Advanced issues in cognitive science and linguistics
Ray Jackendoff
UAB, May 2017

Part 4: A third component of Parallel Architecture: Relational Morphology and the structure of the lexicon

(with Jenny Audring; much of the approach derived from
Geert Booij's Construction Morphology)

Background

"Knowledge of language" = material stored in memory. Fundamental questions:

- What linguistic elements does a speaker store in memory, and in what form?
- How are these elements combined online to create novel utterances?
- How are these elements acquired?

Most linguistic theory emphasizes #2, "infinite use of finite means" (Humboldt), "Basic Property of language" (Berwick and Chomsky).

How morphology fits into PA: Morphological schemas

Basic principle of PA: Linguistic structure has three generative systems and linkages among them:

Phonological structures ⟵ Interface ⟶ Syntactic structures ⟵ Interface ⟶ Conceptual structures

Status of words: Linkage of pieces of semantics, syntax, phonology (small interface rule)

(1) Semantics: [CAT]
Morphosyntax: N
Phonology: /kæt/

Phrase structure rules stated as declarative *schemas*:

(2) Syntax: [$_{VP}$ V – (NP)]

Plural of *cat*:

(3) Semantics: [PLUR (CAT)]
Morphosyntax: [$_{N}$ N, plur]
Phonology: /kæt s/

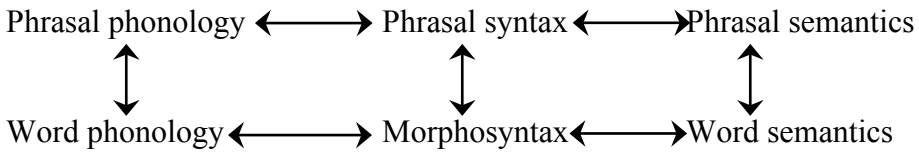
Extract contribution of *cat* to get a schema with variables on all three levels:

(4) Semantics: [PLUR (X)]
Morphosyntax: [$_{N}$ N, plur]
Phonology: /… s/

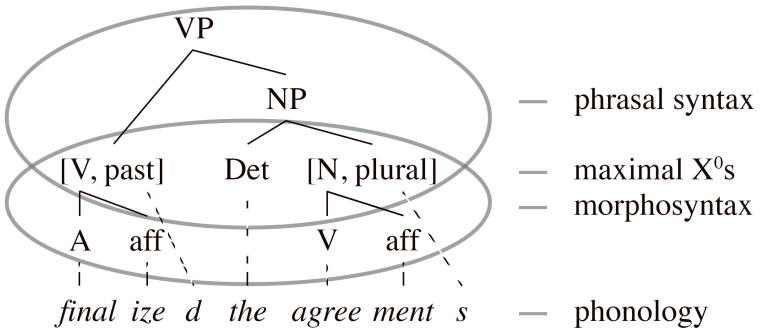
Deviations from 1-1 mapping among the 3 components: Meaningless case and agreement markers, zero morphology, truncation, reduplication, … . Within PA, these are noncanonical links between the levels.

(5) Semantics: [PLUR (SHEEP)]
Morphosyntax: [$_N$ N, plur]
Phonology: /ʃip/

How does morphology relate to the rest of the grammar?



Phrasal syntax and morphosyntax overlap at maximal X^0s.



Inflectional morphosyntax is (mostly) perfectly regular; irregularities lie in noncanonical mappings between morphosyntactic features and phonology.

Schemas vs. rules

PA grammar is stated entirely in terms of declarative schemas (like LFG, HPSG, CxG, CxM). No procedural rules that map an "input" into an "output."

Continuity between words and rules/schemas: both are listed in lexicon, in same format: pieces of linguistic structure. Words have more specific content (e.g. (1)); schemas have variables (e.g. (2) and (4)).

Incompatible with procedural theories of morphology, in which rules are of an entirely different format from the lexicon

Problem: Lots of composite items in language that can't be constructed by general rules:

- In syntax, idioms (*spill the beans*)

In morphology:

- Conventionalized compounds (*football*)
- Irregular phonological relation to base (*sing/sang* vs. *sting/stung*)
- Predictable phonology and meaning but not productive (*widen* vs. **louden*)
- Words with identifiable affix but nonword base (*commotion*, *impetuous*)

Plus frequent regular forms that are shown experimentally to be stored.

Traditional rules both overgenerate (how do you stop **louden*?) and undergenerate (meaning of *football*, form of *impetuous*).

A proposed solution: Lexicalist Hypothesis (Chomsky 1970): Regular patterns belong to syntax, idiosyncratic patterns are due to "redundancy rules" in lexicon. