Abstract
The mechanism of constraint re-ranking basis makes Optimality Theory (OT) a typologically oriented theory of phonology. Permuting the prioritization of constraints on a language-specific basis generates a factorial typology of phonological patterns. Patterns generated by an optimality-theoretic account can be compared with the set of attested patterns found in languages of the world to determine how well the analysis fits the observed typology. The typological predictions of OT are constrained to a large extent by the grounding of markedness constraints in independent facts about speech articulation, perception, and processing. Nevertheless, comprehensive typological analyses within OT commonly overpredict the existence of apparently unattested patterns. These cases of overgeneration may stem from several sources: they may be pathologic, may reflect accidental gaps due to the paucity of a given phenomenon, or may be grounded in diachronic biases in speech perception or production. The phonetic underpinnings and factorial typology of three phenomena are discussed in this article: syllable structure, the distribution of contour tones, and the loci of laryngeal neutralization.

1. Introduction
One of the primary goals of phonological theory is to capture the range of variation observed in sound systems of the world’s languages. While most theories take this goal seriously, the formalism of Optimality Theory (OT) (Prince and Smolensky 2004) is particularly well suited to making explicit typological predictions. The fundamental assumption of OT that constraint ranking varies from language to language has provided fertile ground for typological research in phonology.

2. The Formalism of Optimality Theory
Unlike earlier rule-based generative treatments of phonology, OT (Prince and Smolensky 2004), in its original conception, assumes a direct mapping between underlying and surface forms. The crucial element in this mapping operation is a set of hierarchically ranked constraints that evaluate a series of potential surface forms, or candidate forms, corresponding to the underlying form. Constraints are assumed to be innate and universal
in Prince and Smolensky’s (2004) original formation of the theory, although this assumption is logically independent of the constraint-based formalism employed in OT (see Hayes 1999 and Hayes and Wilson forthcoming for approaches within OT that do not assume innate constraints). Constraints are of two basic types. Some constraints are sensitive to the well-formedness of candidate forms, banning dispreferred or marked structures or requiring felicitous or unmarked properties. Other constraints are sensitive to the mapping, or correspondence, between underlying forms and candidate forms or between surface forms. Ideally, all constraints find principled motivation in independent properties of speech production, perception, and processing (see Boersma 1998 and the articles in Hayes et al. 2004 for a representative sample of OT work employing grounded constraints). For example, constraints may ban sequences of sounds that are difficult to articulate or difficult to perceive, while other constraints may license the preservation of contrasts in environments important for word recognition, for example, in initial position. The OT literature on constraint grounding builds on the natural phonology program (Stampe 1972; Vennemann 1974) and appeals to results of phonetic research (see Ohala 1997 and Kingston 2007 for overviews of the role of phonetics in explaining phonological patterns) as well as speech error (e.g. Schwartz et al. 1994; Dell et al. 1997) and child acquisition (e.g. Smith 1973; Vihman 1978) studies. By imposing this grounding requirement on constraints, OT avoids a proliferation of unprincipled constraints and the unattested patterns they would generate (see Section 7 for further discussion of generation in OT). The actual surface form is determined by evaluating how well the candidates satisfy the constraints, which, in many cases, impose competing demands. The ranking of constraints varies from language to language, thereby accounting for cross-linguistic variation in surface forms.

To take a relatively simple example of constraint evaluation and language-specific constraint ranking, we consider the case of syllable structure, which has been the subject of study since long before the advent of OT (e.g. Greenberg 1978, Ito 1986; see Blevins 1995 for a cross-linguistic overview of syllabification patterns). Many languages, for example, Hawaiian and Cayuvava, only allow open syllables, whereas other languages, for example, English, also permit syllables closed by a consonant. The dispreference for closed syllables is attributed to a constraint against coda consonants, *CODA, where the asterisk indicates that a given structure is prohibited. A competing constraint, FAITHC, requires surface faithfulness to underlying suffix, that is, requires that underlying consonants survive to the surface. Given an underlying string consisting of a consonant final root followed by a consonant initial suffix, that is, CVC–CV, there are at least two possible candidates in competition to surface. One of the two intervocalic consonants could be deleted, a strategy that would satisfy the ban on coda consonants but that would run afoul of FAITHC due to its
missing consonant. Consonant deletion is attributed to the ranking of \( \ast \text{CODA} \) above \( \text{FAITHC} \), as depicted in (1) by means of a tableau, the standard method of formalizing OT analyses.

