Compendium of Medeiros' research on Fibonacci and X-bar

David P Medeiros
University of Arizona
March 10, 2013

Table of Contents

Part I

- Background
  - 3rd factor in biology
  - Golden mathematics
  - The Rabbit sequence

- The study of complex systems seems to affirm the Thompson-Turing claim that "some physical processes are of very general occurrence."
- Notably, those involving Fibonacci-based "golden" forms, ubiquitous in nature.
- This lends immediate interest to the observation that the repeated structural motif in the human syntactic system, the X-bar schema, is likewise a "golden" form (Medeiros 2008),
- and leads us to inquire whether whatever is behind the natural ubiquity of such phenomena, in other domains, might possibly be at work here as well.
- If so, this peculiar aspect of human phrase structure would fall under Chomsky's (2005) "third factor", a fact about language which is neither encoded in the particulars of our genome, nor learned from the environment, but determined by domain-general principles beyond the organism.

"Golden" Mathematics

- The Golden number (aka the golden ratio, the golden section, the golden mean)
- is *x*, such that $x^2 - x - 1 = 0$ $x/(x+1) = 1/x$
- 1.6180339… a constant called *Phi*.
- (or sometimes its reciprocal, 0.618… *phi*)
- The Golden angle, associated with the dominant spiral mode of phyllotaxis, is just *phi* measured out on a circle.
- Phi is the *most irrational* number.
- Intimately linked with the Fibonacci sequence and the Golden String (~Fibonacci word).

The “golden rule”

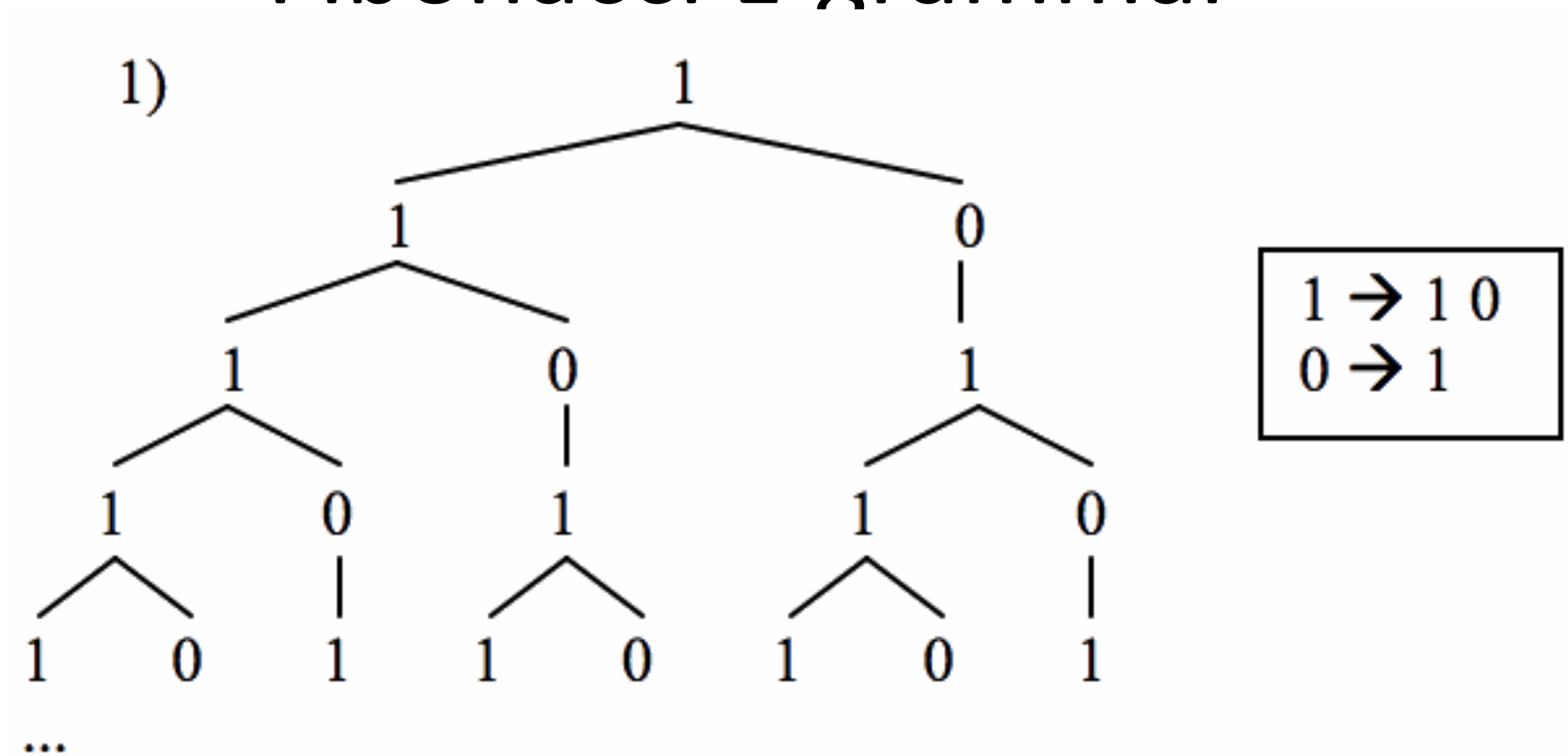
- $\mathbf{a_{n+2} = a_{n+1} + a_{n+0}}$ Fib #s, addition
- $\mathbf{s_{n+2} = s_{n+1} + s_{n+0}}$ Golden String, concatenation
- $x^2 = x^1 + x^0$ Polynomial giving Phi
- $\mathbf{x^{n+2} = x^{n+1} + x^{n+0}}$ (times x^n)
- $\mathbf{SO_{n+2} = SO_{n+1} + SO_{n+0}}$ Golden syntactic recurrence in X-bar format

Fib #: $a_n = a_{n-1} + a_{n-2}$ *1,1,2,3,5,8,13…*
Fib word: $s_n = s_{n-1}\ s_{n-2}$ *1011010110110…*

- The Fibonacci numbers, and the closely linked Fibonacci word (aka Golden String) in particular, are important topics in the study of symbolic dynamics, physics, theoretical computer science, etc.
- These patterns have a number of 'special' mathematical properties.

Fib #'s, addition:	Fib words, concatenation:
1,1,2,3,5,8,13,…	(0),1,10,101, 10110, 10110101,
• 1+1=2	0, 1: 10
• 1+2=3	1, 10: 101
• 2+3=5	101, 10: 10110

Fibonacci L-grammar



- 101101011011010110101101101011011010110 101101101011010110110101...
- http://personal.maths.surrey.ac.uk/ext/R.Knott/Fibonacci/fibrab.qt

The Golden String/Fibonacci Word/ Rabbit Sequence

- This structure is a quasiperiodic binary sequence whose long-range order reflects the mathematics of the Fibonacci numbers.
- By "quasiperiodic", we mean that it never exactly repeats (such sequences are merely periodic), but neither is it random. In fact it is ***self-similar***.
- Its linear organization reflects hierarchical groupings of the kind indicated in the tree above.
- This structure is reflected in many places in nature, from the organization of spin glasses (Binder 2008), to the oscillations of multiperiod variable stars including UW Herculis (Escudero 2003).

Self-generating procedure for the Golden String

{examine the value at a pointer.

If val=1, append 10 to the end of the string.

If val=0, append 1 to the end of the string.

{ Move the pointer one space right.

Repeat.

- Begin with just the first two digits of the GS (10), with pointer on the second digit (0, underlined and bolded):

1 **0**

- The pointer is at 0, so we add 1 to the end and move the pointer.

1 0 **1**

- Now the pointer is on 1; we add 10 and move the pointer.

1 0 1 **1** 0

- And so on:

1011**0**10, 10110**1**01, 101101**0**110, 10110101**1**01, 101101011**1**0110, etc.

The Golden string encodes its own computation

- The Golden String has a fascinating property of 'vertical' self-similarity at many scales.
- The sequence encodes the very procedure used to compute the sequence…
- Idea: perhaps this is significant in light of the **double articulation** of language noted since antiquity: its dual life as a linear outer form and a hierarchically structured inner form (strings and trees, basically).
- This object, in a sense, *brings its own double articulation with it*; the projection of a syntactic form from its sequence is inherently already there.
- In other word: there's already a tree in this string.

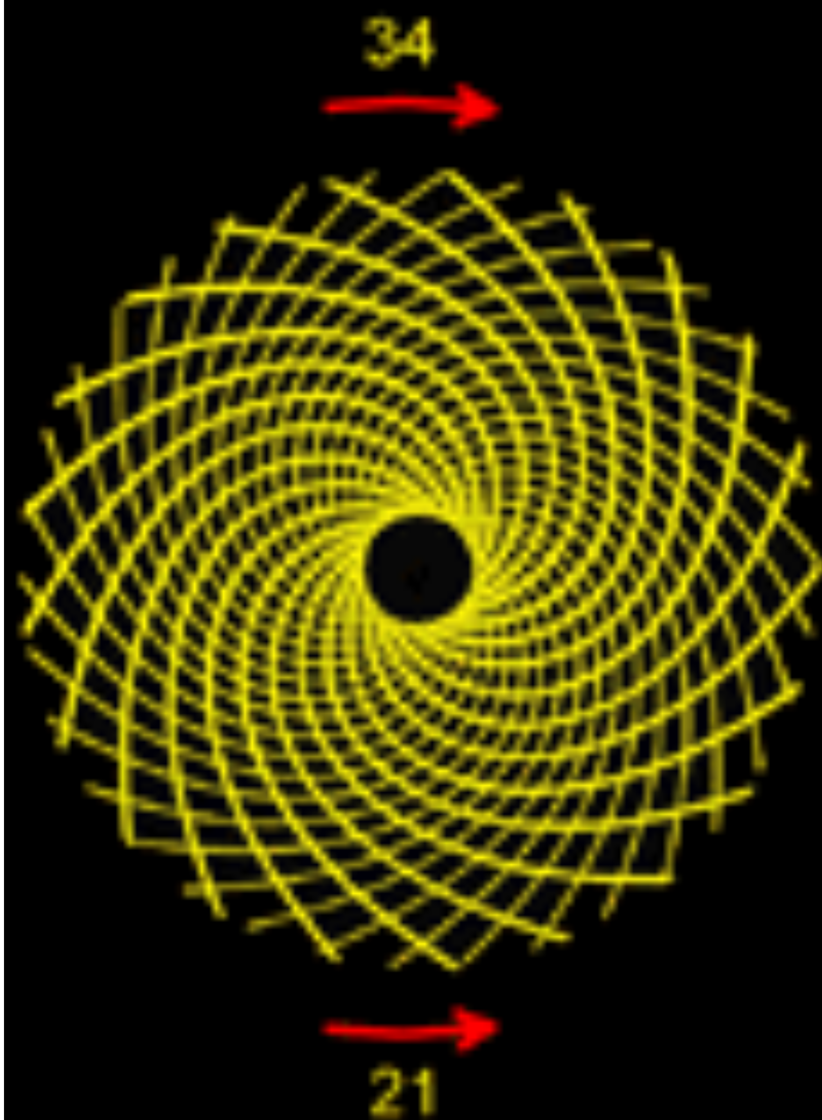
Part II

- Fibonacci/golden properties in Nature
  - Phyllotaxis
  - Brain
  - Penrose Quasicrystals

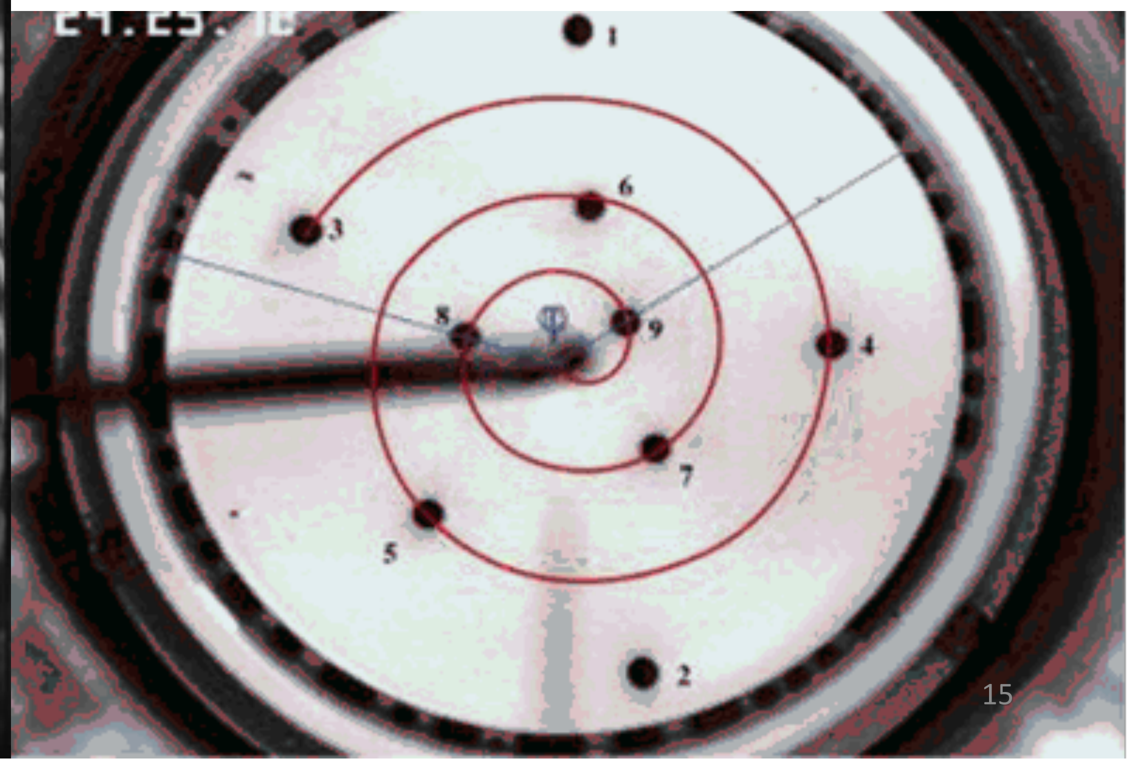
Intriguing cases of “golden” structure:

- The deep robustness of Fibonacci-based organization in phyllotaxis (across > 90% of plant species, Jean 1994)
- This pattern is known to “arise from self-organization in an iterative process […] the ordering is explained as due to the system’s trend to avoid rational (periodic) organization, thus leading to a convergence towards the golden mean.” Douady & Couder 1992
- Asymmetry in mammalian bronchial structure "consistent with a process of morphogenetic self-similarity described by Fibonacci scaling" (Goldberger, West, Dresselhaus, Bhargava 1985)

1, 1, 2, 3, 5, 8, 13, 21, 34…



- As Douady and Couder (1992, 1996) showed, the familiar Fib-spiraling arrangements in phyllotaxis can be explained by purely physical self-organization at the shoot apical meristem.
- They reproduced identical patterns with mutually repelling droplets of ferro-fluid; in essence, if the droplets fall into the center of the dish fast enough to be repelled by more than one previous drop, the Fibonacci organization is virtually inevitable.
- Computer simulations reveal the same pattern emerges robustly even if the repulsion forces scale differently with distance (i.e., under different laws of physics).



Golden frequency ratios in the brain

- Evidence that golden mathematics plays a special role on the "hardware" side of cognition:

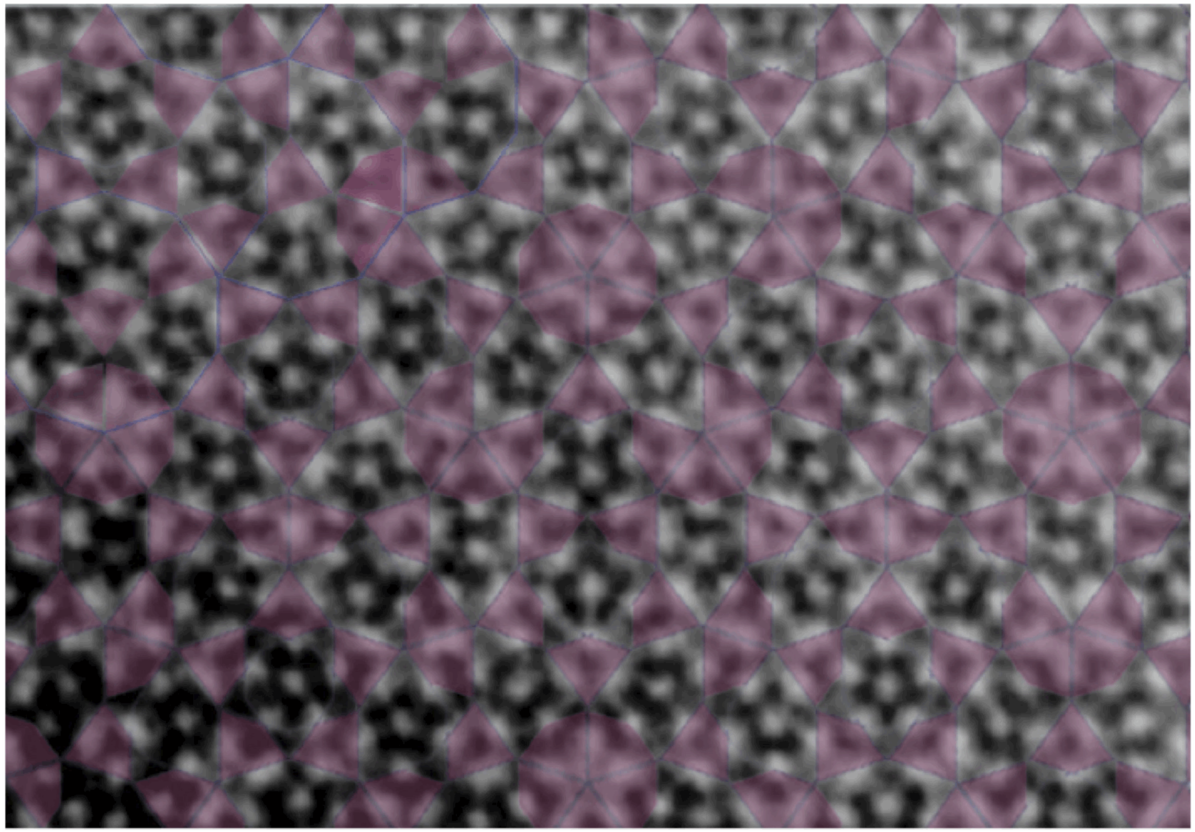
Roopun et al. (2008), "*Temporal interactions between cortical rhythms*":

- "The modal peak frequencies fall into distinct bands, with approximately twice as many bands as expected from a natural log distribution. Instead, the bands appear approximately distributed according to 'phi' (the 'golden mean') rather than 'e' – a constant commonly associated with the organisation of complex natural systems (Atela et al., 2002)."
- "[…] in using phi as a common ratio between adjacent frequencies in the EEG spectrum (Figure 1), the neocortex appears to have found a way to pack as many, minimally interfering frequency bands as possible into the available frequency space."

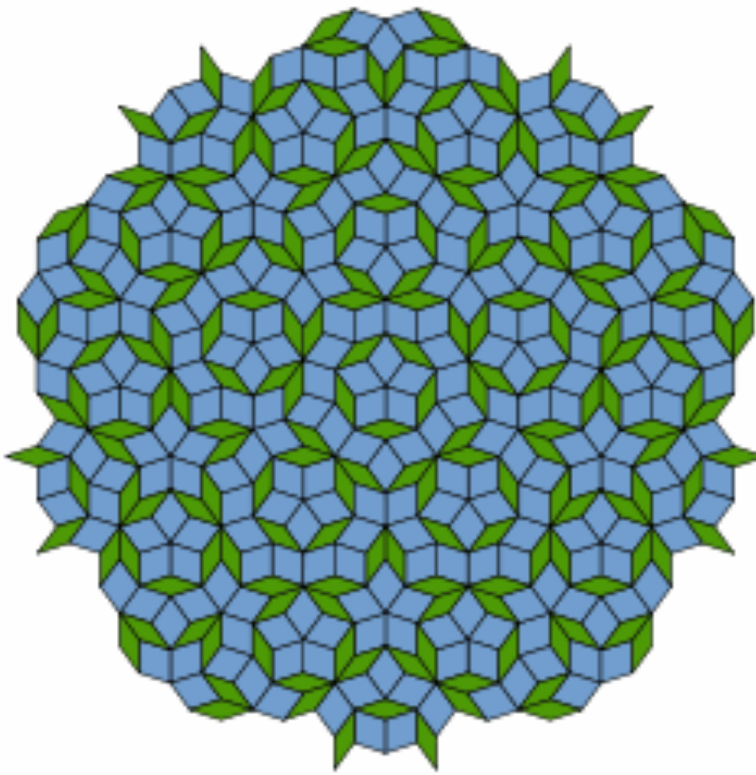
Golden quasicrystals

- A class of crystals with Bragg diffraction showing forbidden (e.g. five-fold) symmetry has been successfully modeled in terms of 3-dimensional Penrose tilings. Although Penrose constructed his tilings with two or more shapes with intricate edge-matching rules, it has been shown that the same geometry can be achieved through uniform, overlapping decagonal tiles.
- "[A] quasiperiodic tiling can be forced using only a single type of tile, and furthermore we show that matching rules can be discarded. Instead, maximizing the density of a chosen cluster of tiles suffices to produce a quasiperiodic tiling. If one imagines the tile cluster to represent some energetically preferred atomic cluster, then minimizing the free energy would naturally maximize the cluster density." (Steinhardt & Jeong 1996: 431)

"A new picture of quasicrystals emerges in which the structure is determined entirely by a single repeating cluster which overlaps (shares atoms with) neighbor clusters according to simple energetics." Gummelt(1996).



"Figure 3: Superposition of a perfect decagon tiling on the high angle annular dark-field (HAADF) lattice image of water-quenched $Al_{72}Ni_{20}Co_8$ obtained by the high angle annular dark field method by Saitoh *et al.*"



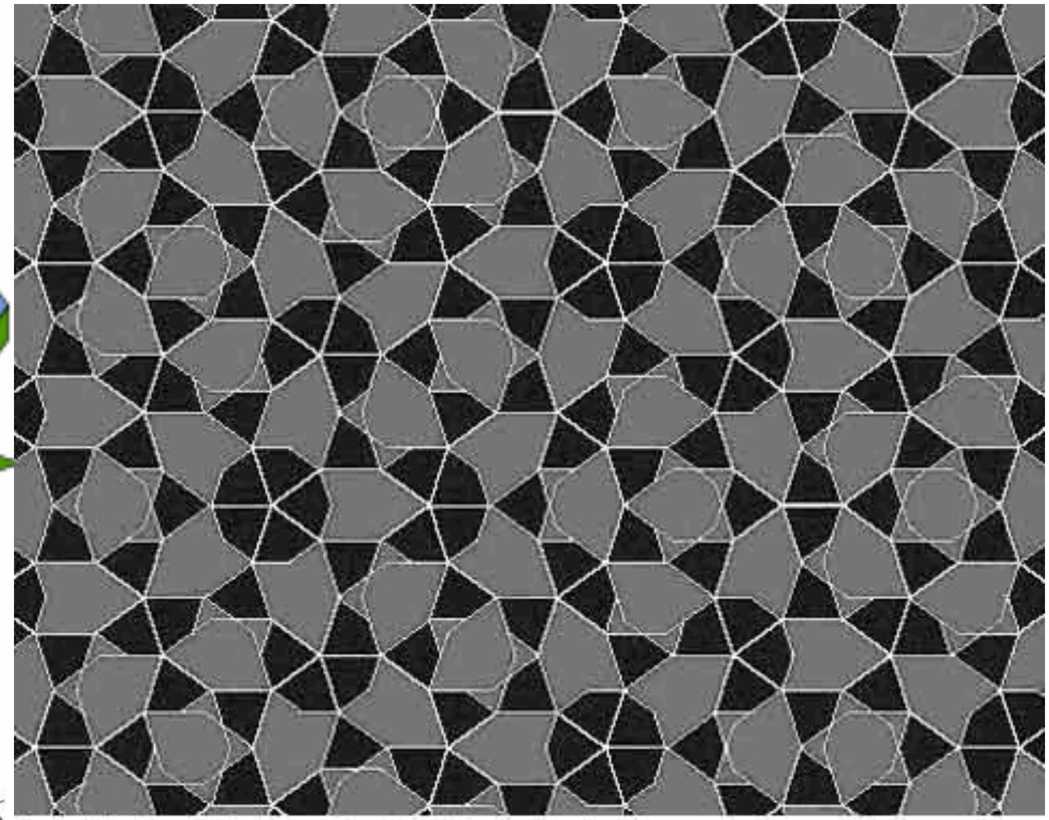
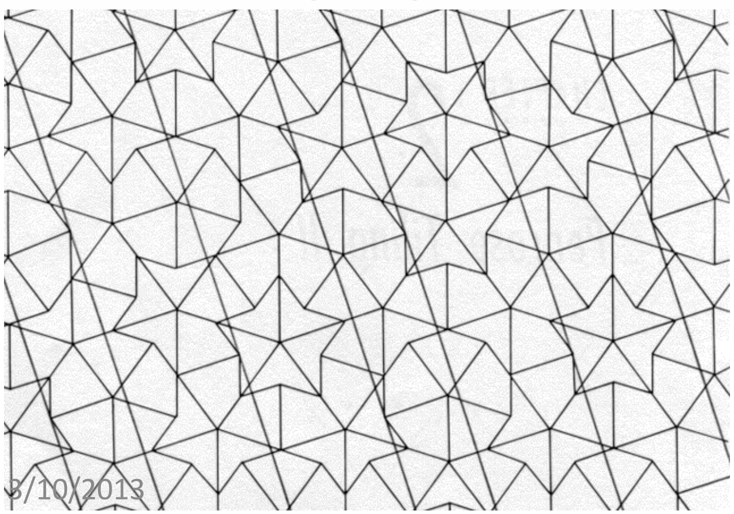


Figure 15. Gummelt's tiling of the plane with overlapping decagons



- Above left: Penrose tiling with rhombi
- Below left: Ammann bars: wide/narrow gaps follow the Golden String
- Above right: Penrose tiling with a single kind of (overlapping) decagon, best model for real quasicrystals.

Part III

- What are grammars?
- Saddy’s results

“Where do grammars come from?”

- This was the question posed in Doug Saddy’s talk here in Spring 2012.
- He presented intriguing experimental results suggesting that the so-called **Fibonacci L-grammar** “entrains the brain” in a special way, more so than any other kind of long-range order he examined.

Saddy's results

- In experiments by J Douglas Saddy and his group at the Centre for Integrative Neuroscience and Neurodynamics,
- He played subjects a long stretch of this Fibonacci grammar sequence, as a string of ba/bi syllables:
  - *ba bi ba ba bi ba bi ba ba bi ba ba bi...*
- As well as other sequences, like the Thue Morse (1 → 01; 0 → 10) and Feigenbaum (1→ 10, 0 → 11) grammars
- Then had subjects decide which of two short samples played subsequently best matched what they had heard before.
- Subjects recognized the Fibonacci sequence more accurately than any other.

More than just a sequence...

- He also tested the Fibonacci and other grammars against “pseudo-clones”,
- Composed of “legal” substrings, assembled legally, but with a randomized long-range order.
- Again Fibonacci fared best; remarkably, subjects could tell it apart from even sequences built from 5- or 8-bit long substrings of it (*10110 & 10110101*).

Remarkable

- This is a remarkable fact; the long-range order of the sequence is strictly beyond the capabilities of finite-state (Markovian) processes to describe.
- Even more stunning: in computer science, the Fibonacci word is known to be **the worst case** for the application of many efficient pattern recognition algorithms (Knuth, Morris, & Pratt 1977, Aho 1990, a.o.).
- Whatever the reason for subjects recognizing this string best, it seems it's not about local statistical regularities (e.g., recognizing *n*-grams).

Saddy's conclusions

- *Safe conclusion*: Humans can detect and discriminate domains of self-embedding in recursive strings.
- *Prudent conclusion*: Investigating the correspondence between cortical activity and the ability to recognise and manipulate structure hidden in recursive signals may point to common mechanisms underlying complex cognitive processes.
- *Bold conclusion*: The physical conditions which yield the L-system governed patterns in sunflowers and brain morphology appear to also govern aspects of optimal neural signal properties involved in information processing.
- *Binky's conclusion*: Some of the defining properties of human grammars follow from physical principles governing certain attributes of a system – they reflect a natural law.

  (Saddy 2012: 58)

Where to go from here?

- To test the hypothesis that the spectral properties of the Fibonacci/X-bar grammar are what make it "special"
- The idea is to test how subjects perform with distinct sequences/grammars with some but not all of its special spectral properties.
- This requires a jump to ternary strings (over an alphabet 0, 1, 2), the first place where these properties can be dissociated.

Tribonacci

This grammar is Endocentric and of Pisot type, but not Polygonal.

$2 \rightarrow 2\ 1$

$1 \rightarrow 2\ 0$

$0 \rightarrow 2$

212021221202121202122120212021221202121202122120212212
021212021221202120212212021212021221202121202122120212
021221202121202122120212212021212021221202120212212021
212021221202120212212021212021221202122120212120212212
021202122120212120212212021212021221202120212212021212
021221202122120212120212212021202122120212120212212021
221202121202122120212021221202121202122120212120212212
021202122120...

Heptagon

This grammar is Polygonal, but neither Endocentric nor Pisot.

$2 \rightarrow 2\ 1$

$1 \rightarrow 0\ 2$

$0 \rightarrow 1$

2102121022102121210212102210221021210221021212102121210
0212102210212121021210221022102121022102210212102210210
2121021210221022102121022102121210212121021210221021210
2102121210212102210212121021210221022102121022102121210
0212121021210221021212102121022102210212102210221021210
0221021212102121022102210212102210221021210221021212100
2121022102210212102210212121021212102121022102121210210
210221022102

Pisot # 1.46557…

This grammar is of Pisot type, but is neither Endocentric nor Polygonal.

$2 \rightarrow 2\ 1$

$1 \rightarrow 0$

$0 \rightarrow 2$

2102212102102210221210221210210221210210221022121021
0221021221022121021022102212102212102102212102102210
2212102212102102212102102210221210210221022121022121
0221210210221022121021022102212102212102102210221210
2212102102210221210210221022121022121021022102212102
1022121021022102212102102210221210221210210221210210
2212102102210221210221210210221022121022121021022121
0221210210221210221

Part IV

- The X-bar schema

The X-bar schema.

Of all the ways that syntactic structure could be built up, one particular ‘growth solution’ seems to dominate in natural language.

This is the so-called *X-bar schema*:

$XP = [ZP\ [X^0\ YP]]$

In words: a phrase of any type (a verb phrase, noun phrase, whatever, thus an XP(hrase)) is built around a head (X^0), with asymmetrically arranged ‘slots’ for two additional phrases of the same shape.

NP = [The barbarians’ [destruction [of Rome]]

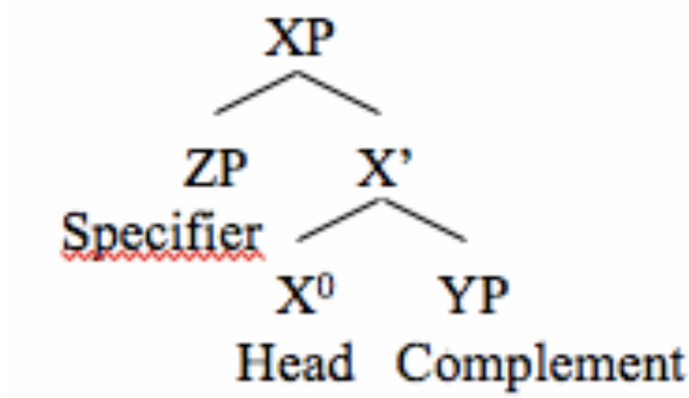
Sentence = [The barbarians [destroyed [Rome]].

The phrasal off-branches (YP, ZP) may be expanded indefinitely:

[The ravening hordes of barbarians] [destroyed [the gleaming city on the hill]], and so on.

A none-too-innocently-chosen example

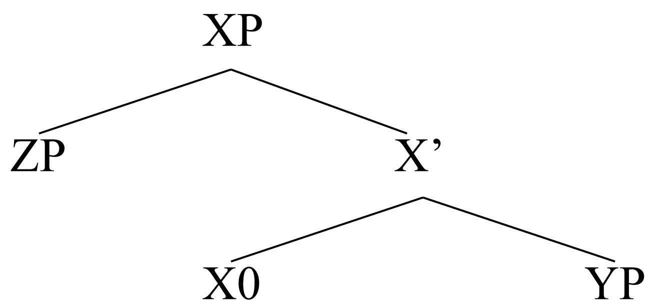
- As an illustration of one kind of pattern of recurrence, consider the familiar "X-bar schema" (Chomsky 1970, Jackendoff 1977).



- This is a "recipe" for structure building: to build a phrase (XP), combine a terminal (X^0) with a phrase (YP), then combine the result (X') with another phrase (ZP).

Background: [Spec [Hd Comp]]

- Since Chomsky (1970), it is widely held that the syntactic structures of natural language are constructed around the 'X-bar molecule' shown below:



- The claim is that one finds phrases (XPs) of only the following shapes:
  a. $XP = X^0$
  b. $XP = [X^0\ YP]$
  c. $XP = [ZP\ [X^0\ YP]]$
- One does not find 'exocentric' phrases such as
  *XP = [YP ZP] (*contra* Starke 2004)
- Nor phrases with more than a single complement and specifier:
  *XP = [WP [ZP [X^0 YP]]] (*contra* Chomsky 1995a)
- Nor phrases in which the head (X^0) is not at the 'bottom':
  *XP = [X^0 [YP ZP]] (but see Moro 2000, Pereltsvaig 2006, below)

The Golden Phrase

- I suggest that we should add to the family of related “golden” mathematical objects (the golden number/section/mean; the golden angle, the golden string)
- The “Golden Phrase”, i.e. the X-bar schema.
- In essence, this phrasal shape is the expression, in binary-branching syntactic trees, of the very same Fibonacci theme.
- This kind of phrasal organization has a number of “special” properties…

The Golden Phrase is special

- In what follows, I will point to three considerations which pick out this kind of phrase structure as special:
  - (1) The X-bar schema is the simplest kind of syntactic (multi-)fractal.
  - (2) The X-bar schema is the minimal semantic generator, the first shape to unlock the full set of predicate-argument meanings.
  - (3) X-bar grammar, with specifier-head-complement order, yields strings related to the Golden String (infinite Fibonacci word); it has the lowest ambiguity among "binary generators of binary".

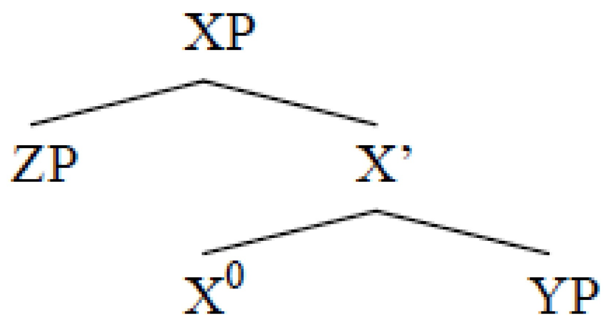
- “Arguably, this configurational schema, known as X-bar theory, is ***the only kind of structure that syntactic representations exploit***. Other structural options, such as adjuncts to phrases, multiple specifiers of a single head, etc., have been experimented with in various ways but Cartographic research has, for the most part, eschewed these options, retaining only the core structures afforded by the X-bar schema. Indeed, Cinque (1999) argues forcefully against the adjunction of adverbials[...] The core structural relations defined by X-bar theory seem to be ***not only necessary, but sufficient to characterize syntactic structure.***”
- (Shlonsky 2010: 2; emphasis added - DPM)

The second step consisted in the recognition that the various classes of adverbs (more accurately, AdvPs) are also ordered among each other in a syntactic hierarchy, and that this hierarchy turns out to match exactly the hierarchy of Mood, Tense, Modality, Aspect and Voice heads, as can be seen if we juxtapose the two hierarchies:

(7)

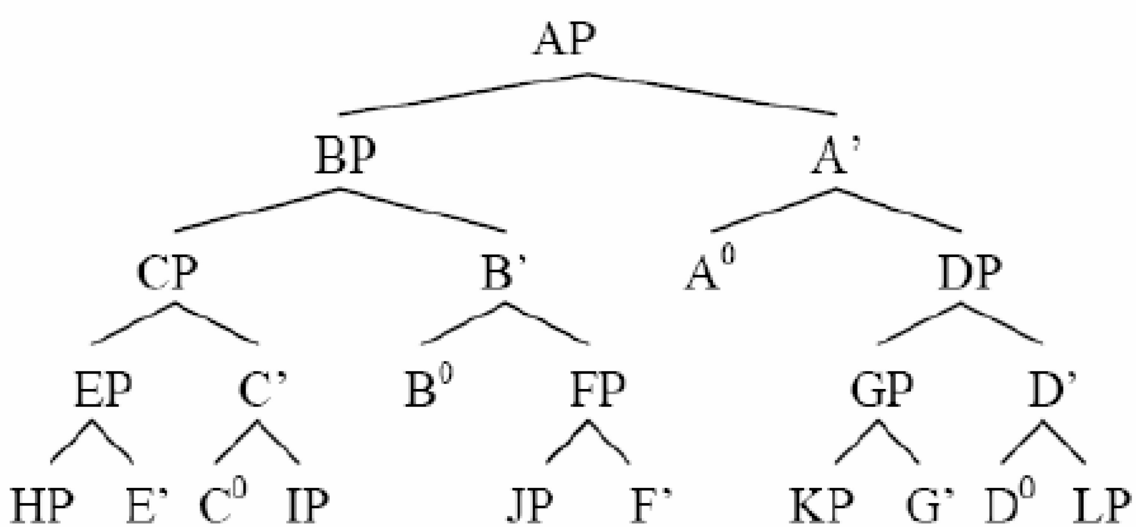
a	b
$\text{Mood}_{\text{speech act}}$	$\text{AdvP}_{\text{speech act}}$ (frankly,..)
$\text{Mood}_{\text{evaluative}}$	$\text{AdvP}_{\text{evaluative}}$ (fortunately,..)
$\text{Mood}_{\text{evidential}}$	$\text{AdvP}_{\text{evidential}}$ (allegedly,..)
$\text{Mod}_{\text{epistemic}}$	$\text{AdvP}_{\text{epistemic}}$ (probably,..)
$\text{Tense}_{\text{past/future}}$	$\text{AdvP}_{\text{past/future}}$ (then,..)
$\text{Mod}_{\text{necessity}}$	$\text{AdvP}_{\text{necessity}}$ (necessarily,..)
$\text{Mod}_{\text{possibility}}$	$\text{AdvP}_{\text{possibility}}$ (possibly,..)
$\text{Aspect}_{\text{habitual}}$	$\text{AdvP}_{\text{habitual}}$ (usually,..)
$\text{Aspect}_{\text{repetitive}}$	$\text{AdvP}_{\text{repetitive}}$ (again,..)
$\text{Aspect}_{\text{frequentative}}$	$\text{AdvP}_{\text{frequentative}}$ (frequently,..)
$\text{Mod}_{\text{volition}}$	$\text{AdvP}_{\text{volition}}$ (willingly,..)
$\text{Aspect}_{\text{celerative}}$	$\text{AdvP}_{\text{celerative}}$ (quickly,..)
$\text{Tense}_{\text{anterior}}$	$\text{AdvP}_{\text{anterior}}$ (already)
$\text{Aspect}_{\text{terminative}}$	$\text{AdvP}_{\text{terminative}}$ (no longer,..)
$\text{Aspect}_{\text{continuative}}$	$\text{AdvP}_{\text{continuative}}$ (still,..)
$\text{Aspect}_{\text{continuous}}$	$\text{AdvP}_{\text{continuous}}$ (always,..)
$\text{Aspect}_{\text{retrospective}}$	$\text{AdvP}_{\text{retrospective}}$ (just,..)
$\text{Aspect}_{\text{durative}}$	$\text{AdvP}_{\text{durative}}$ (briefly,..)
$\text{Aspect}_{\text{prospective}}$	$\text{AdvP}_{\text{prospective}}$ (imminently,..)
$\text{Mod}_{\text{obligation}}$	$\text{AdvP}_{\text{obligation}}$ (obligatorily,..)
$\text{Aspect}_{\text{frustrative}}$	$\text{AdvP}_{\text{frustrative}}$ (in vain,..)
$\text{Aspect}_{\text{completive}}$	$\text{AdvP}_{\text{completive}}$ (partially,..)
$\text{Voice}_{\text{passive}}$	$\text{AdvP}_{\text{manner}}$ (well,..)
Verb	**Verb**

(1)



<u>Why is language that way?</u>

This structure might follow from more general principles. It is suggestive that the X-bar pattern generates the **Fibonacci numbers**, which arise in many natural systems.



XP	X'	X^0
1	0	0
1	1	0
2	1	1
3	2	1
5	3	2

Fibonacci #s: 1,1,2,3,5,8,13…

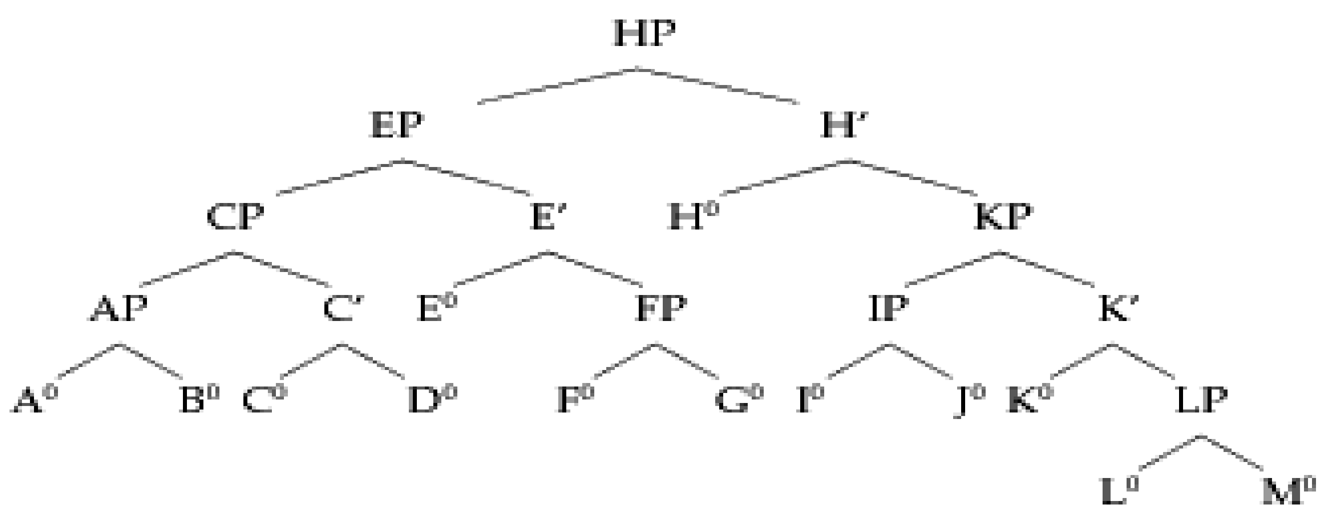
Consider how X-bar analyses may be assigned to strings (where an XP can be expanded as X0, [X0 YP], [ZP [X0 YP]]).

For each string length, there are a number of X-bar branching structures that could underlie it:

ABCD: [A [B [C [D]]]], [[AB][CD]], etc.

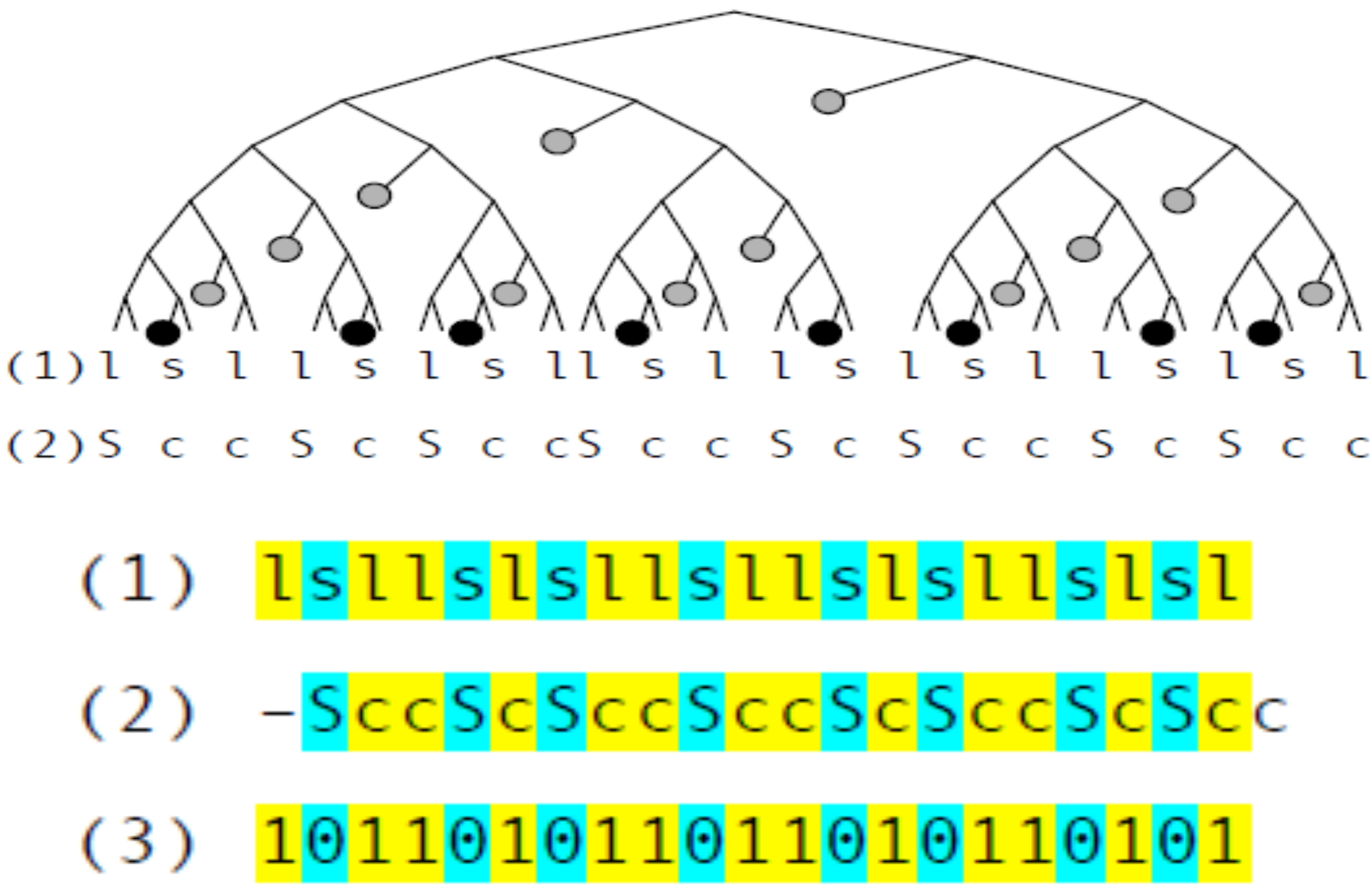
For a given string length, there is a minimal depth any X-bar analysis must have. E.g., string length 3 requires depth 2, as does 4; once you get to 5, you need a tree that has depth 3.

Fibonacci string lengths are special, in that they are the first length to push the required tree-depth one deeper (thus, 6- and 7-long strings also 'fit' in depth 3 trees; 8 is the first length to necessarily push the tree 4 deep) . Below: 13-long string (ABCDEFGHIJKLM), forcing depth-5 X-bar analysis.



- http://www.mcs.surrey.ac.uk/Personal/R.Knott/Fibonacci/fibrab.qt

`



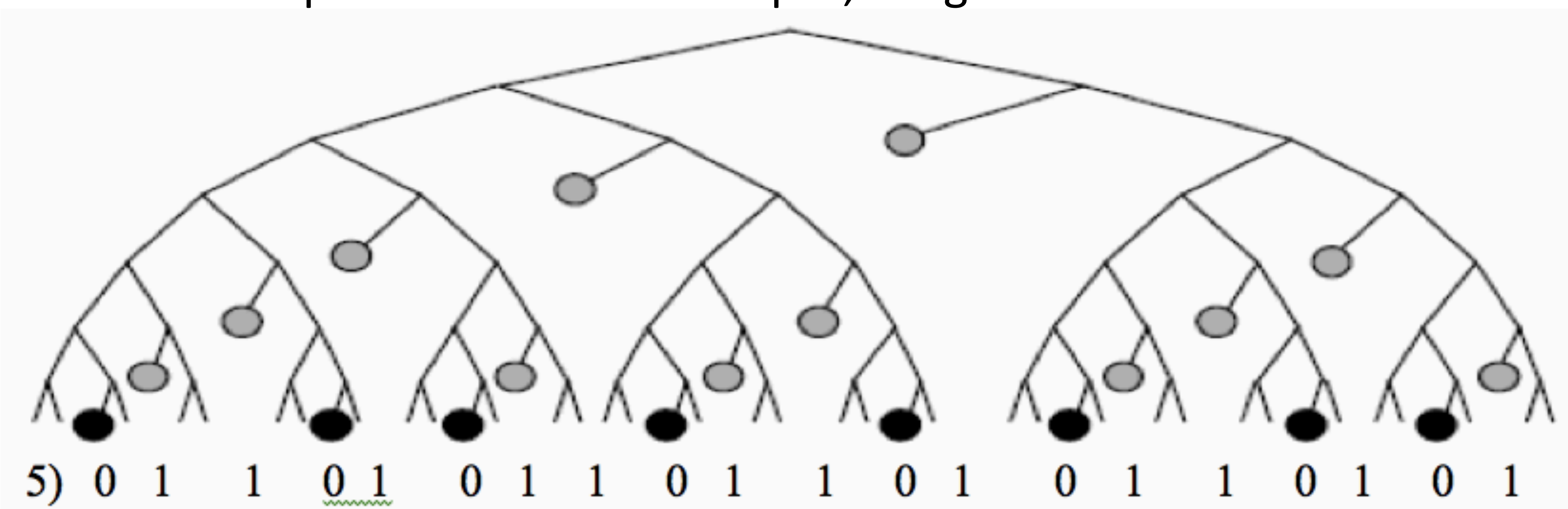
(1) lsllslsllsllslsllslsl

(2) -SccScSccSccScSccScScc

(3) 101101011011010110101

- So: the sequences of large/small open categories on the bottom 'frontier' of the partial expansion of the maximal X-bar tree follow the golden sequence exactly.
- …in fact, are successive Fibonacci length portions of that object.
- The classification of those open categories as specifiers or complements is a slightly different pattern.
- The spec/comp sequence also follows the golden string pattern, but starts at index 2.
- Likewise, if one examines the sequence of head positions, marked for whether they are introduced in the last generation (bottom line of the tree) or not, one also finds the GS starting at index 2.

- According to Chomsky & Halle (1968) (see also Bresnan 1971, 1972, Cinque 1993)
- The deepest part of a syntactic tree gets maximal stress (Nuclear Stress Rule)...
- Reading sequence of stress maxima (black heads) and non-maxima (grey heads) from left to right in an idealized X-bar tree expanded to uniform depth, we get the Fibonacci word:

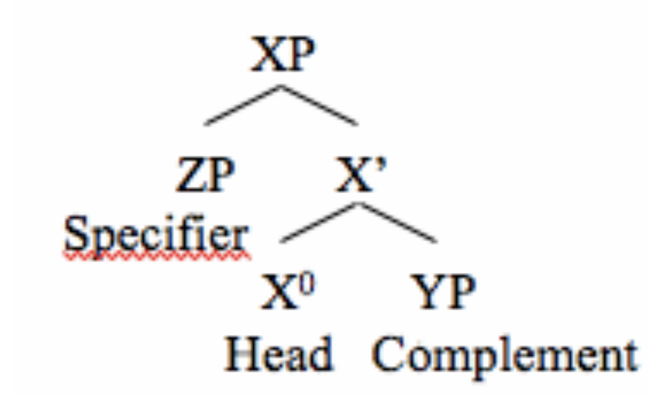


X-bar form without labels

- A traditional way to describe this particular pattern is with phrase structure rules (PSRs), as below:

  XP → ZP X'

  X' → X^0 YP



- Ignoring labels, we can write this as:

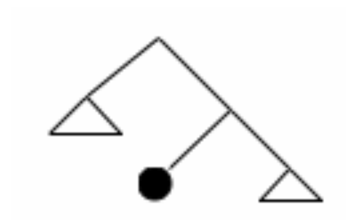
  XP → XP X'

  X' → X^0 XP

- Or, even more simply and abstractly, using 0s to represent terminals, and higher numbers (1, 2) to represent distinct kinds of non-terminals:

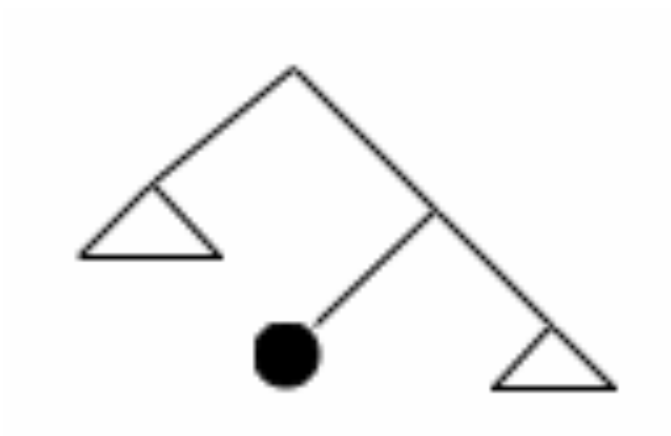
  2 → 2 1

  1 → 0 2



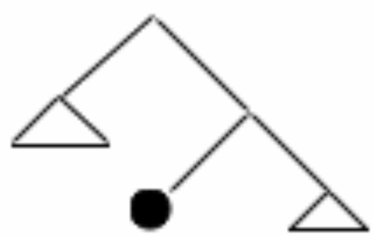
Bare recursion

- Of course, the X-bar schema is more than just a structure-building pattern; it also incorporates the further notion of *labeling* or *headedness*.
- In what follows, I ignore this aspect, considering only the recurrence pattern.
- On this view, the X-bar schema resolves as a simpler object that can be depicted as below:



X-bar pattern as **matrix**

$$\begin{array}{c|cc} & \text{XP out} & \text{X' out} \\ \hline \text{XP in} & 1 & 1 \\ \text{X' in} & 1 & 0 \end{array} \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$



Rows and columns are associated with the kinds of non-terminals.

