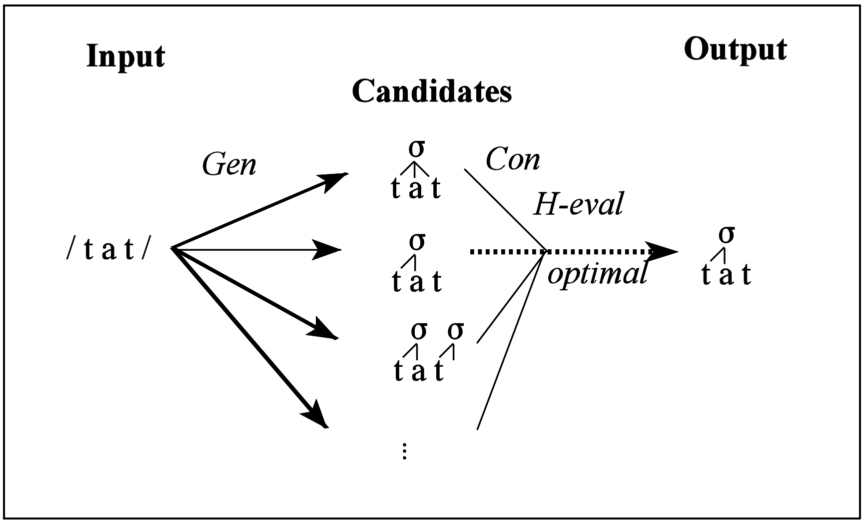
On the Internal Structure of the Constraint Component *Con* of UG

Paul Smolensky

Department of Cognitive Science; Johns Hopkins University

UCLA, April 7, 1995

I. The Architecture of Optimality Theory (Prince & Smolensky 91, 93, forthcoming)



N.B. This 'input' is *not* the input to a *comprehension* process.
'Input' ~ 'underlying form'
In syntax, the 'input' is typically interpretive (simplified d-structure/LF) information.
The phonetic string is part of the 'output', not the 'input.'

(1) **Universal Components of Grammar**:
  a. Input Set
  b. *Con*: Constraint Set
  c. *Gen:* Candidate Set for each input
  d. *H-eval:* Formal procedure for evaluating Harmony/optimality

(2) **Language-particular Component of Grammar** (*l'idéal*)
  Constraint Ranking — *c'est tout*

(3) **OT is a theory** of how one level/component of a structural description is projected from another: optimal satisfaction of ranked and violable constraints
  a. Massively parallel OT:
    input = 'the' input to the grammar, output = 'the' output of the whole grammar
  b. Massively serial OT ('harmonic serialism,' Prince & Smolensky 1993; harmonic phonology, Goldsmith 1990):
    input_k = representation at step k of the derivation
    output_k = representation at step $k+1$ of the derivation = input_{k+1}
  c. N-level OT:
    i Phonology
      input_1 = input to lexical phonology
      output_1 = output of lexical phonology =
      input_2 = input to post-lexical phonology
      output_2 =output of post-lexical phonology
    ii Syntax (ancient history)
      input_1 = d-struc
      output_1 = s-struc = input_P = input_L
      output_L = LF
      output_P = PF

(4) **Main Question** today: Is *Con* simply a set, or does it have internal structure?

(5) **Main Claim**: *Con* is richly structured, possessing at least the following three types of structure:

a. **Parametrized Families** (Part II) *many constraints are specific instantiations of a general schema*

b. **Harmony Scales by Constraint Subhierarchies**: (Part III)

i. **Prominence Alignment**: *an operation in meta-UG generating universal hierarchies of constraints from two scales of prominence (often, one structural and the other substantive)*

ii. **Universal Markedness Subhierarchies**: *hierarchies that establish the universal relative harmonies ('markedness') of elements/configurations*

iii. **Local Conjunction Subhierarchies** *are built from a general operation in UG:*

c. **Local Conjunction**: (Part IV) *an operation in UG by which two constraints governing substructures of a given local domain are conjoined into a higher-ranked constraint.*

Somewhat extended examples:

sonority profiles in syllables (Part V)

***wh*-extraction** (Part VI)

(6) **Syntax examples**, except where otherwise noted, are from:

Legendre, G., Wilson, C., Smolensky, P., Homer, K., & Raymond, W. (1995). Optimality in *wh*-chains. *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, J. Beckman, S. Urbanczyk, & L. Walsh, eds. Amherst, MA: GLSA, University of Massachusetts. In press.

and related on-going work of Legendre, Smolensky, and Wilson.

