Constraining OT: Primitive Optimality Theory

Jason Eisner, U. Penn — MIT Ling-Lunch, 18 Sep 1997
jeisner@linc.cis.upenn.edu

1. A much-asked question.

What constraints does OT allow? i.e., What is the substance of the theory?

(1) Some clearly *bad constraints* (but what makes them bad?):
  a. PALINDROMIC: The candidate reads the same backwards as forwards.
  b. FTQUINT: Feet are quintary (5 syllables or moras).
  c. MEMBEROF(*a, aardvark, aardvarks, aardwolf, aardwolves, Aaron* ...): Candidate must be in the specified set of surface forms.
  d. MATCHESOUTPUTOFSPE: The output matches the result of applying Chomsky & Halle (1968) to the input.

(2) Some clearly *okay constraints* (but what makes them okay?):
  a. CLASH-ATR: Low vowels may not bear the ATR feature.
  b. ONSET: Every syllable must start with a consonant.

(3) Some questionable constraints, by the standards of derivational phonology:
  a. FTBIN: Feet are binary (2 syllables or moras).
  b. ALIGN-L(Foot, PrWd): The sum of all distances from left edges of feet to the left edge of the PrWd is minimized. (For consequences see (34).)
  c. Half the constraints that first-year phonology students make up.

Reasons to try to formalize OT, rather than allowing *ad hoc* English constraints:

(4) a. Results in an explicit, falsifiable theory of UG
  b. Simplifies that theory, exposing formal similarities among constraints
  c. Enables computational work (e.g., Eisner 1997b)
    (tools for linguists; algorithms for generation, parsing, acquisition; theorems on expressive power)
  d. Constrains linguistic description
  e. Aids descriptive work by providing well-motivated and well-formalized constraints and representations
    (many constraints given informally in the literature, including GA, do not specify how to count violations in all circumstances)

The formalization sketched in this talk is called **OTP**—OT with primitive constraints.

(5) Identifying such core constraints is at the center of the OT program:
  "The danger, therefore, lies in ... clinging to a conception of Universal Grammar as little more than a loose organizing framework for grammars. A much stronger stance, in close accord with the thrust of recent work, is available ... *Universal Grammar can supply the very substance from which grammars are built: a set of highly general constraints, which, through ranking, interact* to produce the elaborate particularity of individual languages." (Prince & Smolensky 1993, p. 198)
  (see also Smolensky 1995, Green 1994)

2. The search for core mechanisms

Suppose we had a set **Con** of core constraints for phonology—simple mechanisms that could be used to build up all the basic phonological phenomena. What would it look like?

Ask: What formal devices are regularly used by constraints in the literature?

(6) a. NASVOI (Itô, Mester, & Padgett 1996)
    "Every nasal segment must be linked to some voicing feature."
  b. ONSET (Prince & Smolensky 1993)
    ALIGN(σ, L, C, L) } (equivalent) (McCarthy & Prince 1993)
    "Every syllable must begin with (be left-aligned with) some consonant."
  c. *Common thread:* "Every ... some."
    $\forall\alpha$, $\exists\beta$ such that α and β stand in such-and-such local relationship.

If we allow α and β to be edges (as one option), we only need one kind of local relationship—*temporal coocurrence*:

(7) The primitive **implication** family.
  $\boxed{\alpha \rightarrow \beta}$ means: $\forall\alpha$, $\exists\beta$ such that α and β coincide temporally.

(8) Rewrite (6):
  a. $nas \rightarrow voi$: $\forall nas$, $\exists voi$ such that nas and voi coincide temporally.
  b. ${}_\sigma[\rightarrow {}_C[$: $\forall {}_\sigma[$, $\exists {}_C[$ such that ${}_\sigma[$ and ${}_C[$ coincide temporally.

Thus we can regard alignment as "edge licensing." (Or licensing is "feature alignment.") We can also mix references to edges and interiors:

(9) $F \rightarrow]_\mu$: Every foot must cross a mora boundary. (No degenerate feet.)
  (= MIN-2m: Green & Kenstowicz 1995)

Like GA, primitive implication is formal rather than substantive:
  ONSET: ${}_\sigma[\rightarrow {}_C[$, NOONSET: ${}_\sigma[\rightarrow {}_V[$, CODA: $]_\sigma \rightarrow]_C$, and NOCODA: $]_\sigma \rightarrow]_V$ are all equally easy to express using this family. So as in other theories, UG must still state that ONSET and NOCODA are strongly preferred by human grammars. (The dispreferred constraints may still be possible: e.g., Hammond 1995 proposes a NOONSET constraint for stressless syllables. See Green 1994 on metaconstraints.)

McCarthy & Prince (1993) have previously noted that alignment plays a unifying role, and have suggested that it's *the* core mechanism for all of phonology:

(10) a. "These examples only hint at the generality of the phenomenon to be explored

here, which extends to include all the various ways that constituents may be enjoined to share an edge in prosody and morphology. Data like these have been given widely disparate treatments in the literature ..." (p. 1)

b. "Taken together with $\overline{\mathrm{X}}$-like restrictions on immediate domination and interpreted within the appropriate theory of constraint satisfaction, *GA provides a mechanism for completely specifying a class of formal languages that*, when substantive parameters are set, *ought to be all-but-coextensive with possible human languages.*" (p. 2)

A second constraint family:
Above, we unified feature licensing and alignment.
The opposite of feature licensing is feature clash.
The opposite of alignment is disalignment, i.e., edge clash.

(11) a. *[low, ATR] (Cole & Kisseberth 1994)
"Low features are incompatible with ATR features."
b. NONFINALITY = *ALIGN($PrWd$, R, F, R) (e.g., Buckley 1995)
"Prosodic words may not be right-aligned with feet."

(12) The primitive **clash** family.
$\boxed{\alpha \perp \beta}$ means: $\forall\alpha$, $\nexists\beta$ such that α and β coincide temporally. [cf. (7)]
Equivalently: $\forall\alpha\ \forall\beta$, α and β are temporally disjoint.

(13) Rewrite (11):
a. *low* $\perp$ ATR: All *low* and ATR features are temporally disjoint.
b. $]_{PrWd} \perp]_F$: Each $]_{PrWd}$ does not coincide with (fall on) any $]_F$.

Again, this formulation suggests we can mix edges and interiors, and we can:

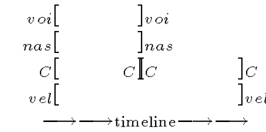
(14) $F \perp {}_M[$: A foot may not cross a morpheme boundary.
(= TAUTO-F, Crowhurst 1994)
(In fact, (14) is more plausible than Crowhurst's formulation, $*\,{}_F[\sigma\ {}_M[\ \sigma\ ]_F$.
It would be surprising to find a language that crucially blocked ${}_M[$ only where Crowhurst states, while still allowing it to interrupt a syllable or a ternary foot.)

Null hypothesis: These two families of local **primitive constraints**—implication and clash—are the *only* ones needed.

$\alpha \rightarrow \beta$ says that α's attract β's.
$\alpha \perp \beta$ says that α's repel β's.

3. What representations are being constrained?

The primitive constraints are easiest to interpret if we assume that *ŋk* is represented as in (15b), not (15a). This representation is inspired by Optimal Domains Theory (Cole & Kisseberth 1994) and Correspondence Theory (McCarthy & Prince 1995).

