What Constraints Should OT Allow?

Jason Eisner, U. Penn — LSA in Chicago, 4 Jan 1997 [corrected 9 Jan]
jeisner@linc.cis.upenn.edu

1. Why bother asking?

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?

(1) 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) 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.

(3) 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.

(4) 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 (35).)
   c. Half the constraints that first-year phonology students make up.

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

(5) a. Results in an explicit, falsifiable theory of UG
   b. Constrains linguistic description
   c. Enables computational work
      (tools for linguists; algorithms for generation (Eisner, in preparation), parsing, acquisition; theorems on expressive power)
   d. Exposes formal similarities among constraints
      (e.g., locality properties: “x projects y” = “y needed to license x” ≈ “do not link x to −y segments”; “[voiced, gl]” = “no implosives”)
e. Clarifies predictions made by descriptive work
   (many constraints given informally in the literature, including GA, do not
   specify how to count violations in all circumstances)

f. Aids descriptive work by providing well-motivated and well-formalized
   constraints and representations

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

### 2. The search for core mechanisms

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

(6) a. **NasVoi**
   “Every nasal segment must be linked to some voicing feature.”
   (Itô, Mester, & Padgett 1996)

   b. **Onset**
   \[ \text{ALIGN}(\sigma, \text{L}, \text{C}, \text{L}) \] (equivalent)
   (Prince & Smolensky 1993) (McCarthy & Prince 1993)
   “Every syllable must begin with (be left-aligned with) some consonant.”

   c. **Common thread**: “Every … some.”
   \[
   \forall \alpha, \exists \beta \text{ such that } \alpha \text{ and } \beta \text{ stand in such-and-such local relationship.}
   \]

If we allow \( \alpha \) and \( \beta \) to be edges (as one option), we only need one kind of local
relationship—temporal cooccurrence:

(7) The primitive **implication** family.
\[
\alpha \rightarrow \beta \]
means: \( \forall \alpha, \exists \beta \text{ such that } \alpha \text{ and } \beta \text{ coincide temporally.} \)

(8) Rewrite (6):
   a. **nasVoi**: \( \forall \text{nas}, \exists \text{voi} \text{ such that } \text{nas} \text{ and } \text{voi} \text{ coincide temporally.} \)
   b. **σ[→C]**: \( \forall \sigma[, \exists C[ \text{ such that } \sigma[ \text{ and } C[ \text{ coincide temporally.} \]

   **Note that** **Onset**: **σ[→C]**, **NoOnset**: **σ[→V]**, **CODA**: **σ[→C]**, and **NoCODA**: **σ[→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.)

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

(9) **F →[β]**: Every foot must cross a mora boundary. (No degenerate feet.)
   \( = \text{Min-2m: Green & Kenstowicz 1995} \)

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 X-like restrictions on immediate domination and inter-
preted within the appropriate theory of constraint satisfaction, GA provides a
mechanism for completely specifying a class of formal languages that, when sub-
stantive parameters are set, ought to be all-but-co-extensive 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.

\[ \alpha \perp \beta \] means: \( \forall \alpha, \beta \) such that \( \alpha \) and \( \beta \) coincide temporally. \[ \text{[cf. (7)]} \]

Equivalently: \( \forall \alpha \forall \beta, \alpha \) and \( \beta \) are temporally disjoint.

(13) Rewrite (11):

a. low \( \perp \) atr: All low and atr features are temporally disjoint.

b. \( \lfloor \text{PrWd} \rfloor \perp \lfloor F \rfloor \): Each \( \lfloor \text{PrWd} \rfloor \) does not coincide with (fall on) any \( \lfloor F \rfloor \).

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

(14) \( F \perp \text{M} \[ \rfloor \): A foot may not cross a morpheme boundary.

(= Tauto-F, Crowhurst 1994)

(In fact, (14) is more plausible than Crowhurst’s formulation. *\[ F \sigma \text{M} \sigma \rfloor F .