II. Parametrized Families (all constraints are present in all languages; languages do not 'set parameters')

(7) **Faithfulness constraints**:

Gen generates candidates which underparse ('deletion') and overparse ('epenthesis')

a. PARSE(X,Y) : X must be parsed into higher structure Y — X = input material *and* structural material

i PARSE(RT,σ); PARSE(Feat,RT); PARSE(μ,σ), PARSE(σ,F)

ii Operator chains: PARSE(wh), PARSE(top), PARSESCOPE

b. UNDERLYING(Y): Y must be present in input ~ FILL(X): X must be filled by underlying material

i Phonology: FILL(Ons); UNDERLYING(μ); UNDERLYING(Feat)

ii Syntax: UNDERLYING(DP): Overt DPs are input material (⇒ *resumptive pronoun)

OBHD: Head of XP must be filled (with a chain element). (Grimshaw 93)

(8) **Pure structural constraints**

a. Prosodic structure: ONS, NOCODA, FOOTBINARITY

b. Operator chains:

*ABSORB: No absorption of one Q operator by another.

BAR: A chain link must not cross a barrier.

(9) **Association constraints**: StructuralRole/ContentFiller

a. Phonology: *Nuc/*t*; *PL/Lab; $*\acute{\sigma}_\mu$; *[+Hi][–ATR]

b. Chains: GOV(t): t is [head] governed

(10) **Alignment constraints**: ALIGN($][_X$, $][_Y$): The right/left edge of every X must coincide with the right/left edge of some Y.

a. Prosodic phonology:

i ALIGN($]_{stem}$, $]_\sigma$): /maɽ/ → .mar.𝕋𝔸. (*.ma.ɽA. = {$[ma]_\sigma$ $[ɽ\}_{stem}$ $𝔸]_\sigma$) (Lardil, Prince & Smolensky 93 §7)

ii ALIGN($[_{suffix}$, $]_{PrWd}$); /na+piro+··· / → .na.TA.pi.ro.··· (Axininca Campa, McCarthy & Prince 93)

iii ALIGN($[_F$,$[_{PrWd}$): /σσσσσ/ → $[\sigma\sigma]_F$ $[\sigma\sigma]_F$ σ, *σ$[\sigma\sigma]_F$ $[\sigma\sigma]_F$, 'L→R' footing (Mester & Padgett 93)

b. Tonal/featural phonology: harmonizing feature F ('spreading' right): ALIGN($]_F$,$]_{PrWd}$)

c. Second-position phenomena (phonology, morphology, syntax?): NONINITIAL(X), ALIGN($[_X$,$[_Y$)

III. Harmony Scales (Prince & Smolensky 93)

(11) **Prominence Alignment**

a. A structural dimension of prominence Nuc > Ons is aligned with a phonetic dimension of prominence, sonority $a > y > l > n > t$

b. producing a non-binary universal constraint:

NUC-H: $\acute{a} \succ \acute{y} \succ \cdots \succ t$, i.e., Nuc/$a \succ$ Nuc/$y \succ \cdots \succ$ Nuc/t

which is realized through a *universal sub-hierarchy of binary constraints*:

Nuc Hierarchy: *Nuc/$t \gg \cdots \gg$ *Nuc/$y \gg$ *Nuc/a

c. Analogously, assume a universal subhierarchy for Onset

Ons Hierarchy: *Ons/$a \gg$ *Ons/$y \gg \cdots \gg$ *Ons/t

d. Each sub-hierarchy is universally ranked internally, but the two inter-digitate in language-particular ways.

e. Other constraints (crucially, PARSE, FILL, NOCODA, ONS) are inserted into these sub-hierarchies in language-particular positions.

f. The result is a typology of sonority-based inventories of possible Onsets, Nuclei, and Codas, about which can be proved:

g. **Thm: Onset/Coda Licensing Asymmetry** (Prince & Smolensky 93:§8; (258), p. 160)

There are languages in which some possible onsets are not possible codas, but no languages in which some possible codas are not possible onsets.