(1) Coda deletion due to the ranking: \( \ast \text{CODA} \gg \text{FAITHC} \)

<table>
<thead>
<tr>
<th>/CVC-CV/</th>
<th>( \ast \text{Coda} )</th>
<th>( \text{FAITHC} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVCV</td>
<td>( \ast )</td>
<td>*</td>
</tr>
<tr>
<td>CVCCV</td>
<td>( \ast ! )</td>
<td></td>
</tr>
</tbody>
</table>

In tableaux, the underlying (or ‘input’) form customarily appears in the top left corner followed in the same column by competing candidate forms corresponding to that underlying form. After the first column, each constraint is given its own column where constraints to the left outrank those to their right. Each cell under a constraint name depicts the number of times the candidate in the corresponding row violates that constraint, where each violation is indicated by an asterisk. A candidate that commits one or more violations of the top-ranked constraint is eliminated from contention, indicated by an exclamation point after the fatal violation, if there is at least one other surviving candidate that commits fewer violations of that constraint. Candidates are gradually eliminated as constraint evaluation proceeds from higher to lower ranked constraints until there is only one candidate remaining. This winning candidate is the surface form, traditionally indicated by a pointing finger \( \ast \). Once a candidate is weeded out due to a violation of a high ranked constraint, it cannot be re-deemed no matter how well it satisfies lower-ranked constraints. Cells that occur to the right of a fatal violation for a given candidate as well as cells occurring to the right of the constraint that has enabled determination of a unique winner are shaded to indicate their irrelevance to the process of selecting a winning candidate.

In (1), the second candidate, the one containing both consonants in the cluster, is eliminated from contention due to its violation of the anti-coda constraint. This leaves the first candidate, the one in which a consonant has been deleted, as the winner even though it commits a violation of lowly ranked faithfulness.

As an alternative to consonant deletion, both members of the consonant cluster could be preserved, thereby honoring faithfulness but not the ban on coda consonants. The ranking of \( \ast \text{CODA} \) and \( \text{FAITHC} \) relative to each other in a given language determines that language’s response to the potential consonant cluster. If \( \ast \text{CODA} \) outranks \( \text{FAITHC} \), one of the consonants does not survive to the surface as shown above in (1), whereas if \( \text{FAITHC} \) is ranked above \( \ast \text{CODA} \), both consonants emerge unscathed, as shown in (2).

(2) Coda preservation due to the ranking: \( \text{FAITHC} \gg \ast \text{CODA} \)
Of course, other constraints come into play in the evaluation of additional candidate forms corresponding to the input. For example, in languages in which *CODA is ranked above FaithC, another constraint must determine which of the intervocalic consonants is deleted in order to satisfy *CODA.

The typology of syllabification can also be expanded to include syllables with complex onsets and codas (see Blevins 1995 for a typology of syllabification patterns). Many languages, for example, Sedang, allow simple one consonant codas, but do not permit multiple consonants in coda position, a pattern that is captured by a highly ranked *COMPLEX-CODA constraint. Other languages, for example, Klamath, tolerate complex codas but not complex onsets, a situation resulting from a highly ranked *COMPLEX-ONSET. Syllabification constraints invoked to account for the cross-linguistic patterns also find support from the language acquisition process, in which syllable types avoided in many languages also tend to be acquired relatively late by children speaking languages with richer syllable inventories (see Levelt and van de Vijver 2004).

3. Re-ranking and Typology in OT

Cross-linguistic variation in rankings is one of the cornerstones of OT and makes the theory inherently well suited to tackle typological issues in phonology. An OT analysis of a given phenomenon can be evaluated on the basis of how closely the patterns predicted through constraint re-ranking, as in our syllabification example earlier, fit patterns attested in languages of the world. Ideally, the patterns generated through permutation of the constraint rankings exactly match the set of patterns found cross-linguistically. In reality, this ideal fit is seldom achieved. Rather, analyses typically overgenerate by predicting patterns that either do not occur cross-linguistically or at least have not been identified yet. Overgeneration is often difficult to assess critically, as unattested patterns predicted to occur by an analysis potentially reflect accidental gaps due to insufficient cross-linguistic data as opposed to pathologic patterns that could not reasonably be expected to occur. More seriously, analyses may undergenerate by failing to capture patterns that have been found in one or more languages. The issue of overgeneration and undergeneration in OT is discussed in Section 7.

4. Conspiracies in OT

One of the virtues of OT in the typological domain is its prediction of ‘conspiracies’: instances where different languages employ diverse strategies...
that conspire to avoid the same ill-formed configuration. OT explicitly predicts the existence of such conspiracies due to its formal separation of constraints banning a particular structure from constraints sensitive to potential ways of eliminating that prohibited structure. For example, the constraint against coda consonants merely penalizes candidates that have a coda consonant without determining the path a language will take to avoid codas. There are several routes a language could take to avoid coda consonants ranging from deletion of the first consonant (CVC1C2V → CVC2V), to deletion of the second consonant (CVC1C2V → CVC1V), to insertion of an epenthetic vowel (CVC1C2V → CVC1VC2V), to changing a consonant into a vowel (CVC1C2V → CVVC2V). The choice among these options is governed by the relative ranking of other constraints, for example, an anti-epenthesis constraint banning vowels that are not present underlyingly, an anti-deletion constraint requiring that underlying consonants surface, a constraint banning the conversion of an underlying consonant to a vowel on the surface. A key feature of the OT analysis is its treatment of all of these processes as strategies serving a common goal driven by the same constraint against codas. This differs from traditional rule-based approaches, which collapse the ill-formed configuration and the process eliminating the ill-formed structure into a single rule. The result is a series of formally distinct rules, each one capturing a different strategy for avoiding codas. The common goal uniting the various coda elimination processes is not transparent in the rule-based analysis and the set of possible rules predicted by the theory is not logically limited to those repairing ill-formed structures. For example, it is not clear what precludes the existence of a rule inserting a coda consonant. In OT, such an operation could only occur in order to satisfy a highly ranked constraint on syllable structure. If there were no constraint that penalized codaless syllables, the theory would never predict coda insertion. Crucially, such a constraint (at least in non-final position) would be ungrounded, as there would be no phonetic or functional reason why a closed syllable would be preferred to an open syllable. The prevocalic consonant in an open syllable is easily identified due to the salient acoustic transitions provided by the following vowel. In contrast, the coda consonant occurring in a closed syllable is less readily identifiable (see Jun 2004; Steriade 1999, 2001 for discussion of this asymmetry in the context of OT). The dispreference for closed syllables thus amounts to a perceptually grounded dispreference for coda consonants.