It would be surprising to find a language that crucially blocked \( \text{M} \rfloor \) 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 \( \alpha \)’s attract \( \beta \)’s.

\[ \alpha \perp \beta \] says that \( \alpha \)’s repel \( \beta \)’s.

3. What representations are being constrained?

The primitive constraints are easiest to interpret if we assume that \( yk \) 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.  

\[
\begin{array}{ccc}
\text{voi} & \text{nas} / \text{vel} \\
\text{C} & \text{C} & \text{C}
\end{array}
\]

b.  

\[
\begin{array}{ccc}
\text{esi} & \text{esi} \\
\text{væs} & \text{væs} \\
\text{C} & \text{C} & \text{C} & \text{C}
\end{array}
\]

(16) Key characteristics of the new representation:

a. Constituents float along a timeline.

Example constituents: nas (auto-semantic), 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 auto-segments with width, such as [nas], think of phonetic gestures. (15b), which begins with simultaneous nas [lower the velum] and esi [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. 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. 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.

b. However, Gen requires that edge brackets come in matched pairs.

c. Gen also does not allow distinct constituents of the same type (e.g., two syllables or two lab auto-segments) to overlap. (Elements on the same tier never link to each other.)

d. 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).)
For computational purposes, regard each candidate timeline as a total ordering over a set of edge brackets.

The lexicon provides an **underspecified timeline**—an ordering over a set of *input* edge brackets. In general this is only a partial ordering, so input constituents may be floating with respect to each other (e.g., floating tones, templatic morphemes). The candidate set consists of all possible fully specified versions of this underspecified timeline.

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

a. Contour tones:

```
   u[         
L[       ]L \\
   v[         
```

b. Geminates (long vowels are similar):

```
   c[         ]c \\
   v[         ]v 
```

4. **Formal definition of the constraints**

(19) Formal statement of the primitive constraint families:

a. \( \alpha \rightarrow \beta \): Each \( \alpha \) temporally overlaps some \( \beta \).

   **Scoring**: Each \( \alpha \) without a \( \beta \) incurs one violation mark.

b. \( \alpha \perp \beta \): Each \( \alpha \) temporally overlaps no \( \beta \).

   **Scoring**: Each overlap incurs one violation mark.

(20) What can \( \alpha \) and \( \beta \) be?

a. *Edges* such as \( [tow] \) 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:

   \[
   [tow] \cap [atr] = \emptyset
   \]

   i.e., the above candidate satisfies \( low \perp atr \) but violates \( low \rightarrow atr \).

c. *Conjunctions and disjunctions* as in (21).

   (Dispreferred in analyses, on grounds of their greater complexity—they refer to more features.)

(21) Occasionally, must allow the following generalized forms of (19). I propose to limit conjunction/disjunction to these configurations only.

a. \( (\alpha_1 \text{ and } \alpha_2 \text{ and } \ldots) \rightarrow (\beta_1 \text{ or } \beta_2 \text{ or } \ldots) \)

   **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. \( (\alpha_1 \text{ and } \alpha_2 \text{ and } \ldots) \perp (\beta_1 \text{ and } \beta_2 \text{ and } \ldots) \)
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).)

5. 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 point to the Rutgers Optimality Archive at http://ruccs.rutgers.edu/roa.html; they are not further listed in the bibliography of this handout.)

Key to unfamiliar notation:

- \( \textit{feat} \) version of feature on output tier
- \( \textit{feat} \) version of feature on input tier (underline denotes “underlyin’” material)
- \( \mu_s \) strong mora, containing onset and nucleus (Zec 1988).
- \( \mu_w \) weak mora, containing coda if any (Zec 1988).
- \( x \) a 2ndary stress mark over a stress-bearing unit (first layer of the grid)
- \( \bar{x} \) a word-primary stress mark (second layer of the grid)
- \( \textit{Seg} \) segmental root node (alternatively, C or V), as distinguished from morphological root \( \textit{Root} \)

Some implication constraints from the literature.

