# OPTIMALITY

**Alan Prince**
Center for Complex Systems
Brandeis University

**Paul Smolensky**
Center for Opto-electronic Computing
University of Colorado

(1) **Optimization in Grammar Choice** (Chomsky, 1951,1965)

UG(D) => $G_1$,G2,..... All $G_i$'s agree on D.
Eval { $G^i$ } —> G* G* contains the real rules of L.

Chomsky's idea was to find a coding for the operations and structures of grammar such that the length (i.e. shortness) of the encoding of G was the measure of G's value.
—For phonology, we must determine Eval(Lexicon)+ Eval (Rules).
Otherwise, we can drive Eval(Rules) to optimum by no Rules.

(2) **Optimization as a grammatical mechanism.**
Simple Case:
Do-$\alpha$ (IN) —> $OUT_1$, $OUT_2$,.....
WF { $OUT_i$ } —> OUT* The real output.
More generally,
B( IN, Do-$\alpha$, OUT*) > B (IN,Do-$\alpha$, $OUT_j$) "Greatest Benefit."

(3) **The rap against optimization.**
a. Have to use numbers — parameter city; and use counting.
b. Many messy tradeoffs if t> k , a>i, what then of ti vs. ka?
c. Optimization is computationally intractable, in general.

(4) **Counter-rap.**
c'. It is not incumbent upon grammar to compute. Well-definition is all; not efficiency or even algorithmicity (Chomsky).
a'. Actually ORDER, not counting, is the key.
b'. We will argue that many cases of apparent tradeoff can be handled without tradeoff calculation, by a principle of combining orders.

(5) **Optimization in a grammatical mechanism**: the case of Berber Syllabification (Dell & El-Medlaoui 1987, 1988)

(6) ANY segment may be syllabic:

| | | | | | |
|---|---|---|---|---|---|
| **rA.tK.tI** | **yA.rB.bI** | **tF.tKt** | **A.tXk** | **txZ.nAkkw** | **tzMt** |
| **tM.zh** | **tR.gLt** | **il.di** | | **tR.bA** | |

(7) ONSETS: syllables must have onsets, except phrase-initially.

(8) SONORITY: syllabification is sensitive to relative sonority.
a. tzMt — *tZmt vs. tM.zh — *tmZh
b. rat.lult — *ra.tL.wLt ie. lul — *LwL

c. D&E establish the following 8-point hierarchy:

a > ui > rl > mn > z > fsx > bdg > ptk

low V > high V > liq > nasal > voiced fric > v'less fric > voiced stop > v'less stop

(9) **D&E's Algorithm** (DEA)

Form core syllables (X)Y ("CV"), where X is any segment and Y is successively replaced by the features describing the steps of the sonority scale; in descending order.
Iterate from Left to Right for each fixing of the nuclear variable Y.

(10) Secondary phenomena (to be ignored):
a. obstruents de-syllabified #_ and _#.
b. sonorant C's optionally desyllabify _#.
c. tautomorphemic geminates never begin an onset.

(11) EXAMPLES:

| | | | | | | |
|---|---|---|---|---|---|---|
| /ratlult/ | –> | {rA}tlult | –> | {rA}t{lU}lt | –> | {rA}t{lU}{lT} |
| /t-xzn-t/ | –> | tx{zN}t | –> | {tX}{zN}t | | 'thou stored' |
| /t-xzn-a-s/ | –> | txz{nA}t | –> | t{xZ}{nA}t | | 'she stored for him' |

(12) DEA serves to *optimize* the quality of syllables: high sonority items like to be nuclei.

(13) But DEA *per se* makes no contact with this. E.g. Why *descend* the scale?

(14) **Harmonic Syllabification** (HS).
Form the best core syllable from free material in the domain, good ($\sigma$) = good (nuc).
Iterate until you can do more, HA ($\Sigma$) = $\Sigma$ .

(15) **This is a fully 'Harmonic' process**: ('Harmony' from Smolensky, 1986.)
Go to the most well-formed available state.

(16) Note that HS is different from DEA. HS has no directional bias, does not sensibly run in a direction (LR) to deal with sonority plateaux.