(15) a.
```
   voi
nas/
 |/
 C   C
  \ /
  vel
```
b.


(The timeline is really just an ordered set of edge brackets. Thus, only horizontal order matters in the drawing above. Ignore spacing and vertical order.)

(16) Key characteristics of the new representation:
a. Constituents float along a *timeline*.
Example constituents: *nas* (autosegmental), μ (prosodic), **x** (stress mark), *Stem* (morphological), *H*-domain (feature domain)
b. The timeline is continuous, not divided into segments.
c. *All* constituents have width and edges. Thus we can refer naturally to the edges of syllables (or morphemes) whose segmental features are scattered across multiple tiers and perhaps shared with other syllables (cf. Itô & Mester 1994).
d. For autosegments with width, such as [*nas*], think of phonetic gestures. (15b), which begins with simultaneous ${}_{nas}[$ (= lower the velum) and ${}_{voi}[$ (= begin vibration of the vocal folds). The primitive constraints can only affect the *order* of bracket edges; it is up to the phonetic component to determine actual durations.
e. Association or Correspondence of two constituents is indicated by having them overlap. (Independently proposed by Bird & Klein 1990.) E.g., the velar gesture in candidate (15b) spans both consonants.
f. No need for faithfulness constraints on the insertion, deletion, or relocation of association lines (cf. Kirchner 1993, Myers 1994, Féry 1994).
g. No need for (inviolable) well-formedness constraints against gapping or crossing of associations (cf. Kirchner 1993, Féry 1994, Oostendorp 1995).
h. No need for Correspondence indices.

(17) The behavior of **Gen**:
a. *Hypothesis:* **Gen** can't do anything fancy, like palindromes—nothing that the primitive constraints couldn't also handle. So for convenience, let's make Gen as simple as possible, and let undominated constraints clean up the huge unbridled candidate set that results.
b. Gen places constituents *freely* along the continuous timeline.
That is, as far as Gen is concerned, brackets may land anywhere.
Conditions such as the prosodic hierarchy are enforced by undominated primitive constraints, not by Gen.
c. However, Gen requires that edge brackets come in matched pairs.
d. Gen also does not allow distinct constituents of the same type (e.g., two syllables or two *lab* autosegments) to overlap. (Elements on the same tier never link to each other.)
e. Gen is free only with regard to output material. It is forced to place a copy

of the input material into every candidate, on its own tier, for purposes of I-O Correspondence. (Cf. Containment (Prince & Smolensky 1993), Strict Consistency Constraint (Polgardi 1995).)

(18) Because the timeline is continuous rather than divided into segments, brackets can fall in mid-segment:

a. Contour tones:

```
H[        ]H                 H
      L[        ]L         .' \   L
   V[     ]V              .    \ / .
                              vowel
```

b. Geminates (long vowels are similar):

```
σ[                 ]σ        ]σ          σ   σ
C[    ]C   C[          ]C               /|\ /|
   V[      ]V    V[      ]V             C V C V
```

4. Input and output

I-O Correspondence (between input and output features): Signaled by alignment between input and output tiers.

Correspondence relations with and without spreading:

(19) a. $_{voi}[\quad]_{voi}$ / $\underline{_{voi}[}\quad \underline{]_{voi}}$ — Perfect faithfulness

b. $\underline{_{voi}[}\quad \underline{]_{voi}}$ — Violates Max-IO (Parse): $\underline{voi} \rightarrow voi$

c. $_{voi}[\quad]_{voi}$ — Violates Dep-IO (Fill): $voi \rightarrow \underline{voi}$

d. $_{voi}[\qquad\qquad]_{voi}$ / $\underline{_{voi}[}\quad \underline{]_{voi}}$ — Like (a), this spread version satisfies Parse & Fill, which only require overlap. Spreading may be required to satisfy some other constraint. On the other hand, various constraints can be invoked against spreading: either $voi \perp \underline{]_{voi}}$ or $]_{voi} \rightarrow \underline{]_{voi}}$ or $\underline{]_{voi}} \rightarrow]_{voi}$ could be used to block (d). (Cf. Yip, 1994:21,fn. 11, on MSeg vs. *Insert Structure)

Thus, the timeline mechanism unifies Correspondence relations with autosegmental associations. Both are encoded by overlap on the constituent timeline. This fleshes out a proposal of McCarthy & Prince (1995):

(20) "The re-casting of autosegmental association in terms of correspondence relations may be expected to have consequences for the analysis of tonal, harmonic, and related phenomena. We do not explore these ideas here, though they are clearly worth developing." (p. 22)

In general, the lexicon and morphology might not completely specify the input tiers. In this case, candidates may differ even on their input tiers—so long as all candidates are consistent with what the lexicon and morphology do specify.

(21) a. *phonologically conditioned allomorphy:* candidates try different allomorphs on the input timeline, and the constraints decide what works best.

b. *floating tones and features:* the lexicon specifies only that H falls somewhere on the input. Different candidates try different locations for it in the input. The output may or may not correspond.

c. *floating morphemes, templatic morphology:* morphology specifies the order of underlying segments within each morpheme, but lets the morphemes overlap so that their segments intermix freely on the input tiers. These segments may not be preserved in the output.

d. *epenthesis* ($\underline{CC} \Rightarrow CVC$): The lexicon does not specify whether input segments are adjacent, so they may be pushed apart.

```
        V[           ]V
C[         ]C   C[      ]C
C[         ]C   C[      ]C   (underlined)
```

e. *syncope* ($\underline{CVC} \Rightarrow CC$): The lexicon does not specify whether input segments have positive width, so they may be crushed to zero width.

```
C[        ]C        ]C
C[        ]C        ]C   (underlined)
        V[ ]V            (underlined)
```

The crushing of $\underline{V}$, when there is no surface V, allows the C's to be adjacent. This is encouraged by $]_{\text{Segment}} \rightarrow {}_{\text{Segment}}[$ and expected by assimilation constraints.

(Only on the input tier may constituents have zero width.)

In short, the lexicon and morphology provide an **underspecified timeline**—an ordering over a set of *input* edge brackets. In general this is only a partial ordering. The candidate set consists of all possible fully specified versions of this input material, annotated in all possible ways with output constituents.

5. Formal definition of the constraints

(22) Formal statement of the primitive constraint families:

a. $\alpha \rightarrow \beta$: Each α temporally overlaps some β.
Scoring: Each α without a β incurs one violation mark.

b. $\alpha \perp \beta$: Each α temporally overlaps no β.
Scoring: Each overlap incurs one violation mark.

(23) What can α and β be?

a. *Edges* such as $_{low}[$ or $]_{low}$.
b. *Interiors* such as *low*.

Denote only the interior of a constituent, *without its edges.*

Thus, *low* and ATR do not overlap here: $\quad {}_{ATR}[\quad]_{ATR}$ / ${}_{low}[\quad]_{low}$

I.e., the above candidate satisfies *low* $\perp$ ATR but violates *low* $\rightarrow$ ATR.

c. *Conjunctions* and *disjunctions* as in (24).
(Dispreferred in analyses, on grounds of their greater complexity—they refer to more features.)