Rows can be thought of as inputs to phrase structure rules;

Columns are the outputs.

Only non-terminals are represented, and we ignore linear order.

.

Let’s make that clearer:

$$\begin{array}{cc} & \begin{array}{cc} \text{XP out} & \text{X' out} \end{array} \\ \begin{array}{c} \text{XP in} \\ \text{X' in} \end{array} & \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

This representation doesn’t record linear order.

Terminals are expressed only indirectly here.

A non-terminal introduces a terminal if its associated row sums to less than 2.

E.g. the second type of non-terminal (X’) introduces a single terminal because its row adds up to 1; the first (XP) row sums to 2, indicating it immediately dominates no terminals.

Part V

- Expressive Power
- The X-bar schema is the “minimal semantic generator”

X-bar schema and semantic expressive power

- There is another reason to think that the X-bar form is "special"
- Related to its expressive power when mapped to semantic interpretation.
- The X-bar schema is the *minimal semantic generator.*
- In the sense that predications of any internal structure, stated with predicates of arbitrary adicity,
- Can be expressed in a X-bar syntactic form (utilizing the syntactic equivalent of Schoenfinkelization/Currying).
- But no simpler form will do.
- In other words, the X-bar schema is *just right*: just big enough to get the job done (i.e. to express any kind of predication) -- having a larger phrasal shape doesn't buy you any additional expressive power.

Bicomplex predication

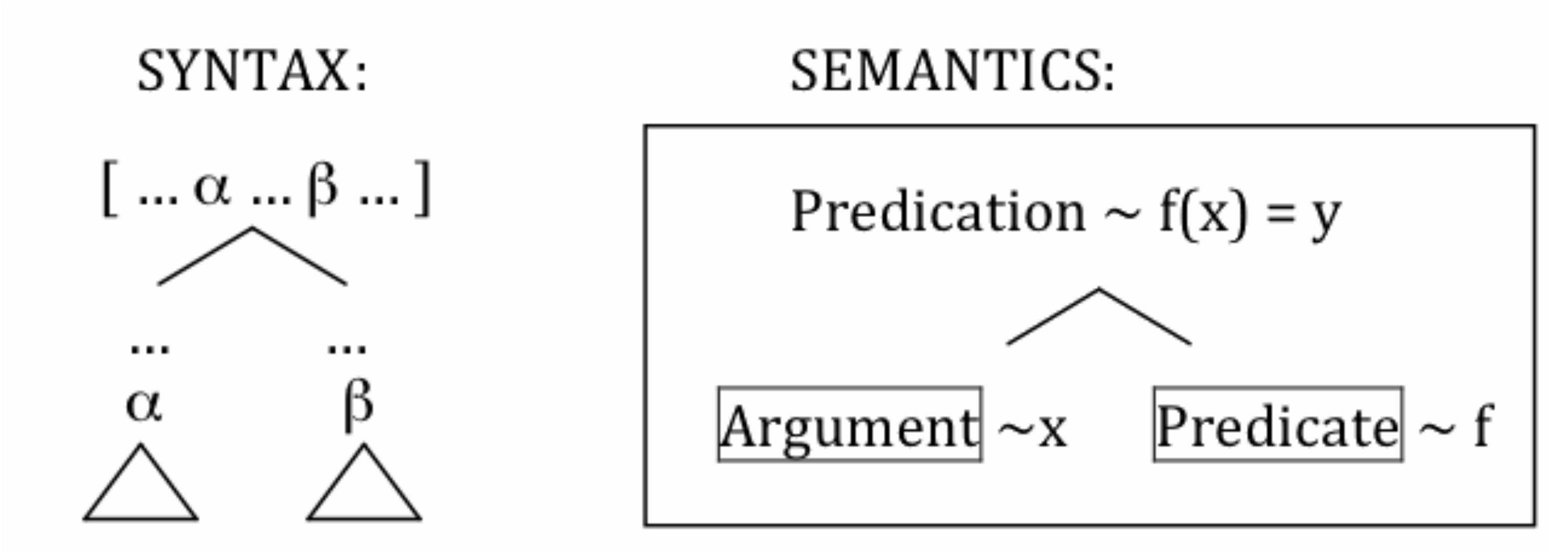
- Note the following about predicate-argument structure in natural language:
- Both predicates and arguments may contain further predicate-argument structure:

Simple: [$_{Arg}$ The people] [$_{Pred}$ know].

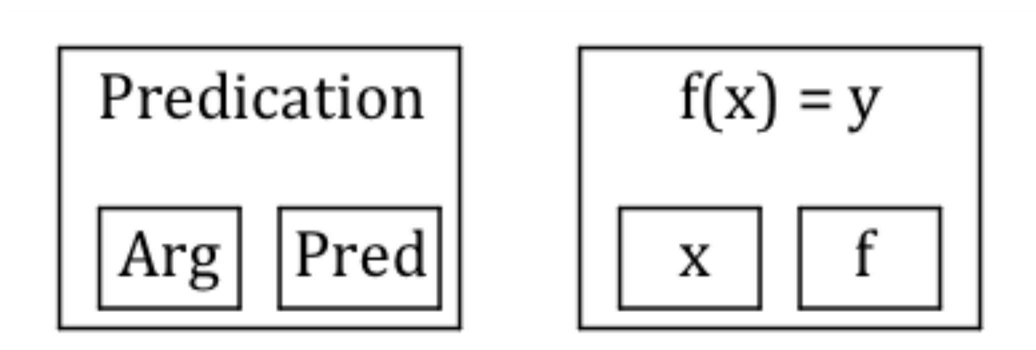
Complex Arg: [$_{Arg}$ The people [$_{Arg}$ you] [$_{Pred}$ met t]]
[$_{Pred}$ know].

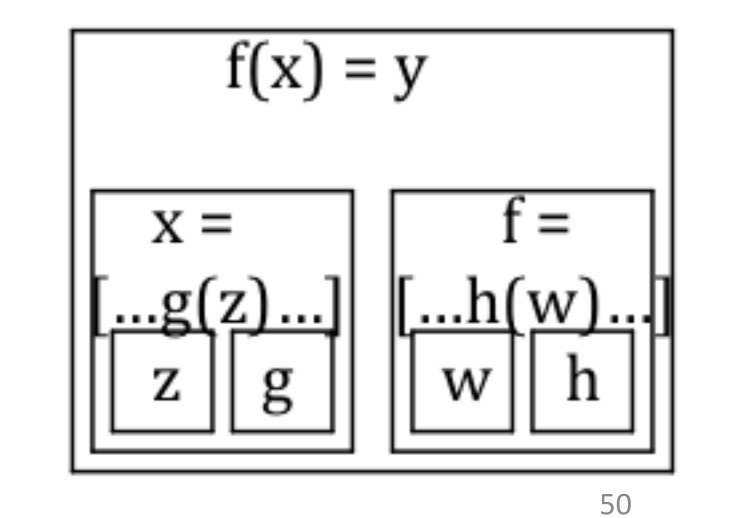
Complex Pred: [$_{Arg}$ The people]
[$_{Pred}$ know [$_{Arg}$ you] [$_{Pred}$ exist t]].

Bi-Complex: [$_{Arg}$ The people [$_{Arg}$ you] [$_{Pred}$ met t]]
[$_{Pred}$ know [$_{Arg}$ you] [$_{Pred}$ exist t]].



- I adopt the minimal notion of compositionality depicted above.
- Bicomplexity in semantic composition requires bicomplex syntax.
- Bicomplex syntax: two(+) growth points per molecule.
- So, at least as complex as X-bar (or D-bar?).

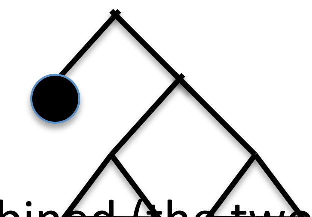




(A)Symmetry, Fractal Semantics?

- Conceivably, the structural asymmetry inherent in the dominant X-bar form is "recruited" for the semantic asymmetry between predicates and arguments.
- Making it more *useful*, in a sense: not only do you get bicomplex semantics, you get **asymmetric bicomplexity**, a basis for Fregean semantic asymmetry.
- A crucial case here is the structure of the copula, argued to be this (Moro 2000, Pereltsvaig 2006):

[cop [$_{SC}$ XP YP]]



- This is a (partial) manifestation of X-bar's obscure sibling, the apparently rare D-bar configuration.
- Crucial property: symmetry between the two objects combined (the two growth points in the phrasal shape).
- Here, we have a symmetric syntactic form just where we need it to construct a symmetric meaning (equation).

Part VI

- Fractals
- X-bar schema is the simplest syntactic schema generating a line fractal
- In fact a multi-fractal
- With "golden" Hausdorff dimension

Next: Fractals & the Cantor Set

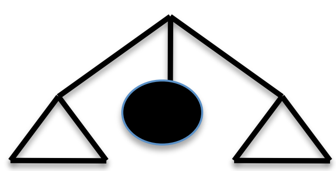
Fractals are self-similar objects of non-whole-number dimension; their "size" depends on the scale at which they are measured.

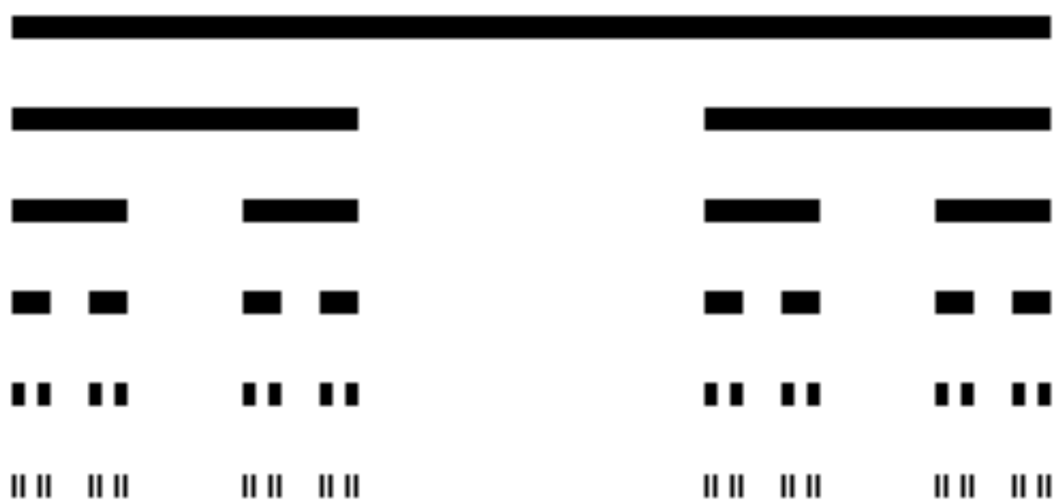
The Cantor set is formed from a line segment, by removing the middle third, then middle thirds of the remainders...

This is the *simplest* fractal:

- Background dimension cannot be lower than a 1-dimensional line.
- Division in thirds is the first scheme yielding a fractal.

The self-similarity here invites a kind of *phrasal* analysis: within each "generation", there are two copies of the whole, and one "dead end" (deleted segment ~terminal):

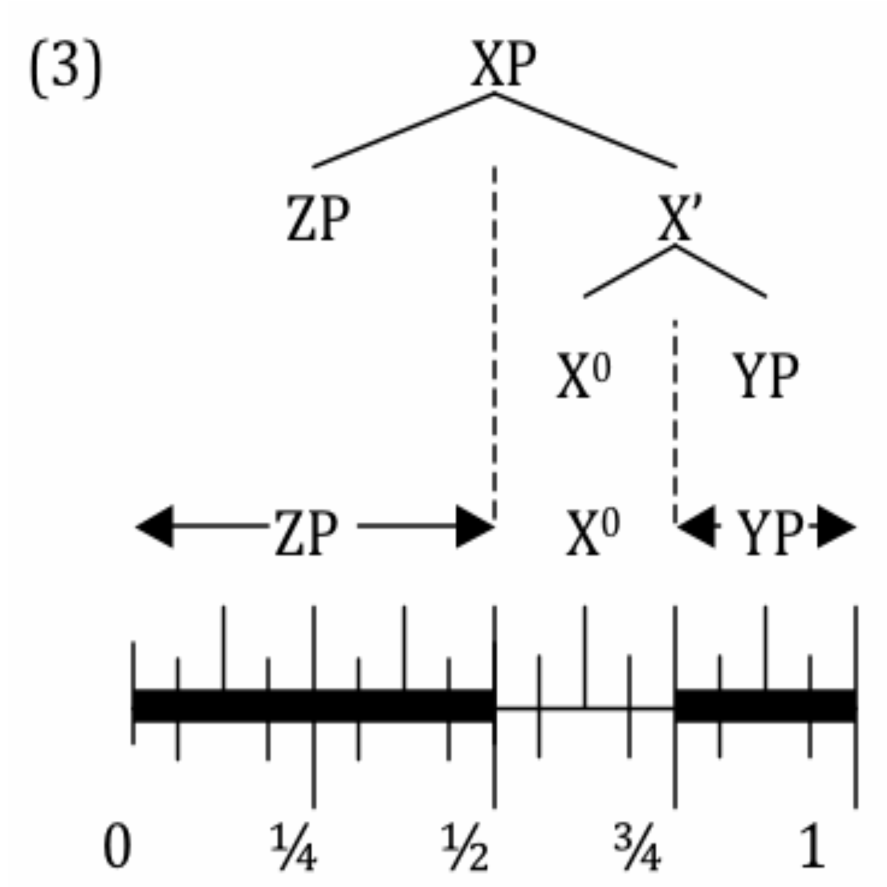




53

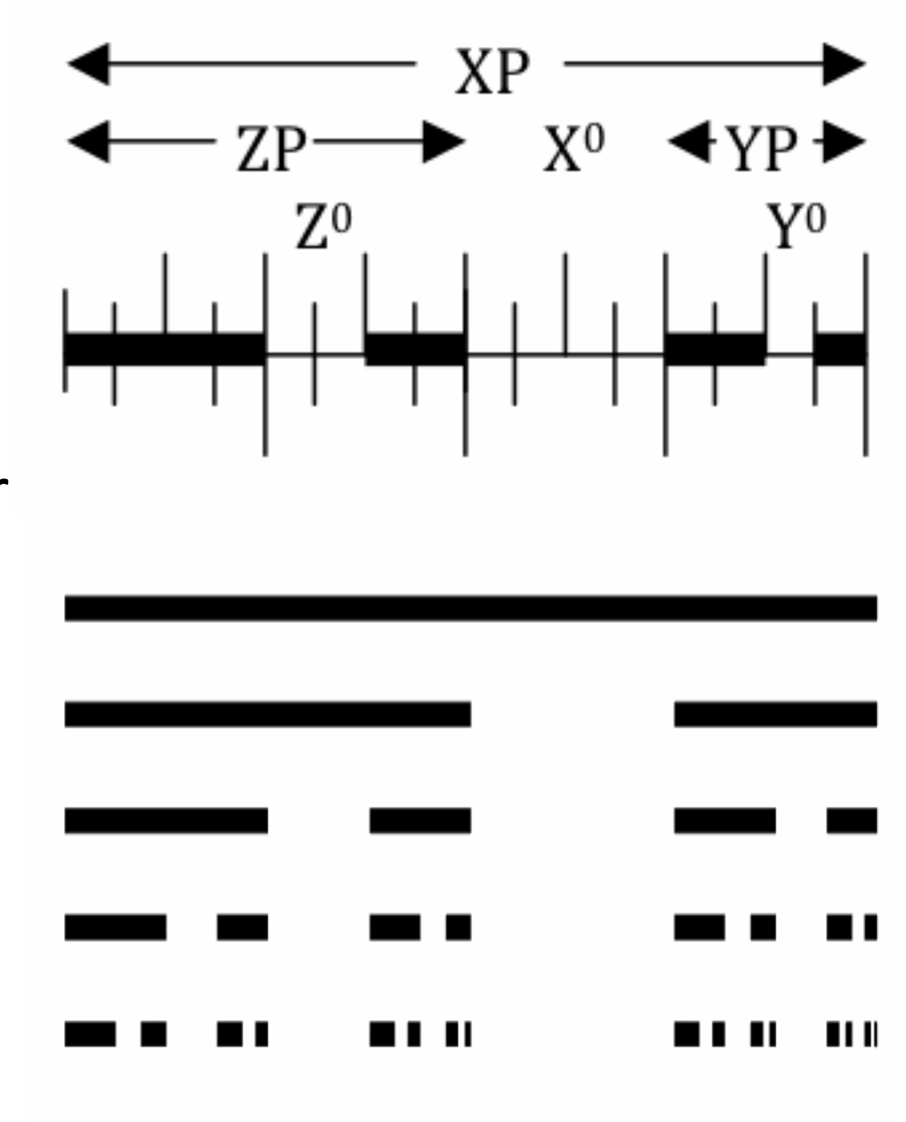
X-bar schema as (multi)fractal

- Consider mapping the X-bar schema to a line segment,
- Such that binary branching in the syntactic form corresponds to geometric halving,
- And heads/terminals corresponding to deleting a line segment.



And so on: fractal structure

- Of course, ZP and YP themselves have the same internal structure as XP:



- Continued indefinitely, this produces an asymmetric (or two scale) Cantor set.
- Each generation has one ½ and one ¼ scale copy of the whole.

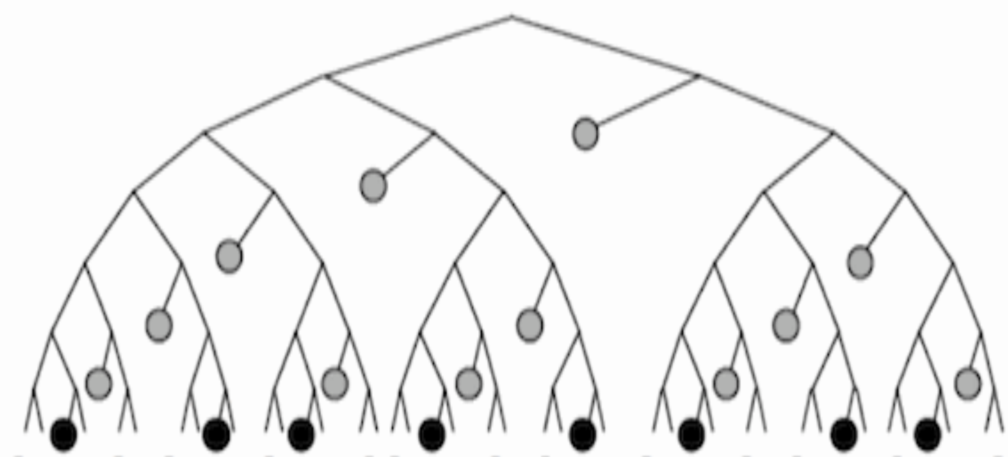
Asymmetric Cantor set ~ X-bar tiling

As a fractal, this has a number of properties worth mentioning.

It is the simplest kind of syntactic fractal; i.e. it is the smallest kind of self-similar binary-branching object whose non-terminal image on the line is neither the full line, nor a single point.

Its (Hausdorff) dimension is $\log_2$(Phi) ~.694

It is actually a multi-fractal (each unit of structure contains two copies of the whole, at different scales (½ & ¼).



Part VII

- Binary generators of binary
- Specifier-head-complement linearization of X-bar is in the set of such with lowest static ambiguity.
- And may have lowest dynamic ambiguity.

Conditions for a “language-like” phrase structure system.
Binary alphabet; 1 $\rightarrow$ *x y*, 0 $\rightarrow$ *z w*; *x,y,z,w* in {0, 1, *}

A: Termination. At least one of x, y, z, w is terminal (/null). Thinking of these systems as a (highly abstract!) basis for something like language, we want them to be discrete, built around lexical atoms. Thus, at least one branch in the system must introduce a terminal.

B. Completeness. At least one of x..w is 1, and at least one is 0. In other words, each non-terminal occurs an infinite number of times as the tree is infinitely expanded.

C. Mixed Loop. 1 $\rightarrow$ ...0... and/or 0 $\rightarrow$ 1. Along the same lines as the above condition, we want to ensure that there are not two disjoint loops, with 1s only occurring under other 1s, 0s only occurring as descendants of 0s. If that were true, since there will only be one start symbol, one or the other loop will never be introduced, hence is spurious.

D.? Bi-complexity. Not both 1 $\rightarrow$ ...*... and 0 $\rightarrow$ ...*... We probably also want to rule out things that are unary-branching; it seems they can have absolute recoverability (at least if they are not just unary, but unidirectionally, branching), but they clearly lack the ‘semantic power’ of something like X-bar. For instance, they allow no mapping of constituents to predicate/argument structure such that both predicates and arguments can each contain further predicate/argument pairs. For the same reason, we will rule out systems including a rule 0/1 $\rightarrow$ * *, since that will produce a ‘double headed spine’.

The table below is all conceivable binary phrase structure systems; those ruled out as “unreasonable” are color-coded for the condition they violate.

What’s so special about this one, the Fibonacci grammar associated with Spec-Head-Comp X-bar?

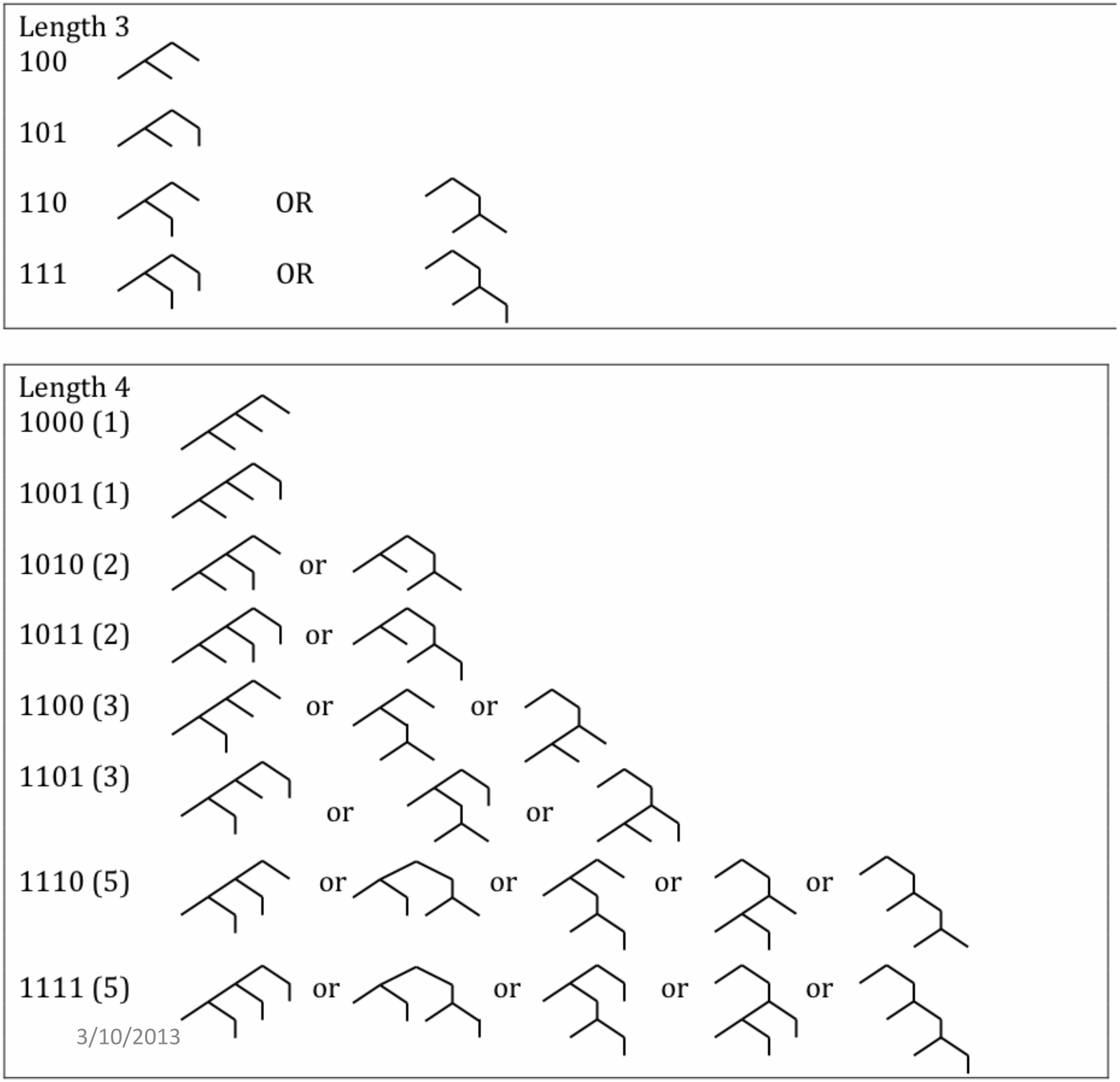
0: 0 0 1: 0 0	0: 0 1 1: 0 0	0: 0 * 1: 0 0	0: 1 0 1: 0 0	0: 1 1 1: 0 0	0: 1 * 1: 0 0	0: * 0 1: 0 0	0: * 1 1: 0 0	0: * * 1: 0 0
0: 0 0 1: 0 1	0: 0 1 1: 0 1	0: 0 * 1: 0 1	0: 1 0 1: 0 1	0: 1 1 1: 0 1	0: 1 * 1: 0 1	0: * 0 1: 0 1	0: * 1 1: 0 1	0: * * 1: 0 1
0: 0 0 1: 0 *	0: 0 1 1: 0 *	0: 0 * 1: 0 *	0: 1 0 1: 0 *	0: 1 1 1: 0 *	0: 1 * 1: 0 *	0: * 0 1: 0 *	0: * 1 1: 0 *	0: * * 1: 0 *
0: 0 0 1: 1 0	0: 0 1 1: 1 0	0: 0 * 1: 1 0	0: 1 0 1: 1 0	0: 1 1 1: 1 0	0: 1 * 1: 1 0	0: * 0 1: 1 0	0: * 1 1: 1 0	0: * * 1: 1 0
0: 0 0 1: 1 1	0: 0 1 1: 1 1	0: 0 * 1: 1 1	0: 1 0 1: 1 1	0: 1 1 1: 1 1	0: 1 * 1: 1 1	0: * 0 1: 1 1	0: * 1 1: 1 1	0: * * 1: 1 1
0: 0 0 1: 1 *	0: 0 1 1: 1 *	0: 0 * 1: 1 *	0: 1 0 1: 1 *	0: 1 1 1: 1 *	0: 1 * 1: 1 *	0: * 0 1: 1 *	0: * 1 1: 1 *	0: * * 1: 1 *
0: 0 0 1: * 0	0: 0 1 1: * 0	0: 0 * 1: * 0	0: 1 0 1: * 0	0: 1 1 1: * 0	0: 1 * 1: * 0	0: * 0 1: * 0	0: * 1 1: * 0	0: * * 1: * 0
0: 0 0 1: * 1	0: 0 1 1: * 1	0: 0 * 1: * 1	0: 1 0 1: * 1	0: 1 1 1: * 1	0: 1 * 1: * 1	0: * 0 1: * 1	0: * 1 1: * 1	0: * * 1: * 1
0: 0 0 1: * *	0: 0 1 1: * *	0: 0 * 1: * *	0: 1 0 1: * *	0: 1 1 1: * *	0: 1 * 1: * *	0: * 0 1: * *	0: * 1 1: * *	0: * * 1: * *

Static ambiguity of binary generators of binary understood as term-rewriting systems

- In terms of ambiguity of complete output strings (with * null), X-bar as understood above is one of 16 possibilities of its 'size' (the others are alternative linearizations of X-bar and the other 'reasonable' 3-type systems).
- In that group, the possibilities fall into three equivalence classes:
- The class containing the GS/X-bar form has the lowest ambiguity for the cases I've worked out: for a string length n, there are two unambiguous strings, two maximally ambiguous (full Catalan number of analyses) strings, and some number of intermediately-ambiguous strings.
- Another class (D-bar) accepts every string of a given length, and assigns the full Catalan number of analyses to each.
- The third class (Spine of Spines) accepts only a single string of each length, but assigns an infinite number of possible analyses to it.

Why SHC?

- Mirror order linearization (comp-head-specifier, CHS) of X-bar structures would yield strings which are *backwards* portions of the Golden String; thus SHC X-bar and CHS X-bar have equivalent structural ambiguity in a static sense, over complete strings.
- But SHC X-bar is more useful for *dynamic* recovery of structure from strings.
- It is easier to compare an incoming sequence, bit-by-bit against a known standard, from an invariant beginning up to a variable ending (as for SHC X-bar), rather than from variable ending backwards to invariant beginning (as for mirror-order CHS X-bar).
- On the other hand, 'mixed' X-bar linearizations (SCH, HCS) would fail to form the relevant golden sequence.

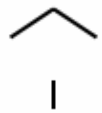


Left: strings accepted by Fibonacci grammar 1→ 10, 0 → 1 (~X-bar), and analyses assigned to each; lengths 3 and 4.

D-bar as term-rewriting system: full Catalan ambiguity.

$1 \rightarrow 0\ 0$

$0 \rightarrow 1$



This grammatical system accepts any binary string, and assigns the full Cataln number of possible analyses to each.

Length 3

000

001

010

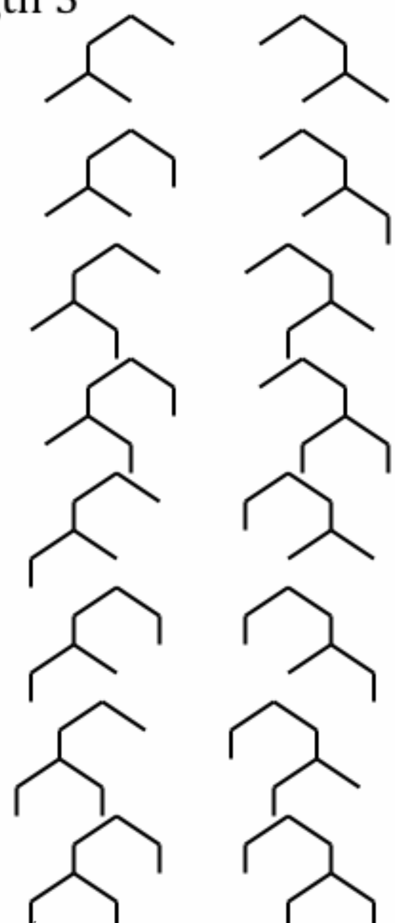
011

100

101

110

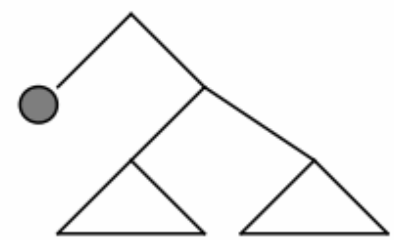
111



$2 \rightarrow 0\ 1$

$1 \rightarrow 2\ 2$

$$\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$$



Part VIII

- Catalogue of discrete infinite phrasal patterns

Discrete infinity

- It is an old insight that natural language is a system of “discrete infinity”
- von Humboldt’s “infinite use of finite means”
- The *finite means* being discrete atomic elements: words, morphemes, features.
- And *infinite use* implicating recursion.

Studying discrete infinity

- In what follows, I report some results obtained from a study of generalized discrete infinity.
- I examine the distinct binary-branching (Kayne 1984, 1994) recurrence patterns that could form the basis for discrete infinity.
  - i.e., self-similar arrangements of terminals and non-terminals
  - Terminals being the "discrete" part, non-terminals the "infinite" part.
- Goal: describe and classify the possibilities and their properties.

Alternative phrasal arrangements…

- What follows is concerned with showing that the X-bar schema has a lot to recommend it, when compared against other ways that phrases might be assembled.
- We'll therefore need to consider what else is possible -- what else could phrases look like?
  - Take phrases to be recursive 'recipes' for structure building.
  - Assume binary branching (Kayne 1984).
  - Phrases enable discrete infinity; concretely, they contain **terminals**, and **non-terminals**.

Maximal expansion: L-system treatment

- I will investigate the properties of different patterns by idealizing them as rigid local tree-building schemata,
- And seeing what happens when they are expanded maximally.
- This amounts to treating them as Lindenmayer (L-) systems, pioneered by Lindenmayer (1968) to investigate plant growth "algorithmic botany".
- L-systems are like familiar PSGs, but all rewrite rules apply obligatorily and simultaneously.
- I will consider patterns that are rigidly uniform (corresponding to deterministic context free (D0L) systems).
- There is a rich literature on this, much of it irrelevant to my work – largely because I ignore linear order, and much work on L-systems is concerned with words and word sequences.

Next: cataloguing the possibilities

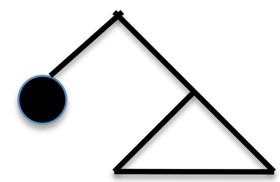
- With this in hand, let's turn to cataloguing the various possibilities for discrete infinite patterning.
- The possibilities are naturally partitioned by the number of non-terminal types they are defined over.
- The simplest class has a single non-terminal.

Simplest ‘molecule’ of structure: one level of embedding.

If we restrict possibilities to a single layer of syntactic combination, only one shape yields discrete infinity:

- The Spine, Phrase = [terminal Phrase].

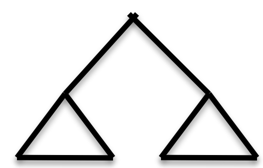
  1 → 0 1



The other naïve possibilities,

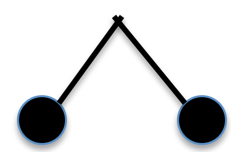
- Phrase = [Phrase Phrase]

  1 → 1 1



- or Phrase = [terminal terminal],

  1 → 0 0

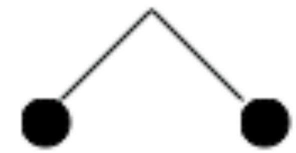


obviously could not serve as bases for a language-like system.

One non-terminal

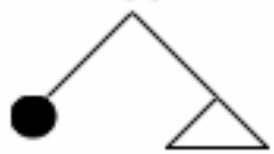
- There is really only one discrete infinite pattern with one non-terminal type: the **Spine**, below.
- The **Pair** (above right) is discrete but not infinite; the **Bush** (below right) is infinite but not discrete.

The Pair

PSR	Matrix	Tree
$1 \rightarrow 0\ 0$	$\begin{bmatrix} 0 \end{bmatrix}$	

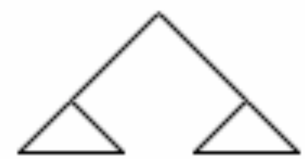
Characteristic polynomial: $x - 0$
Growth factor: 0

The Spine

PSR	Matrix	Tree
$1 \rightarrow 0\ 1$	$\begin{bmatrix} 1 \end{bmatrix}$	

Characteristic polynomial: $x - 1$
Growth factor: 1

The Bush

PSR	Matrix	Tree
$1 \rightarrow 1\ 1$	$\begin{bmatrix} 2 \end{bmatrix}$	

Characteristic polynomial: $x - 2$
Growth factor: 2

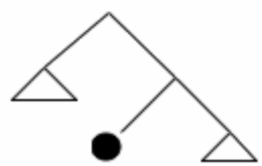
Two non-terminals

- D-bar and X-bar have "high-headed" variants -- really the same recurrence pattern, oriented differently with respect to the root.
- Again, the matrix formulation allows us to express this nicely; the related patterns have **similar** matrices, in the algebraic sense.

(41) X-bar

PSRs: $2 \rightarrow 2\ 1$, $1 \rightarrow 0\ 2$

Matrix: $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$



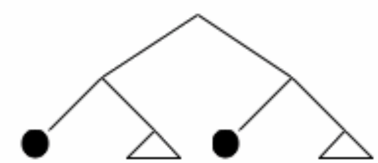
Characteristic polynomial: $x^2 - x - 1$

Growth factor: $\varphi \sim 1.618$

(43) D-bar

PSRs: $2 \rightarrow 1\ 1$, $1 \rightarrow 0\ 2$

Matrix: $\begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$



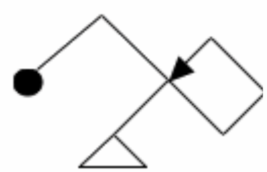
Characteristic polynomial: $x^2 - 2$

Growth factor: $\sqrt{2} \sim 1.414$

(42) High-headed X-bar

PSRs: $2 \rightarrow 0\ 1$, $1 \rightarrow 1\ 2$

Matrix: $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$



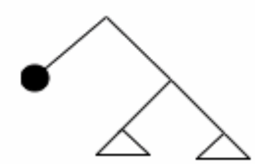
Characteristic polynomial: $x^2 - x - 1$

Growth factor: $\varphi \sim 1.618$

(44) High-headed D-bar

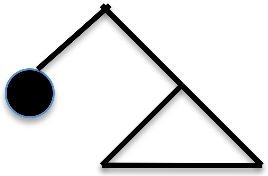
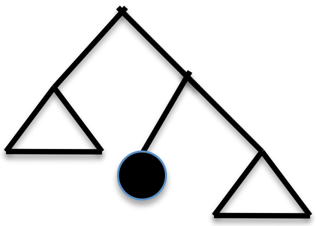
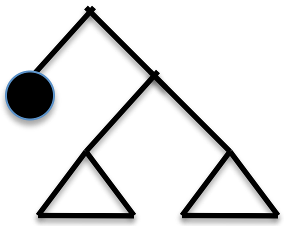
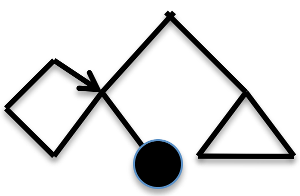
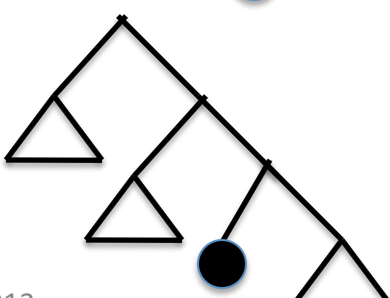
PSRs: $2 \rightarrow 0\ 1$, $1 \rightarrow 2\ 2$

$\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$



Characteristic polynomial: $x^2 - 2$

Growth factor: $\sqrt{2} \sim 1.414$

System	*Tree*	*Matrix*	*Recurrence relation*	*Growth Factor*
Spine		$\begin{bmatrix} 1 \end{bmatrix}$	$a_n = a_{n-1}$	1
X-bar		$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$	$a_n = a_{n-1} + a_{n-2}$	Phi ~ 1.618
D-bar		$\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$	$a_n = 2a_{n-2}$	√2 ~ 1.414
Spine of Spines		$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$	? (complicated)	1
3-bar		$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$	$a_n = a_{n-1}+a_{n-2}+a_{n-3}$	~1.839 “Tribonacci constant”

Finally, brief survey of 3 non-terminals

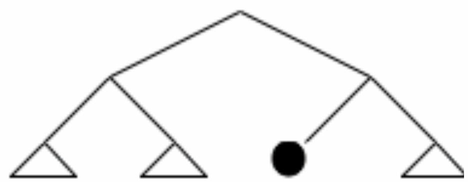
- Moving up the scale of pattern complexity, the next class (built with three kinds of non-terminal) has 57 distinct members.
- In the next slides, I illustrate just a few of these.

Further 3-type systems

- Here are some more examples from this class:

(57) 2 Power of 3

$$3 \rightarrow 1\ 2 \qquad 2 \rightarrow 0\ 3 \qquad 1 \rightarrow 3\ 3 \qquad \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

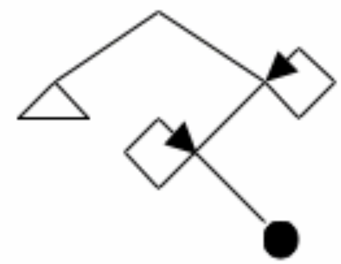


Characteristic polynomial: $x^3 - 3x = 0$

Growth factor: $\sqrt{3} = 1.732...$

(66) K Spine of Spines of Spines

$$3 \dashrightarrow 2\ 3 \qquad 2 \dashrightarrow 1\ 2 \qquad 1 \dashrightarrow 0\ 1 \qquad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

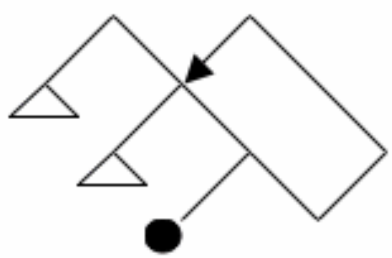


Characteristic polynomial: $x^3 - 3x^2 + 3x - 1$

Growth factor: 1

(63) 39

$$3 \dashrightarrow 2\ 3 \qquad 2 \dashrightarrow 1\ 3 \qquad 1 \dashrightarrow 0\ 2 \qquad \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

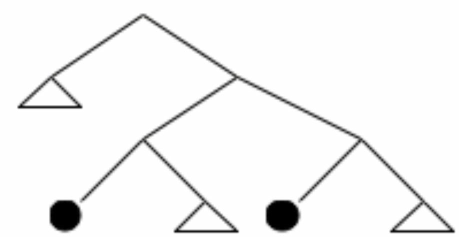


Characteristic polynomial: $x^3 - x^2 - 2x + 1$

Growth Factor: 1.8019…

(62) 29

$$3 \dashrightarrow 2\ 3 \qquad 2 \dashrightarrow 1\ 1 \qquad 1 \dashrightarrow 0\ 3 \qquad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & 0 & 0 \end{bmatrix}$$



Characteristic polynomial: $x^3 - x^2 - 2$

Growth Factor: 1.6956...

Prime growth factors by class

- # Non-terminals Values
- 1: (2), 1, (0)

  In this simplest kind of phrasal patterning, there is only one option, the Spine, with growth factor 1. I include in parentheses the "illegitimate" Bush, with growth factor 2, and Pair, with growth factor 0.
- 2: 1.61803, 1.41421
- With 2 non-terminals, we find two "prime" growth values, associated with the X-bar and D-bar families. I do not list again growth factors for "composite" systems like the Spine of Spines, as they are drawn from lower classes (in this case, the Spine of Spines has growth factor 1).
- 3: 1.83929, 1.80194, 1.76929, 1.75488, 1.73205, 1.69562 (*2), 1.58740, 1.52138, 1.46557, 1.32472, 1.25992
- These are the "new" growth values from the class beyond X-bar, including 3-bar (with the largest growth value, the tribonacci constant). Note that two distinct families of patterns have the same growth value (1.69562).

Growth factors with 3 non-terminals

Systems in family	Polynomial	Growth factor	Special notes
7, 32, 34	$x^3 - 2$	1.2599	$\sqrt[3]{2}$, non-Pisot
13, 25, 35	$x^3 - x - 1$	1.3247	Plastic number *rho*, the smallest Pisot #[23]
22, 28, 33	$x^3 - x^2 - 1$	1.4656	Pisot #
1, 31, 41	$x^3 - x - 2$	1.5214	Non-Pisot
5, 30, 37	$x^3 - 4$	1.5874	
3, 20, 29	$x^3 - x^2 - 2$	1.6956	Non-Pisot
4, 18,27	$x^3 - x^2 - 2$	1.6956	" "[24]
2, 38, 40	$x^3 - 3x$	1.7321	$\sqrt{3}$, Non-Pisot
17, 19, 26	$x^3 - 2x^2 + x - 1$	1.7548	Pisot #; plastic number rho squared
6, 12, 24	$x^3 - 2x - 2$	1.7693	Non-pisot #
11, 16, 39	$x^3 - x^2 - 2x + 1$	1.8019	Non-Pisot, three distinct real roots
9, 10, 21	$x^3 - x^2 - x - 1$	1.8393	Pisot #, the "tribonacci" constant

Growth factors for prime systems over 4 non-terminals

- 4: 1.92756, 1.92129, 1.91439, 1.90517, 1.89932, 1.89718, 1.89329, 1.88721, 1.88320, 1.87939*, 1.87371, 1.87018, 1.86676, 1.86371, 1.85356, 1.85163, 1.84776, 1.83509, 1.82462, 1.82105, 1.81917, 1.81712, 1.80843, 1.79891, 1.79632, 1.79431*, 1.79004, 1.78537, 1.74840, 1.74553, 1.72775, 1.72534, 1.72208, 1.71667, 1.71064, 1.70211, 1.69028, 1.68377, 1.68179, 1.67170, 1.66980, 1.65440, 1.65289, 1.64293, 1.60049, 1.56638, 1.55898, 1.55377, 1.54369, 1.51288, 1.49453, 1.49022, 1.44225, 1.39534, 1.38028, 1.35321, 1.27202, 1.22074, 1.18921
- [59 distinct]
- These are the new “prime” growth factors from the set of phrasal patterns with 4 non-terminal types. Collecting them allows a check on whether the set of systems has been fully reduced (eliminating redundancies and degeneracies); “prime” values should appear in a number of systems that is a multiple of 4 (cf remarks above on the “doubled” value 1.69562 in the three-types set, appearing in two distinct families; it shows up with 6 systems rather than 3, as for the other values).

Count: all the prime systems below occur in multiples of four; only those multiply represented are marked as such (x8 or x12).

Eigenvalue (growth factor)	Polynomial(s)	Count/notes
1.92756	$x^4-x^3-x^2-x-1$	Pisot, limit point
1.92129	$x^4-x^3-x^2-2x+1$	
1.91439	$x^4-x^3-2x^2+x-1$	
1.90517	$x^4-x^3-2x^2+1$	Pisot; "the smallest limit point of the set of univoque Pisot numbers" x8
1.89932	x^4-2x^2-2x-2	non-uniformly discrete spectrum
1.89718	$x^4-2x^3+x^2-x-1$	
1.89329	x^4-2x^2-3x	
1.88721	x^4-3x^2-2	
1.88320	$x^4-2x^3+x^2-2x+1$	Salem
1.87939	$x^4-x^3-3x^2+2x+1$ & x^4-3x^2-x	This value – the 9-gon system – is the only one in this class with two distinct polynomial representations. x8
1.87371	x^4-x^3-2x-2	x8; non-uniformly discrete spectrum
1.87018	x^4-3x^2-2x+2	
1.86676	x^4-2x^3+x-1	Pisot, limit point. x12

- k

1.86371	x^4-x^3-3x	
1.85356	$x^4-x^3-x^2-2$	x8
1.85163	x^4-x^3-4x+2	
1.84776	x^4-4x^2+2	(octagon system)
1.83509	x^4-4x-4	
1.82462	$x^4-x^3-2x^2+2x-2$	x8
1.82105	$x^4-x^3-x^2-2x+2$	x8
1.81917	$x^4-3x^3+3x^2-x-1$	
1.81712	x^4-6x	
1.80843	x^4-x^2-3x-2	
1.79891	x^4-2x^2-4	x8
1.79632	x^4-x^2-4x	x8
1.79431	x^4-x^3-2x-1	
1.79004	$x^4-2x^3+x^2-2$	x12
1.78537	x^4-2x^2-x-2	
1.74840	x^4-x^3-4	x12
1.74553	x^4-x^2-3x-1	
1.72775	x^4-4x-2	
1.72534	x^4-x^3-x-2	x8

• L

1.72208	$x^4-x^3-x^2-x+1$	Salem number
1.71667	x^4-2x^3+2x-2	
1.71064	x^4-2x^2-x-1	
1.70211	x^4-3x^2-x+2	
1.69028	$x^4-x^3-2x^2+2x-1$	
1.68377	x^4-2x^2-2x+1	
1.68179	x^4-8	
1.67170	x^4-x^2-3x	x8
1.66980	$x^4-x^3-x^2+x-2$	x8
1.65440	x^4-x^3-3x+2	
1.65289	x^4-2x^2-2	x8
1.64293	x^4-2x-4	x8
1.60049	x^4-x^2-4	
1.56638	x^4-x^2-x-2	
1.55898	x^4-x^3-2x+1	
1.55377	x^4-2x^2-1	
1.54369	x^4-x^3-2	x12
1.51288	$x^4-x^3-x^2+x-1$	x8
1.49453	x^4-2x-2	x8
1.49022	x^4-2x^2-x+1	
1.44225	x^4-3x	
1.39534	x^4-2x-1	
1.38028	x^4-x^3-1	Pisot – next smallest after plastic number rho
1.35321	x^4-x-2	
1.27202	x^4-x^2-1	
1.22074	x^4-x-1	
1.18921	x^4-2	

Part XI

- Matrix treatment of phrasal patterns
  - Maximal growth is iterated matrix multiplication
  - Growth factor as dominant eigenvalue
  - Characteristic polynomials

A notion of syntactic "growth"

- As we saw on the last slide, the X-bar pattern "grows" more nodes per line than the alternative (HH D-bar) as it is expanded.
- I've explored elsewhere some reasons to think that faster growth in this sense is a desirable property (all else equal); I won't review that here.
- For present purposes, let's just assume that growth in this sense is something worth investigating.
- How can we quantify this notion of growth, and what are the growth properties of the conceivable discrete infinite recurrence patterns?

Growth factor

- Intuitively, we want to find a “growth factor” G for each pattern.
- This number describes how the number of nodes on one line of the tree relates to the number of nodes on the previous line.
- We take G to be (basically) the limit of the ratio of the number of nodes on line *n*, to the number of nodes on line *n*-1, as *n* gets large.
  – Thus, nodes (n) = G * nodes (n-1)

G is the **dominant eigenvalue** of the phrasal recurrence matrix.

- Here, expressing phrasal recurrence patterns as matrices brings its first rewards.
- Matrices can be interpreted in several different ways; a natural and important interpretation is as a linear mapping.
- Under this interpretation, the *n* x *n* (square) matrices we'll be considering (expressing how *n* kinds of non-terminals link to each other), transform a point in *n*-dimensional space into another point in *n*-space.

Phrasal growth ~ iterated matrix multiplication

- Take **A** to be the relevant phrase structure matrix
- Take $\mathbf{x_i}$ to be a column vector expressing the number of each kind of non-terminal on the *i*th line of the tree.
- (we identify the non-terminals with the coordinate axes: the number of non-terminals of a given type is expressed as distance along the associated axis).
- Then maximal expansion of the pattern is simply iterated matrix multiplication.

$$\mathbf{A\,x_i = x_{i+1}}$$

Syntactic growth is iterated mapping

- The syntactic problem we have been considering (how do phrasal patterns grow?)
- Now resolves as a geometric one:
- Given some input vector – a point in *n*-space,
- Where does that vector go as the mapping iterates?
- Thinking of things this way lets us see why G is the dominant eigenvalue.

Eigenvalues and eigenvectors

- An important property of a square matrix is its set of eigenvectors and eigenvalues.
- In general, the transformation of *n*-space induced by matrix multiplication is quite complicated.
- The eigenvectors represent points of stability amidst the complexity of the mapping:
- They are the vectors that, under the transformation, retain their direction.
- i.e., for eigenvector $v = ax + by + cz\ldots$

$$\mathbf{Av} = \lambda \mathbf{v} \qquad (= \lambda ax + \lambda by + \lambda cz\ldots)$$

- The scaling factor λ is the eigenvalue associated with that eigenvector.

Why G is the dominant eigenvalue

- Suppose the starting vector is *ax + by* …
- We can rewrite it in terms of the eigenvectors (a standard and powerful technique):

  $ax + by \ldots = cv_1 + dv_2 \ldots$

  v_i an eigenvector with eigenvalue λ_i
- Multiplication by the matrix *n* times has a particularly nice expression in terms of the eigenvectors:

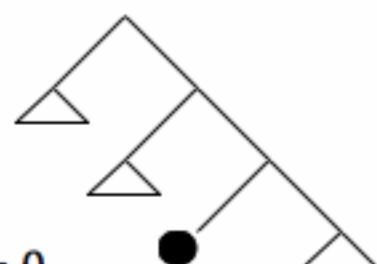
  $\boldsymbol{\lambda_1^n c v_1 + \lambda_2^n d v_2 \ldots}$
- Suppose λ_1 is the largest (i.e. dominant) eigenvalue; as *n* increases, the sum of component vectors converges on $\boldsymbol{\lambda_1^n c v_1}$ (for non-zero c).
- Thus, $x_n \sim \boldsymbol{\lambda_1} x_{\mathbf{n-1}}$; λ_1 is the desired quantity G.

Best growth: generalized X-bar

- The highest growth factor in the 2 non-terminal class belongs to X-bar: the golden mean, associated with the Fibonacci numbers.
- The largest growth factor with three non-terminals is the "tribonacci constant", in the generalized X-bar format in this class (an X-bar like pattern with two specifiers per phrase).

(58) 9_3-bar (Generalized X-bar format with two specifiers)

$$\begin{array}{l} 3 \rightarrow 2\ 3 \\ 2 \rightarrow 1\ 3 \\ 1 \rightarrow 0\ 3 \end{array} \quad \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$



Characteristic polynomial: $x^3 - x^2 - x - 1 = 0$

Growth factor: the "Tribonacci" constant, ~1.839...

Composite systems

- With two non-terminals, we find our first examples of "composite" systems:
- Patterns that have "smaller" subpatterns (i.e., subtrees generated with less than the full set of non-terminals)

(45) Spine of Spines

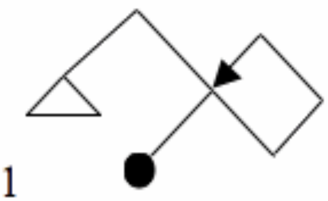
PSRs:

$2 \rightarrow 2\ 1$

$1 \rightarrow 0\ 1$

Matrix:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$



Characteristic polynomial: $x^2 - 2x + 1$

Growth factor: 1

(46) Spine of Pairs

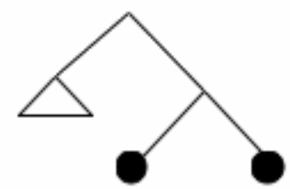
PSRs:

$2 \rightarrow 2\ 1$

$1 \rightarrow 0\ 0$

Matrix:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$



Characteristic polynomial: $x^2 - x$

Growth factor: 1

Factorization of composite systems

- For example, consider the Spine of Pairs (top right).
- This pattern is composed of Pairs (bottom right) substituted within a Spine (center right).
- Its polynomial is the product of the polynomials of its components:

  $x^2 - x = (x - 1)*x$
- Its roots are those of its components; G is the largest among these roots.

(46) Spine of Pairs

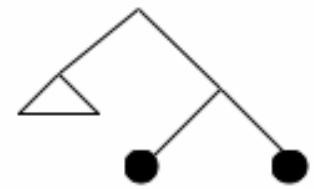
PSRs:

$2 \rightarrow 2\ 1$

$1 \rightarrow 0\ 0$

Matrix:

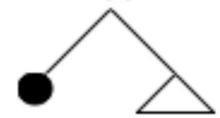
$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$



Characteristic polynomial: $x^2 - x$

Growth factor: 1

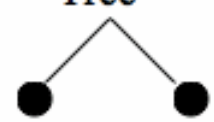
The Spine

PSR	Matrix	Tree
$1 \rightarrow 0\ 1$	$\begin{bmatrix} 1 \end{bmatrix}$	

Characteristic polynomial: $x - 1$

Growth factor: 1

The Pair

PSR	Matrix	Tree
$1 \rightarrow 0\ 0$	$\begin{bmatrix} 0 \end{bmatrix}$	

Characteristic polynomial: $x - 0$

Growth factor: 0

Growth in composite systems

- Composite systems are composed of simpler patterns, one substituted inside another.
- (This is opposed to "prime" patterns, irreducible in terms of simpler patterns)
- How does the growth of the larger pattern relate to the growth of its component sub-patterns?
- Here again, the matrix formulation provides the answer:
- **The growth factor of the larger pattern is just the largest of the growth factors among its components.**
- This is so, because the characteristic polynomial of a composite system is the product of the polynomials of its component systems.
- When multiplying polynomials, roots are preserved.

Even more abstract

- A further abstraction will help in understanding this (and other) pattern(s).
- Namely, we can represent the recurrence relations by means of a ***matrix***:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

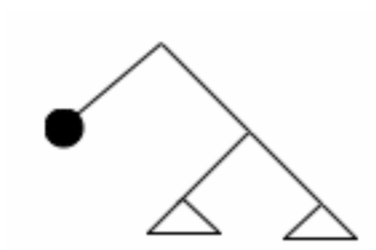
- The rows and columns are associated with the distinct kinds of non-terminal.
- The rows correspond to non-terminal inputs to the PSRs; the columns represent non-terminals in the output of each PSR.

An alternative

- Let's compare the X-bar pattern with a superficially very similar pattern,
- Which also constructs each phrase from a terminal and two further phrases.
  - X-bar: Phrase = [Phrase [terminal Phrase]]
  - **HH D-bar: Phrase = [terminal [Phrase Phrase]]**

$XP \rightarrow X^0\ X'$

$X' \rightarrow XP\ XP$



$$\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$$

Same ingredients, different recipe: different result.

- Consider what happens as we "inflate" these patterns, expanding them maximally:

2) 9 generations of X-bar (Phrase → [Phrase [terminal Phrase]])

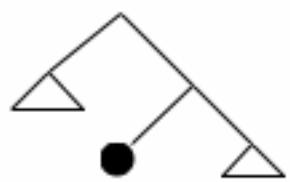
3) 9 generations of HH D-Bar (Phrase → [terminal [Phrase Phrase]])

Back to comparing X-bar and HH D-bar

- This insight lets us capture the difference in growth between X-bar, and the HH D-bar alternative, very simply and directly.

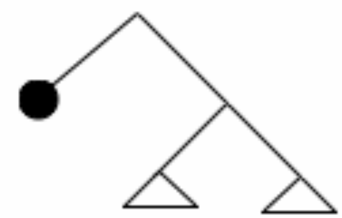
X-bar:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$



HH D-bar:

$$\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$$



G: $\varphi \sim 1.618$ (the golden mean) $\quad \sqrt{2} \sim 1.414$

An example

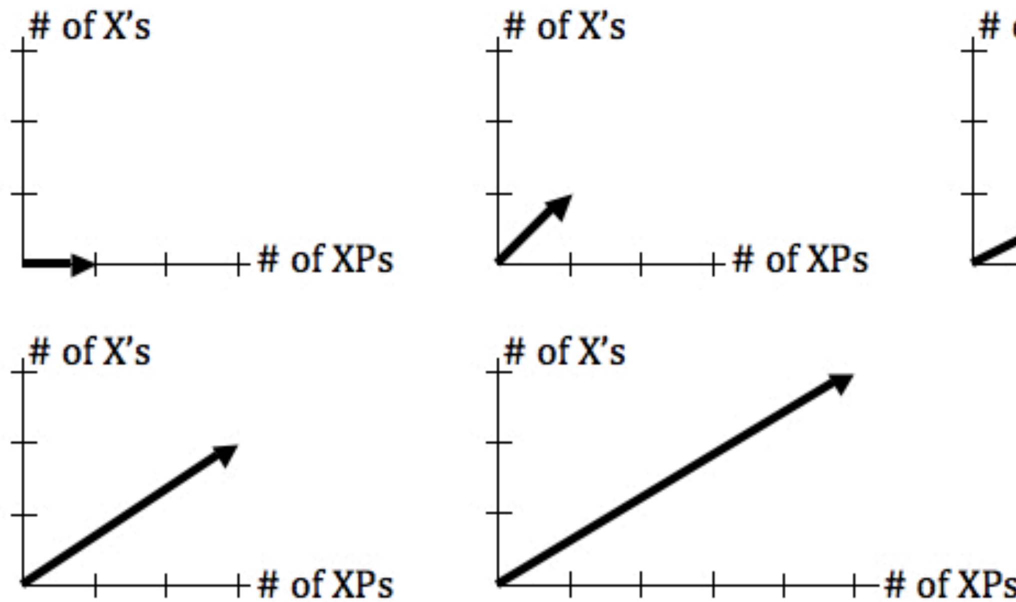
- Growth of the X-bar pattern in these terms:
- At the root, there is a single XP-type non-terminal; thus the initial vector x_0 is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

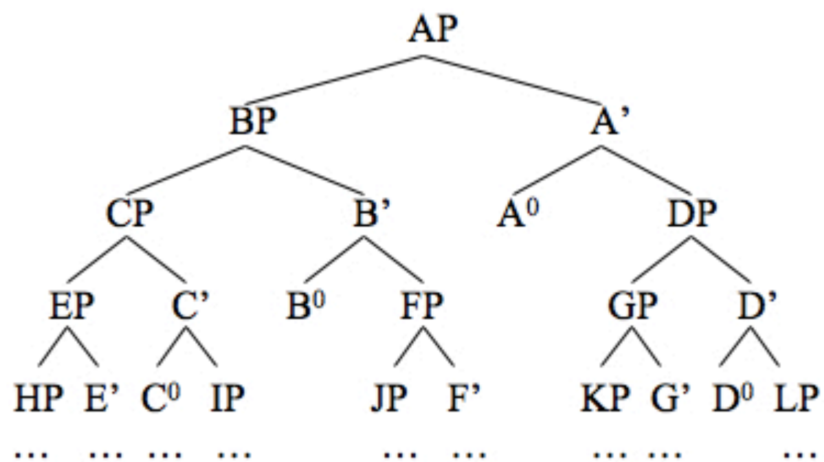
$$\mathbf{A\, x_0 = x_1} \qquad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- That is, the next line (x_1) contains one XP-type non-terminal, and one X'-type.
- Continuing, we get the following sequence of vectors, representing the number of non-terminals on successive lines of the tree:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} \begin{bmatrix} 8 \\ 5 \end{bmatrix} \begin{bmatrix} 13 \\ 8 \end{bmatrix} \begin{bmatrix} 21 \\ 13 \end{bmatrix} \begin{bmatrix} 34 \\ 21 \end{bmatrix}$$

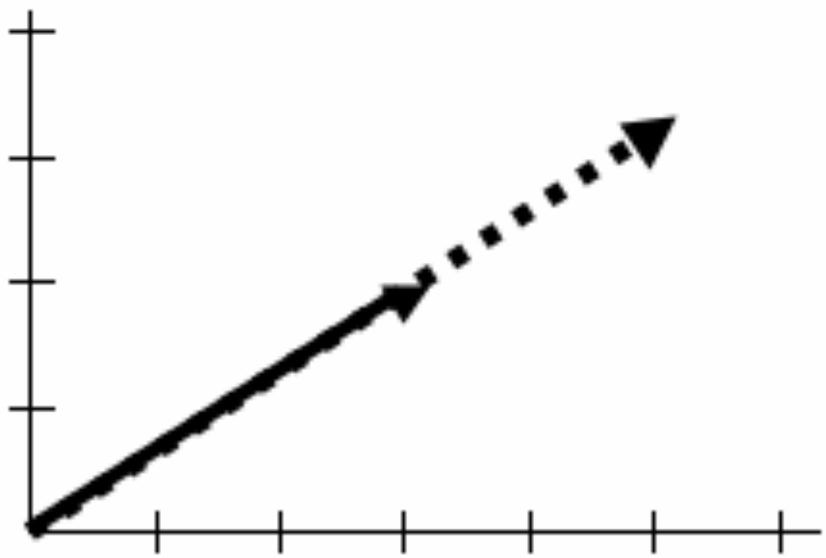
X-bar growth illustrated with vectors





XP	X'
1	0
1	1
2	1
3	2
5	3
Fib(n)	Fib(n-1)

X-bar (dominant) eigenvector



$$Ax = \lambda x \qquad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1.618... \\ 1 \end{bmatrix} = \begin{bmatrix} 2.618... \\ 1.618... \end{bmatrix} = 1.618... \begin{bmatrix} 1.618... \\ 1 \end{bmatrix}$$

One more bit of math

- Associated with each square matrix is a **characteristic polynomial**.
- Among other important properties, the characteristic polynomial has as its **roots** (solutions when it's set equal to 0) the **eigenvalues** of the matrix.
- The X-bar pattern has characteristic polynomial $x^2 - x - 1$; for HH D-bar, it's $x^2 - 2$.

Characteristic polynomials and linear recurrence relations.

- The X-bar phrasal form generates the Fibonacci numbers 1,1,2,3,5,8,13...
- In numbers of each type of non-terminal, on successive lines of the tree.
- For example,
  - there is 1 XP at the root,
  - 1 XP on the next line (its Spec),
  - 2 XPs on the line after that (Comp, and Spec of Spec),
  - 3 on the next line (Spec of Spec of Spec, comp of Spec, spec of Comp)
  - 5 on the next (Spec of Spec of Spec of Spec, Comp of Spec of Spec, comp of comp, spec of comp of spec, etc.
- Fib numbers obey the linear recurrence $\mathbf{a_n = a_{n-1} + a_{n-2}}$.
- Characteristic polynomial of the X-bar matrix is x^2-x-1.
- Setting equal to zero and manipulating a bit, this is $\mathbf{x^n = x^{n-1} + x^{n-2}}$.
- The characteristic polynomial encodes the linear recurrence governing the count of categories on successive lines, with indices in the additive recurrence corresponding to powers in the polynomial.
- This is quite general, holding as well for other patterns.

Part X

- Spectral classes
  - Classifying phrasal matrices in terms of their set of eigenvalues
  - Endocentric
  - Pisot
  - Polygonal

Special members of the infinite zoo

- As we can see, as more non-terminals are allowed, ever more patterns become available.
- Is there anything of interest to be said, beyond mere "stamp-collecting"?
- Two special classes:
  - **Endocentric** (generalized X-bar systems).
    - Have highest growth factor for given # of non-terminals.
    - Characteristic polynomial $x^n - x^{n-1} - \ldots - x^1 - x^0$.
    - Head, complement, and some number of specifiers; complement and specifiers isomorphic to the root.
  - **Polygonal** systems (see next slide)

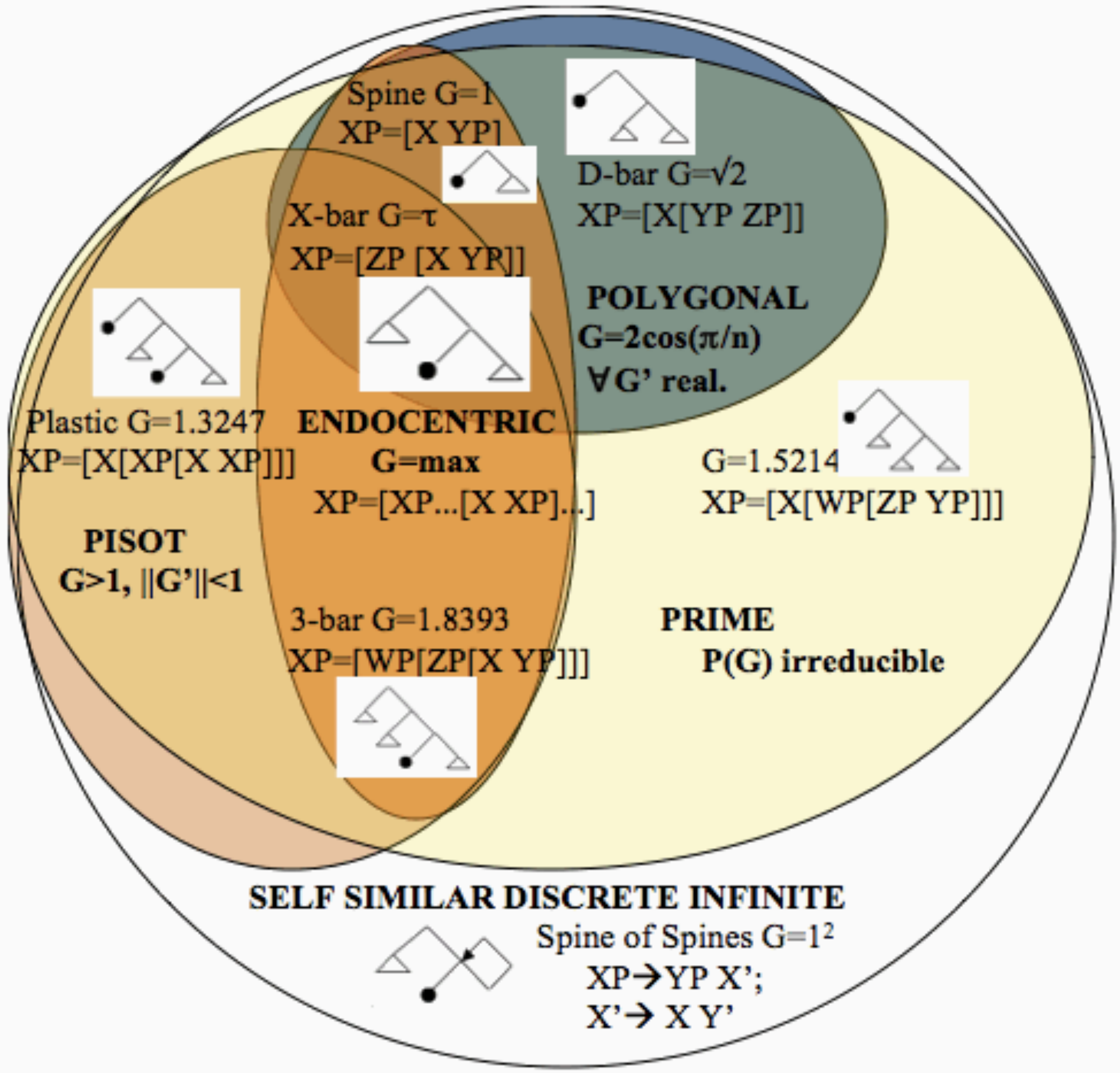


Figure 1. Intersection of spectral classes of phrasal recurrence patterns. Examples of patterns in each of the classes are indicated. The X-bar schema, whose dominant eigenvalue G is the so-called Golden Number τ, is the only pattern lying in all three of the special classes described here (it is Endocentric, with maximal G for its degree; of Pisot type, with unique structural purity, and Polygonal, with real growth).

- Let us call the *degree* of a phrasal pattern the degree of its characteristic polynomial. Let *G* represent the dominant (Perron-Frobenius) eigenvalue (i.e., the largest, necessarily real, root of the characteristic polynomial), and *G′* stand for an arbitrary Galois conjugate (a distinct eigenvalue; equivalently, a distinct root of the characteristic polynomial). Among the "Prime" systems (those whose matrix forms have irreducible characteristic polynomials), the three classes of interest are as follows:

- *The Endocentric class*. These forms can be described as generalized X-bar schemata; intuitively, each combines a head at the deepest level with a complement phrase, and then combines the result with some number of specifier phrases one at a time, to make a full phrase. These forms have the largest G for systems of their degree; there is one such system of each degree (up to permutation of which non-terminal is chosen as the root). As discussed in Medeiros (2008, 2012), such patterns are likely to be of significance, in light of the thesis of "economy of command" developed in those works. Relative to other patterns of their degree, these structural formats support the minimum number of c-command relations.

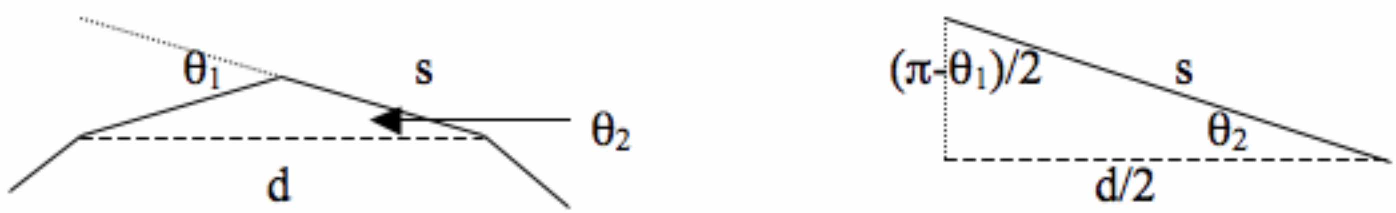
- *The Pisot class*. These patterns have a G that is a so-called Pisot-Vijayaraghavan number, an algebraic integer (i.e. solution to a polynomial with integer coefficients) that is strictly greater than 1, and whose Galois conjugates (here, the various G') are all of magnitude strictly less than 1. Although discovered only in the 20th century, these numbers have been the focus of considerable interest in domains like number theory, harmonic analysis, and crystallography (see for example the works collected in Moody 1997). In syntactic terms, these patterns are likely to be significant because they have a kind of structural purity; all eigenvectors (interpreted as a particular structural 'theme', a stable ratio among non-terminals) save the dominant one vanish as the pattern is grown. All non-Pisot systems, in turn, have infinite growing 'warts' of structure distinct from the dominant theme.

- *The Polygonal class.* Finally, there is a class of patterns whose G is of the form 2cos(π/n). These forms are polygonal: their G is the ratio of the shortest internal diagonal to a side in a regular polygon. They, and only they, have all G' that are real numbers, and all and only their matrices are diagonalizable (i.e., similar to a diagonal matrix, with all entries off the main diagonal 0). Diagonalizable matrices are significant in a number of applications, in part because they are particularly "well-behaved"; the non-diagonalizable matrices corresponding to all non-polygonal matrices are called defective. The odd polygonal systems (whose G relates to the geometry of a polygon with an odd number of sides) furthermore have symmetric matrices (i.e., where arbitrary element $a_{ij} = a_{ji}$). These are a special subcase of Hermitian matrices (equal to their conjugate transpose); the matrix operators that represent physical observables in quantum theory are always Hermitian. In terms of the interpretation of eigenvectors as stable syntactic configurations, in polygonal systems the growth of the pattern reflects real scaling of each of its components, thus an inhomogeneous dilation. On the other hand, all non-polygonal systems have eigenvectors (stable configurations) associated with some complex G' whose growth involves a kind of rotation.

Polygonal phrasal patterns

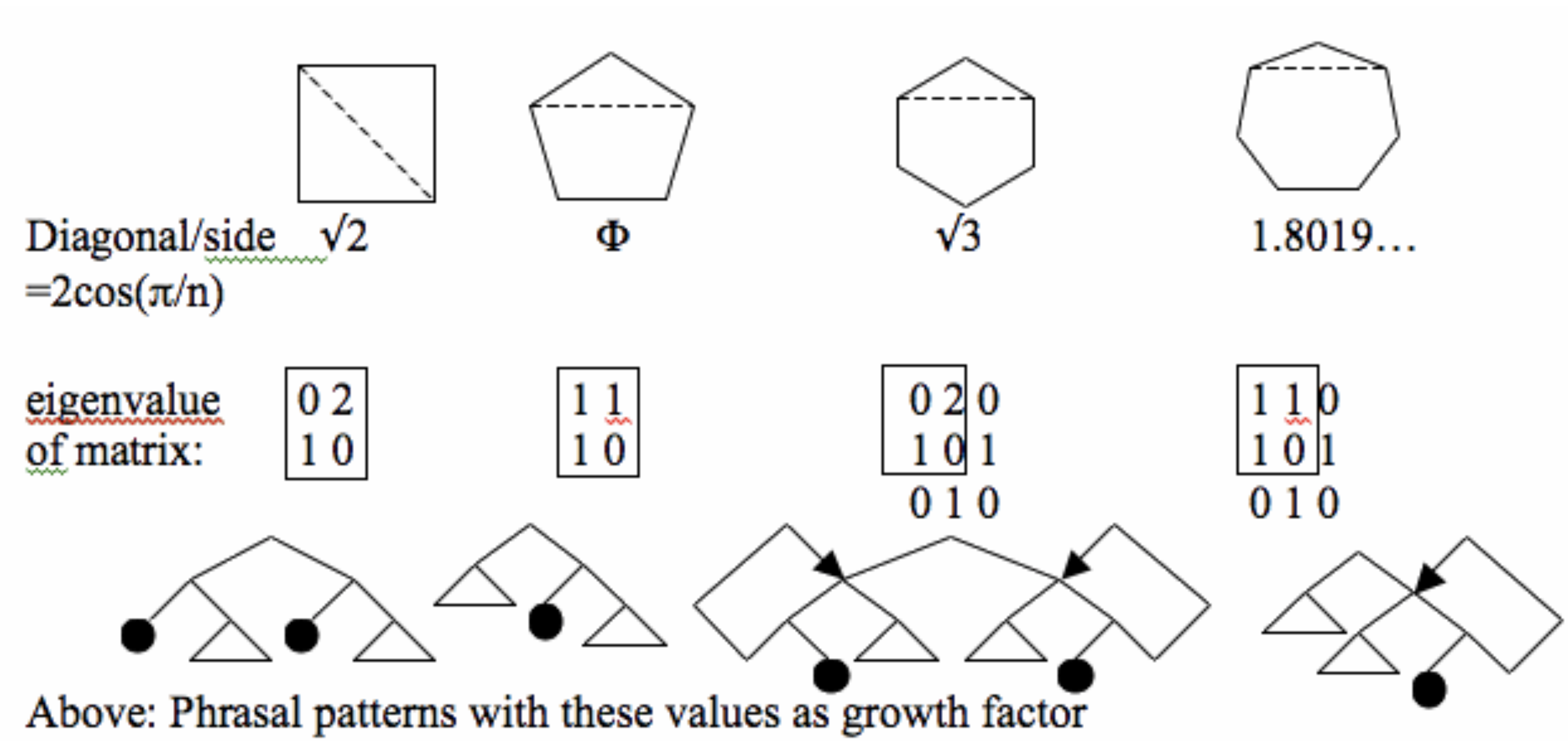
- There is a very special class of 'polygonal' prime patterns, whose G is of the form 2cos(π/n).
- This number expresses the ratio of the shortest internal diagonal to a side, in a regular n-gon.

To see why, note that the exterior angle θ_1 is $2\pi/n$, because the exterior angles for the whole n-gon sum to a complete circle. Then the interior angle is $\pi - 2\pi/n$ (the exterior and interior angles together form a straight angle). Half of that angle $(\pi/2 - \pi/n)$ forms the complement of angle θ_2. Thus $\theta_2 + \pi/2 - \pi/n = \pi/2$, and so $\theta_2 = \pi/n$.



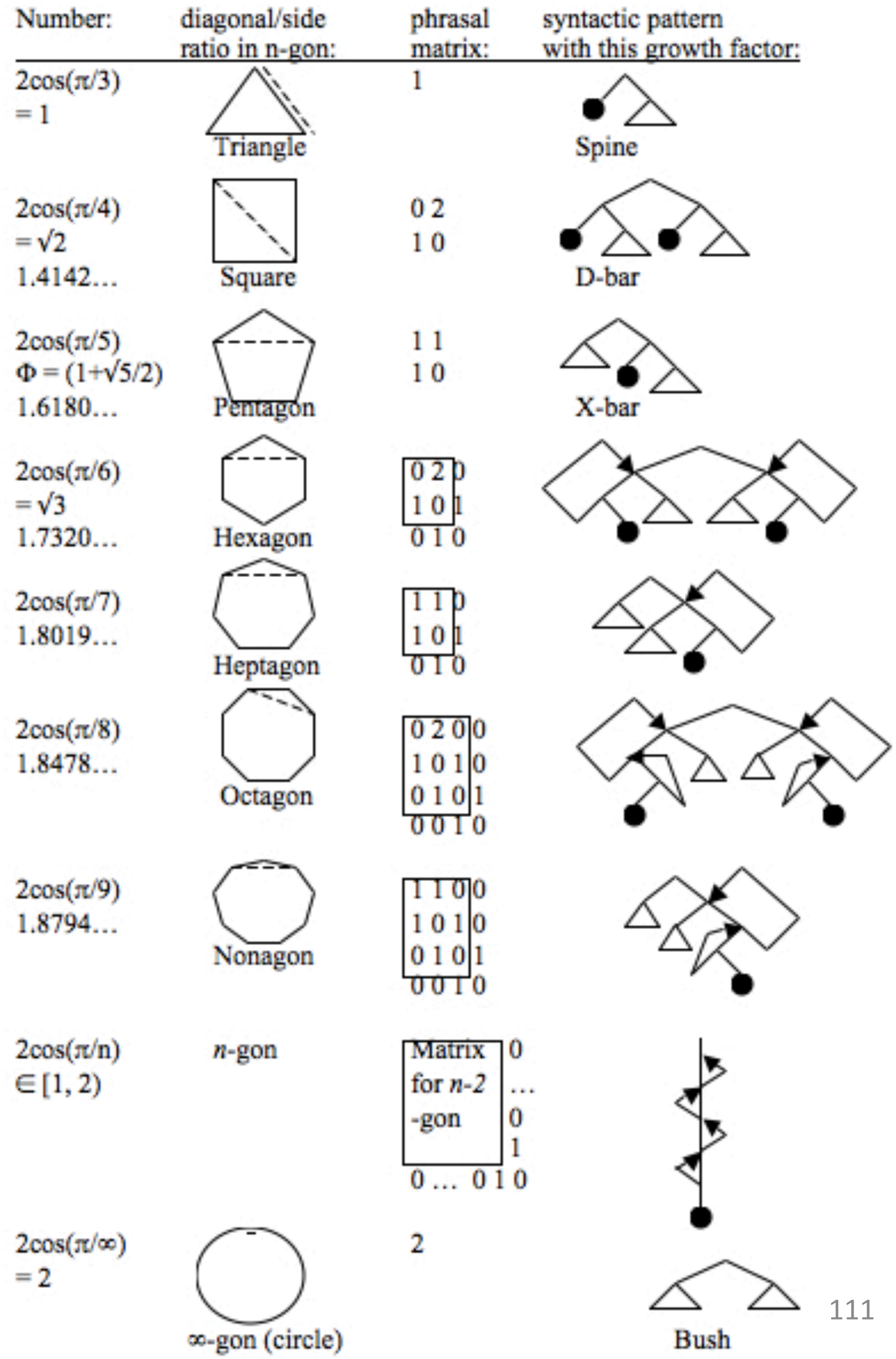
The cosine of angle $\theta_2 = \cos(\pi/n)$ then expresses the ratio of half a diagonal (d/2) to a side (s; see the triangle at right above); we double this to get the diagonal-to-side ratio stated above, $2*\cos(\pi/n)$.