(12) **General Technique:** Typology by re-ranking, subject to universal domination conditions relating constraints on association (e.g., of Nuc to segments of varying sonority).

(13) **Sub-hierarchy Interruption and Constraint Encapsulation:**

a. May get many constraints, but analysis shows that ranking- (i.e., language-) specific groups of constraints can be encapsulated into single, equivalent, constraints; e.g.,

b. POSS-NUC(π_{Nuc}): Segments with sonority less than π_{Nuc} may not be parsed as nuclei.

c. This encapsulates the portion of the Peak Hierarchy:

[*Nuc/$t \gg \cdots \gg$ *Nuc/τ]

from the top down the point where it is **interrupted** by the lower-ranked of PARSE, FILL^{Nuc}.

d. This encapsulated constraint has a parameter (π_{Nuc}, the sonority level just above that of τ).

e. The parameter value is determined entirely by the language-particular ranking.

(14) **Universal markedness sub-hierarchies**

a. Coronal unmarkedness: PL/Lab $\prec$ PL/Cor universally; achieved by:

*PL/Lab $\gg$ *PL/Cor — UG

b. Entails the implicational universal over segmental inventories, Lab $\Rightarrow$ Cor.

c. Entails that Cor is 'invisible to phonological processes', relative to Lab, deriving many effects of radical underspecification theory, without the stipulation, 'if unmarked then unspecified'.

(15) **Sub-hierarchies from ...**

IV. Local Conjunction (Smolensky 93)

(16) **Local *vs.* non-local violations**

a. Indistinguishable to {*PL/Lab, NoCoda}: .tab̲.da. and .tad̲.b̲a.; both incur: {*PL/Lab, *NoCoda}

b. But there are languages with Labials, and Codas, but no Labials in Coda position; Codas frequently license only Cor or no place at all.

c. Idea: two constraint violations are worse when they occur in the same location: constraint interactions can be stronger locally than non-locally

d. **The Local Conjunction of C_1 and C_2 in domain *D***, $C_1 \&_l C_2$, is violated when there is some domain of type *D* in which both C_1 and C_2 are violated.

e. Universally, $C_1 \&_l C_2 \gg C_1, C_2$

f. Above case: *PL/Lab $\&_l$ NoCoda

g. **Self-conjunction**: when $C_1 = C_2 \equiv C$, $C_1 \&_l C_2 \equiv C^2$ is violated when there is some domain of type *D* in which both C is violated twice.

(17) **Banning the worst of the worst; the SHarC property**

a. The Coda licensing example illustrates a very general pattern, 'banning the worst-of-the-worst' (Prince & Smolensky 93 §9) in which C_1 or C_2 is violated in optimal forms, but not both.

b. Local Conjunction is a general (if brute force) solution to this problem.

(18) **Local self-conjunction: Power hierarchies**

a. If Local Self-Conjunction is performed recursively with constraint C, we get a universal sub-hierarchy:

$$\cdots \gg C^3 \gg C^2 \gg C$$

b. In the tableaux, with a C^3 violation we ignore the C^2 and C violations that necessarily come along with it. It can be proved that the ignored violations can never affect optimality.

(19) **Two examples**:

a. Sonority profiles in syllables

b. Economy in *wh*-extraction

V. The OCP and Optimal Sonority Contours of Syllables (Clements 90; Itô & Mester 94, Prince 84, 94)

(20) **Sonority from features**

a.

syllabic	–	–	–	–	+
vocoid	–	–	–	+	+
approximant	–	–	+	+	+
sonorant	–	+	+	+	+
	O[bstr]	N[asal]	L[iq]	G[lide]	V[owel]

b. Constraints will refer to: **V** ≡ + values for all these features; **C** = – values

c. In OT, feature geometry → feature classes targetted by constraints (Padgett 1994)

(21) **Non-derivational autosegmental phonology: Feature domains**

a. The span of feature value F along a tier defines an *F-domain* $\mathcal{D}_F$

b. A domain $\mathcal{D}_F$ has a unique *head* F^0.