(22) “Same edge” implication:

a. \textbf{Features}

1. \( \text{\textit{raised}} \rightarrow \text{\textit{upper}} \)  \( \text{ALIGN[R][U]} \). Bradshaw ROA-93j.

b. \textbf{Prosody}

1. \( \text{\textit{Pr-Wd}} \rightarrow \text{\textit{w}} \)  \( \text{ALIGN: Wd} = \text{\textit{w}} \). Myers, ROA-6.

2. \( \text{\textit{F}} \rightarrow \text{\textit{\mu_w}} \)  \( \text{IAMBIC QUANTITY: In a rhythmic unit (W S), S is heavy. Hung, ROA-24.} \)

3. \( \text{\textit{Pr-Wd}} \rightarrow \text{\textit{\mu_w}} \)  \( \text{ALIGN-H: Align(Pr-Wd, R, heavy syllable, R). Kager, ROA-70.} \)
4. $x \rightarrow \ell$ Foot-Form (trochaic): If there is a head, it is on the $L$. Hung, ROA-9. Trochaic: Align($\ell$, $L$, Foot, $L$). Kager, ROA-35.

5. $\ell \rightarrow x$ Align(Ft, $L$; Head(Ft), $L$). Bermudez-Otero, ROA-136.

6. $\mid_{p, W_d} \rightarrow \mid x$ Final-Str: Align(domain, $R$, $\ell$, $R$). Kager, ROA-35.

7. $\mid F \rightarrow \mid \sigma$ Fill: Respect the usual prosodic hierarchy, without catalexis. Inkelas, ROA-39. (Take catalexis to be $\ell[\sigma[\cdots]\sigma \cdots] F$, and assume another constraint $\mid F \perp \mid \sigma$.)

c. Feature-prosody interaction

1. $\ell \rightarrow \mid c$ 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 $\ell \rightarrow \mid \sigma$.)

2. $x \rightarrow \mid v$ NoOnset: Stressless syllables do not have onsets. Hammond, ROA-58.

3. $\mid p \rightarrow \mid_{p, W_d}$ Align(H tone, $L$, P Wd, $L$). Myers, ROA-6.

4. $\mid \rho, \rightarrow \mid_{\text{son}, \text{et al.}}$ HNuC: A higher sonority nucleus is more harmonic than one of lower sonority. Féry, ROA-34, following P&S 1993.

5. $\mid \rho, \rightarrow \mid v$ Project($\overline{N}, V$): Nucleus must be a vowel. Oostendorp, ROA-84.

6. $\mid \sigma \rightarrow \mid \alpha$ Strong Onset: Syllables begin with a closure $\alpha$. Bakovic, ROA-96.

7. $(\mid \sigma \text{ and } \mid_{hi}) \rightarrow \mid_{\text{back}} \ast \cdots \mid \sigma$. Kenstowicz, ROA-103.

8. $(\mid_{\text{low}} \text{ and } \mid \sigma) \rightarrow \mid x$ No [a]: [a] is not allowed in unstressed open syllables. Kager, ROA-93a.

9. $(\mid_{hi} \text{ and } \mid \sigma) \rightarrow (\mid x \text{ or } \mid_{\text{back}})$ No [i]: [i] is not allowed in unstressed open syllables. Kager, ROA-93a.

d. I-O relationships

1. $\mid r \rightarrow \mid_{\text{ATR}}$ Left-Hd: The leftmost tone bearer of a tone span must be a head. Myers, ROA-6.

2. $\mid_{\text{ATR}} \rightarrow \mid_{\text{ATR-dom}}$ BA-rt: Align(Anchor-s, $R$; [atri]-domain, $R$). Cole & Kisseberth, ROA-22.

e. Morphophonology

1. $\mid_{\text{plural}} \rightarrow \mid_{\text{son}}$ Son[Pl]: Plurals end in a sonorant. Golston & Wiese, ROA-100.