5. Implicational Hierarchies and OT

An important typological feature that must be handled by any phonological theory is the existence of implicational relationships. For example, the presence of closed syllables in a language implies the occurrence of open syllables. The converse implication is not supported: there are many languages
with open syllables that lack closed syllables. As we have seen, this particular implication is handled in OT by assuming a perceptually driven constraint against coda consonants but not its ungrounded counterpart banning codaless syllables. Given this constraint inventory, only two possible patterns are generated: a language with no-coda consonants, a pattern resulting from highly ranked \*CODA, and a language with both closed and open syllables, a distribution attributed to the prioritization of the constraint requiring that underlying consonants be preserved. It is impossible to generate a language that requires coda insertion everywhere as no constraint mandates that all syllables should be closed.

Beyond this simple binary implicational relationship holding of closed and open syllables, there are other more complex implications sensitive to hierarchies. For example, the selection of syllabic peaks adheres to a hierarchy that closely mirrors sonority scales (e.g. Steriade 1982; Selkirk 1984; Clements 1990; Parker 2002). All languages preferentially choose more sonorous segments as syllable nuclei over less sonorous sounds. Using a relatively coarse hierarchy like that depicted in (3), vowels are preferred as syllable peaks over liquids; liquids are preferred over nasals, which, in turn, are preferable to fricatives, which, in turn, make better peaks than stops. Languages differ in the cut-off point they impose between sounds that could potentially be syllable nuclei and those that are not permissible peaks, as shown in (3) (Blevins 1995).

(3) Implication hierarchy of syllabification of syllable peaks

For example, English tolerates vowels as peaks, liquids as peaks, for example, in the second syllable of little and butter, and nasals as peaks, for example, in the second syllable of button and prism. No syllable, however, may have a fricative or stop as a syllable nucleus in English. Hawai’ian is more restrictive in the segments that can serve as nuclei: only vowels may be syllable peaks. Imdlawn Tashihiyt Berber, on the other hand, allows any type of sound, even stops, to be syllable nuclei. There are no languages that tolerate less sonorous sounds as nuclei but disallow more sonorous nuclei. The occurrence of relatively non-sonorous syllable peaks thus implies the existence of more sonorous peaks.

Prince and Smolensky (2004) capture the distribution of syllable nuclei by positing a series of constraints governing the ability of different sounds to be peaks. Constraints take the form \*P/x, meaning that segment type \(x\) is prohibited as a nucleus, where \(x\) is a variable ranging over different sounds. Thus, if we collapse for the sake of illustration, some of the
distinctions made by Prince and Smolensky, one constraint, *P/V, prohibits vocalic nuclei, another constraint, *P/L, bans liquids from being peaks, yet another constraint, *P/N, penalizes nasals serving as syllable nuclei, while other constraints ban syllabic fricatives (*P/S) and syllabic stops (*P/T). In order to capture the implicational nature of the syllabicity hierarchy, Prince and Smolensky assume that the ranking of the peak constraints adheres to a universal hierarchy, where constraints banning less sonorous nuclei outrank constraints banning more sonorous nuclei, as in (4).

(4) Constraint hierarchy for syllabification of syllable peaks

\[ \begin{array}{cccc}
*P/T & >> & *P/S & >> & *P/N & >> & *P/L & >> & *P/V \\
\end{array} \]

They assume that the syllabicity hierarchy is ultimately projected from a scale of intrinsic prominence, which turns out to be most closely linked to the phonetic property of intensity (Ladefoged 1975; Parker 2002). Crucially, if one did not assume that the ranking of syllabicity constraints is fixed, we would incorrectly predict a number of unattested patterns. For example, if *P/V could be ranked above *P/T, we would generate a language in which stops could be syllable peaks but vowels could not be. Such a language is unattested and should thus not be predicted by an analysis to occur. Other competing constraints are interleaved with constraints on syllabicity thereby producing the observed cross-linguistic variation in syllable peak formation. For example, *CODA is often better satisfied in many positions by creating another syllable. Thus, the string [kot] violates *CODA, if the [t] is syllabified as a coda to the syllable headed by [o]. If, on the other hand, [t] is made into a nucleus, that is, [ko.t], a violation of *P/T is incurred.