(24) Occasionally, must allow the following generalized forms of (22). I propose to limit conjunction/disjunction to these configurations only.

a. (α_1 and α_2 and ...) $\rightarrow$ (β_1 or β_2 or ...)
Scoring: Violated once by each set of objects $\{A_1, A_2, \ldots\}$ of types $\alpha_1, \alpha_2, \ldots$ respectively that all overlap on the timeline and whose intersection does not overlap any object of type $\beta_1, \beta_2, \ldots$.

b. (α_1 and α_2 and ...) $\perp$ (β_1 and β_2 and ...)
Scoring: Violated once by each set of objects $\{A_1, A_2, \ldots, B_1, B_2, \ldots\}$ of types $\alpha_1, \alpha_2, \ldots, \beta_1, \beta_2, \ldots$ respectively that all overlap on the timeline.
(Could also be notated: $\alpha_1 \perp \alpha_2 \perp \cdots \perp \beta_1 \perp \beta_2 \perp \cdots$.)

Each violation mark is still triggered individually by a bad *local* condition in the candidate, e.g., a moment on the timeline when certain edges are present and others are not.

Note that some constraints require crisp alignment of edges (${}_x[\rightarrow {}_y[$), while others are weaker and require only overlap ($x \rightarrow y$), allowing spreading. (Cf. the violable CRISPEDGE constraint of Itô & Mester (1994).)

6. Some further example constraints from the literature

This section illustrates how all the types of primitive constraints are ubiquitous across different areas of phonology.

My apologies in advance for any errors or mischaracterizations in these lists. Some of these translations to OTP are not exact, but appear to act correctly on the data in the papers cited. Also, note that sometimes there is more than one way to paraphrase a constraint.

("ROA" citations (http://ruccs.rutgers.edu/roa.html) not further listed in the bibliography.)

Key to unfamiliar notation:

feat	version of feature on output tier
<u>*feat*</u>	version of feature on input tier (underline denotes "underlyin' " material)
μ_s	strong mora, containing onset and nucleus (Zec 1988).
μ_w	weak mora, containing coda if any (Zec 1988). (One could also use explicit constituents Ons, Nuc, Coda.)
x	a 2ndary stress mark over a stress-bearing unit (first layer of the grid)
X	a word-primary stress mark (second layer of the grid)
Seg	segmental root node (alternatively, C or V), as distinguished from morphological root <u>*Root*</u>

Some implication constraints from the literature.

(25) **"Same edge" implication:**

a. **Features**

1. $]_{raised} \rightarrow]_{upper}$ — ALIGN[R][U]. Bradshaw ROA-93j.

b. **Prosody**

1. $]_{PrWd} \rightarrow]_\sigma$ — ALIGN: Wd] = σ]. Myers, ROA-6.
2. $]_F \rightarrow]_{\mu_w}$ — IAMBIC QUANTITY: In a rhythmic unit (W S), S is heavy. Hung, ROA-24.
3. $]_{PrWd} \rightarrow]_{\mu_w}$ — ALIGN-H: Align(*PrWd*, R, heavy syllable, R). Kager, ROA-70.
4. $\mathbf{x}[\rightarrow {}_F[$ — FOOT-FORM (trochaic): If there is a head, it is on the L. Hung, ROA-9. TROCHAIC: Align($\acute{\sigma}$, L, Foot, L). Kager, ROA-35.
5. ${}_F[\rightarrow \mathbf{x}[$ — ALIGN(Ft, L; Head(Ft), L). Bermudez-Otero, ROA-136.
6. $]_{PrWd} \rightarrow]_{\mathtt{X}}$ — FINAL-STR: Align(domain, R, $\acute{\sigma}$, R). Kager, ROA-35.
7. $]_F \rightarrow]_\sigma$ — FILL: Respect the usual prosodic hierarchy, without catalexis. Inkelas, ROA-39. (Take catalexis to be ${}_F[\ {}_\sigma[\cdots]_\sigma \cdots]_F$, and assume another constraint $]_F \perp \sigma$.)

c. **Feature-prosody interaction**

1. ${}_F[\rightarrow {}_C[$ — ALIGN(Ft, L, Onset): The left edge of a foot must always be aligned to the onset of the first syllable in the foot. Goedemans, ROA-26. (Assume we also have ${}_F[\rightarrow {}_\sigma[$.)
2. $\mathbf{x}[\rightarrow {}_V[$ — NOONSET: Stressless syllables do not have onsets. Hammond, ROA-58.
3. ${}_H[\rightarrow {}_{PrWd}[$ — ALIGN(H tone, L, PWd, L). Myers, ROA-6.
4. $]_{\mu_s} \rightarrow]_{son}$, et al. — HNUC: A higher sonority nucleus is more harmonic than one of lower sonority. Féry, ROA-34, following P&S 1993.
5. $]_{\mu_s} \rightarrow]_V$ — PROJECT($\overline{\overline{\mathrm{N}}}$, V): Nucleus must be a vowel. Oostendorp, ROA-84.
6. ${}_\sigma[\rightarrow {}_{A_0}[$ — STRONG ONSET: Syllables begin with a closure A_0. Bakovic, ROA-96.
7. ($]_\sigma$ and $]_{hi}$) $\rightarrow]_{back}$ — $*\ldots \mathrm{i}]_\sigma$. Kenstowicz, ROA-103.
8. ($]_{low}$ and $]_\sigma$) $\rightarrow]_{\mathtt{X}}$ — NO [a]: [a] is not allowed in unstressed open syllables. Kager, ROA-93a.
9. ($]_{hi}$ and $]_\sigma$) $\rightarrow$ ($]_{\mathtt{X}}$ or $]_{back}$) — NO [i]: [i] is not allowed in unstressed open syllables. Kager, ROA-93a.

d. **I-O relationships**

1. $_{H}[\rightarrow \underline{_{H}}[$ — LEFT-HD: The leftmost tone bearer of a tone span must be a head. Myers, ROA-6.
2. $]_{\underline{ATR}} \rightarrow]_{ATRdom}$ — BA-rt: Align(Anchor-s, R; [ATR]-domain, R). Cole & Kisseberth, ROA-22.

e. **Morphophonology**
1. $]_{Plural} \rightarrow]_{son}$ — SON]PL: Plurals end in a sonorant. Golston & Wiese, ROA-100.
2. $\underline{_{M}}[\rightarrow {_{F}}[$ — MORPHEME-FOOT-LEFT: Align(Morpheme, L, Foot, L), where "a *single* violation is assessed for every morpheme which does not meet this requirement." Crowhurst, ROA-19. See also Kager, ROA-35; Bermudez-Otero, ROA-136.
3. $\underline{_{Root}}[\rightarrow {_{PrWd}}[$ — ALIGN-WD: Align(root, Left; PrWd, Left). Cohn & McCarthy, ROA-25.
4. $\underline{_{Root}}[\rightarrow {_{\sigma}}[$, etc. — ALIGN(Root, σ; L,R): "Align root morpheme boundaries with syllable bondaries at both edges." Yip, ROA-14.
5. $_{Red}[\rightarrow {_{F}}[$, $]_{Red} \rightarrow]_{F}$ — Red = Foot. ROA-16. Carleton & Myers, ROA-16. (Also need $Red \perp {_{F}}[$.)