2. $\mid F \rightarrow \ell$ 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. \( \text{Align-WD: Align(root, Left; PrWd, Left). Cohn \\
& McCarthy, ROA-25.} \)

4. \( \text{Align(\text{Root, } \sigma; \text{ L,R): \text{“Align root morpheme boundaries with syllable boundaries at both edges.” } Yip, ROA-14.} \)

5. \( \text{Red = Foot. ROA-16. Carleton \& Myers, ROA-16. (Also need } \text{Red } \perp \text{ Pr[.]}} \)

(23) “\text{Opposite edge}” implication:

a. Features

1. \( ]_{\text{ lax}} \rightarrow \_[w[ \)

\( \text{Project(lax, } \overline{N}): \text{Lax vowels are followed by additional weight (coda consonant or 2nd half of a diphthong). Oosten-

dorp, ROA-84.} \)

2. \( \_[w[ \rightarrow ]_{\text{ lax}} \)

\( \text{Project(} \overline{N}, \text{ lax): Only lax vowels are followed by additional weight (as if tense vowels bore their own). Oosten-

dorp, ROA-84.} \)

3. \( (\_\text{rel} \text{ and } c[ ) \rightarrow (\_\text{ent} \text{ or } ]_{\text{ voi}}) \)

No kC. Bradshaw, ROA-93j.

b. Prosody

1. \( ]x[ \rightarrow \_[ \)

\( \text{Rhythm: A stressed element must be followed by an unstressed element. Hung, ROA-9. (Also need } ]x[ \perp x[.} \)

2. \( (\_\sigma \text{ and } \_\sigma[ ) \)

\( \text{NoLapse: No adjacent unstressed syllables. Anttila, ROA-63.} \)

3. \( (\_\sigma \text{ and } \_\sigma[ ) \)

\( \text{Lapse: Adjacent unstressed syllables are separated by a foot boundary. Green, } \)

\( \text{ROA-45.} \)

c. I-O relationships

1. \( u[ \rightarrow u \)

\( \text{Local: An output TBU bearing tone t must be adjacent to [input] TBU b, where } b[ \text{also} ] \text{ bears } t. \)

\( \text{Bickmore (credited to Myers), ROA-161. (Only right spreading actually appears. Note the variation } u[ \rightarrow (u[ \text{ or } u[.} \)

\( \text{)} \)

d. Morphophonology

1. \( \text{Align-SFX: Align(Affix, I, PrWd, R). McCarthy \& Prince, ROA-7.} \)

(24) “\text{Interior}” implication:

a. Features

1. \( rd \rightarrow \text{back} \)

\( \text{Round } \rightarrow \text{Back. Cole \& Kisseberth, ROA-98.} \)

2. \( \text{nas } \rightarrow \text{voi} \)

\( \text{NasVoi. Itó, Mester, \& Padgett, ROA-38; Yip, ROA-81.} \)
3. \( V \rightarrow ATR\dom \) WSA-I: Align([\(\text{atr}\)]-\(\text{dom}\), I; Word, L). Cole & Kisseberth, ROA-22. (This gets the correct, gradient effect of spreading as far as possible.)

4. \( nas \rightarrow Seg, \text{etc.} \) Features like \( nas \) surface only if linked to a (faithful or epenthetic) segmental root. Zoll, ROA-143.

5. \( ATR \rightarrow ATR\dom \) Not explicitly mentioned in Cole & Kisseberth, ROA-22, but clearly needed there.

6. \( \sigma \rightarrow (H \text{ or } L) \) Max-ET: Every TBU must have a corresponding tone. McCarthy & Prince (1995). Spec(Tone): Every TBU has a tone. Zoll, ROA-143, after Prince & Smolensky (1993).

7. \( V \rightarrow (\text{front or round or low}) \) Color: A vowel is [front] or [round] if it is [-low]. Kirchner, ROA-4.

8. \( C \rightarrow (\text{cor or lab or dors}) \) \( C \rightarrow F_C \): A [+cons] root dominates a consonantal place feature. Oostendorp, ROA-84.