Some polygonal phrasal forms



Polygonal phrasal patterns

- The diagram at right gives several representations of these patterns.
  - The growth factor G
  - Relevant polygon
  - Matrix expressing syntactic recurrence among non-terminals
  - Tree diagram



Diagonalizable iff polygonal

- It turns out that alone among all the patterns we get,
- All and only the polygonal ones have **diagonalizable** matrices.
  - Ones *similar* to a diagonal matrix,
    - where all entries off the main diagonal are 0.
  - Matrix A is diagonalizable iff $A = PDP^{-1}$.
  - If so, then $\mathbf{A}^{\mathbf{n}} = P\mathbf{D}^{\mathbf{n}}P^{-1}$.
- Diagonalizability of matrices is considered an important property in many applications.

Symmetric (Hermitian) iff odd polygonal

- Moroever, the *odd* polygonal systems (those with growth factor related to a regular polygon with an odd number of sides: triangle, pentagon, heptagon, etc.)
- Have symmetric matrices ($a_{ij} = a_{ji}$).
  - This is an even nicer property, also important in physics.
- A symmetric matrix with real entries is a special case of a Hermitian matrix.
  - In quantum theory, physical observables always correspond to Hermitian operators.

Symmetric and 'almost symmetric' matrices in reorientations of polygonal systems

Octagon: (notice, "almost" symmetric: 2 corresponds to 1, but otherwise symmetric)

$$\begin{bmatrix} 0&2&0&0 \\ 1&0&1&0 \\ 0&1&0&1 \\ 0&0&1&0 \end{bmatrix} \quad \begin{bmatrix} 0&1&1&0 \\ 2&0&0&0 \\ 1&0&0&1 \\ 0&0&1&0 \end{bmatrix} \quad \begin{bmatrix} 0&1&1&0 \\ 1&0&0&1 \\ 1&0&0&0 \\ 0&2&0&0 \end{bmatrix} \quad \begin{bmatrix} 0&1&0&0 \\ 1&0&1&0 \\ 0&1&0&1 \\ 0&0&2&0 \end{bmatrix}$$

Nonagon:

$$\begin{bmatrix} 1&1&0&0 \\ 1&0&1&0 \\ 0&1&0&1 \\ 0&0&1&0 \end{bmatrix} \quad \begin{bmatrix} 0&1&1&0 \\ 1&1&0&0 \\ 1&0&0&1 \\ 0&0&1&0 \end{bmatrix} \quad \begin{bmatrix} 0&1&1&0 \\ 1&0&0&1 \\ 1&0&0&0 \\ 0&1&0&1 \end{bmatrix} \quad \begin{bmatrix} 0&1&0&0 \\ 1&0&1&0 \\ 0&1&0&1 \\ 0&0&1&1 \end{bmatrix}$$

Real symmetric (Hermitian)

Characteristic polynomials of polygonal systems

- There is a nice relationship between the polynomials for these patterns and the so-called **Lucas** and **Fibonacci polynomials**.
- These are defined similarly to the additive series of the same name:
- Fibonacci: 1, 1, 2, 3, 5, 8, 13...
- Lucas: 2, 1, 3, 4, 7, 11, 19, 28....
- Different "seeds" fed to the same recurrence relation $a_n = a_{n-1} + a_{n-2}$

Polynomial recurrence: $P_n = x*P_{n-1} + P_{n-2}$

- Lucas Polynomials:

2

x

$x^2 + 2$

$x^3 + 3x$

$x^4 + 4x^2 + 2$

$x^5 + 5x^3 + 5x$

$x^6 + 6x^4 + 9x^2 + 2$

…

- Fibonacci Polynomials:

1

x

$x^2 + 1$

$x^3 + 2x$

$x^4 + 3x^2 + 1$

$x^5 + 4x^3 + 3x$

…

Lucas polynomials and even polygonals: $P_n = \mathrm{Alt}(L_n)$

Lucas polynomials:

2

x

$x^2 + 2$

$x^3 + 3x$

$x^4 + 4x^2 + 2$

$x^5 + 5x^3 + 5x$

$x^6 + 6x^4 + 9x^2 + 2$

…

Characteristic polynomial for even polygonal patterns:

$x^2 – 2$

$x^3 – 3x$

$x^4 – 4x^2 + 2$

$x^5 – 5x^3 + 5x$

$x^6 – 6x^4 + 9x^2 – 2$

…

Fibonacci polynomials and odd polygonals: $P_n = \text{alt}(F_n) - \text{alt}(F_{n-1})$

Fibonacci Fn	Characteristic polynomials of odd polygonal phrasal patterns:	
1		
x	$x - 1$	Spine, triangle
$x^2 + 1$	$x^2 - x - 1$	X-bar, pentagon
$x^3 + 2x$	$x^3 - x^2 - 2x + 1$	heptagon
$x^4 + 3x^2 + 1$	$x^4 - x^3 - 3x^2 + 2x + 1$	nonagon
$x^5 + 4x^3 + 3x$	$x^5 - x^4 - 4x^3 + 3x^2 + 3x - 1$	undecagon
…		

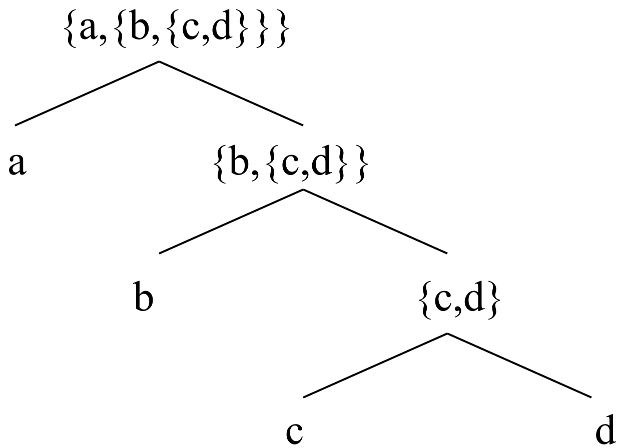
Part XI

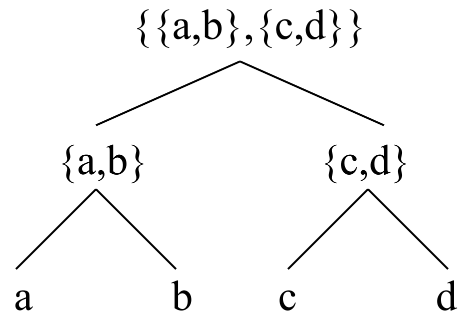
- Economy of command in phrase structure
  - Endocentric structure minimizes c-command relations

Optimal packing… in syntax?

- In phyllotaxis, the dominant Fibonacci pattern provides a dynamically optimal packing solution, spreading new elements as far apart as possible as the plant grows.
- Medeiros (2008, 2012) develops the idea that the Fibonacci-related X-bar schema is a dynamically optimal solution to a derivational problem: **Minimizing c-command relations**.
- C-command relations are the scaffolding for long-distance dependencies of several kinds (binding, agreement, linearization, scope).
- Thinking of c-command as a search process, trees with fewer/shorter c-command relations are preferred; they minimize the space searched.
- Familiar concerns of *locality* in syntax are extended beyond choosing the shortest available dependency (the usual way of thinking); instead, locality directly informs structure-building.

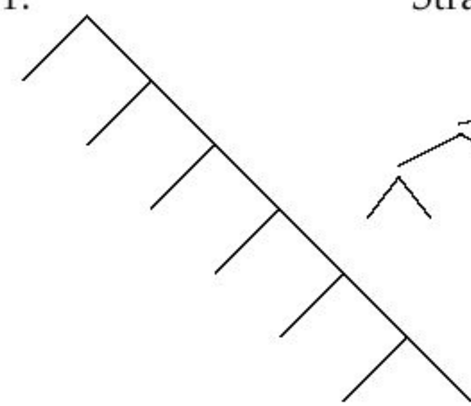
"Optimal packing" in phrase structure: C-command & Dominance



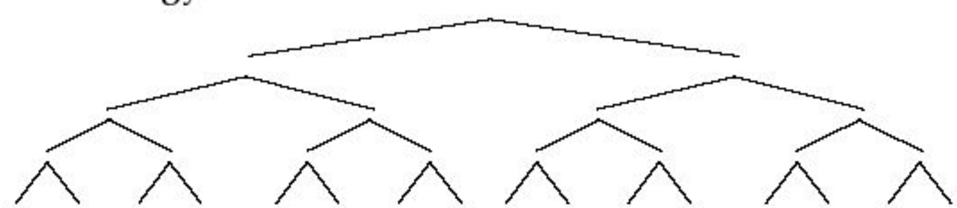


$\Sigma = 12$ $\qquad$ $\Sigma = 10$

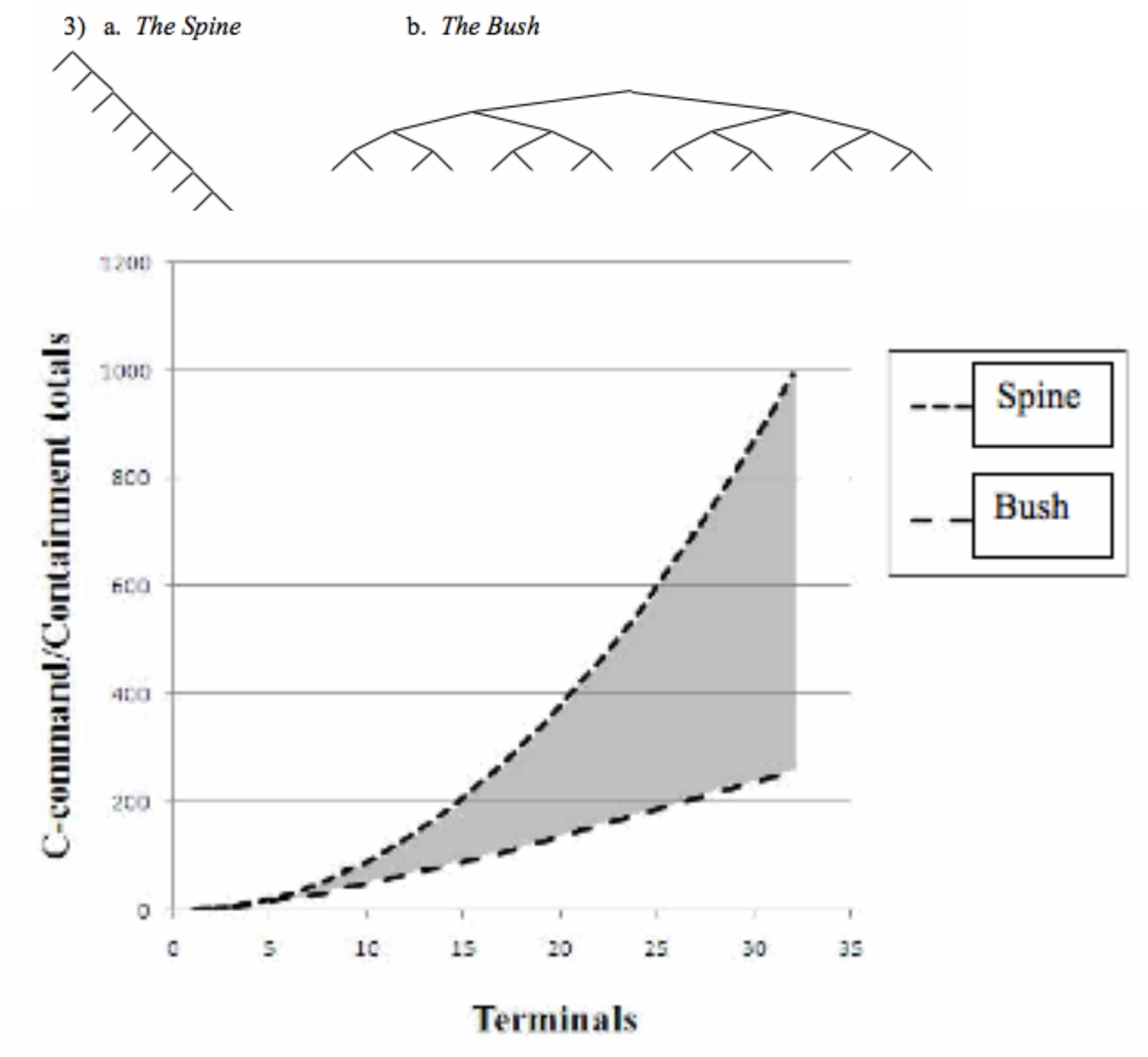
Strategy 1:



Strategy 2:



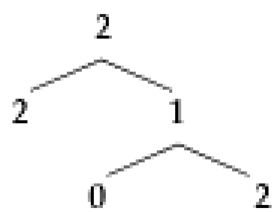
Divergence in total c-command relations for Spine vs Bush



Some PS patterns to compare

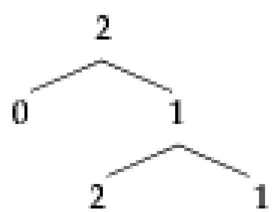
a. 2 → 2 1 ('X-bar')
1 → 0 2

b.



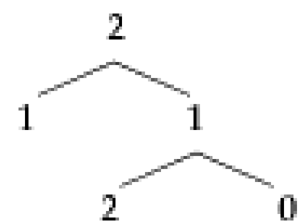
a. 2 → 1 0 ('high-headed X-bar')
1 → 2 1

b.



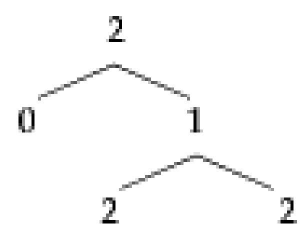
a. 2 → 1 1 ('D-bar')
1 → 2 0

b.



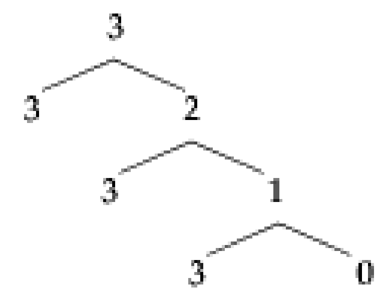
a. 2 → 1 0 ('high-headed D-bar')
1 → 2 2

b.

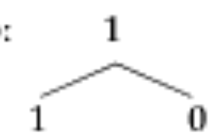


a. 3 → 3 2 ('3-bar')
2 → 3 1
1 → 3 0

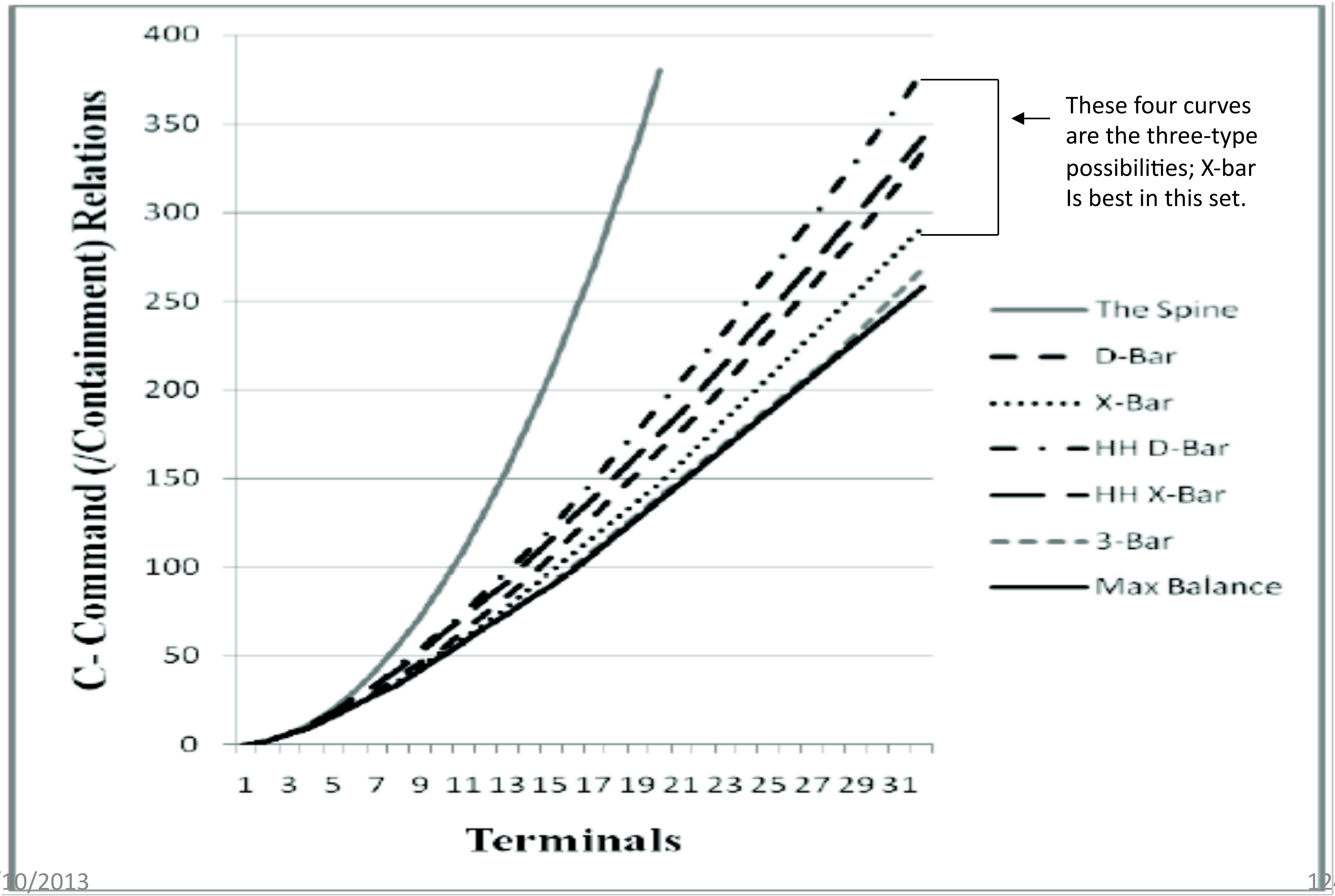
b.



1 → 1 0 ('spine'):



Comparison of ‘best trees’



Prediction: 'projections'

- The considerations above do not uniquely select X-bar as THE optimal solution.
- Instead, it is a member of the *class* of optimal solutions.
- That class is interesting; the general form is isomorphic to 'generalized X-bar'.
- Each phrase has a single head at the bottom, and slots for some number of other phrases.
- In other words, the optimal solution looks like an X-bar style projection with a head, complement, and zero or more specifiers.
- This is not forced by the 'rules of the game'; I consider all ways of self-similarly combining lexical items into larger structures.
- One interpretation of this result is that it may explain *why* natural language has the principle of projection -- because phrasal composition via 'projections' (in a purely structural sense) is optimal.
- See Medeiros (2008) for more details.

Part XII

- Cephalotaxis
  - Comparing phrasal organizations in terms of “head packing”
  - Number of heads grown by depth of maximal expansion

Cephalotaxis

- The next few slides compare the ‘head growth’ of various conceivable phrasal organizations.
- As we allow the structural molecule to be larger, many more possibilities become available:
  - The X-bar class contains 6 possibilities.
  - The next larger class contains 57 possibilities (the best performers are graphed here).
  - The class beyond that has 743 possible phrasal formats; I haven’t gone beyond enumerating them.

Squeezing from one to higher dimensions under a global ‘radial’ gradient

- In phyllotaxis, we have a one-dimensional stream of information -- periodicity of budding at the apical meristem -- mapping into a higher-dimensional distribution (the arrangements of the florets in a seedhead, say).
- There are two ‘forces’ at work here:
- the growth itself pushing the meristem forward and separating the old growths -- modeled by D&C with a radial magnetic field gradient,
- And mutual repulsion among the individual elements, tending to spread them apart near the point of origin.

Squeezing from one to higher dimensions under a global 'rootward' gradient

- In syntax, we have another mapping between a one-dimensional object (the surface form, a string of words), and a higher-dimensional object (the branching syntactic structure).
- I believe there is a global gradient here as well, effectively 'pulling' elements toward the root of the syntactic tree ("economy of command"; see Medeiros 2008, forthcoming, amounts to a preference for shallower trees).
- And a kind of mutual repulsion as well.

Optimal packing of constituents

- Suppose (for reasons related to reducing c-command and dominance totals) that it is desirable to 'fill' the available binary-branching space efficiently/densely with lexical items.
- The problem is not straightforward: if we imagine the structure growing from the root down, there are two desires which directly conflict:

(i) place lexical items as close to the root as possible,

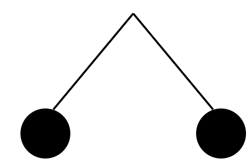
(ii) but leave adequate room (branching space) for further growth.

- These conflict because lexical items are terminals; they 'close off' all further growth along their path.

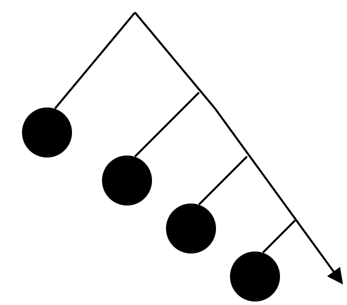
Comparison of phrase structural 'growth'

Consider some of the options:

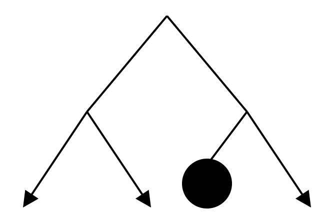
- If one gets 'greedy' and places two lexical items right below the root (locally maximizing the density of terminals), no further growth is possible.



- If one places one terminal beneath each branching point, then only one branch is available for further growth; a uni-branching tree (the Spine) results.
- This is only a good solution for very small trees.

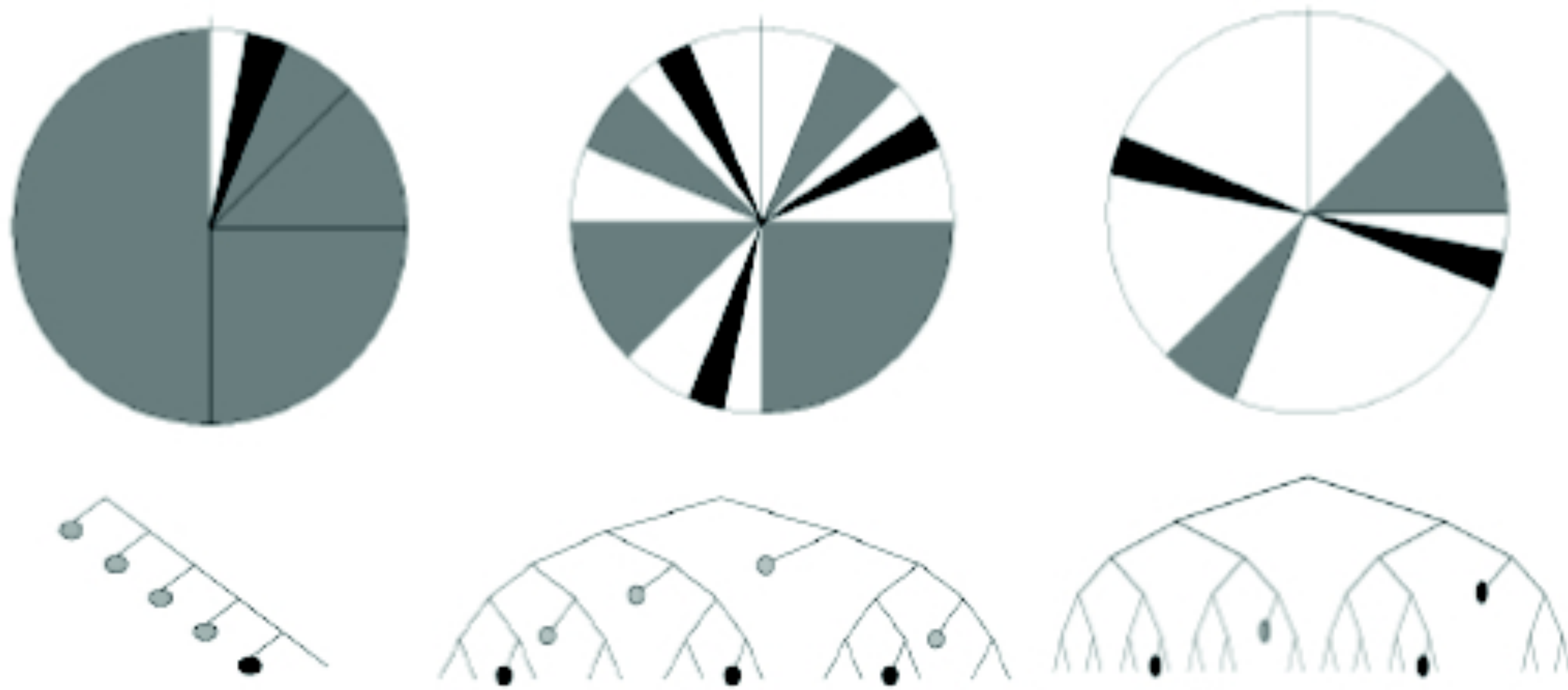


- If one 'delays satisfaction' and pushes terminals even farther down the tree, more room for growth is left.



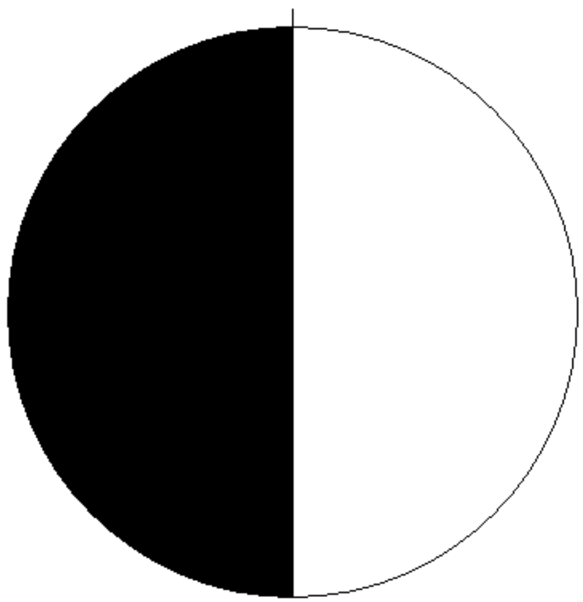
Goldilocks growth

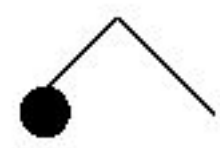
(9a) $XP = [X^0\ YP]$ (b) $XP = [ZP\ [X^0\ YP]]$ (c)$XP = [WP\ [ZP\ [X^0\ YP]]]$



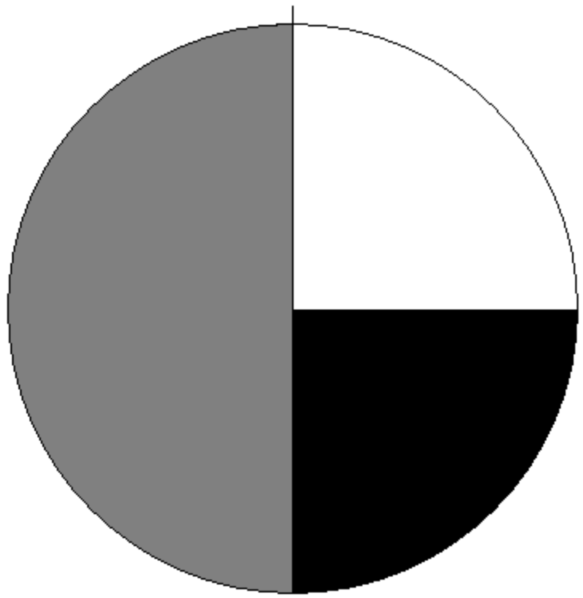
The Spine (9a) grows too fast, choking off room for further growth.
3-Bar (9c) grows too slowly, leaving the space sparsely populated.
X-bar (9b) is 'just right' (at this size scale, anyway).