Let us briefly consider three schematic analyses of phonological hierarchies appealing to phonetically driven scales: one governing the distribution of contour tones (Gordon 1998, 2001), one sensitive to the preservation of laryngeal contrasts (Steriade 1999), and a third one operative in sonority-sensitive stress systems (De Lacy 2004).

5.1. PHONETIC SCALES I: CONTOUR TONES

Gordon (1998, 2001) posits a hierarchy of constraints capturing the ability of different syllable types to support contour tones (see also Zhang 2002,
2004 for a different scale of phonetically driven constraints on tone). Syllables containing long vowels (CVV) are best suited to realizing contour tones, followed by syllables consisting of a short vowel followed by a sonorant coda (CVR), followed by syllables containing a short vowel plus obstruent coda (CVO), followed by open syllables containing a short vowel (CV).

The hierarchy of contour tones is a function of the ability of different segment types to support tonal information. Vowels are best suited to carrying tones followed by sonorant consonants, and then obstruents. Long vowels are best for supporting contour tones as they provide the longest backdrop of maximal sonority for realizing the full tonal excursion associated with a contour. Many languages, 25 in Gordon’s (2001) survey, thus only allow contour tones on CVV syllables. CVR syllables provide the next best docking site for a tonal transition. Accordingly, many languages (29 in Gordon’s study) tolerate contours on both CVV and on CVR, but not on other syllable types. CVO is considerably less effective for realizing contour tones, only slightly better than CV. It is thus not surprising that very few languages (only three in Gordon’s survey) preferentially allow contour tones on CVO but not on CV. As predicted by the phonetic considerations guiding the realization of tonal contours, no languages preferentially allow contour tones on syllables that are relatively poor conveyors of tone but disallow contours on syllable types better suited to realizing tone. The hierarchy of contour tone bearing ability and the resulting phonological patterns corresponding to different cut-off points in the hierarchy are given in (5).

(5) Implicational hierarchy of contour tones

A series of constraints can account for the implicational scale of contour tones. A constraint against contour tones on open syllables with short vowels is at the top of the constraint hierarchy, followed in turn by a constraint banning contour tones on short voweled syllables closed by an obstruent, then a constraint prohibiting contours on short voweled syllables closed by a sonorant, and finally a constraint against contours on syllables containing a long vowel. Interlanguage variation in the cut-off points in the hierarchy between syllables that allow contours and those that do not is attributed to competing faithfulness constraints that act to preserve underlying contour tones. Depending on the ranking of faithfulness relative to the contour tone well-formedness constraints, different patterns are generated. For example, if faithfulness is outranked by all the constraints
against contour tones, contours are not permissible on any syllable types. If faithfulness is prioritized over only the constraint against contours on CV, a language with contour tones on all syllable types except CV is generated. If faithfulness outranks both the constraint against contours on CV and the constraint banning contours on CVO, the resulting pattern permits contours on CVV and CVR but not on CVO or CV. If faithfulness takes precedence over all anti-contour constraints except the one referring to CVV, we get a language that tolerates contours only on CVV. Finally, if faithfulness is ranked above all anti-contour tone constraints, contours are permitted on all syllable types. The schematic ranking of faithfulness relative to constraints governing tone-to-syllable associations and the resulting tonal patterns are shown in (6).

(6) Interaction of faithfulness with contour tone restrictions

The interaction between the constraint requiring faithful realization of tones and the constraints against contour tones on different syllable types can be illustrated with data from Kiowa, an American Indian language of North America. In Kiowa (Watkins 1984), contour tones (a falling tone only) are permitted on long vowels (VV) and on syllables closed by a sonorant (VR) but not on open syllables containing a short vowel (V) or on syllables containing a short vowel followed by an obstruent coda (VO). The restriction against contour tones on V and VO is evident in a productive process of vowel shortening targeting long vowels in closed syllables. If the shortened long vowel is followed by an obstruent coda, the second element of the contour tone is deleted leaving a simple high tone, as in the form [kʰút] ‘pull off’ perfective from underlying [kʰût-l-t]. Contour tone simplification does not take place in syllables containing a shortened vowel followed by a coda sonorant, as in the first syllable of the form [kʰîñm] ‘cough’ imperfective from underlying [kʰîtnm].

A few constraints are necessary to account for the Kiowa tone data. First, the constraints against contour tones on V and VO rimes must be ranked above the constraint requiring preservation of underlying tones, while the constraints banning contours on VR and VV must be ranked below the faithfulness constraint. Second, a phonotactic constraint is needed to drive the process of vowel shortening in closed syllables, a common cross-linguistic phenomenon (see Maddieson 1985; Gordon 2007). This constraint may be simply formulated as *OVERLONG, which bans rimes containing greater than two timing positions or moras. A rime containing
a long vowel plus a coda consonant (or two coda consonants) violates *Overlong, as the long vowel is associated with two timing positions while the coda adds an additional timing slot. *Overlong must outrank preserve tone as well as a constraint requiring that underlying vowel length be preserved, preserve length. We thus have the following ranking, where constraints that are not crucially ranked with respect to each other (i.e. where reversing their ranking does not lead to a different output) are separated by dashed lines in tableaux: *Overlong, *[TT]V, *[TT]VO >> preserve tone, preserve length >> *[TT]VR, *[TT]VV. The tableau in (7) illustrates these rankings for the forms [kʰút] ‘pull off’ and [kʰinmɔ] ‘cough’.