9. \( (ATR\dom \text{ and } V) \rightarrow ATR \) Express: Express[ATR]. Cole & Kisseberth, ROA-22.

b. Prosody
1. \( \mu \rightarrow \sigma \) Parse \( \mu \): Every mora must be parsed into a syllable. Myers, ROA-6.

2. \( \mu_w \rightarrow x \) 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(\( \sigma \)): A syllable must be associated with a [high tone]. Myers, ROA-6.

2. \( V \rightarrow \text{Nuc} \) \( V \rightarrow \sigma \): A vowel must be a syllable head. Green, ROA-8.

3. \( \text{Nuc} \rightarrow \text{son} \) \( \sigma \rightarrow R \): A syllable head must be at least a resonant. Green, ROA-8.

4. \( \text{round} \rightarrow (\text{back or stress}) \) MaV(Pro) (Marked Vowel (Prominent)): Umlauted vowels fall in prominent syllables. Féry, ROA-84.

5. \( x \rightarrow ( \text{lo or hi or front or back}) \) Non-Head(\( \sigma \)): Stressed schwa is prohibited. Cohn & McCarthy, ROA-25.

d. I-O relationships
1. \( \hat{H} \rightarrow H, \text{etc.} \) Parse(T): A tone must be parsed. Myers, ROA-6.

3. \( \textit{lab} \rightarrow \textit{lab}, \text{etc.} \) *Ins(F): Do not insert features. Kirchner, ROA-4. Dep, McCarthy & Prince 1995.

4. \( \mu \rightarrow \mu \) WeightIdent: If an input vowel is bimoraic, then so is the correspondent output vowel. Pater, ROA-107. See also WeightIdent, Alderete, ROA-131.


6. \( (x \text{ and } \text{Affix}) \rightarrow x \) Head-MaxAffix: Specializes Head-Max to affixes. Alderete, ROA-131.

7. \( (\text{Seg} \text{ and } x) \rightarrow \text{Seg} \) 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. \( (\text{nas} \text{ and } x) \rightarrow \text{nas}, \text{etc.} \) HeadSyll-Max(F): No features are deleted from (parsed?) segments in the head syllable. Yip, ROA-159.

9. \( (\mu \text{ and } x) \rightarrow \mu, \text{etc.} \) Head-Wt-Ident: No lengthening or shortening of stressed syllables. Alderete, ROA-131.

10. \( H \rightarrow (H \text{ or } L) \) TPFaith: Preserve tonal prominence profile. Tranel, ROA-72; Zoll, ROA-143.

e. Morphophonology

1. \( MWd \rightarrow x \) 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 \( Lx \approx \text{Pr} \).) Kennedy, ROA-139.

2. \( M \rightarrow PrWd \) MorPax: At least one element of a morpheme is incorporated into a prosodic word. Oostendorp, ROA-84.

3. \( Root \rightarrow F \) Ft-Root: The root must overlap with a foot. Buckley, ROA-93c.

(25) “Mixed” implication:

a. Features

1. \( \text{upper} \rightarrow [\ ] \) Minimal Tone Association (MTA): [+upper] must be linked to more than one TBU. Bradshaw, ROA-93j.

2. \( (\text{a}_e \text{ and } \text{a}_i[ ] \rightarrow \text{pal} \) NoAff: Disallows non-palatal affricates. Bakovic, ROA-96.

3. \( (\text{c}_e \text{ and } \text{c}[ ] \rightarrow (\text{con or durs} \ldots ) \) Contact: Coda should share place with the following Onset [if any]. Kenstowicz, ROA-30.