(26) **"Opposite edge" implication:**

a. **Features**
1. $]_{lax} \rightarrow {_{\mu_w}}[$ — PROJECT(lax, $\overline{\text{N}}$): Lax vowels are followed by additional weight (coda consonant or 2nd half of a diphthong). Oostendorp, ROA-84.
2. $_{\mu_w}[\rightarrow]_{lax}$ — PROJECT($\overline{\text{N}}$, lax): Only lax vowels are followed by additional weight (as if tense vowels bore their own). Oostendorp, ROA-84.
3. $( ]_{vel}$ and $_{C}[ ) \rightarrow ( ]_{cont}$ or $]_{voi})$ — No kC. Bradshaw, ROA-93j.

b. **Prosody**
1. $]_{\mathbf{x}} \rightarrow {_{\mu}}[$ — RHYTHM: A stressed element must be followed by an unstressed element. Hung, ROA-9. (Also need $]_{\mathbf{x}} \perp \mathbf{x}[$.)
2. $( ]_{\sigma}$ and $_{\sigma}[ ) \rightarrow ( ]_{\mathbf{x}}$ or $\mathbf{x}[ )$ — NOLAPSE: No adjacent unstressed syllables. Anttila, ROA-63.
3. $( ]_{\sigma}$ and $_{\sigma}[ ) \rightarrow ( ]_{\mathbf{x}}$ or $\mathbf{x}[$ or $]_{F}$ or $_{F}[ )$ — LAPSE: Adjacent unstressed syllables are separated by a foot boundary. Green, ROA-45.

c. **I-O relationships**
1. $_{H}[\rightarrow]_{\underline{H}}$ — LOCAL: An output TBU bearing tone t must be adjacent to [input] TBU b, where b [also] bears t. Bickmore (credited to Myers), ROA-161. (Only right spreading actually appears. Note the variation $_{H}[\rightarrow (\underline{_{H}}[$ or $]_{\underline{H}})$.)

d. **Morphophonology**
1. $_{Affix}[\rightarrow]_{PrWd}$ — ALIGN-SFX: Align(Affix, L, *PrWd*, R). McCarthy & Prince, ROA-7.

(27) **"Interior" implication:**

a. **Features**
1. $rd \rightarrow back$ — *Round* $\rightarrow$ *Back*. Cole & Kisseberth, ROA-98.
2. $nas \rightarrow voi$ — NASVOI. Itô, Mester, & Padgett, ROA-38; Yip, ROA-81.
3. $V \rightarrow ATRdom$ — WSA-lf: Align([ATR]-dom, L; Word, L). Cole & Kisseberth, ROA-22. (This gets the correct, gradient effect of spreading as far as possible.)
4. $nas \rightarrow Seg$, etc. — Features like *nas* surface only if linked to a (faithful or epenthetic) segmental root. Zoll, ROA-143.
5. $ATR \rightarrow ATRdom$ — Not explicitly mentioned in Cole & Kisseberth, ROA-22, but clearly needed there.
6. $\sigma \rightarrow (H$ or $L)$ — MAX-ET: Every TBU must have a correspondent tone. McCarthy & Prince (1995). SPEC(Tone): Every TBU has a tone. Zoll, ROA-143, after Prince & Smolensky (1993).
7. $V \rightarrow (front$ or $round$ or $low)$ — COLOR: A vowel is [front] or [round] if it is [-low]. Kirchner, ROA-4.
8. $C \rightarrow (cor$ or lab or $dors)$ — $C \rightarrow F_C$: A [+cons] root dominates a consonantal place feature. Oostendorp, ROA-84.
9. $(ATRdom$ and $V) \rightarrow ATR$ — EXPRESS: Express[ATR]. Cole & Kisseberth, ROA-22.

b. **Prosody**
1. $\mu \rightarrow \sigma$ — Parse μ: Every mora must be parsed into a syllable. Myers, ROA-6.
2. $\mu_w \rightarrow \mathbf{x}$ — WEIGHT-TO-STRESS: Heavy syllables are stressed. Hung, ROA-9 (following Prince 1990).
3. $Seg \rightarrow \sigma$ — PARSE(ROOT): Every root node must be associated with a syllable or mora.

c. **Feature-prosody interaction**
1. $\sigma \rightarrow H$ — FILL(σ): A syllable must be associated with a [high tone. Myers, ROA-6.
2. $V \rightarrow Nuc$ — $V \rightarrow \sigma$: A vowel must be a syllable head. Green, ROA-8.
3. $Nuc \rightarrow son$ — $\sigma \rightarrow R$: A syllable head must be at least a resonant. Green, ROA-8.

4.	$round \rightarrow (\, back$ or $stress\,)$	MaV(Pro) (Marked Vowel (Prominent)): Umlauted vowels fall in prominent syllables. Féry, ROA-34.
5.	$\mathtt{x} \rightarrow (\, lo$ or hi or $front$ or $back\,)$	Non-Head(ə): Stressed schwa is prohibited. Cohn & McCarthy, ROA-25.

d. **I-O relationships**

1.	$\underline{H} \rightarrow H$, etc.	Parse(T): A tone must be parsed. Myers, ROA-6.
2.	$\underline{lab} \rightarrow lab$, etc.	MaxPl: Parse underlying place features. Lombardi, ROA-105. Max, McCarthy & Prince 1995.
3.	$lab \rightarrow \underline{lab}$, etc.	*Ins(F): Do not insert features. Kirchner, ROA-4. Dep, McCarthy & Prince 1995.
4.	$\underline{\mu} \rightarrow \mu$	WeightIdent: If an input vowel is bimoraic, then so is the correspondent output vowel. Pater, ROA-107. See also WeightIdent, Alderete, ROA-131.
5.	$\underline{\mathtt{x}} \rightarrow \mathtt{x}$	StressIdent: Parse lexical stress. Pater, ROA-107. Head-Max: Alderete, ROA-131 (from McCarthy 1995).
6.	$(\, \underline{\mathtt{x}}$ and $Affix\,) \rightarrow \mathtt{x}$	Head-Max$_{\text{Affix}}$: Specializes Head-Max to affixes. Alderete, ROA-131.
7.	$(\, Seg$ and $\mathtt{x}\,) \rightarrow \underline{Seg}$	Head-Dep: Every segment contained in a prosodic head in S_2 [output] has a correspondent in S_1 [input]. Roberts-Kohno, ROA-93k.
8.	$(\, \underline{nas}$ and $\mathtt{x}\,) \rightarrow nas$, etc.	HeadSyll-Max(F): No features are deleted from (parsed?) segments in the head syllable. Yip, ROA-159.
9.	$(\, \underline{\mu}$ and $\mathtt{x}\,) \rightarrow \mu$, etc.	Head-Wt-Ident: No lengthening or shortening of stressed syllables. Alderete, ROA-131.
10.	$H \rightarrow (\, \underline{H}$ or $\underline{L}\,)$	TPFaith: Preserve tonal prominence profile. Tranel, ROA-72; Zoll, ROA-143.

e. **Morphophonology**

1.	$\underline{MWd} \rightarrow \mathtt{X}$	HeadProj: $_{MWd}[\ \ldots \text{Head(PWd)} \ldots\ ]_{MWd}$. A lexical head must project a prosodic head: every MWd constituent must include a stressed vowel. (A strengthened replacement for Lx≈Pr.) Kennedy, ROA-139.
2.	$\underline{M} \rightarrow PrWd$	MorPa: At least one element of a morpheme is incorporated into a prosodic word. Oostendorp, ROA-84.
3.	$\underline{Root} \rightarrow F$	Ft-Root: The root must overlap with a foot. Buckley, ROA-93c.