c. Faithful parsing of underlying features requires parsing each feature F

i. as a head F^0: Parse(F, F^0)

ii. into a domain $\mathcal{D}_F$: Parse(F, $\mathcal{D}_F$)

d. Note that if PARSE(F, $\mathcal{D}_F$) is satisfied but PARSE(F, F^0) is violated, F is realized by spreading, so the mark *PARSE(F, F^0) achieves some of what other OT approaches achieve with the mark '*SPREAD'

(22) **Typical implicational universals**

a. Less sonorous onsets are less marked: LV ⇒ OV

b. In unmarked complex onsets, sonority rises monotonically: NOV ⇒ ONV

c. In complex onsets with monotonically rising sonority, less marked onsets have a steady sonority rise: ONV ⇒ OLV.

(23) **Clements' 'dispersion' function**

a. $D \equiv \sum 1/d_i^2$

b. Initial demi-syllables can be ranked with D, the lower D, the less marked.

c. Many of the Greenberg universals can be derived.

d. **Main Question:** can we develop a principled account of syllable structure that gives this result?

(24) **The Alignment theory of syllable structure**

a. ONS ~ ALIGN($[_\sigma$, $[_C$) (McCarthy & Prince 1993)
Here: ONS ≡ ALIGN($[_\sigma$, $[_C$): For each F ∈ **C**, ALIGN($[_\sigma$, $[_F$)

b. CODACOND ~ ALIGN($[_C$, $[\sigma$) (Itô & Mester 1994, Prince 1983, 1994)
Here: CODACOND ≡ ALIGN($[_C$, $[\sigma$): For each F ∈ **C**, ALIGN($[_F$, $[_\sigma$)

(25) **First Result: Onset Hierarchy derived from ALIGN($[_\sigma$, $[_C$) ≡ ONS**

a. OV ≻ NV ≻ LV

Partial Structures (input missing)	ALIGN($[_\sigma$, $[_C$) ≡ ONS
O V syllabic $[_-\ ]\ [_+\ ]$ vocoid $[_-\ ]\ [_+\ ]$ approximant $[_-\ ]\ [_+\ ]$ sonorant $[_-\ ]\ [_+\ ]$	
N V syllabic $[_-\ ]\ [_+\ ]$ vocoid $[_-\ ]\ [_+\ ]$ approximant $[_-\ ]\ [_+\ ]$ sonorant $[_+\quad\ ]$	 *
L V syllabic $[_-\ ]\ [_+\ ]$ vocoid $[_-\ ]\ [_+\ ]$ approximant $[_+\quad\ ]$ sonorant $[_+\quad\ ]$	 * *

b. The alternative structures with $[_+x][_+x]$ in place of $[_+\ x\ x\ ]$ violate the Obligatory Contour Principle (Leben 73, McCarthy 86):
OCP: $*]_F[_F$ = ¬ALIGN($]_F$,$[_F$)
which I here take to be universally undominated; they need not be considered.

c. Ways of stating the result:
   i. Lower-sonority segments make more harmonic onsets.
   ii. Onsets have low sonority.
   iii. Onsets are obstruents (graded).
   iv. Ons/*t* ≻ Ons/*n* ≻ ··· ≻ Ons/*a*

d. This gives desired Harmony scale, e.g., OV ≻ LV, mirroring markedness scale of Clements etc.

e. But can we derive inventories from this scale? Admit NV while excluding LV? Derive the implicational universal LV ⇒ NV?

Candidates	ALIGN($[_\sigma$, $[_C$) ≡ ONS	PARSE (F,$\mathcal{D}_F$)	Candidates	ALIGN($[_\sigma$, $[_C$) ≡ ONS	PARSE (F,$\mathcal{D}_F$)
/LV/ → . L V. syllabic [–] [+] vocoid [–] [+] approximant [+] sonorant [+] – + – + + + + + /L V/	* *		/NV/ → . N V. [–] [+] [–] [+] [–] [+] [+] – + – + – + + + /N V/	*	
. N V. syllabic [–] [+] vocoid [–] [+] approximant [–] [+] sonorant [+] – + – + + + + + /L V/	*	*	. O V. [–] [+] [–] [+] [–] [+] [–] [+] – + – + – + + + /N V/		*

f. If ONS ≫ PARSE, Onset sonority is unchecked and all onsets neutralize to OV; inventory is {OV}
   if PARSE ≫ ONS then Faithfulness is all-important and inventory is complete {OV, NV, LV ...}.