(17) Evidence for directionality weak anyway.

| | | |
|---|---|---|
| a. /kiut/ | kiwt, *kyut | but better onset (DE note) |
| b. /rksx/ | R.kSx, *Rk.sX | but no final syllabic obstruent. |
| c. /bayn-n/ | ba.yNn, *bay.nN | but no geminate qua CV. |
| d. /ugmn/ | u.gMn, ug.mN | both ok. |
| e./itbdrin/ | *i.tBd.rin, only it.bD.rin | *DEA fails*. Nb. closed $\sigma$ bad. |

(18) Conclusion: optimization, with grades of well-formedness (representational Harmony), is a mechanism of grammar.

(19) To make this approach fly, we need two things:
a. Ways of evaluating combinations of phonological dimensions where the dimensions themselves needn't have an internal wf'dness structure, but may be scalar. Harmony scale on dim. comb. [post-Praguean contextual markedness: SPE, Kean,....,Archangeli & Pulleyblank.]

b. Ways of combining different relevant dimensional scales.

(20) **Dimensional Combination**. Example: weight and stress. (Prince, 1990)
a. Separate prominence (not markedness) scales: H>L, stress > unstress.
b. Combine implicatively: if H, then stressed. (WSP = =Hayes's 'quantity sensitivity').
Good are: H', L^. Bad is: H^. [unstressed heavy]
c. Weaker implication: if stressed, then Heavy.
Good are H', L^. Bad is: L'. [stressed light]
d. Combination respecting strength of implications:
H',L^ > L' > H^ . This is a Harmonic ordering.

(21) **Combination of Harmony Scales**. Example: 'unbounded' stress systems like: Stress the rightmost heavy syllable, if no heavy then stress the rightmost syllable (R/R). Prince (1983) and Hayes (1991) see these as enhancements of prominence rather than 'unbounded foot' phenomena. We extend this result, thereby limiting feet to the authentic binary rhythmic units.

(22) **The pattern R/R**. Two factors:
a. WEIGHT. Heavy syllables should be stressed.
b. POSITION. Edgemostness: Stress should be near the end.

(23) WEIGHT dominates POSITION. WEIGHT >> POSN. (Nb. One stress.)

(24) WEIGHT must be satisfied, if possible. But there can be several choices:
a. cv.CVV.cv.cv.cvv.cv.cv.
b. cv.cvv.cv.cv.CVV.cv.cv.

(25) Indeterminacies are resolved then by POSN. (24b) is chosen.

(26) No indeterminacy in (cv)* = ...cv.cv.CV.

(27) This is **LEXICOGRAPHIC ORDER** — a generalization of alphabetic order. Alphabetically, strings are ordered first by initial character (==WEIGHT), then ties are broken by the ordering on the second character (==POSN), etc. ptrs >qaaaaaaa. NB:NO TRADEOFF

(28) **LEXICOGRAPHIC PRODUCT** of two orders,
P *>> Q,
total order on the the pairs (p,q) constructed from the orders on P and on Q; first ordering by P-dimension, then by Q-dimension.

(29) Example: alphabetic ordering itself: POSN *>> ALPHA.
(1,2,3,4,....) *>> (a,b,c,...)
1a >1b > 1c >....2a > 2b >2c....

(30) **Claim**: Lexicographic Order is the essential mode of combining Harmony scales.

(31) Berber Syllabification is also a kind of Lexicographic Order: analyses in which /m/'s are nuclei are absolutely preferred to competitors in which /m/'s are not but /z/'s are.
We can say:
x's should be peaks >>y's should be peaks, for x > y in sonority.
With suitable technical development, this might replace the serial (iterative) HS.

(32) **Prominence-driven Stress**. Claim: works by Lexicographic ordering
**PROMINENCE >>POSN**, with development of notions of prominence and position.