(7) An OT analysis of contour tones in Kiowa

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) kʰút</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) kʰ̂ult</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) kʰút</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) kʰinmɔ</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) kʰ̂inmɔ</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) kʰ̂inmɔ</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Looking at the first form in (7), the second candidate (b) is the fully faithful one in which vowel length and the contour tone have been preserved. This candidate is eliminated for its violation of *Overlong. The final candidate (c) shortens the long vowel but is faithful with respect to tone. This tonal faithfulness, however, is fatal because it entails a violation of the constraint against contour tones on VO rimes. This leaves the winner as the candidate (a) with a short vowel and a simple high tone. If the ranking of *Overlong above either preserve tone and preserve length were reversed, one of the failed candidates would erroneously become the winning candidate. Furthermore, if the constraint against contour tones on VO were ranked below preserve tone, the third candidate would incorrectly be predicted to win. It may be noted that other constraints are necessary to rule out other candidates not considered in the tableau for reasons of space. For example, a complete analysis must account for the preservation of the high element in the contour tone rather than the low element, as well as the deletion of the root final /l/ rather than the perfective suffix /-t/.

Turning to the second form, candidate (b) fails due to its violation of *Overlong. Candidate (c) is also eliminated, because it needlessly simplifies the contour tone even though the VR rime is able to support a contour tone. The ranking of preserve tone above the constraint against
contour tones on VR produces this effect. The output form is thus the candidate (a) with a shortened vowel and a contour tone. The reader is referred to Zhang (2002) and Yip (2002, 2007) for further typological discussion of tonal typology and its analysis within OT.

5.2. PHONETIC SCALES II: LARYNGEAL NEUTRALIZATION

Steriade (1999) explores the distribution of laryngeal contrasts in obstruents, that is contrasts involving voicing, aspiration, or glottalization. She develops an account of the neutralization of laryngeal contrasts grounded in perceptibility factors. Laryngeal contrasts are more likely to be lost where the contrast is less likely to be clearly audible. For example, contrasts between voiced and voiceless obstruents are less likely preceding another obstruent than before a sonorant, since a following sonorant provides a far better backdrop for the salient realization of important right edge cues to an obstruent’s laryngeal features (e.g. voice onset time, burst amplitude, fundamental frequency, and first formant frequency) than a following obstruent. Steriade documents a perceptibility hierarchy that mirrors an implicational hierarchy in the distribution of laryngeal contrasts cross-linguistically: a laryngeal contrast is only preserved in a given context, if it is also preserved in other contexts in which the contrast is perceptually more viable. This hierarchy is sensitive both to the sound preceding and following the laryngeal contrast and whether the contrast occurs at the left or right edge of a word. Crucially, for most obstruents, the following sound is more important in determining the perceptual robustness of a contrast, since the more decisive cues to an obstruent’s identity reside at the release of an obstruent rather than during the transition from a preceding sound. The perceptibility scale is thus primarily a function of the right edge context in which the obstruent occurs. Steriade posits a series of constraints banning laryngeal contrasts in environment $x$, where $x$ varies according to the contexts in (8). Taking voicing as a representative laryngeal contrast, languages draw different cut-off points between permissible and impermissible contexts for voicing contrasts. Contrasts are allowed in all contexts to the left of the demarcation point.

(8) Implicational hierarchy of environments for laryngeal neutralization

<table>
<thead>
<tr>
<th>Voicing contrasts preferred</th>
<th>Voicing contrasts dispreferred</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nowhere</td>
<td>intersonorant</td>
</tr>
<tr>
<td>Lenakel</td>
<td>Totontepac</td>
</tr>
</tbody>
</table>

The laryngeal well-formedness constraints obey a universal hierarchy corresponding to the perceptibility scale in (8). A competing constraint,
preserve *[αvoice], requires faithfulness to underlying voicing contrasts. Steriade characterizes the faithfulness constraint in terms of [αF], where F stands for the relevant laryngeal feature, in this case [voice], and α indicates both positively and negatively specified voicing values. This notation captures the fact that the perceptibility of the contrast between voiced and voiceless sounds is at stake rather than the perception of just one or the other. The result of contrast loss is a stop that lacks its own voicing target gesture and instead receives its surface realization as a function of its environment: voiced when adjacent to voiced sounds, voiceless when adjacent to voiceless ones and in final position, where aerodynamic factors encourage devoicing (Westbury and Keating 1986).

By varying the ranking of the faithfulness constraint relative to the constraints banning contrasts, different laryngeal neutralization patterns corresponding to the hierarchy in (8) are produced, as shown in (9).