4. \( (\text{nas}_e \text{ and } \text{c}[ ] \rightarrow \text{voi} \) *NC: No nasal – voiceless obstruent sequences. Pater, ROA-160.
5. \( \text{voi and } c[ \rightarrow ]_{nas} \Rightarrow \ldots \Rightarrow \text{No-NC-LINK, Itô, Mester,} \)
   \& Padgett, ROA-38.

b. **Prosody**
   1. \( F \rightarrow \rho[ \)
   2. \( \text{PrWd} \rightarrow s_{eg}[ \)
      **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 \( p_{rWd} \rightarrow s_{eg}[ \), \( ]_{p_{rWd} \rightarrow s_{eg} \).
   3. \( [\sigma] \text{ and } c[ \rightarrow (]_F \text{ or } F[ \text{ or } F ) \)
      **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. \( (F[ \text{ and } \text{Root} ] \rightarrow c[ \)
      **FT-ONSET**\(^{[71]} \): Align(Ft that is in root, L, C or Root, L). Buckley, ROA-56.
   2. \( (V \text{ and } \mu_{\mu}[ \rightarrow \text{low} \)
      **LOWER:** Long vowels are low. \( V_{\mu_{\mu}} \rightarrow [\text{Low}] \). Cole \& Kisseberth, ROA-98.

d. **I-O relationship**
   1. \( (H \text{ and } \sigma[ ] \rightarrow ]_{H} \)
      **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.

(26) **“Same edge” clash:**

a. **Prosody**
   1. \( [x \perp ]_{PrWd} \)
      *Final Stress. Anttila, ROA-63. Non-Fin(\( \sigma \)).
   2. \( [x \perp ]_{PrWd} \).
      **NonFinality:** Feet should not be word-final.
      Ní Chiosáin, ROA-89 (credited to Spelleti as **WeakEdge**(P-Cat)), et al.

b. **Feature-prosody interaction**
   1. \( \sigma[ \perp ]_{nas} \)
      *OSS/N. Smolensky, ROA-86 (following Prince \& Smolensky 1993).
   2. \( ]_{ta\sigma} \perp ]_{\sigma} \)
      **PROJECT(lax, N):** Lax vowels are followed by additional weight (coda consonant or 2nd half of a diphthong). Oostendorp, ROA-81.
   3. \( ]_{obs} \perp ]_{\mu_{\mu}} \)
   4. \( (]_{C} \text{ and } ]_{\sigma} \perp ]_{lab} \)
      **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. $(\mathbb{P}_{Wd} \text{ and } \mathbb{j}_{w}) \perp j$
   Free-V: $Pr Wd$-final vowels must not be parsed.
   So final heavy syllables are CVC, not CVV.
   Kager, ROA-70.

d. Morphophonology
1. $\underline{M} \perp \underline{low}$ \*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 vo-
   lation by distance. Bickmore, ROA-161.

(27) “Opposite edge” clash:

a. Features
1. $\underline{H} \perp \underline{H}$
   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 do-
   mains are unnecessary). Myers, ROA-6.
2. $\underline{son} \perp \underline{voi}$[\*rg: No sonorant-voiced clusters. Ñí
   Chiosán, ROA-89.
3. $(\underline{nas} \text{ and } c[\phantom{\underline{c}]}) \perp \underline{voi}$[\*NC: No nasal – voiceless obstru-
4. $(\underline{set} \text{ and } \underline{cont}) \perp \underline{lab}$[\*VelCont Lab: No sequence of a ve-
   lar continuant before a labial. Bradshaw,
   ROA-93j.
5. $(\underline{nas} \text{ and } c[\phantom{\underline{c}]}) \perp \underline{voi}$[\*NC-LINK. Itó, Mester, & Padgett,
   ROA-38.

b. Prosody
1. $\underline{x} \perp \underline{x}$[\*CLASH: No adjacent strong beats on the grid.
2. $\underline{F} \perp \underline{F}$[\*FtFt: Feet must not be adjacent. Kager, ROA-
   35.

(28) “Interior” clash:

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

b. Prosody
1. \( \sigma \perp \text{PrWd} \) Monosyllabicity: The fewer syllables, the better. Noske, ROA-109. \( \star \text{Struct(}\sigma\text{)} \): No syllables. Zoll, ROA-143.

c. Feature-prosody interaction
1. \( \mu_w \perp (\text{gl} \text{and} \ldots) \) Coda-h: A /h/ may only occur in an onset. Oostendorp, ROA-84.