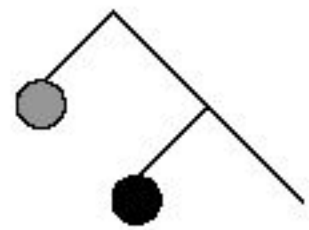
Spine, depth 1



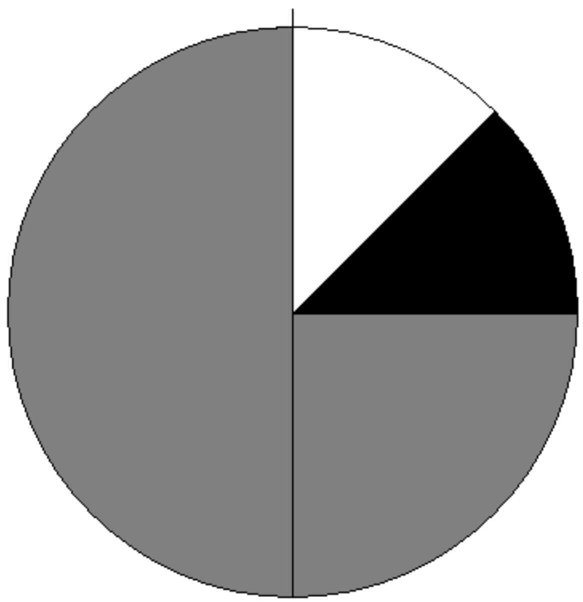


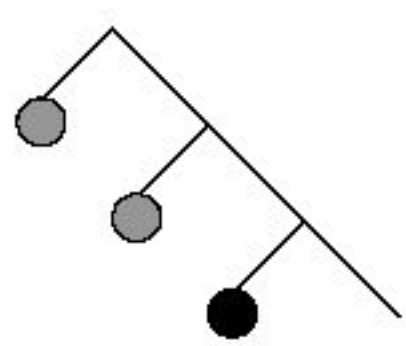
Spine, depth 2



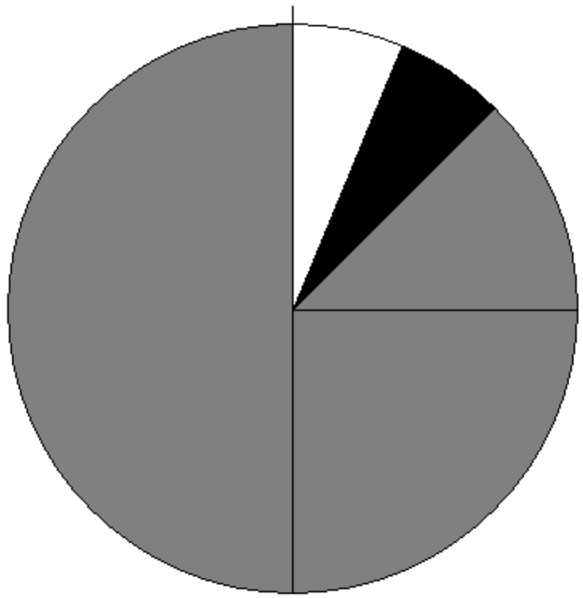


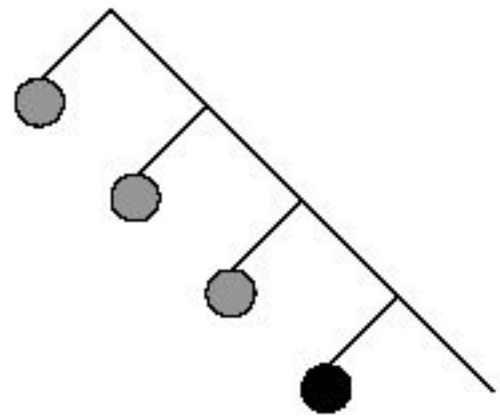
Spine, depth 3



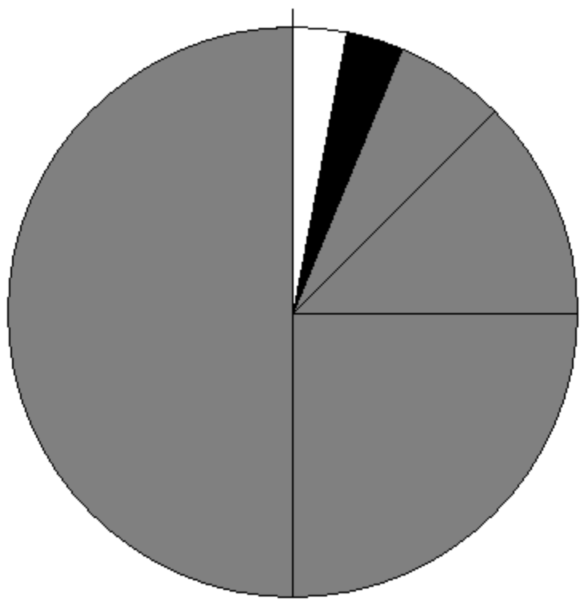


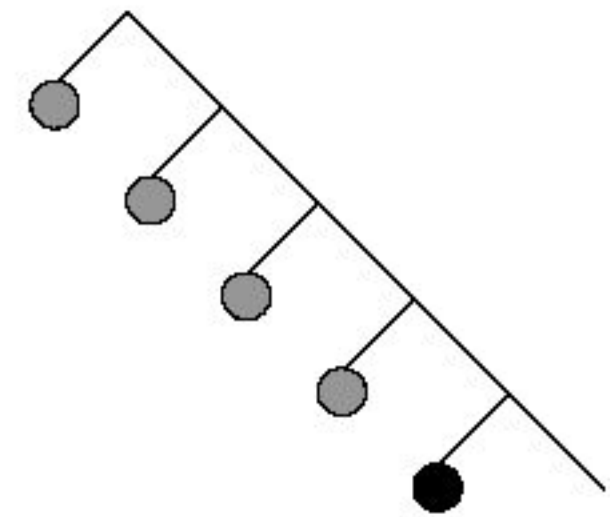
Spine, depth 4



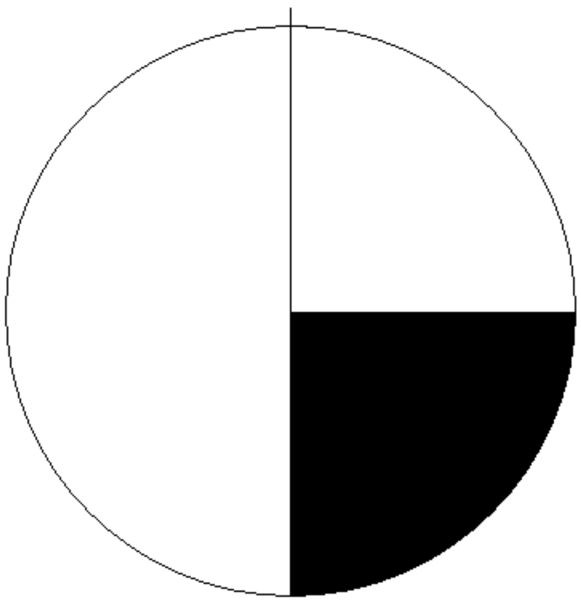


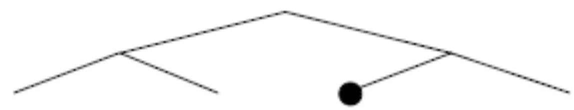
Spine, depth 5



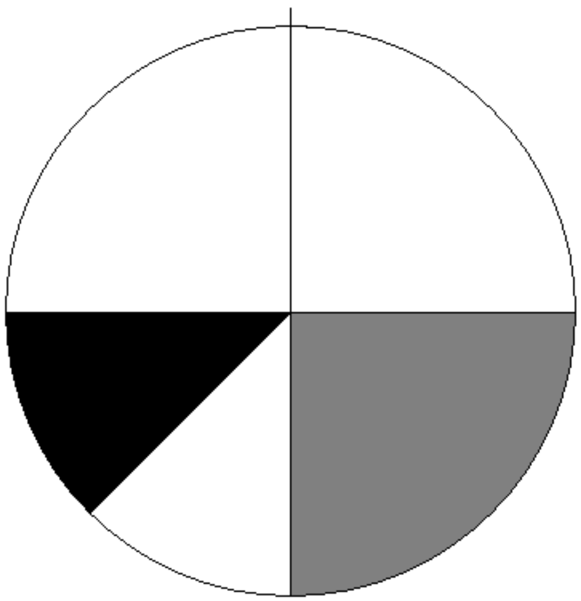


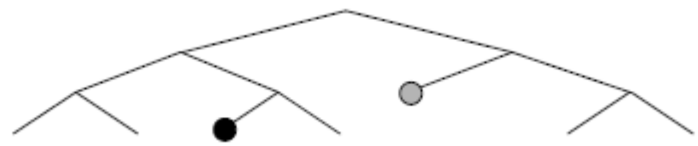
X-bar, depth 2



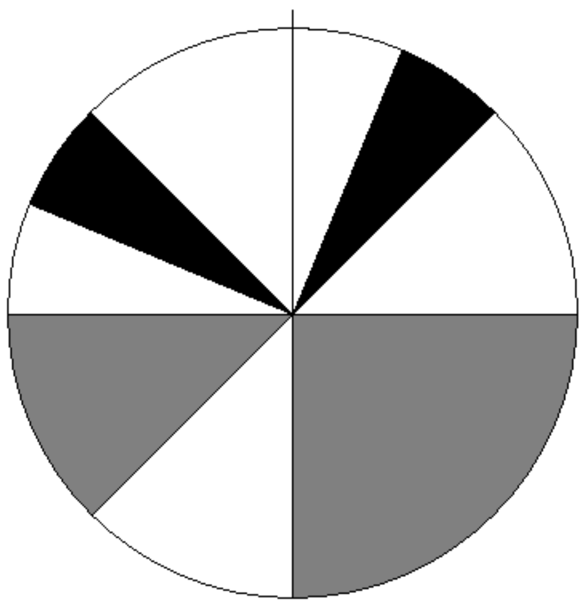


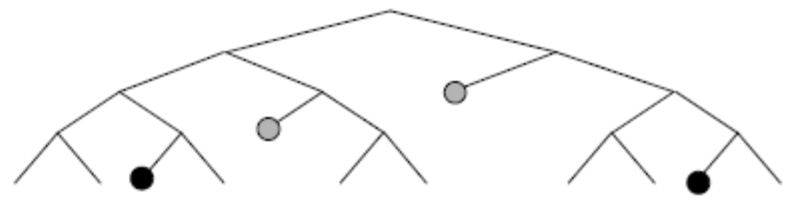
X-bar, depth 3



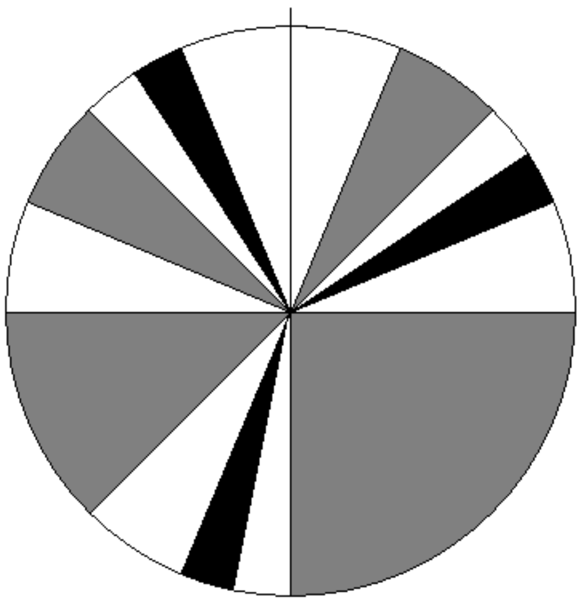


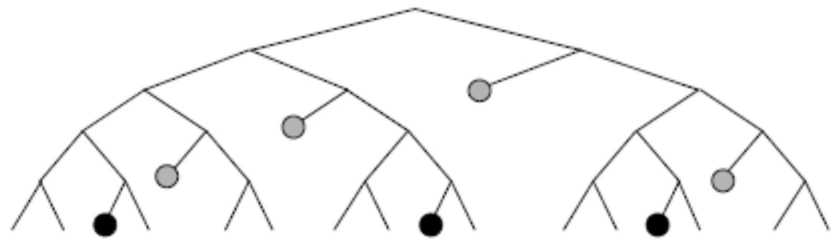
X-bar, depth 4



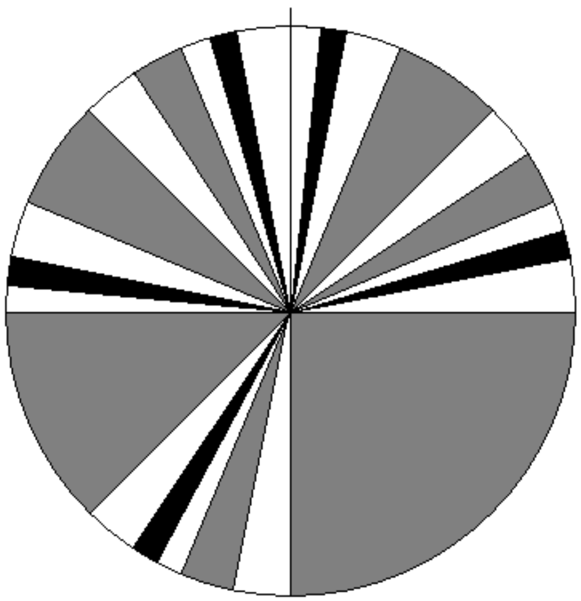


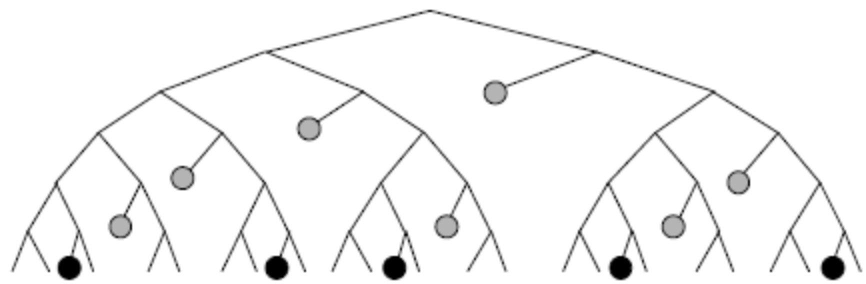
X-bar, depth 5



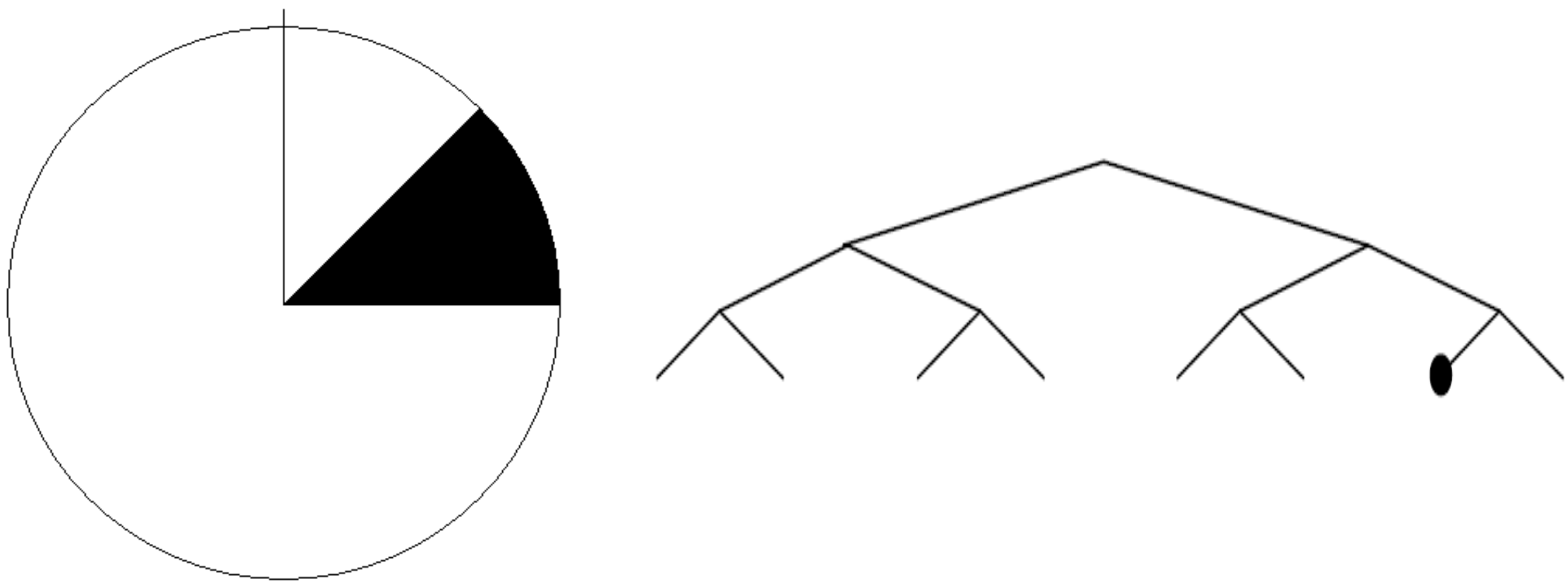


X-bar, depth 6

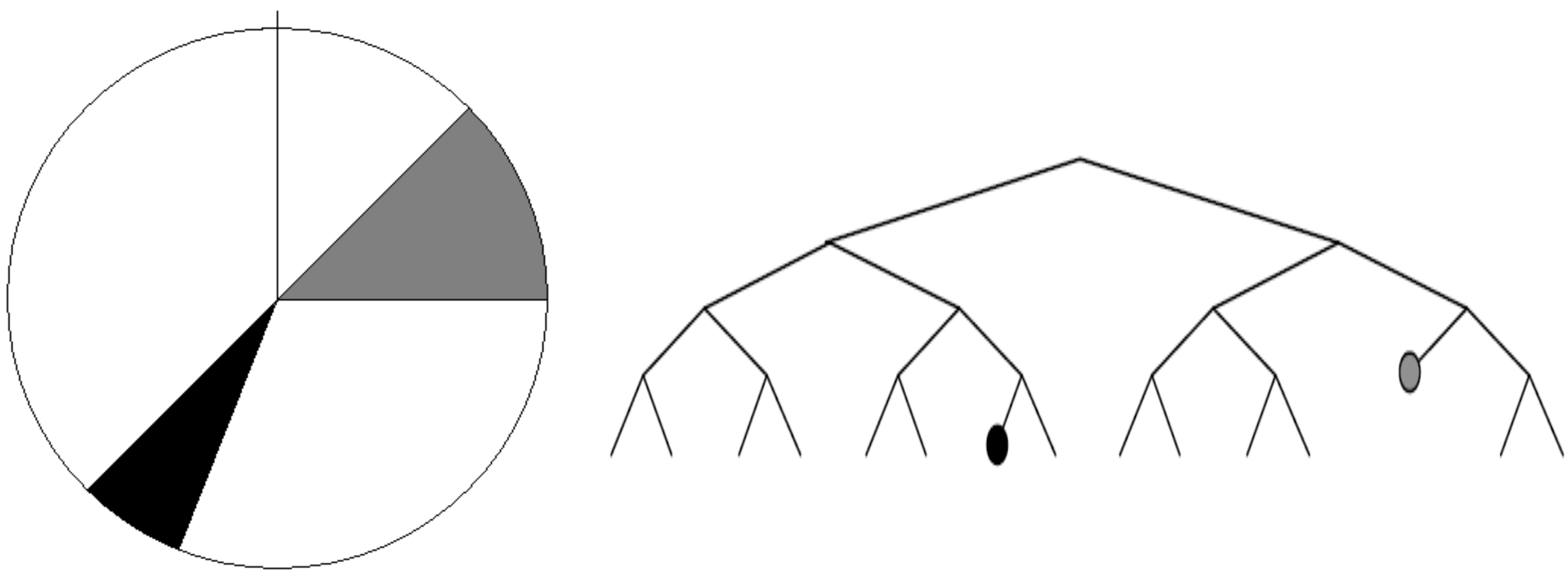




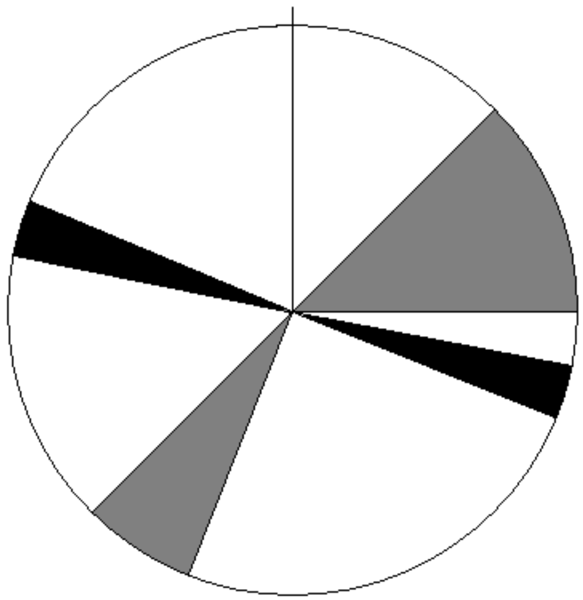
3-bar, depth 3

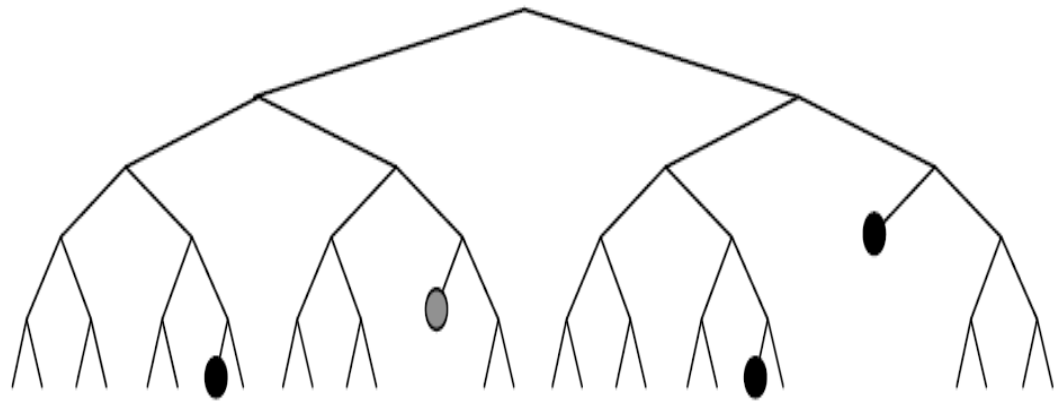


3-bar, depth 4

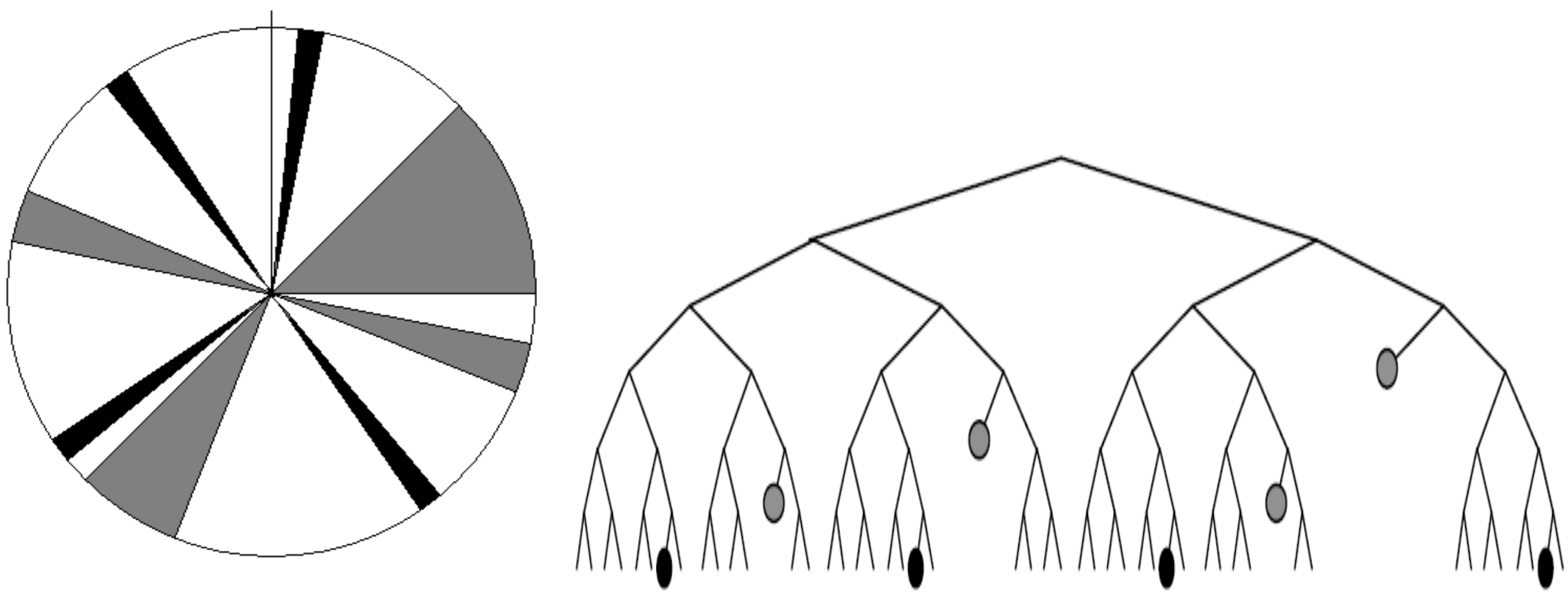


3-bar, depth 5



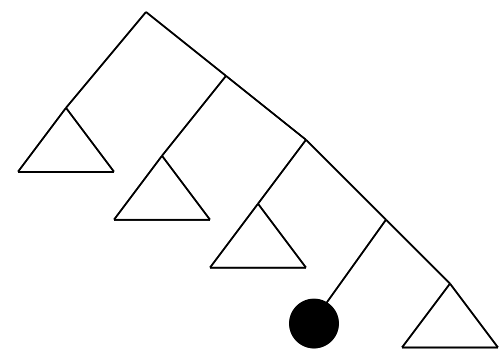


3-bar, depth 6

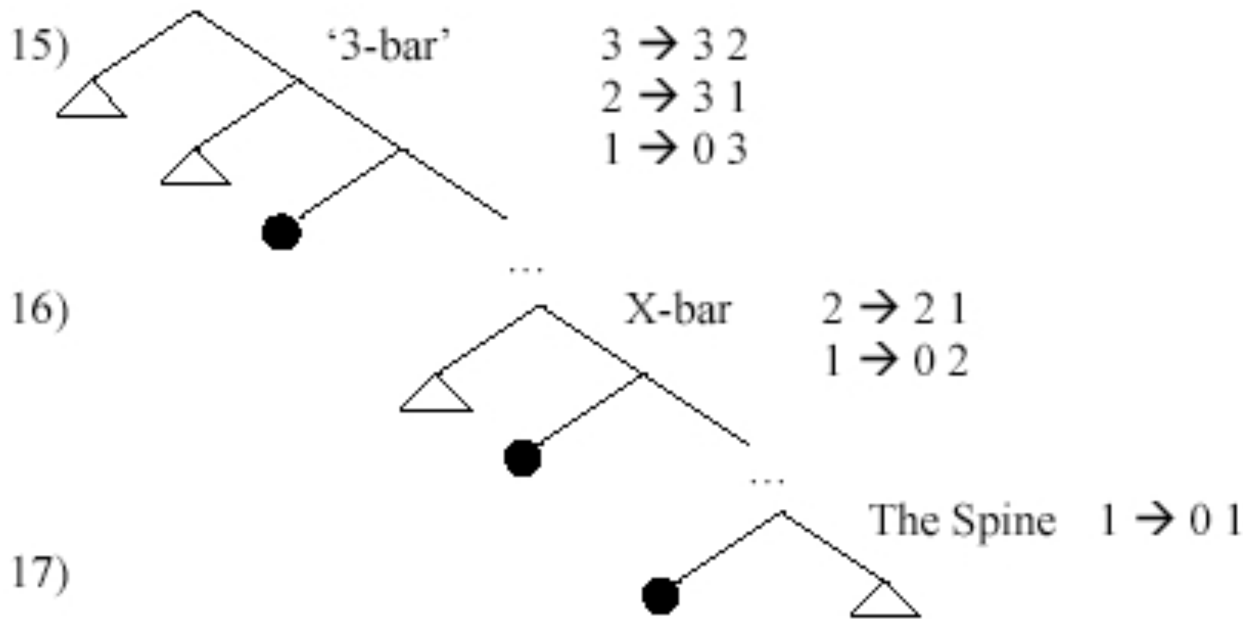


Global pull/local push

- The sort of organization at right represents the flavor of the dynamic mini-max compromise - it eventually grows more heads than any arrangement of the same amount of recursive structure.
- To *globally* maximize heads -- to pack as many as possible as near the root as possible -- *locally* heads are as sparsely distributed (deep down/far apart) as possible.
- Perhaps a case of "dynamic frustration"? (Binder, Uriagereka)



Prediction: if optimal packing matters for phrase structure, we may well expect that which solution is chosen might change as the tree grows. If syntax proceeds by cycles (e.g. phases), I predict something like the following sequence of 'growth modes':



That is, I predict that the 'bottom' of the cycle should contain only head-complement structures [X^0 YP], with single-specifier X-bar [ZP [X^0 YP]] strcuture above that, and multiple specifiers, if found at all, only at the highest level of the cycle.