(9) Ranking of constraints governing voicing contrasts

\[
\begin{align*}
*\{\alpha\text{voice}\} / \{[-\text{son}], \#\} & \rightarrow \{[-\text{son}], \#\} & \text{Khasi} \\
*\{\alpha\text{voice}\} / [+\text{son}] & \rightarrow [-\text{son}] & \text{Arabic} \\
*\{\alpha\text{voice}\} / [+\text{son}] & \rightarrow \# & \text{Hungarian} \\
*\{\alpha\text{voice}\} / [-\text{son}] & \rightarrow [+\text{son}] & \text{Lithuanian} \\
*\{\alpha\text{voice}\} / [+\text{son}] & \rightarrow [+\text{son}] & \text{Totontepec Mixe} \\
*\{\alpha\text{voice}\} / [+\text{son}] & \rightarrow [\#] & \text{Lenakel}
\end{align*}
\]

5.3. PHONETIC SCALES III: SONORITY-SENSITIVE STRESS

As De Lacy (2004) and Gouskova (2004) show, it is also possible to derive implicational scales like those operative for contour tones and laryngeal neutralization by assuming constraints that conflate contiguous elements in a well-formedness hierarchy. In De Lacy’s (2004) analysis of sonority-sensitive stress, the preferential attraction of stress by certain vowel qualities over others in many languages is attributed to a series of constraints (phonetically motivated in terms of loudness; see Ladefoged 1975) banning the stressed syllable of a foot from containing relatively low sonority vowels. The sonority hierarchy from which the constraints are projected is as follows (see also Kenstowicz 1997): low vowels (most sonorous) >> mid vowels >> high vowels >> central vowels (least sonorous). One constraint, *[Hd\text{\_vowel}/\alpha]*, bans central vowels, the least sonorous vowel type, from being the head (i.e. stressed) syllable in a foot. Another constraint, *[Hd\text{\_vowel}/\alpha, i\text{\_u}]*, prohibits stress on both central and high vowels. Yet, another constraint, *[Hd\text{\_vowel}/\alpha, i\text{\_u}, e\text{\_o}]*, bans stress on central, high, and mid vowels. Only elements that are adjacent on the sonority scale may be conflated; thus,
no constraint groups together low vowels and high vowels to the exclusion of mid vowels. The ranking of these constraints, however, are not universally fixed in De Lacy’s analysis unlike in the analyses of tone and laryngeal neutralization discussed earlier. Rather, De Lacy’s sonority-sensitive metrical constraints are ranked on a language-specific basis relative to each other and relative to other metrical constraints, for example, constraints pulling stress toward the right or left edge in a word, constraints against stress lapses (i.e. consecutive unstressed syllables), etc. Depending on which of the sonority-sensitive stress constraints are highly ranked enough to be active in a language, different distinctions in vowel weight are captured. For example, if \( *[\text{Hd}^r_e/\text{o}] \) outranks other metrical constraints, only central vowels are resistant to stress, as in Yil (Martens and Tuominen 1977). If \( *[\text{Hd}^r_e/\text{o}, i^u] \) is ranked above other stress constraints, stress is repelled by both central and high vowels, as in Nganasan (Helimski 1998). Finally, if \( *[\text{Hd}^r_e/\text{o}, i^u, e^o] \) is the highest ranked of the metrical constraints, a system in which low vowels preferentially attract stress is produced, as in Kara (Schlie and Schlie 1993, de Lacy 1997). De Lacy (2004) shows that an analysis relying on freely re-rankable conflated constraints is able to account for certain types of data that universally fixed ranking scales cannot derive (see De Lacy’s analysis for details).

6. Computer-Generated Factorial Typology in OT

A productive area of recent research in phonological theory has concerned the computational modeling of OT. The literature on computational algorithms of OT and their learnability is relatively vast (see, for example, Tesar 1995, 2007; Samek-Lodovici and Prince 1999; Eisner 1997; Hammond 1997; Karttunen 1998; Tesar and Smolensky 2000; Riggle 2004). One of the major advances from a typological perspective has been the development of software that allows for rigorous evaluation of the typological coverage of OT analyses. For example, OTSoft (Hayes et al. 2000) takes a set of user entered constraints and generates the set of output patterns, the factorial typology, resulting from all possible ranking permutations of the constraints. Employing software for calculating the factorial typology generated by an analysis represents an important advance, since even an analysis with a relatively small set of constraints may outstrip the human ability to thoroughly assess the predictions of an analysis. An analysis employing even six constraints has 720 (6!) possible ranking permutations (but not necessarily 720 surface patterns, since it is possible for different rankings to overlap in the patterns they generate), while one with 10 constraints has 3,628,800 rankings! It is often only through computer aided factorial typology generation that undesirable predictions of an analysis become apparent.

The potential for errors in OT analyses derived without computer assistance is exemplified by Karttunen (2006), who shows that the analyses of Finnish stress in Elenbaas (1999), Elenbaas and Kager (1999), and
Kiparsky (2003) fail to derive the correct stress patterns for certain word shapes. Karttunen shows that the shortcoming of these analyses stems from a failure to consider the full range of potential surface candidates corresponding to certain input forms.