(29) "Mixed" clash:
a. Features
1. \( \text{hi} \perp s_{seg}, \text{lo} \perp s_{seg} \) *Multi-Height: No multiply linked height features. Kirchner, ROA-4.
2. \( \text{front} \perp f_{r:ent}, \text{etc.} \) *Spread: Do not insert association lines.
3. \( \text{RdDom} \perp h_i:D_m, \text{etc.} \) Uniformity: The (round-)harmony domain must be monotonic: high or low. Cole & Kisseberth, ROA-98. (Cf. parasitic harmony.)
4. \( (\text{v} \text{and}\text{v}[\text{)] \perp \text{hi}, \text{etc.} \) 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: \( \star F[\sigma M[\sigma]F. \) Crowhurst, ROA-19.
2. \( \mu_s \perp s_{seg} \) *Branch(S)\mu. Walker, ROA-142.
3. \( F \perp \sigma, \text{etc.} \) 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. \( \sigma \perp (\text{c} \text{and} \text{c}[) \) *Complex: Only one element can be in onset or coda position.

c. Feature-prosody interaction
1. \( C \perp \) Geminate: No geminate consonants. Oostendorp, ROA-84.
2. \( \sigma \perp h_i, \text{etc.} \) *Complex(T): A tone-bearer must not be associated with more than one tone. Myers, ROA-6.
3. \( \sigma \perp c, \text{etc.} \) NoComplexOnsetOrRhyme. Noske, ROA-109.
4. \( \mu \perp c \) *Complex: No complex onset or coda. Kestnowercz, ROA-103.
5. \( \text{rime} \perp n_{ax}, \text{etc.} \) Rhyme Harmony: All segments in the rhyme must share any nasal specification. Yip, ROA-81, ROA-135.

d. Morphophonology
1. \( \text{Red} \perp f[], \text{Red} = \text{Foot} \). Carleton & Myers, ROA-16. (Also need \( \text{Red}[ \rightarrow f[] \rightarrow ]F. \) )
2. \( Iab \perp_M [ \)  

\textbf{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. Oostendorp, ROA-84. (Also need \( Iab \perp_M [ \) )

3. \( ( x[ \text{ and } y[ ) \perp Root \)  

\textbf{FtOnset, }\left( \tau \right): \text{Align(Ft that is in root, L, C or Root, L). Buckley, ROA-56.}

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

Note that when Gen constructs a candidate, it is constrained to place specified lexical material on the input tier, although the position of floating lexical material may be freely chosen.

Correspondence relations with and without spreading:

\begin{align*}
(30) & \quad \text{Perfect faithfulness} \\
(31) & \quad \text{Violates MAX-IO (\textsc{ParSe}): } vo_i \rightarrow vo_i \\
(32) & \quad \text{Violates DEP-IO (\textsc{FILL}): } vo_i \rightarrow vo_i \\
(33) & \quad \text{Like (a), this spread version satisfies \textsc{ParSe} & \textsc{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 } vo_i \perp [ vo_i \text{ or } [ vo_i \rightarrow [ vo_i \text{ or } [ vo_i \rightarrow vo_i \text{ could be used to block (d). (Cf. Yip, 1991:21, fin. 11, on MSEG vs. *Insert Structure).}}
\end{align*}

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):

\begin{quote}
(31) “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)
\end{quote}

Epenthesis and syncope rely on the fact that the input specifies only a weak partial order of edge brackets—input brackets are ordered with \( \leq \) at best. This makes (32) and (33) possible.

(32) Epenthesis (\( \text{CC } \Rightarrow \text{CVC} \)): the C's are pushed apart.
6. How about measuring distance?