g. With ONS as a single constraint, can't have 'ONS forces underparsing for LV but does not for NV'

h. Must 'explode' ONS for the Harmony ranking NV ≻ LV to derive the right inventories.

i. To get inventories, Prince & Smolensky 93 had to achieve the Harmony scale Ons/*t* ≻ Ons/*n* ≻ ··· ≻ Ons/*a* not with one graded constraint but rather a subhierarchy of binary constraints:

   *Ons/*a* ≫ ··· ≫ *Ons/*n* ≫ *Ons/*t*

(26) **Power ONS**

a. Form the Local Conjunction Subhierarchy generated by ONS:

ONS-H: $\text{ONS}^4 \gg \text{ONS}^3 \gg \text{ONS}^2 \gg \text{ONS}$

⇔ *Ons/V ≫ *Ons/G ≫ *Ons/L ≫ *Ons/N (Ons/O is completely unmarked)

↑ admits NV but not LV into inventory

$\text{PARSE}(F,\mathcal{D}_F)$

Candidates	ONS^2 = *Ons/L	**PARSE** $(\mathbf{F},\mathcal{D}_\mathbf{F})$	ONS = *Ons/N	Candidates	ONS^2 = *Ons/L	**PARSE** $(\mathbf{F},\mathcal{D}_\mathbf{F})$	ONS = *Ons/N
/LV/ →				/NV/ →			
. L V. syllabic [–] [+] vocoid [–] [+] approximant [+] sonorant [+] – + – + + + + + /L V/	*!			☞ . N V. [–] [+] [–] [+] [–] [+] [+] – + – + – + + + /N V/			*
☞ . N V. syllabic [–] [+] vocoid [–] [+] approximant [–] [+] sonorant [+] – + – + ⟨+⟩ + + + /L V/		*	*	. O V. [–] [+] [–] [+] [–] [+] [–] [+] – + – + – + ⟨+⟩ + /N V/		*!	

b. Note import of *Locality* of conjunction:

.NV.NV. incurs marks {***O** ***O**} while

.LV. incurs mark {$*\mathbf{O}^2$}; .LV. ≺ .NV.NV. even though both violate ONS twice

(27) **The Monotonicity Effect**

a. 'Sonority should rise *monotonically* in the onset': *NOV

Candidates	ALIGN($[_C$, $[_\sigma$) ≡ CODACOND	PARSE (F,$\mathcal{D}_F$)	Candidates	ALIGN($[_C$, $[_\sigma$) ≡ CODACOND	PARSE (F,$\mathcal{D}_F$)
/NOV/ → .N O V. syllabic [–][+] vocoid [–][+] approximant [–][+] sonorant [+][–][+] – – + – – + – – + + – + /N O V/	*		/ONV/ → .O N V. [–][+] [–][+] [–][+] [–][+] – – + – – + – – + – + + /O N V/		
. N L V. syllabic [–][+] vocoid [–][+] approximant [–][+] sonorant [+] – – + – – + – ⟨–⟩ + + ⟨–⟩ + /N O V/		* *	.O L V. [–][+] [–][+] [–][+] [–][+] – – + – – + – ⟨–⟩ + – + + /O N V/		*

b. /NOV/ → .NOV. has a consonantal domain away from the left syllable edge; it violates ALIGN($[_C$, $[_\sigma$) ≡ CODACOND. NOV will be admitted into the inventory only if this constraint is dominated by PARSE (F,$\mathcal{D}_F$), in which case all onset clusters are admitted.

c. On the other hand, /ONV/ → .ONV. meets the alignment constraint so no ranking of PARSE can bar it from the inventory. (Although we next see other constraints which can block both ONV and NOV.)

d. We derive NOV ⇒ ONV, a consequence of NOV ≺ ONV.

(28) **Dispersion**

a. 'Sonority should rise *steadily* in the onset.'

