C_2	C_1	[C_1 >> C_2]
☞[*C_1]		*	*
[*C_2]	*!		

c. [C_1 >> C_2] violated and above C_2

	[C_1 >> C_2]	C_2	C_1
[*C_1]	*!		*
☞ [*C_2]		*	

There are 3 constraints, and therefore a total of 6 possible rankings. There are three positions for [C_1 >> C_2] to rank into in tableau (9a). This tableau therefore represents 3 rankings. In (9b) there are two potions for [C_1 >> C_2] to rank into such that it still below C_2. This tableau therefore represents 2 possible rankings. Tableau (9c) represents only one possible ranking. It then follows that in this model it is predicted that [*C_1] should be chosen as output in 2 out of the 6 possible rankings, and [*C_2] in 4 out of the 6 possible rankings. The table below compares the predictions of a model without preference constraints to a model with preference constraints.

(10) Frequencies predicted by a model with [C_1>>C_2] and a model without [C_1 >> C_2]

Output	**Without [C_1 >> C_2]**	**With [C_1 >> C_2]**
[*C_1]	50	33
[*C_1]	50	67

This example makes it clear how preference constraints introduce a bias in the predicted output patterns. By penalizing outputs that benefit from the non-preferred ranking, preference constraints cause some grammars with the non-preferred ranking to choose as output the candidate that would have been chosen had the preferred ranking been observed. This ability to shift output patterns is a very general and important characteristic of preference constraints.

(11) Output shifting

A preference constraint can cause some grammars that would have yielded a specific output pattern, to yield a less marked output pattern. It "shifts" some grammars from a more marked to a less marked group.

Now that preference constraints have been introduced, they can be applied to problems like the distribution of front rounded and back unrounded vowels. This is done in the next section.

4.2 Front rounded and back unrounded vowels again

At the beginning of this section (§4), it was concluded that the pattern of cross-linguistic distribution of front rounded and back unrounded vowels in the phonemic inventories of the languages in UPSID points to a preference for the following rankings: (i) *BKUNRD, *FRRD >> IDENT(rnd), and (ii) *FRRD >> *BKUNRD. These two preferences can be combined into a single preferred ranking hierarchy:

(12) Preferred vowel roundness hierarchy

*FRRD >> *BKUNRD >> IDENT(rnd)

In §4.1 it was argued that preferred rankings such as these should be expressed by preference constraints. The preference hierarchy in (12) can be expressed by the following three preference constraints:

(13) [*FrRd >> *BkUnRd]

In a language with the ranking *BKUNRD >> *FRRD, assign one violation to every candidate violating *FRRD.

(14) [*BKUNRD >> IDENT(rnd)]

In a language with the ranking IDENT(rnd) >> *BKUNRD, assign one violation mark to every candidate violating *BKUNRD.

(15) [*FRRD >> IDENT(rnd)]

In a language with the ranking IDENT(rnd) >> *FRRD, assign one violation mark to every candidate violating *FRRD.

The violations that these preference constraints will assign depend on the ranking between the three ordinary non-preference constraints. The three ordinary constraints can be ranked in six different ways. In the table below the six possible rankings between the ordinary constraints are paired with (i) the outputs they would have yielded had it not been for preference constraints, and (ii) the preference constraints that are "activated"[11] in each of the six rankings.

(16) Predicted output configurations disregarding the preference constraints, and the preference constraints activated with each of the rankings

	Frt, +rnd	**Bk, -rnd**	**Ranking**	**Activated Preference Constraints**		
				[*FRRD >> *BKUNRD]	[*BKUNRD >> ID(rnd)]	[*FRRD >> ID(rnd)]
(a)	No	No	*FRRD >>*BKUNRD >>ID(rnd)			
(b)			*BKUNRD >>*FRRD >>ID(rnd)[12]			
(c)	No	Yes	*FRRD >>ID(rnd)>> *BKUNRD		✓	
(d)	Yes	Yes	ID(rnd)>> *FRRD >>*BKUNRD		✓	✓
(e)			ID(rnd)>> *BKUNRD >> *FRRD	✓	✓	✓
(f)	Yes	No	*BKUNRD >>ID(rnd)>> *FRRD	✓		✓

There are six constraints in total under consideration here (the three ordinary and the three preference constraints). This means that there is a total of 720 possible rankings to consider. Each of the six rows in the table above therefore represents 120 possible rankings. The default output pattern[13] of each of the lines in table, are represented in the

11 Remember that a preference constraint can only assign a violation if the preferred ranking it demands is not observed. In a grammar with the preferred ranking, the preference constraint is therefore not active and can assign no violation marks.

12 Even though [*FRRD >>*BKUNRD] is activated in this grammar, it will have no influence on the outcome of the inter-candidate competition. Because of the ranking *FRRD >> IDENT(rnd), the optimal candidate for a front rounded vowel input will be one in which roundness has been neutralized. The additional violation of a candidate with a front rounded vowel in terms of [*FRRD >>*BKUNRD], simply confirms the already non-optimal status of such a candidate.

13 The output pattern that would have resulted from that ranking had it not been for the preference constraints.

first and second columns. The ranking represented by each of the rows are guaranteed to have their default output patterns only if there are no preference constraints activated on that line. As soon as there are activated preference constraints, it is predicted that these preference constraints will cause output shifting effects. In fact, the more preference constraints activated in a specific row, the less likely it is that the rankings represented in that row will actually yield their default output pattern.

Comparing table (16) with table (6) above shows that there is a correlation between the frequency with which a specific output pattern is actually attested, the number of rankings that will yield that output pattern as their default, and the number of preference constraints that are activated in these rankings. The most frequently attested pattern (neither front rounded nor back unrounded vowels), would result as the default output pattern of the first two rows in (16) – i.e. from 240 possible rankings. Also, there are no preference constraints activated in these two rows that can cause output shifting. On the other hand, the least frequently attested pattern (only front rounded vowels), would result as the default output of only one row in (16) – namely the row indicated as (f). And, in this row there are two activated preference constraints. These preference constraints will, through output shifting, cause at least some of the 120 rankings represented by this row, to yield an output that corresponds to a less marked (more frequently) attested pattern.

Intuitively, it seems as if this model might make the correct predictions about the cross-linguistic distribution of front rounded and back unrounded vowels. To confirm this, it is necessary to investigate all or the different possible grammars more closely to allow for output shifting effects of the preference constraints.

To inspect each of these 720 different grammars to determine their output patterns would have been time consuming. However, there are easier ways to determine how many rankings will yield specific output patterns. Below the rankings represented by one of the rows in (16) are discussed in detail, namely those ranking represented by the fifth row IDENT(rnd)>> *BKUNRD >>*FRRD. This discussion will serve as an example of how the number of rankings that will yield each output pattern can be determined. The choice of this specific group of rankings is motivated by the fact that this is the most interesting group of rankings – all three preference constraints are activated.[14]

After the discussion of this example, the results of determining the output patterns of all 720 possible rankings are tallied. The predictions about the frequency distribution of back unrounded and front rounded vowels that follow from this, are then compared with the actually observed numbers as a measure of the success of preference constraints.

4.2.1 Disobeying all three preference constraints: IDENT(rnd)>> **BKUNRD >>*FRRD*

In general, if the ranking between *m* constraints is fixed, the number of possible ways in which *n* constraints can be interleaved between the *m* fixed constraints can be calculated by the following formula:

[14] For a discussion of the output patterns that will result from the ranking represented by each of the other six rows in (16), refer to Appendix A.

(17) Number of possible rankings between *m* constraints in a fixed ranking, and *n* constraints that can freely rerank

$$\frac{(m+n)!}{m!}$$

In each of the rows in (16) the ranking between the three ordinary constraints is fixed, i.e. *m* = 3. The three preference constraints can freely rerank, i.e. *n* = 3. Each row thus represents (3 + 3)!/3! = 6!/3! = 720/6 = 120 possible rankings. The default output of the 120 grammars implied by the fifth row in (16), IDENT(rnd)>> *BKUNRD >>*FRRD, is to faithfully preserve the underlying specification for lip rounding on both front rounded and back unrounded vowels. However, as soon as any of the preference constraints outranks faithfulness, output shifting effects will be observed.

When the rankings IDENT(rnd)>> *BKUNRD >>*FRRD is observed, all three preference constraints are activated. The violations that will be afforded in these grammars, are recorded in the tableau below.

(18) Ident(RND)>> *BKUNRD >>*FRRD

	ID(rnd)	*BKUNRD	*FRRD	[*FRRD >> *BKUNRD]	[*BKUNRD >> ID(rnd)]	[*FRRD >> ID(rnd)]
(a) /ø/ ø			*	*		*
(b) e	*					
(c) /ɯ/ ɯ		*			*	
(d) u	*					

When [*BKUNRD >>IDENT(rnd)] outranks IDENT(rnd), roundness on back vowels is obligatory. Disregarding the other two preference constraints, there is only one position above IDENT(rnd) for [*BKUNRD>>IDENT(rnd)], while there are three positions below IDENT(rnd). Of the 120 rankings under consideration here, one quarter will then have [*BKUNRD>>IDENT(rnd)] outranking faithfulness, and three quarters will have faithfulness ranked higher. In 30 of the 120 grammars all back vowels will surface as round, and in 90 unrounded back vowels will be tolerated.

If either or both of [*FRRD>>*BKUNRD] or [*FRRD>>IDENT(rnd)] outrank IDENT(rnd), all front vowels will be unrounded. First consider the cases where unrounded back vowels are not preserved, i.e. grammars with the ranking [*BKUNRD>>IDENT(rnd)] >> IDENT(rnd) >> *BKUNRD >> *FRRD. Now calculate the number of these rankings in which both of [*FRRD>>*BKUNRD] and [*FRRD>>IDENT(rnd)] are ranked below IDENT(rnd). These are languages in which front rounded vowels are tolerated. There are 3 positions for [*FRRD>>*BKUNRD] to be ranked into.

(19) Rankings with IDENT(rnd) >> [*FRRD>>*BKUNRD]

(a) [*BKUNRD>>ID(rnd)] >> ID(rnd) >> [*FRRD>>*BKUNRD] >> *BKUNRD >> *FRRD

(b) [*BKUNRD>>ID(rnd)] >> ID(rnd) >> *BKUNRD >> [*FRRD>>*BKUNRD] >> *FRRD

(c) [*BKUNRD>>ID(rnd)] >> ID(rnd) >> *BKUNRD >> *FRRD >> [*FRRD>>*BKUNRD]

In each of these three sub-hierarchies there are 4 positions below IDENT(rnd) for [*FRRD>>IDENT(rnd)] to be ranked into. This means that in 12 of the 30 rankings in which back unrounded vowels are rounded, rounded front vowels will be preserved. In the remaining 18 of these 30, roundness on front vowels will be neutralized.

Now consider the 90 rankings where unrounded back vowels are preserved, i.e. with [*BKUNRD>>IDENT(rnd)] ranked below IDENT(rnd). Disregarding the other two preference constraints, there are three rankings that comply with this requirement.

(20) Rankings with IDENT(rnd) >> [*BKUNRD>>IDENT(rnd)]

(a) IDENT(rnd) >> [*BKUNRD>>IDENT(rnd)] >> *BKUNRD >> *O

(b) IDENT(rnd) >> *BKUNRD >> [*BKUNRD>>IDENT(rnd)] >> *O

(c) IDENT(rnd) >> *BKUNRD >> *O >> [*BKUNRD>>IDENT(rnd)]

If both of [*FRRD>>*BKUNRD] and [*FRRD>>IDENT(rnd)] rank below IDENT(rnd), both front rounded and back unrounded vowels will be preserved. In each of the sub-hierarchies in (20) there are four positions for [*FRRD>>*BKUNRD] to rank into such that IDENT(rnd) dominates it. This yields 12 sub-hierarchies. One of these is given as an example below.

(21) IDENT(rnd) >> [*FRRD>>*BKUNRD] >> [*BKUNRD>>IDENT(rnd)] >> *BKUNRD >> *FRRD

In this sub-hierarchy there are 5 positions for [*FRRD>>IDENT(rnd)] to be ranked such that IDENT(rnd) outranks it. There are 5 such positions in each of these 12 sub-hierarchies, implying that there are 60 rankings in which both [*FRRD>>*BKUNRD] and [*FRRD>>IDENT(rnd)] will be outranked by IDENT(rnd). In these 60, both back unrounded and front rounded vowels will be preserved. In the remaining 30 of the 90 in which back unrounded vowels will be preserved, front rounded vowels will be neutralized.

The possible outputs of the rankings represented by IDENT(rnd)>> *BKUNRD >>*FRRD are summarized in the table below. The results in the last row (both front rounded and back unrounded vowels) represent the default output of this ranking. The first three lines are the results of output shifting caused by the preference constraints. The preference constraints assign violations only to *BKUNRD-violators and *FRRD-violators. Candidates that violate IDENT(rnd) will never receive violations from a preference constraints – since IDENT(rnd) prefers to be ranked in bottom position. Therefore output shifting will only be observed when one of the preference constraints outrank IDENT(rnd) – only then can a candidate that violate IDENT(rnd) be optimal. The last column in the table therefore indicates where the preference constraints must be ranked relative to IDENT(rnd) for that specific output pattern to result.

(22) Output of a grammar with the sub-hierarchy IDENT(rnd) >> *BKUNRD >> *FRRD

Front round	Back unround	Number of rankings	Conditions on the ranking of Preference Constraints
No	No	18	[*BKUNRD>>ID(rnd)] and at least one of [*FRRD>>*BKUNRD] and [*FRRD>>ID(rnd)] outranks ID(rnd)
No	Yes	30	[*BKUNRD>>ID(rnd)] outranks ID(rnd), but both [*FRRD>>*BKUNRD] and [*FRRD>>ID(rnd)] ranks below ID(rnd)
Yes	Yes	60	All three preference constraints rank below ID(rnd)
Yes	No	12	[*BKUNRD >>ID(rnd)] ranks below ID(rnd), but at least one of [*FRRD>>*BKUNRD] and [*FRRD>>ID(rnd)] outranks ID(rnd)

By similar reasoning the output patterns of the other rankings represented in (16) can also be determined.[15] In the next section the output patterns of all 720 possible rankings are tallied, and the predictions of the preference constraints model about the frequency distribution of front rounded and back unrounded vowels are evaluated.

4.2.2 Final count and evaluation

The results obtained from all 720 possible rankings are summarized in the table below. Recall the hypothesis in (1) (§3.1), that the percentage of the possible rankings yielding a specific output configuration predicts the percentage of languages in which that specific output configuration should be observed. The last column in this table therefore represents the predicted distribution of front rounded and back unrounded vowels under a theory that allows preference constraints.

(23) Predicted distribution patterns of front rounded and back unrounded vowels

Front rounded	Back unrounded	Number of rankings	Percentage
No	No	414	58
No	Yes	108	15
Yes	Yes	132	18
Yes	No	66	9

15 See Appendix A for a detailed discussion of these other rankings.

The largest portion of the possible rankings yield languages with neither front rounded nor back unrounded vowels (the first row). There are two reasons for this. As the table in (16) (§4.2) shows, systems with neither of these groups of sounds are the default output of 240 of the 720 possible rankings – the first two rows in table (16). Table (16) also shows that there are no preference constraints that can interfere with these 240 rankings actually yielding their default outputs patterns. These 240 rankings therefore represent more than half of the 414 rankings in the first row in table (23) above. The remaining 174 of these 414 rankings, yield this output pattern as a result of the output shifting effects of preference constraints. The preference constraints force a portion of the rankings represented by each of the other rows in table (16) to yield inventories from which both front rounded and back unrounded vowels are absent. Table (23) clearly shows how preference constraints favor the preferred rankings, and thereby introduce a statistical bias in the frequency distributions.

The shift in the predicted output pattern that is observed in table (23) is in the right direction. Languages with no front rounded or back unrounded vowels are indeed the most frequent. This shift therefore represents a gain in explanatory power. This gain was achieved through the introduction of preference constraints into the theory. However, the introduction of preference constraints does represent a complication to the theory. It therefore has to be asked whether this gain in explanatory power is significant enough to warrant this complication to the theory. To decide this, it is necessary to compare the predictions made by the preference constraints theory, to the predictions made by an OT grammar without preference constraints.

In the table below this comparison is made. The outputs predicted by the an OT grammar with free reranking of the constraints, by an OT grammar with a fixed ranking between the markedness constraints, and by the preference constraints theory are compared to each other, and to the actual distributional patterns in the UPSID database. (For the predictions of the non-preference constraints models, refer to §3.2.)

(24) Comparison between percentages predicted by different theories

Front round	**Back unround**	**Free ranking**	**Fixed ranking**	**Preference Constriants**	**UPSID**
no	no	33	33	58	72
no	yes	16.5	33	15	18
yes	yes	33	33	18	7
yes	no	16.5	0	9	3

This table shows that the model with preference constraints does indeed account better for the pattern observed in the UPSID database. In fact, it is possible to quantify the degree to which the preference constraints model improves on the two non-preference constraints models. The so-called "sums of the residual squares" analysis is a measurement of the degree to which some model fits the data it is supposed to model. The results of this analysis for each of the options represented in table (24) are given in the table below. The last column is an indication of the improvement of each model relative to the model with free re-ranking.

(25) Sums of the residual squares of each of the three models

Model	**Sum of residual squares**	**Improvement**
Free reranking	2381.5	–
Fixed ranking	2431	-2%
Preference constraints	362	85%

Table (25) shows that the preference constraint model represents an improvement of 87% on the fixed ranking model, and an 85% improvement on the free reranking model. This is a highly significant improvement. It is therefore concluded that the complication to the grammar implied by the addition of preference constraints is warranted by the significant improvement in the explanatory power of the theory.

The UPSID database is only a sample of the human languages. And in spite of the measures taken by Maddieson and his co-workers to make this a representative sample (cf. Maddieson 1984: 5-7), it still remains only a sample. The value of UPSID should therefore be taken as showing the basic patterns of distribution, the basic trends, rather than the absolute numbers in which different patterns are observed. Since the preference constraints option does basically predict the trends observed in UPSID, it is concluded that it does indeed give an accurate view of the actual distributional patterns of roundness on front and back vowels.[16]

5. Preference constraints and fixed hierarchies

There is an implicational relationship between front rounded and back unrounded vowels. The presence of front rounded vowels in the phonemic inventory of some language implies that back unrounded vowels will also be present in that language.[17] Implicational universals like these are typically handled in OT by positing fixed hierarchies between markedness constraints. There are many examples of such fixed hierarchies in the literature: (i) the place markedness hierarchy (Prince 1998, Ito & Mester 1998, Lombardi 1995, Gnanadesikan 1995), the peak and margin hierarchies (Prince & Smolensky 1993), the nasalization hierarchy (Walker 1998).

---

[16] For some discussion on whether it can be expected that precise probabilities should be derivable from grammar, see Guy (1991a & 1991b) who is rather optimistic, and Anttila (1997: 48-9) and Reynolds (1994: 136-7) for a more carefully qualified stance. This paper follows Anttila's and Reynold's viewpoint rather than that of Guy.

[17] Or at least that there is a very strong probability that back unrounded vowels will also be present.

Since the implicational relationship between front rounded and back unrounded vowels are expressed by preference constraints, it is proposed here that all fixed hierarchies should be replaced by preference constraints.

Much is gained in terms of theoretical simplicity if this is done – there is then only one mechanism in the theory to account for implicational relationships instead of two. However, this simplicity also comes at a cost. Implicational relationships expressed by fixed hierarchies, are absolute universals. This can be seen in the predictions that the fixed ranking option made about the vowel roundness example in §3.2. Under this option no language is possible with front rounded vowels but without back unrounded vowels. However, when these implicational relationships are expressed by preference constraints, they are not absolute universals but simply strong statistical tendencies. Since there are some of these implicational relationships that do seem to be absolute universals, replacing all fixed hierarchies by preference constraints might cause problems.

The proposal here is that all fixed hierarchies should indeed be replaced by preference constraints. The vowel roundness example discussed in the previous section (§3), showed that this is sometimes required. If the implicational relationship between front rounded and back unrounded vowels is expressed by a fixed ranking, then the possible (but scarce) group of languages with front rounded but no back unrounded vowels is not possible.

In section §5.1 the potential problems caused by this move are discussed. It is argued that the fact that replacing all fixed hierarchies by preference constraints does lead to the prediction of languages that (possibly) do not exist, is no real problem. Since there are many preference constraints militating against these non-existent output patterns, the likelihood that they will ever be chosen as optimal in a language are statistically very small.

Once it has been determined that all fixed hierarchies should be replaced by preference constraints, one last fact needs to be explained, namely that preference constraints also have wider applicability than fixed hierarchies. Preference constraints can establish a preferred ranking between two constraints that can never go into a fixed ranking. This is discussed in §5.2.

5.1 Preference constraints and absolute implicational universals

In general, if a language allows a segment with certain sonority level to be parsed into syllabic peak positions, all segments of higher sonority will also be tolerated in peak position. This is a classical example of an implicational universal that has traditionally been captured in OT by a fixed ranking between markedness constraints. In fact, Prince and Smolensky (1993: Chapter 8) used this implicational universal to introduce the notion of universally fixed rankings. To explain their idea, assume that all segments can be classified into four sonority levels:

(26) Simplified Sonority Hierarchy

Low sonority ⟶ High sonorioty

Obstruents — Nasals — Liquids — Vowel

To capture the one way implicational universal, Prince and Smolensky propose markedness constraints that militate against the parsing of segments of a certain sonority level into peak position, and then universally fix the ranking between these constraints:

(27) Simplified peak affinity hierarchy

*P/Obstruent >> *P/Nasal >> *P/Liquid

This hierarchy captures the fact that a liquid peak is more harmonic than a nasal peak, which is again more harmonic than an obstruent peak.[18] To simplify the discussion assume that the only way in which violation of these constraints can be avoided is through epenthesis, i.e. at the cost of a DEP violation. If the fixed ranking between these constraints are accepted, then only four possible output patterns are predicted to exist.

(28)

Ranking	Resulting systems of allowed peaks
DEP >> *P/O >> *P/N >> *P/L	{O, L, N, V}
*P/O >> DEP >> *P/N >> *P/L	{L, N, V}
*P/O >> *P/N >> DEP >> *P/L	{L, V}
*P/O >> *P/N >> *P/L >> DEP	{V}

The four possible peak inventories all comply with the implication universal. No peak inventory is possible in which, for instance, nasals are tolerated in peak positions but liquids not. This phenomenon has been dubbed “harmonic completeness” by Prince (1997). Harmonic completeness requires that a system that contains a structure with a certain level of markedness along some dimension, also contains all structures of lesser markedness along the same markedness dimension. Prince defines it as follows:

(29) **Harmonic completeness** (based on Prince 1997)

Let L be the set of all licit output forms of some language. Let HS be some universal harmony scale, and $\alpha, \beta \in$ HS such that $\alpha \succ \beta$ with respect to HS.
L is harmonically complete with regard to HS iff whenever $\beta \in$ L, then $\alpha \in$ L.

---

[18] Unlike the original Prince and Smolensky hierarchy, there is no constraint against parsing a vowel into peak position. This is because no language disallows vowels in syllabic peak position. Clements (1997) raises this same point with regard to Prince and Smolensky’s explanation of Berber syllabification. However, this deviation from Prince and Smolensky is not crucial for the point that is made here. The same point could have been made with inclusion of *P/Vowel.

All four possible syllabic peak inventories under the fixed hierarchy theory are harmonically complete. Whenever an implicational relationship between two or more structures is expressed by a fixed ranking between markedness constraints, it is guaranteed that all possible systems will be harmonically complete.

One example of a harmonically incomplete system that is predicted as impossible under the fixed hierarchy theory, is a language that allows nasals into peak position but not liquids, i.e. a language with {O, N, V} as syllabic peak inventory. Since liquids make more harmonic peaks than nasals, the presence of nasal peaks should imply the presence of liquid peaks. This prediction is most probably correct – most probably such a language does not exist. The ranking that would have resulted in such a system is:

(30) Impossible system – disallowing only liquids from peak position

*P/L >> DEP >> *P/O, *P/N {O, N, V}

If the fixed hierarchy in (27) is replaced by preference constraints, then a language with {O, N, V} as peak inventory is possible. However, there will also be preference constraints that will in many cases force a hierarchy such as that in (30) to allow liquids. The result of this is that the portion of languages that are predicted to have an {O, N, V} syllabic peak inventory is very small. So small, that the chance of ever stumbling across such a language is statistically speaking negligible. The argument here is thus that such a language is indeed possible, but that the odds against ever finding such a language are very high.

Every output pattern possible under some ranking between the constraints is a possible language, even if some output patterns will be extremely rare. Output patterns that are harmonically incomplete are then predicted as possible, but because preference constraints will militate against the rankings that will result in these harmonically incomplete systems, they are predicted to be very scarce.

As an example to illustrate this general point the case of {O, N, V} peak inventories is discussed in detail here. It is accepted that the hierarchy in (27) does state the preferred ranking between the three peak affinity constraints. However, most languages allow only vowels in peak position. This implies that the preferred position for DEP in the hierarchy in (27) is right at the bottom. The full preferred ranking is therefore:

(31) Preferred ranking for syllabic peak affinity

*P/O >> *P/N >> *P/L >> DEP

There are six preference constraints required to express the preferred rankings in (31). These six constraints are stated below. Refer to (7) in §4.1 for an explanation of how to interpret these constraints:

(32) Preference constraints required to express preferred ranking in (31)

[*P/O>>*P/N], [*P/N>>*P/L], [*P/L>>DEP],

[*P/O>>P/L], [*P/O>>DEP], [*P/N>>DEP].

To determine what percentage of languages are predicted to have {O, N, V} syllabic peak inventories it is necessary to determine what portion of all the possible rankings between the six preference constraints in (32) and the four ordinary constraints in (31) will actually result in a language with {O, N, V} as syllabic peak inventory. This is done in the rest of this section.

With the four constraints that are in the preferred ranking relation and the six preference constraints required to express this relation, there is a total of 10 constraints. These constraints can be ranked in a 10! = 3,628,800 different ways. What percentage of this will result in a grammar allowing only {O, N, V} into syllabic peak position?

There are two ways in which such a system can be achieved in an OT grammar with preference constraints. (i) The first way is through grammars in which the ordinary constraints are ranked in such a way that they would have yielded this output pattern. As soon as the preference constraints are added, they will cause some of these rankings to shift into other output patterns. (ii) Some grammars with rankings between the ordinary constraints that would not have yielded {O, N, V} as possible peak inventory, can be shifted into this group of grammars by preference constraints. Each of these two sources for {O, N, V} peaks is discussed separately.

How many grammars with the ranking in (30) will yield systems with {O, N, V} as possible peaks? To answer this the output shifting of the preference constraints must be considered. Two constraints in (30) are not ranked relative to each other (*P/O and *P/N). To determine the violations of the preference constraint [*P/O>>*P/N] these constraints must be ranked. Both rankings implied by (30) should therefore be considered.

(33) Rankings that can result in a syllabic peak system {O, N, V}

(a) *P/L >> DEP >> *P/N >> *P/O

(b) *P/L >> DEP >> *P/O >> *P/N

The grammars represented by (33a) are discussed here as an example. The tableau below shows how these grammars will evaluate three different inputs: an obstruent, a nasal, and a liquid. The violations of each candidate in terms of the preference constraints are also listed, although the preference constraints are not ranked relative to the ordinary constraints. The preference constraint [*P/L>>DEP] is not included since this constraint is not activated. All candidates will therefore vacuously satisfy this constraint and it will have no influence on the choice of optimal candidate.

(34) *P/L >> DEP >> *P/N >> *P/O

			*P/L	DEP	*P/N	*P/O	[*P/O>> *P/N]	[*P/N>> *P/L]	[*P/O>> P/L]	[*P/O>> DEP]	[*P/N>> DEP]
(a)	/O/	.O.				*	*		*	*	
(b)		.OV.		*							
(c)	/L/	.L.	*!								
(d)	☞	.LV.		*							
(e)	/N/	.N.			*			*			*
(f)		.NV.		*							

From this tableau it is clear that, disregarding the preference constraints, these rankings will indeed result in systems allowing only {O, N, V} into peak position. However, taking into account the preference constraints will change this. A system with as possible nuclei {O, N, V} will only result as long as the following conditions hold:

(35) Conditions for {O, N, V} peaks with ranking *P/L >> DEP >> *P/N >> *P/O

(a) DEP >> [*P/O>>*P/N], [*P/O>>*P/L], [*P/O>>DEP]

AND

(b) DEP >> [*P/N>>*P/L], [*P/N>>DEP]

In systems not complying with condition (a), the set of possible peaks will exclude obstruents, i.e. {V, N}. In systems not complying with condition (b), nasals will be excluded from peaks, i.e. {V, O}. Systems complying with neither of these conditions will allow only vowels into peak position, i.e. {V}. The purpose here is then calculating the number of rankings that comply with both conditions (a) and (b). This will be only systems in which all five of the preference constraints in (34) are dominated by DEP. When all five of the preference constraints are ranked below DEP, there are 7 constraints in total below DEP, 2 of which are in a fixed ranking (*P/O >> *P/N), while the other 5 (the preference constraints) can freely rerank. Using the formula in (17) it can be calculated that this yields (2 + 5)!/2! = 2,520 rankings. The sixth preference constraint should be taken into account. Recall that the preference constraint [*P/L>>DEP] is vacuously satisfied by all candidates. It can therefore be ranked in any position without having an influence on the outcome. In a hierarchy of 9 constraints there are 10 possible positions into which a tenth constraint can be ranked. In each of 2,520 rankings there is thus 10 positions for [*P/L>>DEP], meaning that there are 25,200 total rankings complying with both conditions (36a) and (36b). The ranking *P/L >> DEP >> *P/N >> *P/O then contributes 25,200 grammars that will result in a syllabic peak inventory {O, N, V}.

By similar reasoning it can be determined that the ranking represented by (33b) contributes 32,400 grammars that will result in a syllabic peak inventory {O, N, V}.[19] In total this source of {O, N, V} syllabic peaks contributes 32,400 + 25,200 = 57,600 rankings. To determine the total number of rankings that will have this output pattern, it is necessary to consider grammars that are shifted into this group by preference constraints. What are the circumstances under which this will occur?

In the output system that is under discussion here it should be worse to violate DEP than to violate *P/N or *P/O. Preference constraints cannot cause this, since there are no preference constraints that will militate against DEP-violators. It follows that it will be necessary for DEP to outrank both *P/N and *P/O. When this happens, the preference constraints [*P/N>>DEP] and [*P/O>>DEP] are activated, and they will assign additional violations to candidates violating *P/N and *P/O. For a DEP-violator to be worse than a *P/N-violator and *P/O-violator, it is therefore also necessary for DEP to outrank both of these preference constraints. The first necessary condition is therefore:

(36) First necessary condition for an output to be shifted into {O, N, V}

(a) DEP >> *P/O, *P/N

AND

(b) DEP >> [*P/O>>DEP], [*P/N>>DEP]

Comparison with (35) shows that this condition also held for grammars discussed there. In the discussion there, grammars where *P/L outranked DEP were considered. What remains to consider here are the cases with the ranking DEP >> *P/L. When this ranking is observed, the preference constraint [*P/L>>DEP] is activated, and it will assign an additional violation mark to all candidates violating *P/L. The output system that is the aim here, is one in which parsing a liquid into peak position should be worse than epenthesis. The ranking DEP >> *P/L states the opposite. The only way in which to achieve the desired result, is to rank the preference constraint [*P/L>>DEP] above DEP.

(37) Second necessary condition for an output to be shifted into {O, N, V}

(a) DEP >> *P/L[20]

AND

(b) [*P/L>>DEP] >> DEP

These two sets of requirements can be collapsed into a single necessary condition:

---

19 For a discussion of the other ranking represented by (33b), see Appendix B.

20 This is not really a condition to get a peak inventory of {O, N, V}. However, systems with the opposite ranking between these two constraints have already been accounted for. This condition is therefore included here to prevent the situation where certain rankings are counted twice.

(38) Combining first and second conditions for an output to be shifted into {O, N, V}

[*P/L>>DEP] >> DEP >> *P/L, *P/O, *P/N, [*P/O>>DEP], [*P/N>>DEP]

In the ranking in (38) there are three ordinary constraints below DEP that are unranked. These three constraints are in a preferred ranking relation, and therefore their ranking relative to each other can also cause activation of preference constraints. The preference constraints that state the preferred ranking between these constraints are [*P/O>>*P/N], [*P/O>>*P/L], [*P/N>>*P/L]. These preference constraints will therefore, when activated, assign additional violation marks to candidates violating either *P/O or *P/N. The system that is the aim here has as syllabic peaks {O, N, V}. Peaks with nasals or obstruents should thus be tolerated. When these preference constraints are activated, it is therefore necessary that they be ranked below DEP. The three ordinary constraints can be ranked in six different ways. In each of these six ways a different combination of the preference constraints will be activated. Only one of these six possible rankings is discussed here in detail.[21]

(39) Ranking: *P/L >> *P/N >>*P/O

Additional preference constraints activated: [*P/N>>*P/L], [*P/O>>*P/L], [*P/O>>*P/N]

Necessary condition for an output to be shifted into {O, N, V}:

[*P/L>>DEP] >> DEP >> *P/O >> *P/N >>*P/L

>> [*P/O>>DEP], [*P/N>>DEP], [*P/N>>*P/L], [*P/O>>*P/L], [*P/O>>*P/N]

There are 8 constraints below DEP. The ranking between 3 of them is fixed (the ordinary constraints), while the other 5 (the preference constraints) can freely rerank as long as they are ranked below DEP. Using formula (17) it can be computed that this represents (5 + 3)!/3! = 6,720 rankings. This group of rankings therefore contributes a total of 6,720 rankings that will result in a {O, N, V} syllabic peak inventory.

For each of the other 5 groups of rankings that comply with condition (38) a similar calculation can be done. It can thus be determined that the second source of a systems with {O, N, V} syllabic peaks contributes a total 59,520.

The total contribution of both sources is then 57,600 + 59,520 = 117,120. This is out of the total of 10! = 3,628,800 possible rankings. Roughly 3.2% of the possible rankings

21 For a discussion of the five rankings, refer to Appendix C.

will therefore result in a {O, N, V} peak inventory.[22] Even though the preference constraints theory allows for the possibility of this system, it also predicts that the likelihood of such a system actually existing is very small. In fact, even if such a system did exist, the likelihood of stumbling across it is also very small. The fact that no such language is (currently) known, does not mean that it actually does not or cannot exist.

In a theory where fixed hierarchies are replaced by preference constraints harmonic completeness is sacrificed. In some instances this is a preferred result – some systems are in actual fact not harmonically complete even though they lean towards harmonic completeness. The preference constraints theory accounts for this straightforwardly. In some instances the abandoning of harmonic completeness seems to be problematic. Some non-existent patterns are then predicted as possible. However, these non-existent systems will have several preference constraints militating against them. The effect of this is that the actual percentage of possible rankings that will result in these systems are usually negligibly small. In a preference constraint theory harmonic completeness is not an absolute truth about all systems in all languages anymore, but a strong cross-linguistic statistical tendency.

The final conclusion then: It is proposed that all fixed hierarchies should be replaced by preference constraints.

5.2 Can preference constraints do more than fixed hierarchies?

Can a preference constraint only establish a preferred ranking relation between two constraints that were in a fixed ranking in classic OT? Or can it also relate two constraints that would not have been allowed to go into a fixed ranking? This question has been answered implicitly in the discussion in §3 and in §4. In these sections preference constraints were allowed wider scope than the fixed hierarchies of classic OT. In §3 the faithfulness constraint IDENT(rnd) was given a place in the vowel roundness hierarchy. In classic OT the fixed hierarchy would have consisted of only the markedness constraints *FRRD >> *BKUNRD. However, with the introduction of preference constraints, it was possible to extend this hierarchy by placing IDENT(rnd) at the bottom, i.e. *FRRD >> *BKUNRD >> IDENT(rnd). Similarly, in §4 the faithfulness constraint DEP was given a

[22] This number should not be taken as an absolute indication of the percentage of the possible rankings that will result in this peak inventory. Recall that the effects of the MAX and IDENT constraints were not taken into account. These constraints are most probably, like DEP, preferably dominated by the peak affinity constraints. There will then also be preference constraints expressing these relations and many of these preference constraints will also militate against an {O, N, V} outcome. Recall also that only a three level peak affinity hierarchy was considered here. The Berber example from Prince and Smolensky (1993) has shown that it is necessary to make much finer distinctions on the sonority scale, and that this hierarchy should include many more constraints. These additional peak affinity constraints will each have a specific position in the preferred ranking hierarchy under consideration here. This adds more preference constraints and more constraints militating against a {O, N, V} outcome. All of this means that the 3.2% result attained here, is probably too high.

position in the peak affinity hierarchy. In the classic OT version of this hierarchy faithfulness had no place.

Why is it possible for preference constraints to have a wider scope than fixed rankings? A fixed ranking creates an all-or-nothing situation. If two constraints are in a fixed ranking A >> B, then no system with the ranking B >> A is possible at all. If IDENT(rnd) were fixed in bottom position on the vowel roundness hierarchy, no language with front rounded or back unrounded vowels would have been possible. Markedness would always outrank faithfulness, and all front rounded vowels will loose their rounding, while all back unrounded vowels will be rounded. A preferred ranking expressed by a preference constraint is fundamentally different. With preference constraints all rankings are allowed. However, preference constraints militate against the effect of some rankings. Giving IDENT(rnd) a place in the preferred ranking therefore does not exclude the possibility of a ranking where IDENT(rnd) will dominate one or both of the markedness constraints, and therefore it does not exclude the possibility of languages that will allow front rounded or back unrounded vowels. What it does, is to limit the number of these rankings that will actually result in preservation of these marked segments unto the surface. The preference constraint makes a demand on the choice of optimal candidate, but like all other constraints in CON, violation of the preference constraint can be forced by some higher ranked constraint. This seems to be more in line with the general spirit of OT than stipulating an indefeasible fixed ranking between constraints.

6. Alternatives considered

There are other conceivable ways within OT to derive the type of distributional tendencies that motivated the introduction of preference constraints. Two of the most obvious are discussed below, namely constraints in a stringency relation, and ranking of constraints along a continuous ranking scale (Boersma & Hayes 1999) (§4.2). Free reranking of constraints and constraints in a fixed ranking are not discussed here. In §3 these two options were compared to the preference constraints theory, and it was concluded that preference constraints theory is to be preferred above them.

6.1 Stringency relations between constraints

To explain this alternative, a different example is necessary. The distribution of aspiration is similar to that of roundness on vowels. Aspirated stops are marked in general, voiced aspirated stops are more marked than voiceless aspirated stops, and the presence of voiced aspirated stops depends on the presence of voiceless aspirated stops.[23] The preference constraints theory can explain these tendencies in the same way that it explained the tendencies about roundness on vowels (cf. §3). However, these tendencies may also be captured by formulating a general markedness constraint against aspirated stops *[+aspiration] and a special markedness constraint against voiced aspirated stops *[+aspiration, +voice]. These constraints are in a stringency relationship – all forms violating *[+aspiration, +voice] also violate *[+aspiration]. The faithfulness constraint IDENT(asp) will interact with these two constraints, yielding six possible rankings. What are the output patterns predicted by this grammar? To determine this, first consider all the violations afforded by these constraints to the candidates without considering the ranking between the constraints.

(40) [+aspiration] regulated by a stringency relationship between markedness constraints

			*[+asp]	*[+asp, +voi]	IDENT(asp)
(a)	/dh/	dh	*	*	
(b)		d			*
(c)	/th/	th	*		
(d)		t			*

For a language to tolerate voiceless aspirated stops, the ranking IDENT(asp) >> *[+asp] is required. For a language to tolerate voiced aspirated stops the ranking IDENT(asp) >>

[23] Patterns of co-occurrence of voiced and voiceless aspirated stops in the 451 languages in the UPSID database

Voiced aspirated	**Voiceless aspirated**	**Percentage**
No	No	73
No	Yes	23
Yes	Yes	3
Yes	No	<1

Comparison of these numbers with that in Table (2) shows that they are very similar. There is possibly one language (Javanese) that allows voiced aspirated stops to the exclusion of voiceless aspirated stops - i.e. one language in the last row. There is disagreement in the literature about the precise nature of the relevant segments. Kiliaan (1919), Van der Valk (1928) and Poedjosoedarmo (1974) consider these segments to be breathy voiced stops. Horne (1974) can be interpreted in both ways. Ladefoged (1971) and Fagan (1988), on the other hand, consider these stops to be voiceless.

*[+asp], *[+asp, +voi] is required. Now consider the six logically possible rankings between these constraints and the output patterns that follow from each of them.

(41) Output patterns of a grammar with IDENT(asp), *[+asp], *[+asp, +voice]

Ranking	**Voiced aspirates?**	**Voiceless aspirates?**
IDENT(asp) >> *[+asp] >> *[+asp, +voi]	Yes	Yes
IDENT(asp) >> *[+asp, +voi] >> *[+asp]	Yes	Yes
*[+asp] >> *[+asp, +voi] >> IDENT(asp)	No	No
*[+asp] >> IDENT(asp) >> *[+asp, +voi]	No	No
*[+asp, +voi] >> IDENT(asp) >> *[+asp]	No	Yes
*[+asp, +voi] >> *[+asp] >> IDENT(asp)	No	No
Total:	2	3

In order to judge the success of this option, it is necessary to compare it both to the actual UPSID numbers and to the results of a grammar with preference constraints. This is done in the table below.[24]

(42) Comparison with observed frequencies and a preference constraints grammar

	Voiced aspirates	**Voiceless aspirates**	**Preference constraints**	**Stringency related constraints**	**UPSID**
(a)	no	No	58	50	73
(b)	no	Yes	15	17	23
(c)	yes	Yes	18	33	3
(d)	yes	No	9	0	<1

The stringency related constraints do not derive the markedness of aspirated stops – half of the languages are predicted to have these segments, while only about a quarter actually do. In the preference constraints grammar the markedness of aspirated stops is captured.[25] The stringency related constraints do capture the relative markedness of

[24] The UPSID numbers are from previous footnote and the preference constraint numbers are the same as for the vowel rounding grammar discussed in §2. See table (22) for the specific results.

[25] Not quite to the observed degree, but see §3.2 for discussion of whether it should be expected of preference constraints theory to make exact predictions.

voiced aspirated stops – only 33% of languages are predicted to have voiced aspirated stops while 50% are predicted to have voiceless aspirated stops. However, this is still far off the mark. Voiced aspirated stops are highly marked and this is not captured. The preference constraints grammar predicts 27% of languages with voiced aspirated stops, which is at least not worse. Stringency related constraints also capture the dependency of voiced aspirated stops on voiceless aspirated stops – all languages with voiced aspirated stops are predicted to also have voiceless aspirated stops. This account therefore allows only harmonically complete systems, while the preference constraints grammar captures this dependency without excluding a language with only voiced aspirated stops.

The fact that harmonically incomplete systems are excluded by the stringency constraints model is not necessarily problematic in the aspiration example. It is at least questionable whether a language that allows only voiced aspirated stops does actually exist (see footnote 19). However, it has been shown that there are examples of strong universal tendencies that are not always expressed in harmonically complete systems (cf. the vowel roundness case in §3 and the other examples mentioned in §2). A theory that excludes the possibility of harmonically incomplete systems is too restrictive.

This alone is enough reason to abandon the stringency related constraint. However there are other difficulties that will have to be addressed if this option were indeed taken. How will an example like the vowel roundness case of §3 be explained? Since the two marked structures involved in this system (front rounded and back unrounded vowels) have opposite values for the feature [back], there is no easy way in which to formulate a general markedness constraint that will be violated by both.

An explanation based on constraints in a stringency relation is both too strong (in excluding harmonically incomplete systems), and too weak (in that it cannot account for all observed patterns of non-random frequency distribution).

6.2 Replacing categorial rankings with ranking along a continuous scale

In classic OT constraint ranking is categorical. If constraint C_1 outranks constraint C_2 it is not relevant how far C_1 is ranked above C_2. However, there have been proposals to replace categorical ranking with ranking along a continuous scale (cf. Boersma & Hayes 1999; Boersma & Levelt, 1999; Zubritskaya 1995, 1997). A constraint that has a high value relative to other constraints on the ranking scale then corresponds to a higher-ranked constraint. Because ranking is along some scale, the distance between two constraints has meaning and can influence the predictions of the theory. The Gradual Learning Algorithm (GLA) of Boersma and Hayes is the version of this kind of OT grammar that is worked out in the most detail. The discussion below will therefore focus on their model. However, the discussion is in general terms and is applicable to any grammar that views constraint ranking as an ordering along a continuous scale.

In the GLA every constraint has some basic ranking value along a continuous scale. The actual point where a constraint is ranked along the continuous ranking scale is not equivalent to its basic ranking value. The GLA includes a noise component – every time the grammar has to evaluate some set of input-output pairs a (positive or negative) random value is added to the basic ranking value of every constraint. The result of this process determines the precise place where that constraint will be ranked along the continuous scale in the evaluation of that specific set of input-output pairs. The random value added to the ranking value has a normal distribution, with the basic ranking value of the constraint as its mean, and some arbitrarily chosen standard deviation that is set at the same value for all constraints. What is different between constraints then, is their basic ranking values. Their final ranking values will have the same distribution around their basic ranking values.

Because the ranking of a constraint is not rigidly fixed, but rather normally distributed within some range around its basic ranking value, GLA is able to account for intra-language variation. If the basic ranking values of two constraints are sufficiently close to each other that there is significant overlap between their ranking ranges, then it can be expected that their ranking will vary. If the basic ranking values of constraints C_1 and C_2 are close to each other, the random perturbing of ranking might result $C_1 >> C_2$ in some instances and in $C_2 >> C_1$ in others. On the other end of the spectrum, if two constraints are ranked sufficiently far apart that there is no appreciable overlap between their ranges, the ranking between them is practically (even if not in principle) fixed.

An important point is that the degree of overlap between the ranges of two constraints can be manipulated by changing the distance between the basic ranking values of the two constraints. The extent of variation in ranking between two constraints depends on the degree of overlap between their ranges. The degree of overlap in ranking ranges depends on the distance between the basic ranking values. These two facts together imply that GLA has very fine control over the degree of variation.

The GLA is a model of intra-language variation, but it is possible to construct a model of inter-language variation along similar lines that will have the same degree of control over the inter-language variation. One conceivable way is shortly sketched here without working out the details thereof.

Suppose that the process of determining a specific ranking value for a constraint is not something that occurs every time a different set of input-output pairs must be evaluated, but rather something that occurs only once when a language “chooses” its specific grammar from the array of possible grammars. If this were accepted, then cross-linguistic ranking tendencies can be derived. If constraint C_1 cross-linguistically prefers to outrank constraint C_2, but the opposite ranking is possible, then the basic ranking values of these two constraints will be chosen such that C_1’s basic ranking value is higher than C_2’s basic ranking value, but such that there is considerable overlap between the ranges of C_1 and C_2 around their basic ranking values. Because C_1’s basic ranking value is higher than that of C_2, the ranking $C_1 >> C_2$ will be encountered most frequently. As in the GLA the frequency with which each ranking will be chosen can be controlled in the finest detail by controlling the overlap between the ranges.

It is therefore possible to construct a model of grammar along these lines that will be able to model very precisely the types of frequency tendencies that are the focus of this paper. In fact, by fine manipulation of the basic ranking values of different constraints it will be possible to model frequency distributions significantly better than a preference constraints grammar can. If the only consideration was how closely a theory can model observed frequencies, then a model along the lines sketched above would have been a clear choice. Why will such a model be able to make predictions that mimic the real frequency distributions so closely? The reason for this is that it is not a very constrained model. Because the distance between constraints on the ranking scale can be manipulated freely, it is possible to get the grammar to predict any output pattern.

A preference constraints grammar, on the other hand, is a highly constrained theory. To state a preferred ranking between two constraints $C_1 >> C_2$, one preference constraint [$C_1 >> C_2$] will be added to the grammar. With these three constraints there are six possible rankings, four of which will yield an output pattern that will in effect be similar to a $C_1 >> C_2$ ranking, and two that will yield an output pattern similar to $C_2 >> C_1$. A preference constraints grammar predicts that output patterns in accordance with the preferred ranking will be seen in 66% of languages, and output patterns in accordance with the dispreferred raking in 33% of languages. No other prediction is possible in CPR with regard to a two level preferred ranking relation.

The decision between the ranking scale type model and a preference constraints grammar will therefore have to be made not on the grounds of which one can model actually observed patterns the best – a preference constraints grammar is less successful. The general predicting powers of the theories must also be considered. The ranking scale model can predict any conceivable output frequency. This is problematic. There may be certain patterns that just never occur. If such a theory is able to model such non-occurring patterns, it is too powerful a theory. As far as a preference constraints grammar goes, it is quite constrained in the predictions that it can make, and preliminary investigation suggests that it at least predicts all observed tendencies as tendencies – even if it cannot predict the actually observed frequencies exactly.

7. Summary and conclusion

In OT focus has been on accounting for all and only the possible human languages, and not on explaining frequency effects within the set of possible languages. It often happens that something is possible, but encountered very infrequently. There are different ways to think about the purpose of a formal theory of language. It is possible to assume that the theory should only distinguish between the possible and the impossible, and that frequency effects lie outside of the domain of the formal theory (Hale & Reiss 2000). In this paper the opposite view is taken. It is assumed that a formal theory should account for all significant generalizations about language, and therefore that significant frequency effects also fall within the scope of the formal theory of language.

As a first step towards equipping OT to explain frequency effects Anttila's (1995, 1997) model of intra-language variation is extended to also account for inter-language variation. If some output pattern results from a certain portion of possible rankings within CON, it is predicted that this pattern will be attested in that portion of the world's languages. If the theory must account for frequency effects, traditional OT must be enabled to distinguish between the frequently attested and rarely attested. Preference constraints are introduced to do this. If two output patterns are the result of different rankings between two constraints, and one of the two patterns is attested significantly more frequently, one of the two rankings between the constraints is preferred. This preference is expressed by a preference constraint that penalizes candidates that benefit from the dispreferred ranking. Although the ranking between the two constraints is free, the effect of the one ranking is seen more frequently. This paper showed how preference constraints enable OT to account better for many frequency effects.

Preference constraints are quite different from the typical OT constraint. In classic OT markedness constraints see only the candidate they evaluate, and faithfulness constraints the input and the candidate they evaluate. Constraint evaluation is not contingent upon the rest of the constraint set, or the ranking between other constraints. Preference constraints are fundamentally different. The violation marks that they assign crucially depend on the ranking between constraints. Allowing constraints with more global vision may have repercussions elsewhere in the theory. This paper has not explored all the possible repercussions, and this is an area that still needs further consideration. What are the implications of allowing constraints whose evaluation is ranking dependent into CON? Are there undesirable consequences? If so, can these be controlled for? Can preference constraints state a preferred ranking between any two constraints? Can the general idea of allowing ranking contingent constraint evaluation be applied towards other problems in phonology?

References

Alber, Birgit & Ingo Plag. (1999) *Epenthesis, Deletion and the Emergence of the Optimal Syllable in Creole*. Ms., University of Marburg and University of Hannover. [Available in ROA.]

Anttila, Arto. (1995) *Deriving Variation from Grammar: A Study of Finish Genitives*. Ms., Stanford University.

Anttila, Arto. (1997) Deriving variation from grammar. In Frans Hinskens, Roeland van Hout & W. Leo Wetzles (eds.) *Variation, Change and Phonological Theory*. Amsterdam: John Benjamins. p. 35-68.

Anttila, Arto and Young-mee Yu Cho. (1998) Variation and change in Optimality Theory. *Lingua*, 104:31-56.

Bell, Alan. (1978) Syllabic consonants. In Joseph H. Greenberg (ed.) *Universals of Human Language. Volume 2: Phonology*. Palo Alto: Stanford University Press. p. 153-201.

Boersma, Paul & Bruce Hayes. (1999) *Empirical tests for the Gradual Learning Algorithm*. Ms., UCLA and University of Amsterdam. [Available on ROA.]

Boersma, Paul & Clara Levelt. (1999) *Gradual Constraint-Ranking Learning Algorithm Predicts Acquisition Order*. [Available on ROA.]

Cho, Young-mee Yu. (1998) Language change as reranking of constraints. In Richard M. Hogg and Linda van Bergen (eds.) *Historical Linguistics 1995. Volume 2*. Amsterdam: John Benjamins. p. 45-62.

Clements, G.N. (1997) Berber syllabification: derivations or constraints? In Roca, Iggy (ed.) *Derivations and Constraints in Phonology*. Oxford: Clarendon. p. 289-330.

Comrie, Bernard. (1989) *Language Universals and Linguistic Typology: Syntax and Morphology. Second Edition*. Chicago: University of Chicago Press.

Croft, William. (1990) *Typology and Universals*. Cambridge, MA: Cambridge University Press.

Crothers, John. (1978) Typology and universals of vowel systems. In Joseph H Greenberg (ed.) *Universals of Human Language. Volume 2: Phonology*. Stanford: Stanford University Press. p. 93-152.

Fagan, Joel L. (1988) Javanese intervocalic stop phonemes: the light/heavy distinction. In Richard McGinn (ed.) *Studies in Austronesian Linguistics*. Athens, OH: Center for International Studies, Ohio University. p. 173-200.

Ferguson, Charles A. (1966) Assumptions about nasals: a sample study in phonological universals. In Joseph H. Greenberg (ed.) *Universals of Language. 2nd edition*. Cambridge: The MIT Press. p. 53-60.

Flemming, Edward. (1995) *Auditory Representations in Phonology*. Ph.D. dissertation, UCLA, Los Angeles.

Gnanadesikan, Amalia E. (1995) *Markedness and Faithfulness Constraints in Child Phonology*. MS., University of Massachusetts, Amherst.

Gnanadesikan, Amalia E. (1998) *Phonology with Ternary Scales*. Ph.D. dissertation, University of Massachusetts, Amherst.

Greenberg, Joseph H. (1966) *Language Universals with Special Reference to Feature Hierarchies*. The Hague: Mouton.

Greenberg, Joseph H. (ed.) (1978) *Universals of Human Language. Volume 2: Phonology*. Stanford: Stanford University Press.

Guy, G.R. (1991a) Contextual conditioning in variable phonology. *Language Variation and Change*, 3:223-239.

Guy. G.R. (1991b) Explanation in variable phonology. *Language Variation and Change*, 3:1-22.

Hale, Mark & Charles Reiss. (2000) Phonology as Cognition. Ms., Concordia University. [Available on ROA.]

Hock, H.H. (1985) Regular metathesis. *Linguistics*, 23(4):529-546.

Horne, Elinor C. (1974) *Beginning Javanese*. New Haven: Yale University Press.

Hume, Elizabeth. (to appear) The role of perceptibility in consonant/consonant metathesis. In Susan Blake, Eun-Sook Kim & Kimary Shahin (eds.) *WCCFL XVII Proceedings*. Stanford: CSLI.

Hume, Elizabeth. (1998) Metathesis in phonological theory: the case of Leti. *Lingua*, 104:147-86.

Kiliaan, H.N. (1919) *Javaansche Spraakkunst*. ‘s-Gravenhage: Martinus Nijhoff.

Kirchner, Robert. 1998. *An Effort-Based Approach to Consonant Lenition*. Ph.D. dissertation, UCLA, Los Angeles. [Available on ROA.]

Ladefoged, Peter. (1971) *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.

Ladefoged, Peter & Ian Maddieson. (1996) *The Sounds of the World’s Languages*. Cambridge: Blackwell.

Lombardi, Linda. (1995) *Why Place and Voice are Different: Constraint Interactions and Featural Faithfulness in Optimality Theory*. MS, University of Maryland.

Maddieson, Ian. (1984) *Patterns of Sounds*. Cambridge: Cambridge University Press.

Maddieson, Ian & Kristin Precoda. (1992) *UPSID and PHONEME (version 1.1)*. Los Angeles: UCLA.

McCarthy, John J. (to appear) Sympathy, cumulativity and the Duke-of-York gambit. In Féry, Caroline & Van de Vijver, Ruben (eds.) (to appear) *The Optimal Syllable*. Cambridge: Cambridge University Press.

McCarthy, John J. (2000) Sympathy and phonological opacity. *Phonology*, 16:331-399.

Newmark, L. (1957) *Structural Grammar of Albanian*. Bloomington: Indiana University.

Padgett, Jaye. (1995) Feature class theory. In Jill Beckman, Suzanne Urbanczyk and Laura Walsh Dickey (eds.) *University of Massachusetts Occasional Papers in Linguistics 18*. Amherst: GLSA. p. 385-420.

Paradis, Carole. (1988) Towards a theory of constraint violations. *McGill Working Papers in Linguistics*, 5:1-43.

Poedjosoedarmo, Gloria Risser. (1974) *Role Structure in Javanese*. Ph.D. dissertation, Cornell University.

Poulos, George & Louwrens, Louis J. (1994) *A Linguistic Analysis of Northern Sotho*. Pretoria: Via Afrika.

Prince, Alan. (1997) *Topics in OT 3: Harmonic Completeness, AP Order, Chain Shifts.* Lecture notes for LSA Institute 1997.

Prince, Alan. (1998) *Two lectures on Optimality Theory*. Handout of paper presented to Phonology Forum 1998, Kobe University. [http://ling.Rutgers.edu/reqources/archive/kobe-all.pdf.]

Prince, Alan & Paul Smolensky. (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms. Rutgers University and University of Colorado, Boulder. To appear, Cambridge, MA: MIT Press.

Reiss, Charles. (2000) *Deriving An Implicational Universal In Two Theories of Phonology*. Ms. Concordia University. [Available on ROA.]

Reynolds, W.T. (1994) *Variation and Phonological Theory*. Ph.D. thesis, University of Pennsylvania.

Rice, Keren. (1987) The function of structure preservation: derived environments. In Bernadette Plunkett and Joyce McDonough, eds. *Proceedings of the North East Linguistic Society 17*. Amherst: GLSA. pp. 501-19.

Samek-Lodovici, Vieri. (1992) *Universal Constraints and Morphological Gemination: a Crosslinguistic Study*. Ms., Brandeis University, Waltham, MA. [Available on ROA.]

Saporta, Sol. (1966) Phoneme distribution and language universals. In Joseph H. Greenberg (ed.) *Universals of language. Second edition*. Cambridge, MA: MIT Press.

Schiefer, L. and K. Kotten. (1986) Amplitude envelope and perception of breathy stops in Hindi. In M.P.R. Van den Broecke and A. Cohen. (eds.) *Proceedings of Tenth International Congress of Phonetic Sciences*. Dordrecht: Foris, 459-462.

Singh, Rajendra and Pieter Muysken. (1995) Wanted: a debate in pidgin/creole phonology. *Journal of Pidgin and Creole Languages* 10, 157-169.

Stevens, Kenneth N. (1998) *Acoustic Phonetics*. Cambrudge: MIT Press.

Thackston, Wheeler M. (1993) *An Introduction to Persian. 3rd ed*. Bethesda: Iranbooks.

Ultan, Russell. (1978) A typological view of metathesis. In Joseph H. Greenberg (ed.) *Universals of Human Language. Volume 2: Phonology*. Stanford: Stanford University Press. p. 376-402.

Van der Valk, J.W. (1928) *Beknopte Handleiding om de Javaansche Taal te Leeren Spreken*. Leiden: E.J. Brill.

Vennemann, Theo. (1988) *Preference Laws for Syllable Structure and the Explanation of Sound Change*. Berlin: Mouton de Gruyter.

Walker, Rachel. (1998) *Nasalization, Neutral Segments, and Opacity Effects*. Ph.D. dissertation, University of California, Santa Cruz.

Yip, Moira. (1988) The Obligatory Contour Principle and phonological rules: a loss of identity. *Linguistic Inquiry*, 19:65-100.

Zubritskaya, Katya. (1995) Markedness and sound change in OT. In Jill Beckman (ed.) *Proceedings of the NELS 25*. Amherst: GLSA. p. 249-64.

Zubritskaya, Katya. (1997) Mechanism of sound change in Optimality Theory. *Language Variation and Change*, 9:121-148.

Appendix A: Computing the output patterns of a grammar of the distribution of roundness on front and back vowels (see §4.2)

This appendix contains a discussion of the groups of grammars represented by (16a) through (16e). In this discussion it is shown how the possible output patterns of each of these groups of rankings can be determined, and also how it can be determined how many of the possible rankings will actually result in each of the possible output patterns. For a discussion of the grammars represented by (16f), refer to §4.2.

*A.1 All three preference constraints obeyed: *FRRD >>*BKUNRD >>IDENT(rnd)*

Since the preferred order is observed, not one of the preference constraints will be activated. All forms will vacuously satisfy these constraints, and they will therefore have no influence on the choice of the optimal forms. Since in all 120 of these rankings the markedness constraints outrank the faithfulness constraint, all of them will yield grammars from which both back unrounded and front rounded vowels are absent.

(43) No back unrounded or front rounded vowels = 120 rankings

			*FRRD	*BKUNRD	ID(rnd)	[*FRRD>> *BKUNRD]	[*BKUNRD>> ID(rnd)]	[*FRRD>> ID(rnd)]
(a)	/ø/	ø	*!					
(b)		☞ e			*			
(c)	/ɯ/	ɯ		*!				
(d)		☞ u			*			

*A.2 Disobeying only [*FRRD>>*BKUNRD]: *BKUNRD >> *FRRD >> IDENT(rnd)*

In these grammars, the preference constraint [*FRRD>>*BKUNRD] will be activated, and any candidate that violates *FRRD, will also be afforded a violation in terms of this preference constraint. Both of the other two preference constraints are obeyed, and all candidates will therefore vacuously satisfy them. The extra violation of all front rounded vowels in terms of the preference constraint [*FRRD>>*BKUNRD] will in actual fact have no influence on the choice of optimal candidate in these grammars. Because of the ranking *FRRD >> IDENT(rnd) a front rounded vowel will in any case be neutralized. Adding an extra violation to this candidate will simply confirm its non-optimal status. The ranking *BKUNRD >> IDENT(rnd) also assures that all back unrounded vowels are neutralized. Therefore, these grammars yield languages with no front rounded or back unrounded vowels.

(44) No front rounded or back unrounded vowels = 120 rankings

			*BKUNRD	*FRRD	ID(rnd)	[*FRRD>> *BKUNRD]	[*BKUNRD>> ID(rnd)]	[*FRRD>> ID(rnd)]
(a)	/ø/	ø		*!		*		
(b)		☞ e			*			
(c)	/ɯ/	ɯ	*!					
(d)		☞ u			*			

*A.3 Disobeying only [*BKUNRD>>IDENT(rnd)]: *FRRD >> IDENT(rnd) >> *BKUNRD*

The violations that will be afforded in grammars with this ranking, are indicated in the tableau below.

(45) *FRRD >> IDENT(rnd) >> *BKUNRD

			*FRRD	ID(rnd)	*BKUNRD	[*FRRD>> *BKUNRD]	[*BKUNRD>> ID(rnd)]	[*FRRD>> ID(rnd)]
(a)	/ø/	ø	*!					
(b)		☞ e		*				
(c)	/ɯ/	ɯ			*		*	
(d)		u		*				

Without taking into account the effect of preference constraints, these grammars would have resulted in phonemic inventories with back unrounded vowels, but no front rounded vowels. Since no preference constraint is violated by candidates (a) and (b), the choice of optimal candidate for a front rounded vowel input will not be influenced by addition of preference constraints. All of these grammars will result in inventories from which front rounded vowels are absent.

Because of the ranking IDENT(rnd) >> *BKUNRD, the default result of these grammars will be to faithfully parse underlying back unrounded vowels. However, the preference constraint [*BKUNRD>>IDENT(rnd)] is activated in these grammars with the result that all candidates with unrounded back vowels will receive an additional violation in terms of this constraint. Consequently, unrounded back vowels will also be neutralized in some of the grammars, namely in those where [*BKUNRD>>IDENT(rnd)] is ranked above IDENT(rnd). Disregarding the other two preference constraints leaves two possible positions higher than IDENT(rnd) for [*BKUNRD>>IDENT(rnd)] to be ranked into:

(46) Ranking positions of [*BkUnRd>>Ident(rnd)] that will result of neutralization of back unrounded vowels

(a) [*BkUnRd >> Ident(rnd)] >> *FrRd >> Ident(rnd) >> *BkUnRd

(b) *FrRd >> [*BkUnRd >> Ident(rnd)] >> Ident(rnd) >> *BkUnRd

Each of these two sub-hierarchies has 4 constraints in a fixed ranking. In each of them, the other two preference constraints can freely rerank – since they are not activated and therefore vacuously satisfied by all candidates. Using formula (17), it can then be calculated that each of (46a) and (46b) represents (4 + 2)!/4! = 30. The preference constraint [*BkUnRd>>Ident(rnd)] is responsible for shifting 60 of these grammars into the group that yields inventories from which both back unrounded and front rounded vowels are absent. In the remaining 60 rankings, back unrounded vowels will be faithfully parsed unto the surface.

The results of these rankings are summarized in the table below.

(47) Output of a grammar with the sub-hierarchy *FrRd >> Ident(rnd) >> *BkUnRd

Neither front rounded nor badck unrouded vowels	60
Back unrounded, but no front rounded vowels	60
Front rounded, but no back unrounded vowels	0
Both Front rounded and back unrounded vowels	0

*A.4 Disobeying [*BkUnRd >>Ident(rnd)] and [*FrRd >>Ident(rnd)]: Ident(rnd) >> *FrRd >> *BkUnRd*

The violations that will be afforded in grammars with this ranking, are indicated in the tableau below. [*BkUnRd>>Ident(rnd)] and [*FrRd>>Ident(rnd)] are both activated.

(48) Ident(rnd) >> *FrRd >> *BkUnRd

	Id(rnd)	*FrRd	*BkUnRd	[*FrRd>> *BkUnRd]	[*BkUnRd>> Id(rnd)]	[*FrRd>> Id(rnd)]
(a) /ø/ ø		*				*
(b) e	*					
(c) /ɯ/ ɯ			*		*	
(d) u	*					

Disregarding the preference constraints, these grammars would have resulted in inventories with both back unrounded and front rounded vowels. However, there is one preference constraint militating against back unrounded vowels, and also one against front rounded vowels. Consequently, depending on the ranking of the preference constraints, some of the grammars will be shifted into the group with neither front rounded nor back unrounded vowels, some into the group with only back unrounded vowels, and some into the group with only front rounded vowels.

When both of the activated preference constraints ([*BkUnRd>>Ident(rnd)] and [*FrRd>>Ident(rnd)]) are ranked below Ident(rnd), grammars with both front rounded and back unrounded vowels will result. To calculate how many rankings will comply with this requirement, first disregard all the preference constraints except for [*FrRd >>Ident(rnd)]. There are three possible rankings in which this constraint will be dominated by Ident(rnd).

(49) Rankings with [*FrRd>> Ident(rnd)] dominated by Ident(rnd)

(a) Ident(rnd) >> [*FrRd>> Ident(rnd)] >> *FrRd >> *BkUnRd

(b) Ident(rnd) >> *FrRd >> [*FrRd>> Ident(rnd)] >> *BkUnRd

(c) Ident(rnd) >> *FrRd >> *BkUnRd >> >> [*FrRd>> Ident(rnd)]

In each of these three sub-hierarchies there are 4 possible positions for [*BkUnRd>>Ident(rnd)] to be ranked into, such that it is also dominated by Ident(rnd), yielding 12 rankings in which both of the activated constraints are ranked below the faithfulness constraint. The preference constraint [*FrRd>>*BkUnRd] can then be ranked into 6 different positions in each of these 12, to give a total of 72 rankings that will result in preservation of both back unrounded and front rounded vowels.

When both [*BkUnRd>>Ident(rnd)] and [*FrRd>>Ident(rnd)] outrank Ident(rnd), languages with neither back unrounded nor front rounded vowels will result. Disregarding [*FrRd>>*BkUnRd] there are two rankings that comply with this requirement:

(50) Rankings with [*FrRd>>Ident(rnd)] and [*BkUnRd >>Ident(rnd)] dominating Ident(rnd)

(a) [*FrRd>>Id(rnd)] >> [*BkUnRd>>Id(rnd)] >> Id(rnd) >> *FrRd >> *BkUnRd

(b) [*BkUnRd>>Id(rnd)] >> [*FrRd>>Id(rnd)] >> Id(rnd) >> *FRRd >> *BkUnRd

Adding [*FrRd>>*BkUnRd] into these two sub-hierarchies results in 12 possible rankings that will lead to neutralization of aspiration on all stops.

When the ranking [*FrRd>>Ident(rnd)] >> Ident(rnd) >> [*BkUnRd>>Ident(rnd)] is observed, grammars with neutralization of front rounded vowels but preservation of back unrounded vowels will result. Disregarding [*FrRd>>*BkUnRd] there are three rankings complying with this requirement.

(51) Rankings with [*FRRD>>IDENT(rnd)]>>IDENT(rnd) >>[*BKUNRD>>IDENT(rnd)]

(a) [*FRRD>>ID(rnd)] >> ID(rnd) >> [*BKUNRD>>ID(rnd)] >> *FRRD >> *BKUNRD

(b) [*FRRD>>ID(rnd)] >> ID(rnd) >> *FRRD >> [*BKUNRD>>ID(rnd)] >> *BKUNRD

(c) [*FRRD>>ID(rnd)] >> ID(rnd) >> *FRRD >> *BKUNRD >> [*BKUNRD>>ID(rnd)]

There are six possible positions for [*FRRD>>*BKUNRD] to be ranked in each of these three sub-hierarchies. This means that there are 18 rankings in which back unrounded vowels will be preserved, but front rounded vowels neutralized. By exactly similar reasoning, it can be calculated that there are also 18 rankings in which front rounded vowels will be preserved, but back unrounded vowels neutralized.

The possible outputs of the rankings in this section are summarized in the table below:

(52) Outputs of a grammar with the sub-hierarchy IDENT(rnd) >>*FRRD >> *BKUNRD

Neither front rounded nor badck unrouded vowels	12
Back unrounded, but no front rounded vowels	18
Front rounded, but no back unrounded vowels	18
Both Front rounded and back unrounded vowels	72

*A.5 Disobeying [*FRRD>>*BKUNRD] and [*FRRD>>IDENT(rnd)]: *BKUNRD >> IDENT(rnd) >> *O*

The violations that will be afforded in grammars with this ranking, are recorded in the tableau below. Both [*FRRD>>*BKUNRD] and [*FRRD>>IDENT(rnd)] are activated.

(53) *BKUNRD >> IDENT(rnd) >> *FRRD

	*BKUNRD	ID(rnd)	*FRRD	[*FRRD>> *BKUNRD]	[*BKUNRD>> ID(rnd)]	[*FRRD>> ID(rnd)]
(a) /ø/ ø			*	*		*
(b) e		*				
(c) /ɯ/ ɯ	*!					
(d) ☞ u		*				

If the preference constraints are disregarded, these grammars will yield inventories with front rounded vowels, but no back unrounded vowels. However, there are two separate preference constraints militating against front rounded vowels. Rankings in which any of these two outrank faithfulness, will be shifted into the group of languages with neither front rounded, nor back unrounded vowels.

Since neither of candidates (c) or (d) violates any of the preference constraints, these constraints will have no influence on the choice of optimal candidate for a back unrounded vowel input. It follows then that all of the possible rankings in this group will result in inventories without back unrounded vowels.