(33) **Hindi** (modern) (Kelkar, 1968; Hayes, 1991)
"Stress falls on the heaviest available syllable, and in the event of a tie, the rightmost nonfinal candidate wins." (Hayes)

(34) Heaviness: CvvC, CvCC ($\mu\mu\mu$) > Cvv, CvC ($\mu\mu$) > Cv ($\mu$)

(35) Heaviest is best:
kiDHAR jaNAAB as.BAAB ru.pi.AA REEZ.gaa.rii

(36) Among equals, last nonfinal is best:
[$\mu$]⊢ sa.MI.ti
[$\mu\mu$]⊢ ru.KAA.yaa PUS.ta.kee roo.ZAA.naa
[$\mu\mu\mu$]⊢ AAS.maa.jaah aas.MAAN.jaah

(37) Analyis: WEIGHT >>POSN
But POSN is constructed from NONFINALITY, and RTMOSTNESS.

(38) NONFINALITY >>RTMOSTNESS

(39) WEIGHT >>NONFINALITY >>RTMOSTNESS.

(40) **Piraha.** (Everett & Everett,84; Everett.,88; Levin, 85; Davis, 88; Hammond, 90; Hayes, 91).

"Stress the rightmost token of the heaviest syllable type in the last 3 syllables of the word."
(41) Parallels: 'rightmost heavy or rightmost syllable'. 'HEAVINESS' >>EDGEMOSTNESS

(42) What is 'Heaviest'? C = voiceless, G = voiced cons.

| | | |
|---|---|---|
| a. ?**A**.ba.gi | 'toucan' | Cv>Gv |
| b. ?a.ba.**PA** | 'Amapá' | rtmost Cv wins |
| c. so.**AI**.pi | 'nose' | vv > Cv |
| d. **KAA**.gai | 'word' | Cvv>Gvv |
| e. <u>poo</u>.**GAI**.hi.ai | 'banana' | Gvv>vv ; in 3-$\sigma$ window; nb poo. |
| f. ka.pii.ga.ka.ga.ka<u>i.ka</u>.bao.**BAO** | | rtmost: in window: nb. kai |

(43) **Full 'Heaviness' Hierarchy:**

CVV > GVV > VV > CV > GV > V , where C=voiceless, G = voiced.

(44) Evidence that the stress is there:
a. Some speakers treat it as a beat wrt hand-gestures.
b. Some speakers devoice everything after stressed syllable.
c. Some speakers delete everything after stressed syllable.
d. May be some tropism of tone to stress.

(44) **ISSUES:**
A) How is 3 syllable limitation established?
B) How is 'heaviness' hierarchy established?

(45) Ans. to (A). *Domain of Stress is Minimal Prosodic Word, minWd.*
(Cf. Prosodic Circumscription (McCarthy & Prince, 1990),
whereby a prosodically delimited subdomain serves, in lieu of the expected morphological category, as the domain of a process.)

(46) *What's minimal about 3?* Prosodic Hierarchy says Wd contains F contains $\sigma$. So, hierarchically-locally, minWd contains ONE F. Maximal such Wd = F+$\sigma$ (or indeed $\sigma$ + F).

(47) Piraha stress must lie in maximal minWd, for F = $\sigma\sigma$.

(48) Evidence for involvement of Wd category:
Closely-knit phrases show ONLY ONE STRESS when last word is 2 $\sigma$.

| | |
|---|---|
| bao.sai # bii.**SAI** | 'cloth # red' |
| ?a.ba.pa # go.**GI** | 'city-name # where-at' |
| BUT: | |
| ?i.sii.**HOA**. # ai.**BAI**.?i | 'liquid fuel # much' |
| ?a.pa.**PAI** # **?II**.ta.ha | 'head # hurts' |
| ka.H(A)I # ?o.ga.ba.GAI | 'arrow # want' |

**Analysis:** PWd = max minWd is established at ]-edge of morph-word.
(Cf. Chen, Selkirk, Hewitt). Note RL orientation: final wins.

(49) PWd Parses. —Stress falls inside PWd. Must have }#

bao.{sai # bii.SAI}#
{?a.pa.**PAI**}# {**?II**.ta.ha}#
{ka.H(A)I}# ?o.{ga.ba.GAI}#

(50) ANS. to (44B). 'Heaviness' Hierarchy.
CVV > GVV > VV > CV > GV > V
a. xVV > yV in every case.
b. C > G > nil This is just onset goodness! (R. Shaefer).

Prominence Order = Weight >>Onset

(51) Full Story. IN-PWD>> WEIGHT >>ONSET >>RIGHTMOSTNESS.