Gordon (2002) provides an example of computer-generated factorial typology in the area of metrical stress theory. He focuses on weight-insensitive stress systems in which the location of stress is predictable on phonological grounds without reference to the internal structure, or weight, of syllables. Some examples of weight-insensitive stress patterns include languages that place stress on the initial syllable of words (e.g. Chitimacha), those that stress the penultimate syllable (e.g. Nahuatl), those with second syllable stress (e.g. Koryak), and those with final stress (e.g. Atayal). Also, falling under the heading of weight-insensitive systems are languages with two stresses per word, for example, initial and penultimate (e.g. Lower Sorbian), and languages that stress alternating syllables counting from either the right or left edge of words, for example, odd-numbered syllables counting from the left (e.g. Maranungku) or even-numbered syllables counting from the right (e.g. Cavineña).

Gordon posits a set of 12 constraints designed to account for all known species of weight-insensitive stress. Permuting the constraints such that all of the 479,001,600 possible rankings were tested would have been an intractable task without the aid of a computer. Submitting the 12 constraints to the factorial typology component of OTSoft generates a total of 79 patterns, a small subset of the nearly 11 trillion logically possible stress systems for words containing between 1 and 8 syllables. All of the stress patterns found in an extensive cross-linguistic survey of weight-insensitive stress systems are generated by Gordon’s analysis. The account also generates a number of stress systems that are apparently unattested in languages of the world, although most of the non-existent systems contain elements that are found in other languages, but not in the exact combination observed in the unattested systems. For example, one generated pattern situates stress on both the second and the final syllable of the word, conflating two stress locations observed in isolation.

7. The Evaluation of Overgeneration in an OT Analysis

While the failure of an analysis to offer complete empirical coverage of the attested phonological patterns is clearly a shortcoming of that analysis, it is more difficult to evaluate the generation of unattested patterns that is virtually guaranteed to occur in any exhaustive treatment of a phenomenon. One possibility is that unattested patterns reflect accidental gaps due either to lacunae in our typological knowledge of a phenomenon or to a paucity of languages providing probative distributional data for the examined property.

An example of the former type of accidental gap is plausibly provided by secondary stress patterns. Gordon’s (2002) account of stress predicts a
number of systems with secondary stresses that appear to be unattested. However, many grammars report only on the location of primary stress, the perceptually most prominent type of stress. It is thus conceivable that further study of many of the languages described in those grammars will reveal the existence of secondary stress in addition to primary stress and that the patterns discovered in these languages will fill some of the gaps in the factorial typology generated by Gordon’s constraint set.

For many phenomena, discrepancies between attested patterns and predicted patterns might simply be due to the rarity of a phenomenon. For example, although Steriade (1999) fully fleshes out the typology generated by varying the ranking of faithfulness relative to a constraint banning voicing contrasts, her account also extends to neutralization of the contrast between voiceless stops and ejectives. We may thus assume a parallel set of constraints banning contrasts between voiceless stops and ejective stops, since this contrast also relies heavily on the stop release and thus the following context. As expected, typological inspection of ejectives suggests a hierarchy of sites in which ejective and voiceless stops are contrasted. Thus, Kabardian tolerates ejectives in a full range of environments including before obstruents, word finally, and before sonorants. Hupa, on the other hand, restricts ejectives to word final and presonorant position. Navajo only permits ejectives in presonorant position. Kabardian, Hupa, and Navajo thus instantiate three patterns in the continuum of perceptibility established through evaluation of voicing. I am not aware, however, of any languages that only tolerate ejectives between sonorants, the ejective analog to the Totontepec Mixe voicing pattern in (8). Nor am I aware of any languages that tolerate ejectives before another obstruent, but only if the ejective is preceded by a sonorant, a distribution parallel to that observed for voicing in Arabic. Furthermore, I am not aware of any languages in which the syllabic nature of a following sonorant consonant influences ejective contrasts, the ejective analog to a pattern found in Russian, whereby a syllabic sonorant consonant before another consonant optionally blocks voicing neutralization in a preceding obstruent while a non-syllabic sonorant in preconsonantal position fails to block neutralization in a preceding obstruent. (Steriade attributes the Russian pattern to the greater length of syllabic sonorants, which allows for a longer and thus more salient realization of transitional cues to voicing.)

Even if such patterns turn out to be unattested for ejectives, these gaps arguably could merely be an artifact of the independent rarity of languages with ejective stops. Only 16.4% of the languages in Maddieson’s (1984) survey of 317 languages possess ejectives compared to 66.9% that possess voiced stops. Given the relative paucity of ejectives cross-linguistically, it is not surprising that certain patterns predicted to occur by the Steriade analysis of laryngeal neutralization appear to be unattested. Overgeneration in the case of ejectives is thus less likely to reflect a weakness in the theory than overgeneration in the case of voiced stops. Steriade also develops a
A perceptually driven account of preaspirated stops, in which the preceding context plays the primary role in determining their distribution. However, fewer divisions along the perceptibility hierarchy appear to be exploited by languages for preaspiration than for voicing. This also is probably not surprising in light of the independent rarity of preaspiration in languages of the world.