(28) **“Mixed” implication:**

a. **Features**

1.	$upper \rightarrow {}_{\mu}[$	Minimal Tone Association (MTA): [+upper] must be linked to more than one TBU. Bradshaw, ROA-93j.
2.	$(\,]_{A_0}$ and $_{A_f}[\,) \rightarrow pal$	NoAff: Disallows non-palatal affricates. Bakovic, ROA-96.
3.	$(\,]_C$ and $_C[\,) \rightarrow (\, cor$ or $dors \ldots)$	Contact: Coda should share place with the following Onset [if any]. Kenstowicz, ROA-30.
4.	$(\,]_{nas}$ and $_C[\,) \rightarrow voi$	*NC̥: No nasal – voiceless obstruent sequences. Pater, ROA-160.
5.	$(\, voi$ and $_C[\,) \rightarrow]_{nas} \gg \ldots$	$\ldots \gg$ No-NC-Link, Itô, Mester, & Padgett, ROA-38.

b. **Prosody**

1.	$F \rightarrow {}_{\mu}[$	Min-2m: A metrical foot contains at least two moras. Green & Kenstowicz, ROA-101.
2.	$PrWd \rightarrow {}_{Seg}[$	Disyll: The left and right edges of the PrWd, must coincide, respectively, with the left and right edges of *different* syllables. Kager, ROA-70. (Also need $_{PrWd}[\rightarrow {}_{Seg}[$, $]_{PrWd} \rightarrow]_{Seg}$.)
3.	$(\,]_{\sigma}$ and $_{\sigma}[\,) \rightarrow (\,]_F$ or $_F[$ or $F\,)$	Parse-2: One of two adjacent stress units should be parsed by a foot. Kager, ROA-35. Parse-Adj-Syll. Alderete, ROA-94.

c. **Feature-prosody interaction**

1.	$(\, _F[$ and $\underline{Root}\,) \rightarrow {}_C[$	FtOnset$^{\{rt\}}$: Align(Ft that is in root, L, C or Root, L). Buckley, ROA-56.
2.	$(\, V$ and $_{\mu_w}[\,) \rightarrow low$	Lower: Long vowels are low. $V_{\mu\mu} \rightarrow [Low]$. Cole & Kisseberth, ROA-98.

d. **I-O relationship**

1.	$(\, H$ and $_{\sigma}[\,) \rightarrow]_{\underline{H}}$	T-Bin: A tone span can have at most one non-head (in a domain); limits spread to one syllable from underlying tone. Myers, ROA-6.

Some clash constraints from the literature.

(29) **“Same edge” clash:**

a. **Prosody**

1. $]_{\mathbf{x}} \perp]_{PrWd}$ — *FINAL STRESS. Anttila, ROA-63. NON-FIN($\acute{\sigma}$). Cohn & McCarthy, ROA-25. Cf. RHYTHM, Hung, ROA-9.
2. $]_F \perp]_{PrWd}$. — NONFINALITY: Feet should not be word-final. Ní Chiosáin, ROA-89 (credited to Spaelti as WEAKEDGE(P-Cat)), et al.

b. **Feature-prosody interaction**

1. ${}_{\sigma}[\perp {}_{nas}[$ — *ONS/N. Smolensky, ROA-86 (following Prince & Smolensky 1993).
2. $]_{lax} \perp]_{\sigma}$ — PROJECT(lax, $\overline{\text{N}}$): Lax vowels are followed by additional weight (coda consonant or 2nd half of a diphthong). Oostendorp, ROA-84.
3. $]_{obs} \perp]_{\mu_w}$ — *OBSNUC. Pater, ROA-107.
4. $(\ ]_C$ and $]_{\sigma}\ ) \perp]_{lab}$ — CODACOND: Syllable-final consonant may not have place features. Lombardi, ROA-105.

c. **I-O relations**

1. ${}_{H}[\perp {}_{\underline{H}}[$ — *ALIGN(H,L)-I/O: High tone in output must not left-align with its position in input. Bickmore, ROA-161.
2. $(\ ]_{PrWd}$ and $]_{\mu_w}\ ) \perp]_V$ — FREE-V: *PrWd*-final vowels must not be parsed. So final heavy syllables are CVC, not CVV. Kager, ROA-70.

d. **Morphophonology**

1. $]_{\underline{M}} \perp]_{low}$ — *a]: No low vowel in a morpheme-final open syllable. Kager, ROA-93c.
2. ${}_{H}[\perp {}_{\underline{M}}[$ — *ALIGN(H, L, Source Morpheme, L) with no violation by distance. Bickmore, ROA-161.

(30) <u>**“Opposite edge” clash:**</u>

a. **Features**

1. $]_H \perp {}_H[$ — OCP: *FF, where F is a parsed [output] feature specification. “Furthermore, we will consider two tones to be adjacent if they are associated by parsed associations with adjacent tone bearers” (so domains are unnecessary). Myers, ROA-6.
2. $]_{son} \perp {}_{voi}[$ — *rg: No sonorant-voiced clusters. Ní Chiosáin, ROA-89.
3. $(\ ]_{nas}$ and ${}_C[\ ) \perp]_{voi}$ — *NC̥: No nasal – voiceless obstruent sequences. Pater, ROA-160.
4. $(\ ]_{vel}$ and $]_{cont}\ ) \perp {}_{lab}[$ — NO VELCONT LAB: No sequence of a velar continuant before a labial. Bradshaw, ROA-93j.
5. $(\ ]_{nas}$ and ${}_C[\ ) \perp voi$ — NO-NC-LINK. Itô, Mester, & Padgett, ROA-38.

b. **Prosody**

1. $]_{\mathbf{x}} \perp {}_{\mathbf{x}}[$ — *CLASH: No adjacent strong beats on the grid. Kager, ROA-35. NOCLASH. Anttila, ROA-63. Cf. RHYTHM, Hung, ROA-9.
2. $]_F \perp {}_F[$ — *FTFT: Feet must not be adjacent. Kager, ROA-35.

(31) <u>**“Interior” clash:**</u>

a. **Features**

1. $voi \perp gl$ — *[voiced, gl]: No implosives. Buckley, ROA-57.
2. $tense \perp low$ — *TENSE-low: No tense low vowels. Benua, ROA-74.
3. $phar \perp dor$ — *MID (no mid vowels): *[Phar, Dor]. Alderete, ROA-94.
4. $hi \perp low$ — Non-occurrence of +hi and +low. Kirchner, ROA-4.
5. $Seg \perp Word$ — *STRUCTURE(Root). Myers, ROA-6.
6. $H \perp Word$ — *Struct(A): There must be no association. Myers, ROA-6.
7. $low \perp Word$ — *[low]. Oostendorp, ROA-84 (following Prince & Smolensky 1993).

b. **Prosody**

1. $\sigma \perp PrWd$ — MONOSYLLABICITY: The fewer syllables, the better. Noske, ROA-109. *STRUC(σ): No syllables. Zoll, ROA-143.

c. **Feature-prosody interaction**

1. $\mu_w \perp (\ gl$ and $\ldots)$ — CODA-h: A /h/ may only occur in an onset. Oostendorp, ROA-84.

(32) <u>**“Mixed” clash:**</u>

a. **Features**

1. $hi \perp {}_{Seg}[$, $lo \perp {}_{Seg}[$ — *MULT-HEIGHT: No multiply linked height features. Kirchner, ROA-4.
2. $front \perp {}_{\underline{front}}[$, etc. — *SPREAD: Do not insert association lines.
3. $RdDom \perp {}_{HiDom}[$, etc. — UNIFORMITY: The (round-)harmony domain must be monotonic: high or low. Cole & Kisseberth, ROA-98. (Cf. parasitic harmony.)
4. $(\ ]_V$ and ${}_V[\ ) \perp hi$, etc. — NOLONGVOWEL: Two adjacent vocalic roots may not be linked to the same material (but diphthongs are allowed). Oostendorp, ROA-84.

b. **Prosody**

1. $F \perp {}_M[$ — Tautomorphemic-Foot: $*{}_F[\sigma\ {}_M[\sigma]_F$. Crowhurst, ROA-19.
2. $\mu_s \perp {}_{Seg}[$ — *Branch(S)μ. Walker, ROA-142.
3. $F \perp {}_\sigma[$, etc. — Unarity: A prosodic category p contains no more than one of the next lower prosodic category p-1. A. Green, ROA-115.
4. ${}_F[\perp \sigma$ — SyllInt: Syllable integrity (violable). Everett, ROA-163.
5. $\sigma \perp (\ ]_C$ and ${}_C[\ )$ — *Complex: Only one element can be in onset or coda position.

c. **Feature-prosody interaction**

1. $C \perp]_\sigma$ — Geminate: No geminate consonants. Oostendorp, ROA-84.
2. $\sigma \perp {}_H[$, etc. — *Complex(T): A tone-bearer must not be associated with more than one tone. Myers, ROA-6.
3. $\sigma \perp {}_C[$ — NoComplexOnsetOrRhyme. Noske, ROA-109.
4. $\mu \perp {}_C[$ — *Complex: No complex onset or coda. Kenstowicz, ROA-103.
5. $rime \perp {}_{nas}[$, etc. — Rhyme Harmony: All segments in the rhyme must share any nasal specification. Yip, ROA-81, ROA-135.

d. **Morphophonology**

1. $Red \perp {}_F[$, — Red = Foot. Carleton & Myers, ROA-16. (Also need ${}_{Red}[\rightarrow {}_F[$, $]_{Red} \rightarrow]_F$.)
2. $lab \perp {}_M[$ — Monolog: The edges of a morphological domain should be crisp; no feature should be linked both to an edge segment of that domain and to an element outside of the domain. Oosetndorp, ROA-84. (Also need $lab \perp]_M$.)
3. $(\ \mathbf{x}[$ and ${}_V[\ ) \perp$ <u>*Root*</u> — FtOnset$^{\{rt\}}$: Align(Ft that is in root, L, C or Root, L). Buckley, ROA-56.

7. How about measuring distance?

Two important differences between ${}_F[\rightarrow {}_{PrWd}[$ and Align(F, L, $PrWd$, L):

- The $\rightarrow$ family *doesn't measure distance.*
  E.g., $(\acute\sigma\sigma)(\acute\sigma\sigma)(\acute\sigma\sigma)\sigma$ violates ${}_F[\rightarrow {}_{PrWd}[$ twice, once for each non-initial foot.
- The $\rightarrow$ family *isn't only used for edges.*

Interestingly, Zoll (1996:137–38) has independently argued that licensing has just those properties (leading to her constraint Coincide(x,y)):

> "There are two properties of licensing which distinguish it from the cases of affixation discussed [in M&P (1993)].
>
> "First, licensing of marked structure never involves an injunction to be as close to a strong position as possible. Rather, licensing *always* constitutes an all-or-nothing proposition whereby marked structures are licit in licensed positions but ill-formed everywhere else."
>
> "The second important difference is that licensing does not strictly involve coincidence of edges or distance from an edge, but is concerned rather with membership in a constituent which may be peripheral ... [e.g.] heavy syllables belong to the first foot."

Q: Is this local version of alignment powerful enough?
A: Perhaps so. For cases where it's really necessary to measure distance, for example to control the width of a feature domain:

(33) a. $\sigma \rightarrow XDom$: X-domain should be as wide as possible (contain many σ's).
b. $\sigma \perp XDom$: X-domain should be as narrow as possible (contain few σ's).

Note that this trick, unlike GA, automatically specifies the units of measurement. It also avoids other definitional problems with GA.

Q: Is Generalized Alignment too powerful?
A: Probably. It's a family of non-local constraints that do addition. That lets us express very non-local, unattested phenomena.

Example of unwarranted power: The GA constraint in (34) wants the floating tone to anchor as close to the <u>center of the word</u> as possible (subject to higher-ranked constraints).

(34) Notes:

1. ´ denotes tone, not stress.
2. The n^{th} column records the degree of misalignment of the n^{th} syllable, at least if GA measures this in syllables rather than segments (or moras: see Mester & Padgett (1993)).
3. Assume that high-ranked faithfulness constraints rule out other candidates. For example, as there is only one floating tone underlyingly, σ́σ́σ́σ́σ́σ́σ́ is ruled out by Dep(H).

σσσσσσσ + [H]	Align(σ, R, H, R)							
a. σ́σσσσσσ	0	*	**	***	****	*****	******	= 21
b. σσ́σσσσσ	*	0	*	**	***	****	*****	= 16
c. σσσ́σσσσ	**	*	0	*	**	***	****	= 13
♡ d. σσσσ́σσσ	***	**	*	0	*	**	***	= 12
e. σσσσσ́σσ	****	***	**	*	0	*	**	= 13
f. σσσσσσ́σ	*****	****	***	**	*	0	*	= 16
g. σσσσσσσ́	******	*****	****	***	**	*	0	= 22

↑ violations contributed by 2nd syllable's misalignment; candidate's total viols

If there were *two* floating tones, they'd want to anchor at 1/4 and 3/4 of the way through the word.

This kind of non-local behavior via GA is unattested to my knowledge. It is also beyond the power of known computational OT methods, in particular the finite-state method of Ellison (1995) and the context-free method of Tesar (1996). The primitive constraints are provably incapable of producing such behavior.

8. How to handle non-local phenomena?

Since OTP uses only the primitive constraints of §5, it claims that *all phonology is* *local.*

Some apparently non-local phenomena can be reanalyzed:

- *Metrical stress.* Most non-local constraints in the literature concern metrical stress, which has received both local and non-local analyses in the past.
  - *Local:* Non-OT, iterative accounts (e.g., Prince 1983, Halle & Vergnaud 1987, Kager 1993, Hayes 1985, 1995).
  - *Non-local:* McCarthy & Prince (1993) propose using Generalized Alignment constraints to measure the distance from each foot to the edge of the word.
  - *Local:* Less powerful alternatives to GA are possible. Could use directional "greedy" versions of primitive constraints like PARSE(σ) or FILL(Root), in which early violations count as decisively worse than later ones. (Cf. Kager (1994), who argues for a greedy ALIGN evaluated "foot by foot.")
  - *Local:* Eisner (1997a) gives an OTP typology of metrical stress. This paper uses a small set of primitive constraints, which are freely reranked to get attested systems. This gives a unified fine-grained account of the following phenomena described by Hayes (1995).
    1. asymmetric foot shape typology
    2. iambic lengthening
    3. unbounded stress
    4. simple word-initial and word-final stress
    5. LR and RL footing, but no clear cases of RL iambs
    6. syllable and foot extrametricality
    7. no cases of final-syllable extrametricality for LR trochees
    8. strong and weak prohibitions on degenerate feet
    9. word-level stress, including prominence-based systems

    The asymmetries above are reduced to (i) the universal onset-coda asymmetry and (ii) the universal tendency of extrametricality to be final.