Candidates	PARSE(F,F^0)	PARSE ($F,\mathcal{D}_F$)	Candidates	PARSE(F,F^0)	PARSE ($F,\mathcal{D}_F$)
/ONV/ →			/OLV/ →		
.O N V.			.O L V.		
syllabic [–][+]			[–][+]		
vocoid [–][+]			[–][+]		
approximant [–][+]			[–][+]		
sonorant [–][+]			[–][+]		
– – +	*		– – +	*	
– – +	*		– – +	*	
– – +	*		– + +	*	
– + +	*		– + +	*	
/O N V/			/O L V/		
. O L V.			.O G V.		
syllabic [–][+]			[–][+]		
vocoid [–][+]			[–][+]		
approximant [–][+]			[–][+]		
sonorant [–][+]			[–][+]		
– – +	*		– – +	*	
– – +	*		– ⟨–⟩ +	*	⟨*⟩
– ⟨–⟩ +	*	⟨*⟩	– + +	*	
– + +	*		– + +	*	
/O N V/			/O L V/		

b. As it stands, neither ONV nor OLV is more harmonic: both have the same marks.

c. But in reality ONV is more marked than OLV: *because the violations of OLV are less locally concentrated.* The *medial segment* of ONV is less harmonic than that of OLV.

d. This is really an OCP effect: informally, the OCP favors non-identity of adjacent elements. Here, with OCP ≡ ¬ALIGN($]_F$, $[_F$) universally undominated, underlying /– –/ projects [– x x] rather than [– x][– x], so the two underlying '–'s cannot both be domain heads: one must violate PARSE(F,F^0).

e. Formalization: Local Conjunction. Let **P** ≡ PARSE(F,F^0). Then Local Conjunction on **P** alone gives: DISPERSION: ··· ≫ $\mathbf{P}^3$ ≫ $\mathbf{P}^2$ ≫ **P**

Candidates	P³	PARSE (F,𝒟F)	P²	P	Candidates	P³	PARSE (F,𝒟F)	P²	P
/ONV/ →					/OLV/ →				
.O N V.					☞ .O L V.				
syllabic [–][+]					[–][+]				
vocoid [–][+]					[–][+]				
approximant [–][+]					[–][+]				
sonorant [–][+]					[–][+]				
– – +	* !				– – +			*	
– – +					– – +				
– – +					– + +			*	
– + +					– + +				
/O N V/				*	/O L V/				
☞ . O L V.					.O G V.				
syllabic [–][+]					[–][+]				
vocoid [–][+]					[–][+]				
approximant [–][+]					[–][+]				
sonorant [–][+]					[–][+]				
– – +			*		– – +				*
– – +					– ⟨–⟩ +		⟨*⟩		
– ⟨–⟩ +		⟨*⟩	*		– + +	* !			
– + +					– + +				
/O N V/					/O L V/				

f. Minimum sonority distance effect: the violations **P**, **P**^{2}, **P**^{3} correspond to sonority differences 3, 2, 1 so where PARSE (F,𝒟F) cuts the subhierarchy corresponds to a sonority *distance*.

(29) **Numerical implementation of strict domination and the dispersion function**

a. Why does Clements' dispersion function have the peculiar form $\sum 1/d_i^2$?

b. "It can be noted in passing that other ways of defining the value of D are possible in principle. ... Other possible versions, involving some simple summation of the distance between members instead of the inverse of the square, prove not to yield the desired complexity rankings..." (p. 305)

c. Numerical implementation of strict dominance: (here, *minimize* –Harmony ~ 'dispersion'):

exponential growth of constraint strength (violation penalty) as climb the hierarchy (Prince & Smolensky 93 §10)

	$\mathbf{P}^4$ ≫	$\mathbf{P}^3$ ≫	$\mathbf{P}^2$ ≫	**P**	∅
–Harmony penalty	1	½	¼	1/8	0
d_i	*0*	*1*	*2*	*3*	*4*
$1/d_i^2$	∞	1	¼	1/9	1/16
ONV		*		*	
0.625		½		1/8	
1.11		1		1/9	
OLV			* *		
0.5			¼+¼		
0.5			¼+¼		

d. ONV ≺ OLV: ONV has lower Harmony because it suffers the greater –Harmony penalty
e. Using simply d_i instead of $1/d_i^2$, ONV = 1 + 3, OLV = 2 + 2; no difference: must have:
weight($\mathbf{P}^3$) + weight(**P**) > weight($\mathbf{P}^2$) × 2
Guaranteed by all *strict domination* weightings.