8. Typology as a Diachronic Phenomenon

There is another possible interpretation of overgeneration in factorial typology. Myers (2002) suggests that certain patterns predicted to occur by an analysis may reflect systematic gaps attributed to phonetic biases against such patterns. Under this view, which is consistent with work by Ohala (1981), Hyman (2001), Blevins and Garrett (2004) and others, the phonetic basis for well-formedness constraints is diachronic and results in large part from misapprehensions of patterns intended by the speaker. According to Myers (2002), unattested patterns predicted to occur by an analysis may be absent, because they are unlikely to arise phonetically. For example, Myers discusses gaps in the typology of strategies for avoiding nasal plus voiceless obstruent clusters, which are phonetically dispreferred, because aerodynamic conditions favor voicing following nasals (Hayes 1999). Pater (1999) shows that several ways of avoiding nasal plus voiceless obstruent clusters are attested cross-linguistically, including voicing of postnasal obstruents, deletion of the nasal, deletion of the voiceless obstruent, and replacement of the nasal with a voiceless obstruent. Pater develops an OT analysis of these strategies in which a well-formedness constraint against nasal plus voiceless obstruent clusters drives the elimination of these clusters, while various faithfulness constraints determine which strategy for eliminating the ill-formed structures is adopted in a given language. In addition to predicting the attested strategies for resolving nasal plus voiceless obstruent clusters, Pater’s analysis also predicts a number of unattested strategies, including lenition of the obstruent to a sonorant, epenthesis of a vowel between members of the cluster, and metathesis of the consonants. Myers suggests that the attested patterns are attributed to historical changes resulting from phonetically plausible misapprehensions on the part of listeners. For example, the overlap of perseverative voicing from the nasal with the voiceless obstruent could easily lead to the obstruent being perceived as voiced. Conversely, coarticulatory overlap could be regressive in which case the nasal might become devoiced. Over time, the devoiced nasal might fail to be perceived as distinct from the following obstruent, in which case the result would either be deletion of the nasal or reinterpretation of the nasal as the first half of a voiceless geminate. Alternatively, the velum lowering gesture associated with the nasal could overlap rightward onto the following obstruent, making the obstruent less perceptible and thus likely to delete.
The unattested patterns, on the other hand, are less likely to arise as misapprehensions on the part of the listener. Although it is possible for an intrusive vowel to occur upon release of a stop, including a nasal stop, this vowel should occur before both voiceless and voiced obstruents, since it is not conditioned by voicing but rather by manner of articulation. Similarly, lenition of the voiceless obstruent is also unlikely to be sensitive to voicing of the preceding nasal. Myers argues that it is not incumbent upon an OT analysis to fail to generate the unattested patterns, since such patterns are precluded on phonetic grounds. Steriade’s (1999) analysis of laryngeal neutralization represents another more acute case of observed repair strategies only being a subset of those predicted by an analysis. Laryngeal neutralization almost universally resolves in favor of voiceless obstruents even though Steriade’s constraints on contrast could also be satisfied through other neutralization strategies, for example, through voicing. Steriade appeals to a combination of articulatory and perceptual factors to account for the bias in favor of voiceless obstruents as the output of laryngeal neutralization; she argues that the devoicing strategy is attributed to a dispreference for implementing an articulatory difficult voicing gesture in an environment where its perceptual salience would be compromised.

Another approach to these overgeneration problems is to assume that the role of phonetics in phonology is purely diachronic rather than synchronic, a view espoused, for example, by Ohala (1981), Blevins (1995), Hyman (2001), and Blevins and Garrett (2004). Hyman (2001), in fact, shows that not all constraints needed to offer synchronic accounts of phenomena are in fact phonetically natural. Rather, a series of historical events, each of which might itself be phonetically natural, might conspire to create a phonetically unnatural synchronic pattern, such as postnasal devoicing. The role of phonetics as a diachronic vs. synchronic factor in sound systems is an ongoing debate in phonological theory (see Blevins 2004; Kiparsky 2006; de Lacy and Kingston 2006, and papers in Hayes et al. 2004 for discussion).

9. Conclusions

Typological research has been an important focus within OT. The fundamental assumption that constraint ranking varies on a language-specific basis makes the prediction that languages will vary in their strategies for repairing ill-formed configurations banned by markedness constraints grounded in phonetic and functional factors. The range of variation predicted by an OT analysis (assessed ideally through computational implementations) can be compared with the range of variation observed in typological surveys in order to evaluate the efficacy of an analysis. Mismatches between predictions made by OT analyses and the observed typologies may stem from several different sources (both synchronic and diachronic), which are the subject of ongoing debate in the literature.
Short Biography

Matthew Gordon is a Professor in the Linguistics Department at University of California, Santa Barbara. His primary areas of research include phonetics, phonology, the interface between the two fields, and the documentation of endangered languages, particularly indigenous languages of the Americas. Much of his work deals with prosodic properties such as stress and intonation. His website is found at http://www.linguistics.ucsb.edu/faculty/gordon/index.html.

Note

* Correspondence address: Matt Gordon, Department of Linguistics, University of California, Santa Barbara, Santa Barbara, CA 93012, USA. Email: mgordon@linguistics.ucsb.edu.

Works Cited