- *Intervocalic phenomena* (e.g., lenition). A constraint like *VsV (Green & Kenstowicz 1995) appears non-local, since [s] must look to *both* sides to decide whether it can surface as s or must become z. However, a local reanalysis is possible.

  Sample reanalysis: For *VsV, say that /s/ *always* wants to surface as [z], but only succeeds in the VsV context. For instance: (*cor* and *cont*) $\rightarrow$ *voi* rules out [s] in favor of [z]. It is outranked by $]_z \rightarrow ( ]_{\underline{voi}}$ or $_V[$), which says that any surface [z] not underlyingly voiced is followed by a vowel, and also by the mirror image of this, so that such a [z] must also be preceded by a vowel. Here $]_z$ abbreviates ($]_{cor}$ and $]_{cont}$ and $]_{voi}$).

However, *reduplication* occupies a special role in phonology, in that it is inherently non-local; it cannot be reanalyzed as local.

Therefore, to handle reduplication in OTP we need a representational trick (similar to Clements 1985). Translate the Correspondence account of McCarthy & Prince (1995) into OTP as follows:

a. As for all relations, OTP can enforce Correspondence only locally, so Correspondent elements must always overlap on the timeline,
b. Thus, I-B faithfulness requires I and B to occupy the same portion of the timeline. (on separate input and output tiers)
c. B-R faithfulness apparently requires R and B to occupy the same portion of the timeline. But this would rule out B-R juncture effects. which require B to precede R or vice-versa. (e.g., enforcement of *VhV in Javanese)
d. So instead require R (on the output tier) and a copy of B (on its own special tier) to occupy the same portion of the timeline.
e. Gen produces only candidates in which this copy of B is perfect. Thus, Gen must know how to do reduplication of morphemes, not just affixation.
f. Now all the non-locality is handled within Gen; the violable constraints remain local.

(35) Some candidates produced by Gen on RED(*bədah*)-e. In Javanese, first candidate wins.

a.	bədah-e	Input tier (used for I-B faithfulness)
	bəda bəda -e	Output tier: passed to phonetics (here violates MAX-IO)
	[Red][Base][Af]	Morphemic tier: mentioned by some constraints
	bəda ✓	Exact copy of base (used for B-R correspondence)
b.	bədah-e	
	bədah bədah-e	Satisfies MAX-IO, but violates surface constraint **VhV*
	[Red][Base][Af]	
	bədah ✓	Exact copy of this candidate's base (enforced by Gen)
c.	bədah-e	
	bədah bəda -e	Satisfies MAX-IO & **VhV*, but not DEP-BR, i.e., $C \rightarrow \underline{\underline{C}}$
	[Red][Base][Af]	
	bəda ✓	Exact copy of this candidate's base (enforced by Gen)
d.	bədah-e	
	bəda bədah-e	Satisfies MAX-IO & **VhV*, but not MAX-BR, i.e, $\underline{\underline{C}} \rightarrow C$
	[Red][Base][Af]	
	bədah ✓	Exact copy of this candidate's base (enforced by Gen)

In a language also requiring I-R faithfulness (McCarthy & Prince's (1995) Full Model), Gen must put two copies on the input tier: bədah bədah-e.

Haplology is a related example that may also be intrinsically non-local. (Yip 1995)

9. Constraints used for stress typology

See Eisner (1997a) for the OTP account of stress typology summarized in §8 above.

Undominated, prosodic-hierarchy constraints:

(36) a. FILL-F: ${}_F[\rightarrow {}_\sigma[\quad , \quad]_F \rightarrow]_\sigma$ (says where feet *can* appear)
"Each foot is strictly built from syllables: it starts and ends on syllable edges (perhaps the edges of different syllables)."

b. PARSE-σ: $\sigma \rightarrow F$ (says where feet *must* appear)
"Every syllable overlaps with (roughly, is 'linked to') some foot."

(37) a. PARSE-F: $F \rightarrow \mathbf{x}$ (says where stress *must* appear)
"Any foot bears stress somewhere (overlaps with at least one stress mark)."

b. FILL-**x**(trochaic): $\mathbf{x}[\rightarrow {}_F[\quad , \quad]\mathbf{x} \rightarrow]_\mu \quad , \quad \mathbf{x} \perp]_\sigma$ (says where stress *can* appear)
"Stress only appears at the start of a foot."
"Stress ends on a mora boundary, so extends over some integral number of moras."
"Stress may not extend across (overlap with) a syllable boundary."

The basic substantive constraints for secondary stress:

(38) SPREAD-**x**(trochaic): $]\mathbf{x} \perp {}_{\mu_w}[$
"Stress shouldn't end immediately before a weak mora (but may spread onto it)."

(39) ANTILAPSE(σ): $(\ ]_\sigma$ and ${}_\sigma[\ ) \rightarrow (\ ]\mathbf{x}$ or $\mathbf{x}[\ )$
"Every syllable boundary coincides with the edge of a stress mark. That is, adjacent syllables must contrast for stress."

(40) WEIGHTEDGE(iambic): $]_F \rightarrow]_{\mu_w}$ (alternatively, $]\mathbf{x} \rightarrow]_{\mu_w}$)
"The stressed (right) edge of a foot should be supported by syllable weight, i.e., by a weak mora."

(41) FILLWEIGHT: ${}_{\mu_w}[\rightarrow (\ \underline{s}[$ or $\underline{{}_{\mu_w}}[\ )$
"Don't lengthen: weak moras on the surface must correspond to underlying segments or weak moras."

(42) STRESSALL: $\sigma \rightarrow \mathbf{x}$ (alternatively, $]_\sigma \rightarrow]_F$ or ${}_\sigma[\rightarrow {}_F[)$
"Other things equal, have as many feet as possible (where feet and stresses are in 1-1 correspondence)."

(43) BRANCH(iambic): $\mathbf{x}[\perp {}_F[$ [compare the iambic version of (37)]
"Just as the right edge of an iambic foot insists on stress, the left edge absolutely rejects it. Hence stress may not consume the entire foot, but must alternate."

Extrametricality and primary stress require additional constraints, given in the paper.

10. Computational issues

Q: Gen produces infinitely many candidates. How do we find the best?
A: By using *intensional descriptions* of the infinite sets. For example, *son* $\rightarrow$ *voi* $\gg \mu_w \perp$ *voi* yields "Utterances in which obstruent codas are voiceless and sonorants are voiced."

If we stick to the primitive constraints, we can use finite-state automata as our intensional descriptions. E.g., the infinite set of candidates that survive constraints 1–5 can be described in finite space with an automaton. Then we use constraint 6 to narrow this set down further, etc.
(Strategy is due to Ellison (1994); Eisner (1997b) gives an efficient version.)