Two important differences between \( r \rightarrow p_{prwd} \) and \( \text{ALIGN}(F, L, Prwd, L) \):

- The \( \rightarrow \) family doesn't measure distance.
  - E.g., \( (\sigma\sigma)(\sigma\sigma)(\sigma\sigma)\sigma \) violates \( r \rightarrow p_{prwd} \) 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 \text{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:

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

Note that this trick, unlike GA, automatically specifies the units of measurement.

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 (35) wants the floating tone to anchor as close to the center of the word as possible (subject to higher-ranked constraints).

(35) 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.

<table>
<thead>
<tr>
<th>$\sigma R H [H]$</th>
<th>ALIGN$\sigma, R, H, R$: each syll must align with the floating tone</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. $\sigma R H$</td>
<td>0 * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * = 21</td>
</tr>
<tr>
<td>b. $\sigma R H$</td>
<td>* 0 * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * = 16</td>
</tr>
<tr>
<td>c. $\sigma R H$</td>
<td>* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * = 13</td>
</tr>
<tr>
<td>$\triangledown$ d. $\sigma R H$</td>
<td>* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * = 12</td>
</tr>
<tr>
<td>e. $\sigma R H$</td>
<td>**** **** **** * * * * * * * * * * * * * * * * * * * * * * * = 13</td>
</tr>
<tr>
<td>f. $\sigma R H$</td>
<td>**** **** **** **** * * * * * * * * * * * * * * * * * * * * * * = 16</td>
</tr>
<tr>
<td>g. $\sigma R H$</td>
<td>**** **** **** **** **** * * * * * * * * * * * * * * * * * * * * = 22</td>
</tr>
</tbody>
</table>

| candidate’s total vio ls violations contributed by 2nd syll’s misalignment |

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.

7. How to handle non-local phenomena?

Since OTP uses only the primitive constraints of §4, 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.
  - **Non-local:** Less powerful alternatives to GA are possible. Could use directional “greedy” versions of primitive constraints like PARSE($\sigma$) 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 (in press) 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: 
  \[
  \text{ cor and cont } \rightarrow \text{ voi rules out [s] in favor of [z].}
  \]
  It is outranked by 
  \[
  \text{ [ei or v] },
  \]
  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 [:] abbreviates 
  \[
  \text{ [ei and voi] }.
  \]
  
  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.

(36) Some candidates produced by Gen on \(\text{RED} (\text{b dahe})\)-e. In Javanese, first candidate wins.
  a. \(\text{boda-e}\) Input tier (used for I-B faithfulness)
  b. \(\text{boda-e}\) Output tier: passed to phonetics (here violates Max-IO)
  c. \(\text{[Red][Base][Af]}\) Morphemic tier: mentioned by some constraints
  d. \(\text{boda}\) Exact copy of base (used for B-R correspondence)
b. Satisfies Max-IO, but violates surface constraint \( *VhV \)
\[
\text{[Red} \text{[[Base][Af]} \text{]} \text{bodah/} \text{]}
\]
Satisfies Max-IO, but violates surface constraint \( *VhV \)
\[
\text{Exact copy of this candidate’s base (enforced by Gen)}
\]

c. Satisfies Max-IO & \( *VhV \), but not Dep-BR, i.e., \( C \rightarrow C \)
\[
\text{[Red} \text{[[Base][Af]} \text{]} \text{boda/} \text{]}
\]
Satisfies Max-IO & \( *VhV \), but not Dep-BR, i.e., \( C \rightarrow C \)
\[
\text{Exact copy of this candidate’s base (enforced by Gen)}
\]

d. Satisfies Max-IO & \( *VhV \), but not Max-BR, i.e., \( C \rightarrow C \)
\[
\text{[Red} \text{[[Base][Af]} \text{]} \text{bodah/} \text{]}
\]
Satisfies Max-IO & \( *VhV \), but not Max-BR, i.e., \( C \rightarrow C \)
\[
\text{Exact copy of this candidate’s base (enforced by Gen)}
\]

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

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

8. 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 (in press) for a detailed reanalysis of stress typology that has some empirical benefits.)
References (exclusive of §5)


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


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


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