If any or both of [*FRRD>>*BKUNRD] and [*FRRD>>IDENT(rnd)] outrank IDENT(rnd), then front unrounded vowels will be neutralized. To determine the number of grammars that will comply with this requirement, it is easier to calculate its complement, i.e. the number of grammars in which both of the activated preference constraints will be ranked below IDENT(rnd). Disregarding all the preference constraints except for [*FRRD>>*BKUNRD], there are two rankings in which [*FRRD>>*BKUNRD] will rank below IDENT(rnd).

(54) Rankings with IDENT(rnd) >> [*FRRD>>*BKUNRD]

(a) *BKUNRD >> IDENT(rnd) >> [*FRRD>>*BKUNRD] >> *FRRD

(b) *BKUNRD >> IDENT(rnd) >> *FRRD >> [*FRRD>>*BKUNRD]

In both of these rankings there are 3 possible positions for [*FRRD>>IDENT(rnd)] to be ranked such that it is also dominated by IDENT(rnd). This means that there are 6 of these sub-hierarchies in which both of the activated preference constraints are ranked below IDENT(rnd). In each of these 6 sub-hierarchies there are 6 possible sites for [*BKUNRD>>IDENT(rnd)] to be ranked into. There are thus a total of 36 rankings in which both of the activated preference constraints will be outranked by IDENT(rnd). In these 36 rankings front rounded vowels will be parsed faithfully unto surface structure. In the complement of this in 120, i.e. in 84 rankings, front rounded vowels will be neutralized.

The results of these rankings are summarized in the table below.

(55) Output of a grammar with the sub-hierarchy *BKUNRD >> IDENT(rnd) >> *FRRD

Neither front rounded nor back unrouded vowels	84
Back unrounded, but no front rounded vowels	0
Front rounded, but no back unrounded vowels	36
Both Front rounded and back unrounded vowels	0

Appendix B: Grammars in (33b) that will result in a {O, N, V} syllabic peak inventory (cf. §5.1)

(56) *P/L >> DEP >> *P/O >> *P/N

			*P/L	DEP	*P/O	*P/N	[*P/O>> *P/N]	[*P/N>> *P/L]	[*P/O>> P/L]	[*P/O>> DEP]	[*P/N>> DEP]
(a)	/O/	.O.			*				*	*	
(b)		.OV.		*							
(c)	/L/	.L.	*!								
(d)	☞	.LV.		*							
(e)	/N/	.N.				*		*			*
(f)		.NV.		*							

In this tableau a system with as possible nuclei {O, N, V} will only result as long as the following conditions hold:

(57) Conditions for {O, N, V} as peaks under ranking *P/L >> DEP >> *P/O >> *P/N:

(a) DEP >> [*P/O>>*P/L], [*P/O>>DEP]

AND

(b) DEP >> [*P/N>>*P/L], [*P/N>>DEP]

Systems not complying with condition (a), will not allow consonants of sonority lower than nasals into peak position, i.e. {N, V}. Systems not complying with condition (b) will not allow nasals into peak position, i.e. {O, V}. Systems complying with neither condition (a) nor condition (b) will allow only vowels into peak position, i.e. {V}. Once again it is necessary to compute the percentage of grammars complying with both conditions (a) and (b), i.e. systems where DEP outranks [*P/O>>*P/N], [*P/O>>DEP], [*P/N>>*P/L], and [*P/N>>DEP]. When these 4 constraints, together with *P/O and *P/N, are ranked below DEP, there is a total of 6 constraints in this position. These 6 constraints can be ranked in 6! = 720 different ways. However, since the ranking between *P/O and *P/N is fixed, the number should be divided by 2, yielding 360 possible rankings. Now the two unviolated preference constraints ([*P/O>>*P/N] and [*P/L>>DEP]) still need to be accounted for. These two constraints are obeyed by all candidates, and therefore their ranking will have no influence on the output. Without these two constraints there are 8 constraints. In a hierarchy of 8 constraints, there are 9 possible positions for a 9th constraint to be ranked into. In the 8 level hierarchy there is therefore 9 possible positions for [*P/L>>DEP]. This then creates a 9 level hierarchy with 10 possible positions for [*P/O>>*P/N]. The earlier total of 360 therefore has to be multiplied by 9 and 10, yielding a final total of 32,400 rankings complying with both conditions (57a) and (57b). The ranking *P/L >> DEP >> *P/O >> *P/N then contributes 32,400 grammars that will result in a syllabic peak inventory {O, N, V}.

Appendix C: Outputs shifted into the {O, N, V} syllabic peak inventory group (cf. §5.1)

This Appendix contains a discussion of the groups of ranking that will as a result of output shifting result in grammars with as peak inventories {O, N, V}. One of these groups of rankings was discussed in detail in the text (see §5.1 and (38)). This Appendix is a discussion of the remaining 5 groups of rankings that fall into this category.

(58) Ranking: *P/O >> *P/N >>*P/L

Additional preference constraints activated: ∅.

Necessary condition for an output to be shifted into {O, N, V}:

[*P/L>>DEP] >> DEP >> *P/O >> *P/N >>*P/L

>> [*P/O>>DEP], [*P/N>>DEP]

There are 5 constraints below DEP. The ranking between three of them is fixed (the three ordinary constraints), while the other two (the preference constraints) can freely rerank as long as they are ranked below DEP. Using formula (17) it can be computed that this represents (3 + 2)!/3! = 20 rankings that comply with the conditions in (38). Only seven constraints have been accounted for thus far – 4 ordinary constraints and 3 preference constraints. There are 3 more preference constraints, however. These preference constraints are not activated and therefore they can freely rerank between the other seven constraints. Each of the 20 rankings accounted for thus far therefore represent rankings with 7 constraints in a fixed ranking, and 3 that can freely rerank. Once again using (17), it is computed that each of these hierarchies therefore represent (7 + 3)/7! = 720 rankings. This group of rankings therefore contributes a total of (20)(720) = 14,400 rankings that will result in a {O, N, V} syllabic peak inventory.

(59) Ranking: *P/O >> *P/L >>*P/N

Additional preference constraints activated: [*P/N>>*P/L]

Necessary condition for an output to be shifted into {O, N, V}:

[*P/L>>DEP] >> DEP >> *P/O >> *P/N >>*P/L

>> [*P/O>>DEP], [*P/N>>DEP], [*P/N>>*P/L]

There are 6 constraints below DEP. The ranking between three of them is fixed (the three ordinary constraints), while the other three (the preference constraints) can freely rerank as long as they are ranked below DEP. Using formula (17) it can be computed that this represents (3 + 3)!/3! = 120 rankings that comply with the conditions in (38). Only eight constraints have been accounted for thus far – 4 ordinary constraints and 4 preference constraints. There are 2 more preference constraints, however. These preference constraints are not activated and therefore they can freely rerank between the other seven constraints. Each of the 120 rankings accounted for thus far therefore represent rankings with 8 constraints in a fixed ranking, and 2 that can freely rerank. Once again using (17), it is computed that each of these hierarchies therefore represent (8 + 2)/8! = 90 rankings. This group of rankings therefore contributes a total of (120)(90) = 10,800 rankings that will result in a {O, N, V} syllabic peak inventory.

(60) Ranking: *P/N >> *P/L >>*P/O

Additional preference constraints activated: [*P/O>>*P/N], [*P/O>>*P/L]

Necessary condition for an output to be shifted into {O, N, V}:

[*P/L>>DEP] >> DEP >> *P/O >> *P/N >>*P/L

>> [*P/O>>DEP], [*P/N>>DEP], [*P/O>>*P/N],

[*P/O>>*P/L]

There are 7 constraints below DEP. The ranking between 3 of them is fixed (the ordinary constraints), while the other 4 (the preference constraints) can freely rerank as long as they are ranked below DEP. Using formula (17) it can be computed that this represents (3 + 4)!/3! = 840 rankings that comply with the conditions in (38). Only 9 constraints have been accounted for thus far – 4 ordinary constraints and 5 preference constraints. There is 1 more preference constraint, however. This preference constraint is not activated and therefore it can freely rerank between the other seven constraints. Each of the 840 rankings accounted for thus far therefore represent 10 rankings. This group of rankings therefore contributes a total of (10)(840) = 8,400 rankings that will result in a {O, N, V} syllabic peak inventory.

(61) Ranking: *P/N >> *P/O >>*P/L

Additional preference constraints activated: [*P/O>>*P/N]

Necessary condition for an output to be shifted into {O, N, V}:

[*P/L>>DEP] >> DEP >> *P/O >> *P/N >>*P/L

>> [*P/O>>DEP], [*P/N>>DEP], [*P/O>>*P/N]

There are 6 constraints below DEP. The ranking between 3 of them is fixed (the ordinary constraints), while the other 3 (the preference constraints) can freely rerank as long as they are ranked below DEP. Using formula (17) it can be computed that this represents (3

+ 3)!/3! = 120 rankings that comply with the conditions in (38). Only 8 constraints have been accounted for thus far – 4 ordinary constraints and 4 preference constraints. There are 2 more preference constraints, however. These preference constraints are not activated and therefore they can freely rerank between the other seven constraints. Each of the 120 rankings accounted for thus far therefore represent rankings with 8 constraints in a fixed ranking, and 2 that can freely rerank. Once again using (17), it is computed that each of these hierarchies therefore represent (8 + 2)/8! = 90 rankings. This group of rankings therefore contributes a total of (90)(120) = 10,800 rankings that will result in a {O, N, V} syllabic peak inventory.

(62) Ranking: *P/L >> *P/O >>*P/N

Additional preference constraints activated: [*P/N>>*P/L], [*P/O>>*P/L]

Necessary condition for an output to be shifted into {O, N, V}:

[*P/L>>DEP] >> DEP >> *P/O >> *P/N >>*P/L

>> [*P/O>>DEP], [*P/N>>DEP], [*P/N>>*P/L],

[*P/O>>*P/L]

There are 7 constraints below DEP. The ranking between 3 of them is fixed (the three ordinary constraints), while the other 4 (the preference constraints) can freely rerank as long as they are ranked below DEP. Using formula (17) it can be computed that this represents (3 + 4)!/3! = 840 rankings that comply with the conditions in (38). Only 9 constraints have been accounted for thus far – 4 ordinary constraints and 5 preference constraints. There is 1 more preference constraint, however. This preference constraint is not activated and therefore it can freely rerank between the other 9 constraints. Each of the 840 rankings accounted for thus far therefore represent 10 different rankings This group of rankings therefore contributes a total of (10)(840) = 8,400 rankings that will result in a {O, N, V} syllabic peak inventory.