Growth mode transitions

- There is some evidence that this is true:
- Pylkkänen (2002), following Larson (1988): multiple v/VP arguments are not ‘piled up’ within the lower lexical VP, but are introduced one-at a time in individual Appl(icative)P(hrase)s.
- And the subject is ‘severed’ from the verb, e,g, Kratzer (1996) and much subsequent work.
- Thus, the single-specifier regime reaches down as low as the ApplPs, but perhaps not into the lowest level (VP, which allows just a complement).
- That is as predicted: we have the Spine at the bottom, X-bar organization higher up.

From one to multiple specifiers

What about the other half of the prediction, that multiple specifiers (if permitted at all) are restricted to the 'top' of the cycle?

That might also be correct. The 'phase edges' have been held to allow multiple specifiers:

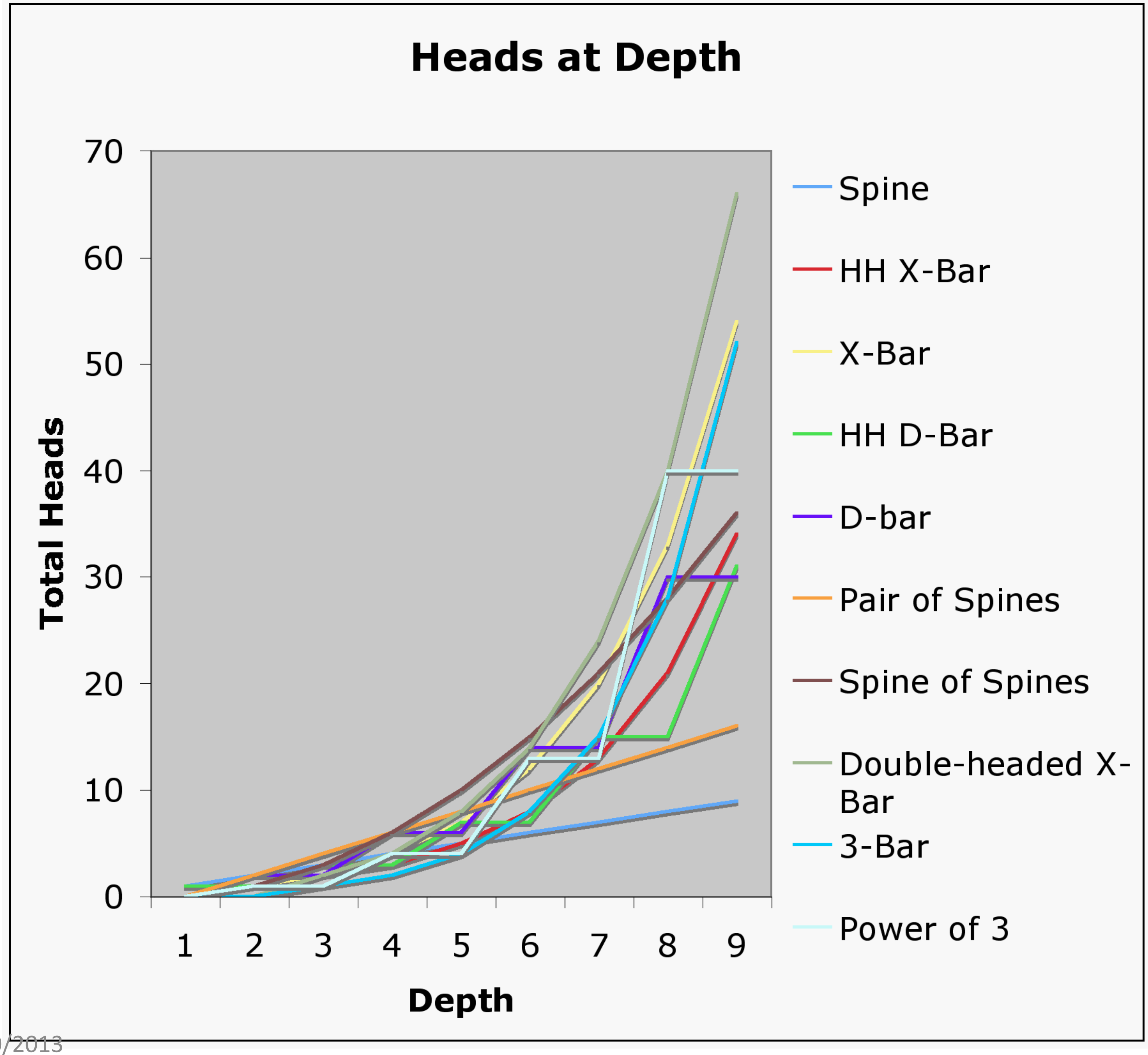
- vP hosts the external argument, and successive-cyclically moving whPs.
- CP has been argued to allow multiple specifiers, in light of 'multiple-wh constructions':
  - Ko sta gdje kupuje?

    Who what where bought

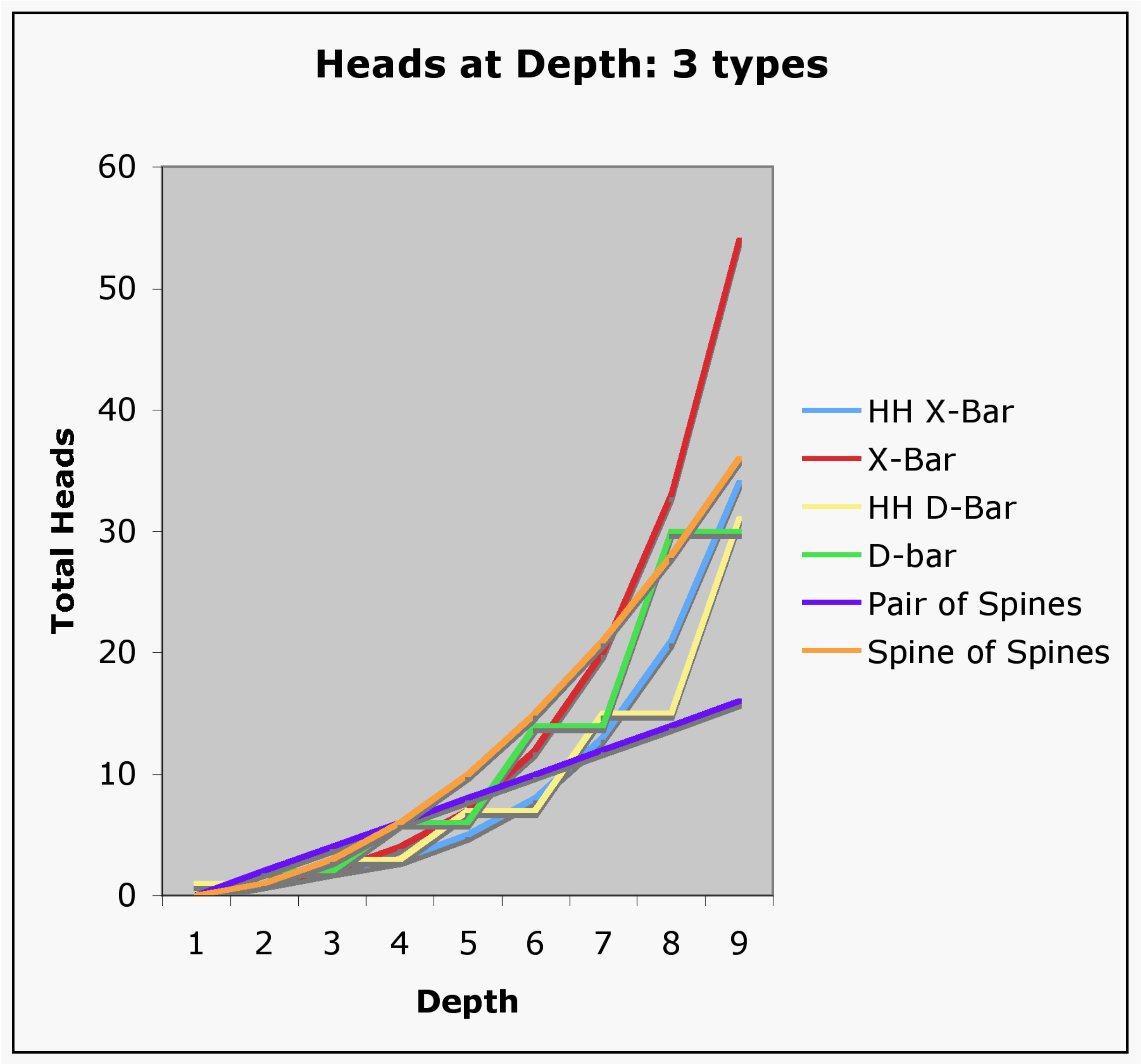
    "Who bought what where?" (Serbo-Croatian, Stepanov 1997: 3)

Multiple specifiers are not apparent elsewhere.

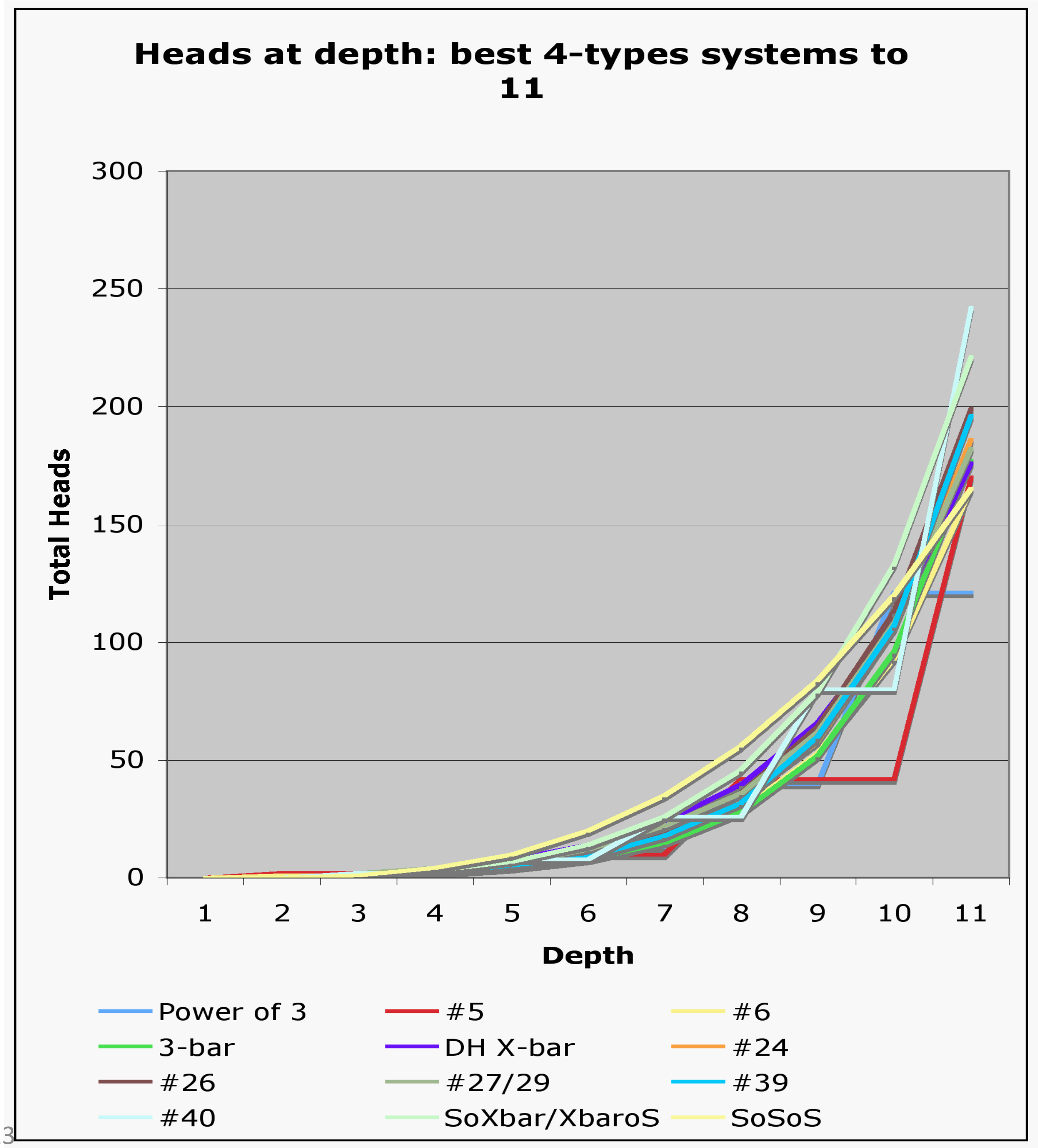
- One exception (thx Yosuke Sato): Japanese 'multiple subject' constructions, analyzed as multiple specifiers of TP; the problem is that TP is 'too early' for multiple specifiers. Leaving the details aside, this observation falls into place if Japanese clauses have more structure due to massive leftward movement, as proposed by Kayne (1994).

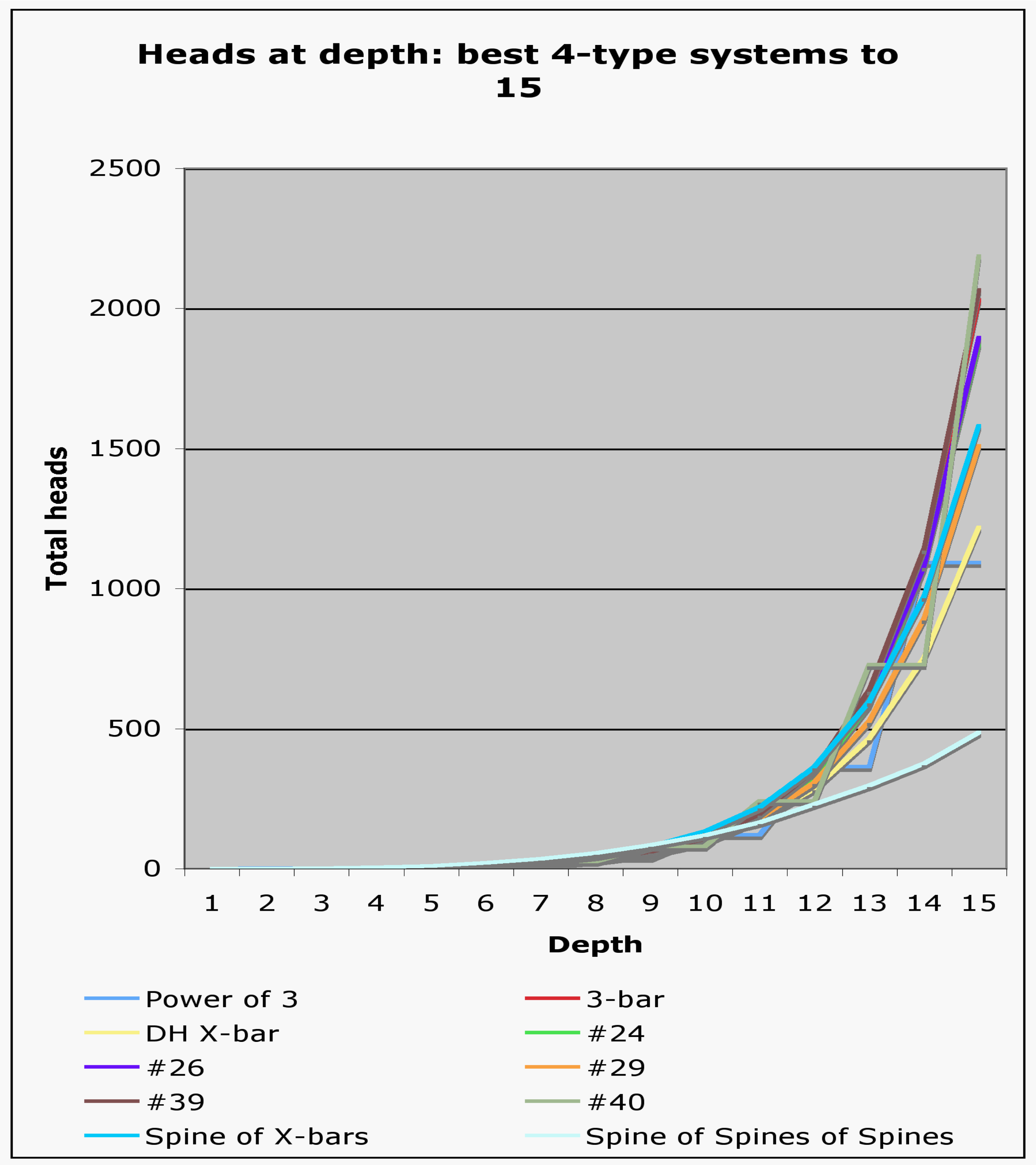


Heads at Depth
Total Heads
0
10
20
30
40
50
60
70
1
2
3
4
5
6
7
8
9
Depth
Spine
HH X-Bar
X-Bar
HH D-Bar
D-bar
Pair of Spines
Spine of Spines
Double-headed X-Bar
3-Bar
Power of 3



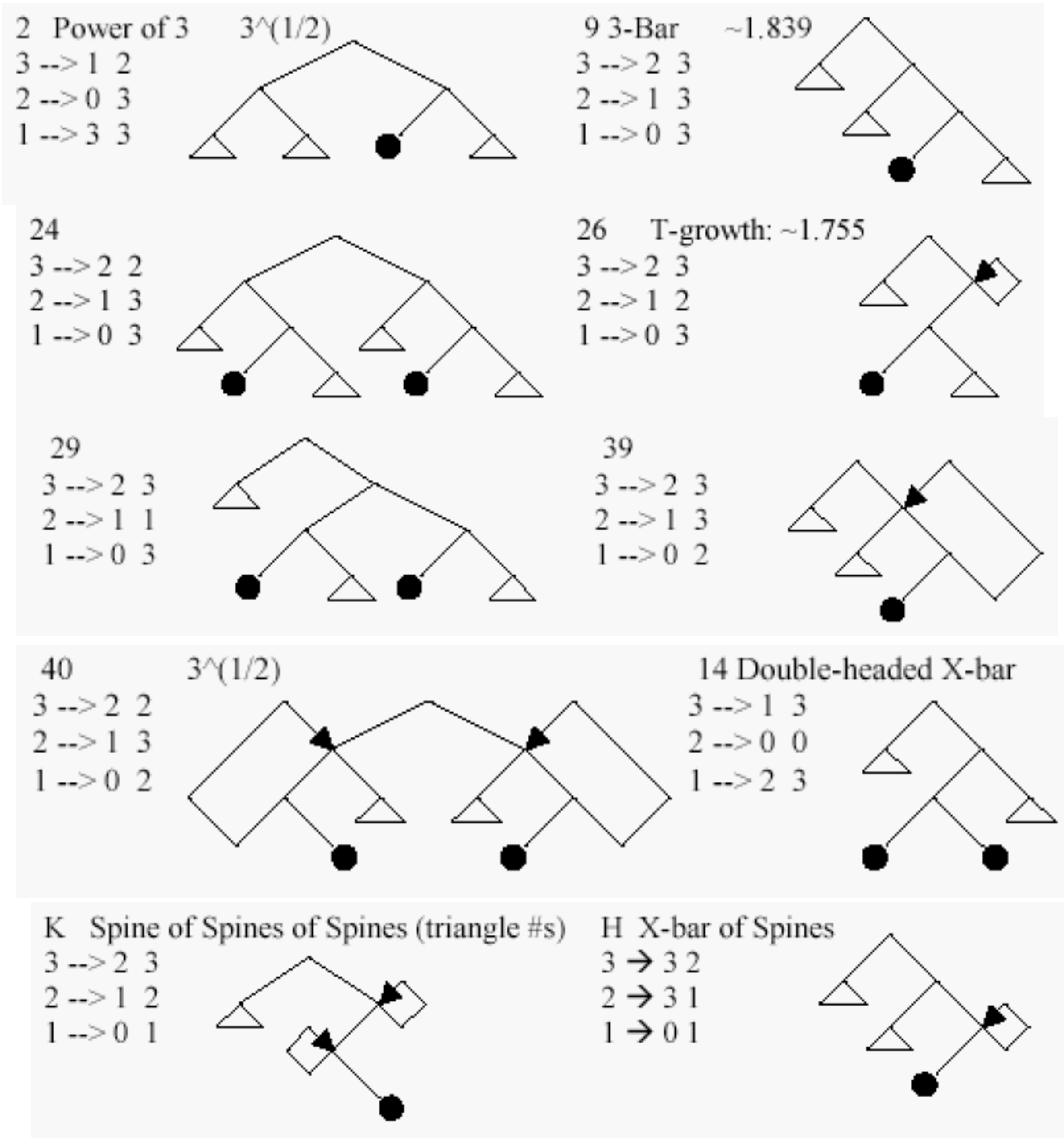
Heads at Depth: 3 types
Total Heads
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9
Depth
HH X-Bar
X-Bar
HH D-Bar
D-bar
Pair of Spines
Spine of Spines


Heads at depth: best 4-types systems to 11
Total Heads
300
250
200
150
100
50
0
1
2
3
4
5
6
7
8
9
10
11
Depth
Power of 3
#5
#6
3-bar
DH X-bar
#24
#26
#27/29
#39
#40
SoXbar/XbaroS
SoSoS


Heads at depth: best 4-type systems to 15
2500
2000
1500
1000
500
0
Total heads
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Depth
Power of 3
3-bar
DH X-bar
#24
#26
#29
#39
#40
Spine of X-bars
Spine of Spines of Spines

4 types

- One point to be made about the 4-type systems is that the competition is quite ‘messy’:
- The eventual winner (3-bar) doesn’t rise above the rest of the field until depth 20.
- For smaller structures, different organizations are optimal for small stretches of the intermediate range.
- Intuitive observation: it looks like those intermediate-best organizations (except the Spine of Spines of Spines) are formed by ‘mixing’ X-bar and 3-bar, or doubling/iterating some sub-portion of the X-bar pattern…



These are the 4-type systems graphed on the previous slides.

Intuitively, all but (K) look to be minimal variations on the X-bar shape, derived by stretching/splitting or doubling some node. Thus (14) replaces the X0 node with a pair of such; (29) doubles the X-bar level, etc.

References:

Binder, Philippe M. 2008. Frustration in complexity. Science 320, 322–323.

Chomsky, Noam. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum, eds., Readings in English Transformational Grammar. Waltham: Ginn, 184-221.

Chomsky, Noam and Morris Halle. 1968. The Sound Pattern of English. New York: Harper and Row.

Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. Linguistic Inquiry 24: 239-297.

Escudero, Juan Garcia. 2003. Fibonacci sequences and the multiperiodicity of the variable star UW Herculis. Chinese Journal of Astronomy and Astrophysics (3)3: 235-240.

Gummelt, Petra. *Geometriae Dedicata* 62, 1 (1996).

Jackendoff, Ray. 1977. X-bar Syntax: A Study of Phrase Structure. Cambridge, MA: MIT Press.

Jean, Roger. 1994. Phyllotaxis: A Systematic Study in Plant Morphogenesis. Cambridge: Cambridge University Press.

Medeiros, David P. 2008. Optimal growth in phrase structure. Biolinguistics 2: 152-95.

Medeiros, David P. 2012. Economy of Command. Ph.D. dissertation, University of Arizona.

Moody, R.V., ed. 1997. The Mathematics of Long-range Aperiodic Order. Dordrecht: Kluwer.

Roopun, Anita, Mark Kramer, Lucy Carracedo, Marcus Kaiser, Ceri Davies, Roger Traub, Nancy Koppell, and Miles Whittington. 2008. Temporal interactions between cortical rhythms. Frontiers in Neuroscience 2(2): 145-154.

Rozenberg, Grzegorz and Arto Salomaa. 1980. The Mathematical Theory of L-systems. New York, NY: Academic Press.

Saddy, Doug. 2012. “Where do grammars come from?” Presentation in UA Linguistics Colloquium series.

Shlonsky, Ur. 2010. The cartographic enterprise in syntax. Language and Linguistics Compass, Wiley Online.