(52) **Maithili** (Jhaa, 1958; Hayes, 1986, 1991)
Stress penultimate, final, antepenultimate, dependent on foot structure AND VV location.

(53) In words with only SHORT VOWELS, stress is penultimate:

SAta pa.TIta dhana.HEra bhina.SEra
ja.han.NEma sun.NARA

This is actually a Moraic-Trochaic (RL) pattern (Hayes)
— weak vowels in open syllables reduce; those in closed do not; nb. no closed final syllables.

(54) LONG VOWEL FINAL: Stress is FINAL for SHORT penult V—

a.dina.TAA gora.min.TII eka.hat.TAA

LONG V in antepenultimate stressed if VV followed by two LIGHT:

ji.muta.BAA.hana kaka.ROO.hari BAA.sana

PENULT LONG V is stressed:

pu.RUU.kha SAA.bhaa ba.raha.MAA.saa

(55) ANALYSIS: Main Stress falls on MOST PROMINENT in final circumscribed window; when there is equality, penult is favored (Non-finality).

(56) Window: 2 units = FF, Fσ, σF. = 2F, if headless feet exist (H86)

(57) **Elementary Prominence Orders:**
FOOT: Foot-heads > Nonheads;
VV: VV > V

(58) Relation: VV >>FOOT

(59) As usual, PROMINENCE >> POSN,
for POSN = NONFINALITY >> RIGHTMOSTNESS

(60) Cases:

| | | |
|---|---|---|
| ...\|CVcv.**CV**cv\| | = = prom/rightmost Fhead. | dhana.HEra |
| ...\|CVC.**CV**cv\| | ditto | ja.han.NEma |
| cvv.\|cv.**CVV**\| | prom wins. | ku(u).ra.NAA |
| ...\|**CVV**.CVcv\| | prom wins. | BAA.sana |
| ....\|**CVV**.cvv\| | = = prom/nonfinality wins. | SAA.bhaa |

(61) Fact Note: Following Hayes, we assume ...cvv.**CVC**.cv#
galaa + benna —> galaBENna but CAA.rillo 'fourth'
VV shortens in nonfinal, non-mainstressed syllables. If wrong, then PROM = heavy.

(62) **Awadhi** (Saksena, 1971; Hayes, 1991). Exercise for auditor: Stress is penultimate, except that in light-heavy#, heavy is stressed.

(63) **Broad Vista**. To run grammar on optimization, we need to be able to say WHEN to perform an operation. I.e. when it is better to act than to do nothing:
B( X—>Y ) > B ( X —>X )

(64) **Basic Harmonic Condition**: H(Y) > H(X). Well-formedness increasing: up Harmony Hill. But not sufficient: else Wd —> Tata.

(65) From universal perspective, need to order Transitions and combinations of Transitions and contexts to obtain markedness implications about processes.

(66) Example. **LANGO** Vowel Harmony — one property thereof. (Pulleyblank, Yest.)
a. ATR Harmony spreads forward from ALL vowels over a SINGLE CONSONANT:
in $V_1$-C-$V_2$ , $V_2$ harmonizes with 1st vowel.

b. But over TWO intervening C's, the trigger V must be high:
only in hi-CC-V do we get spread of ATR.

(67) Appears to involve complex **TRADE-OFF** in virtues of spreading from various sources, and difficultly of doing same in various environments. But there is NO TRADE-OFF.

(68) Approach. Suppose we know the following:
a. Spread ATR from HI > Spread ATR from NonHI.
b. Spread over single C > Spread over CC.

(69) Form the Lexicographic Product SPREAD *>> DISTANCE:

Spread from HI over C > Spread from H over CC
> Spread from NonHi over C > Spread from Non-Hi over CC.

(70) Now, an individual language may insert DO-NOTHING anywhere in the order.
All things preferable will be done, then.

(71) **Insert DO-Nothing** at bottom: any action is better than inaction: no constraint, full ATR harmony.

(72) **Lango**. Insert Do-Nothing above last element in order.
Spreading-from-NonHi-over-CC is worse than Nothing:
But everything else is better. So act accordingly.

###