VI. Economy in Operator Chains

(30) **Economy and the BAR constraint**
a. 'Shorter movements are better than longer ones' (cf. minimalist notions, Chomsky 91 et seq.)
b. Units of measurement for movement?
c. Barriers: Non L-marked (θ-governed) maximal projections. (Chomsky 86)
d. A movement must not cross a barrier.
e. In our OT analysis, no movement, just chains, composed of links.
f. **BAR: A chain link must not cross a barrier.**
g. This constraint will be violated in many grammatical structures.

(31) **Cyclicity and syllable structure**
a. BAR establishes the Harmony scale (β = a barrier):
$[X_i \cdots \beta \cdots \beta \cdots \beta \cdots Y_i] \prec [X_i \cdots \beta \cdots \beta \cdots Y_i] \prec [X_i \cdots \beta \cdots Y_i] \prec [X_i \cdots Y_i]$
b. But for all the same reasons as before, with syllable structure, to get the right inventory of extractions, this scale needs to be implemented with a Local Conjunction power hierarchy on BAR:
MINLINK: ⋯ ≫ BAR^3 ≫ BAR^2 ≫ BAR^1
BAR^k: There must not be k violations of BAR at one link = A link must not cross k barriers.
c. Economy effects are governed by where this subhierarchy is interrupted by other constraints.

d. The MINLINK hierarchy, but not BAR alone, entails that *ceteris paribus*, cyclic movement is preferred:

	BAR^3	BAR^2	BAR^1	BAR
a. Cyclic, 2 links ☞ [X_i ··· β ··· β ··· t_i ··· β ··· Y_i]		*	*	** *
b. Non-cyclic, 1 link [X_i ··· β ··· β ··· ··· β ··· Y_i]	* !			* * *

(of course, *ceteris* is not *paribus* when t_i is not head-governed; see below)

e. Note that the explanation pattern here is exactly the same as that used to explain why ONV ≺ OLV: ONV's 'chain' of sonority-feature-overlap 3–1 has a worse worst case than OLV's, 2–2.

(32) **An OT analysis of questions**

a. Representation of scope:
no movement, just chains consisting of an abstract operator Q in highest spec, binding a variable x in D-stucture position
[English: Q is realized as a wh-phrase and the variable is empty; Chinese: Q remains empty and the variable is realized as a wh-phrase (e.g. Aoun and Li, 93).]

b. Inputs: building blocks, i.e. syntactic categories, clausal boundaries, predicate-argument structure; scope information

c. *Gen* generates a universal set of candidate outputs (in accordance with X'-Theory); marks as overt Q or x; places Q in highest spec; generates candidates which fail to parse some feature of the input.

d. Universal input for questioning a direct object out of a simple clause:
[Q_j [...x_j...]]

e. Candidate set:

i. [Q_j [...wh_j...]] faithful parse (in situ; optimal in, e.g., Chinese)
ii. [wh_j [...t_j...]] faithful parse (fronted; optimal in, e.g., English)
iii. ⟨Q_j⟩ [...NP/⟨wh_j⟩...] unfaithful parse (failure to parse *as a question*; optimal in, e.g., Kwakwala: no object extraction from simple clauses)

f. And analogously for topicalization: operator 'TOp'

(33) **Constraints**

a. *t "No traces" (Grimshaw 93)
b. *Q "No empty Q[uestion]-operators"
c. *TOp "No empty T[opic]-operators"
d. *ABSORB "No absorption of Q-operators"
e. GOV: t must be head-governed by a category non-distinct from [+V].
f. PARSE(X) X = wh, top, scope

(34) **Economy and extraction patterns**

a. Topicalization of an object out of complement of *think* is good, but PARSE(top) ≫ BAR^2 (GOV)
b. topicalization of an object out of a sentential subjects is bad; but BAR^3 ≫ PARSE(top) (*GOV)
c. questioning an object out of a sentential subject is good. PARSE(wh) ≫ BAR^4 (*GOV)