Analogy: In mathematics, we don't work directly with the infinite sum

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \frac{1}{4\cdot 5} + \ldots$$

because that would take forever. Instead we manipulate the notation $\sum_{i=1}^{\infty} \frac{1}{i\cdot(i+1)}$. This lets us draw interesting conclusions without processing the terms one by one:

$$\sum_{i=1}^{\infty} \frac{1}{i\cdot(i+1)} = \sum_{i=1}^{\infty} \frac{1}{i} - \frac{1}{i+1} = \sum_{i=1}^{\infty} \frac{1}{i} - \sum_{i=1}^{\infty} \frac{1}{i+1} = (1 + \sum_{i=2}^{\infty} \frac{1}{i}) - \sum_{i=2}^{\infty} \frac{1}{i} = 1$$

BUT: To find the optimal candidate is NP-hard on the size of the grammar (Eisner 1997b). So while the automaton algorithm above is usually efficient, *any* algorithm will be slow for a pathological grammar. This is unfortunate for learning theories that may blunder into such a grammar and try to test it.

In addition to the algorithm to find the optimal candidate, we can also characterize the expressive power of OTP:

(44) a. *Equal* in power to OTFS, in which Gen is a finite-state transducer and the constraints are arbitrary weighted FSAs. Any formal OTP grammar can be converted to a formal OTFS grammar, and vice-versa.
However, the two grammars may have very different constituent types and constraint granularity. OTP grammars are more fine-grained, so they make stronger predictions about the effect of reranking constraints.

b. *More* power than systems of ordered rewrite rules. The crucial example is due to Bob Frank, Giorgio Satta, and Paul Smolensky—a funny trick that OTP can do but finite-state transducers can't.

c. *Less* power than if Generalized Alignment were allowed. The crucial example is (34)—a funny trick that GA can do but OTP can't.

11. What role do the primitive constraints play in OT?

Three kinds of constraints:
- Primitive: the implication and clash families.
- Compound: Expressible as a monolithic block of primitive constraints in fixed order. (Kennedy (1996) uses blocks of Align constraints.)
- Complex: Any constraint not expressible in this restricted framework.

The balance among these remains to be seen. It is not yet clear what compound or complex constraints are actually needed (and which of the primitive constraints are *not* needed!).

We must also discover which of the formally possible primitive constraints are favored in real languages (on phonetic or other grounds), and what rankings are favored. OTP claims that languages use only local constraints; but it does not say *which* local constraints.

Meanwhile,

- Primitive constraints are "safe to use." They're simple, radically local, and ubiquitous.
- The restricted version of OT allowing *only* primitive constraints—called OTP—is easy to reason within and is computationally tractable.
- OTP is the simplest explanation that stands a chance. Let's refine it against the data, adding new core constraints only as we're forced to.
- If OTP is close to correct, it may be fruitful to reanalyze languages and typologies within OTP. (For concreteness, see Eisner (1997a) for a detailed reanalysis of stress typology that has some empirical benefits.)

References (exclusive of §6)

Bird, Steven, & Ewan Klein. 1990. Phonological events. *Journal of Linguistics* **26**, 33–56.

Buckley, Eugene. 1995. Alignment and constraint domains in Manam stress. Ms. ROA-56.

Clements, G. N. 1985. The problem of transfer in nonlinear morphology. Cornell Working Papers, Cornell University, Ithaca, NY, Fall.

Cole, Jennifer, & Charles Kisseberth. 1994. An optimal domains theory of harmony. *Studies in the Linguistic Sciences* **24**: 2.

Crowhurst, Megan 1994. Prosodic alignment and misalignment in Diyari, Dyirbal, and Gooniyandi: an optimizing approach. *WCCFL* **13**. ROA-19.

Eisner, Jason. 1997a. FootForm decomposed: Using primitive constraints in OT. SCIL 8. MIT Working Papers in Linguistics, Cambridge, MA.

Eisner, Jason. 1997b. Efficient generation in primitive Optimality Theory. Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and the 8th Conference of the European Association for Computational Linguistics, Madrid, July.

Ellison, T. Mark. 1994. Phonological derivation in optimality theory. Proceedings of **COLING**, 1007-1013.

Féry, Caroline. 1994. Umlaut and Inflection in German. Ms. ROA=34.

Green, Thomas and Michael Kenstowicz. 1995. The Lapse constraint. **FLSM** 6,1: 1-14. Bloomington, Ind: Indiana University Linguistics Club. ROA-101.

Green, Thomas. 1994. The conspiracy of completeness. Proceedings of Rutgers Optimality Workshop I. ROA-8.

Halle, Morris and Jean-Roger Vergnaud. 1978. Metrical structures in phonology. Ms., Dept. of Linguistics, MIT.

Hammond, Michael. 1995. Syllable parsing in English and French. Ms. ROA-58.

Hayes, Bruce. 1985. Iambic and trochaic rhythm in stress rules. *BLS* **11**, 429–446.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies.* University of Chicago Press.

Itô, Junko, & Armin Mester. 1994. Realignment. Proceedings of the June 1994 Utrecht Prosodic Morphology Workshop.

Kager, René. 1993. Alternatives to the Iambic-Trochaic Law. *LI* **11**, 381–432.

Kager, René. 1994. Ternary rhythm in alignment theory. Ms. ROA-35.

Kirchner, Robert. 1993. Turkish vowel harmony and disharmony: an optimality theoretic account. Proceedings of Rutgers Optimality Workshop I. ROA-4.

McCarthy, John, & Alan Prince. 1986. Prosodic morphology. Ms., Brandeis University.

McCarthy, John, & Alan Prince. 1993. Generalized alignment. *Yearbook of Morphology*, ed. Geert Booij & Jaap van Marle, pp. 79-153. Kluwer.

McCarthy, John and Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman et al., eds., Papers in Optimality Theory. UMass, Amherst: GLSA. 259–384.

Mester, Armin and Jaye Padgett. Directional syllabification in Generalized Alignment. Ms. ROA-1.

Myers, Scott. 1994. OCP effects in Optimality Theory. Ms. ROA-6.

Polgardi, Krisztina. 1995. Derived Environment Effects and Optimality Theory. Handout, Tilburg "Derivational Residue" Conference. ROA-93i.

Prince, Alan, & Paul Smolensky. 1993. *Optimality theory: constraint interaction in generative grammar.* Technical Reports of the Rutgers University Center for Cognitive Science.

Prince, Alan. 1983. Relating to the Grid. *LI* **14**, 19–100.

Selkirk, Elizabeth. 1980. Prosodic domains in phonology: Sanskrit revisited. In Mark Aranoff and Mary-Louise Kean, eds., *Juncture*, pp. 107–129. Anna Libri, Saratoga, CA.

Smolensky, Paul. 1995. On the structure of the constraint component Con of UG. Talk at UCLA, April 7. ROA-86.

Yip, Moira. 1994. Phonological constraints, optimality, and phonetic realization in Cantonese. Ms. ROA-14.

Yip, Moira. 1995. Identity Avoidance in Phonology and Morphology. Proceedings of the Conference on Morphology and its relation to Syntax and Phonology. University of California at Davis, May. ROA-82.

Zec, Draga. 1988. *Sonority constraints on prosodic structure.* Ph.D. dissertation, Stanford University.

Zoll, Cheryl. 1996. *Parsing Below the Segment in a Constraint Based Framework.* Ph.D. dissertation, University of California at Berkeley. ROA-143.