Chinese	GOV	PARSE (wh)	BAR		PARSE (top)	BAR		*TOp	FILL	*t	*Q	*AB-SORB
			4	3		2	1					
Topicalization out of complement of *think*												
a. ☞ Z_j, ⟦ [V ⟦ e_j [[t_j						⊛$_j$	⊛$_j$			⊛ ⊛		
b. Z_j, ⟦ [V ⟦ [[res$_j$				*$_j$!					*			
c. ⟨top⟩					*!							
d. Z_j, ⟦ [V ⟦ e_j [[res$_j$						*$_j$	*$_j$		*!			
e. Z_j, ⟦ [V ⟦ [[t_j				*$_j$!						*		
f. Top$_j$, ⟦ [V ⟦ [[Z_j				*$_j$!				*				
g. TOp$_j$, ⟦ [V ⟦ e_j [[Z_j						*$_j$	*$_j$	*!				
Topicalization out of sentential subjects												
a'. Z_j, ⟦ [e_j [[t_j	*!					*	*			* *		
b'. Z_j, ⟦ [[[t_j				*!						*		
c'. ☞ ⟨top⟩					⊛							
d'. Top$_j$, ⟦ [[[Z_j				*!				*				
e'. TOp$_j$, ⟦ [e_j [[Zj	*!					*	*	*		*		
Questions out of sentential subjects												
a". Q_j [[e_j [[wh$_j$	*!					* *				*	*	
b". ☞ Q_j [[[[wh$_j$			⊛								⊛	
c". ⟨wh$_j$⟩		*!										
d". wh$_j$ [[e_j [[t_j	*!					* *				**		
e". wh$_j$ [[[[t_j			*							*!		

(35) **Referentiality**

a. REF: Links must be referential.

b. In Chinese: non-referential: manner *how*, reason *why* (Tsai, 94); in English, all adjuncts are non-referential. (In Bulgarian, referential = discourse-linked, e.g., *which book* rather than *what*.)

c. Local Conjunction of REF with the MINLINK hierarchy gives

$$\text{MINLINK}^{[-\text{ref}]}: \ldots \gg \text{BAR}^{[-\text{ref}]\,2} \gg \text{BAR}^{[-\text{ref}]\,1}$$

where, e.g, $\text{BAR}^{[-\text{ref}]\,2} \equiv$ BAR & BAR & REF = 'No non-referential links crossing two barriers'

d. The Local Conjunction operation entails that $\text{BAR}^{[-\text{ref}]\,n} \gg \text{BAR}^{n}$ for each n; in this sense the $\text{MINLINK}^{[-\text{ref}]}$ subhierarchy ranks higher than the MINLINK subhierarchy.

e. This means that non-referential extractions will be more marked than their referential counterparts.

f. When $\text{BAR}^{[-\text{ref}]\,2} \gg$ PARSE $\gg \text{BAR}^{2}$, can have good extraction with [+ref] but bad with [–ref] element:

Chinese	BAR[–ref]		PARSE	BAR[–ref]		PARSE	BAR		BAR	
	4	3	(wh)	2	1	SCOPE	4	3	2	1
Q out of *wh*-islands: referential adjunct										
a. ☞ Q_k [[V ⟦Q_i+e_k [wh_i [wh_k									⊛⊛$_k$	⊛$_i$
b. V ⟦ Q_k Q_i [wh_i [wh_k						*$_k$!				
c. ⟨wh_k⟩			*!							
Q out of *wh*-islands: non-referential adjunct										
a'. Q_l [[V ⟦ Q_i+e_l [wh_i [wh_l				**$_l$!						*$_i$
b'. ☞ V ⟦ Q_l Q_i [wh_i [wh_l				⊛$_l$		⊛$_l$				⊛$_i$
c'. ⟨wh_l⟩			*!							

VII. Summary

(36) ***Con* is richly structured**:

a. Parametrized families (PARSE, FILL, ALIGN)

b. Subhierarchies of constraints establishing a universal Harmony scale along some dimension (*Ons/x)

c. Local Conjunction, which itself generates universal subhierarchies:

i. ONS-H from ALIGN($[_\sigma$, $[_C$) ≡ ONS

ii. DISPERSION from PARSE(F, F^0)

iii. MINLINK from BAR

iv. $\text{MINLINK}^{[-\text{ref}]}$ from BAR and REF

Ackowledgement. Thanks to Junko Itô, Armin Mester, Alan Prince, and Draga Zec for very helpful conversations.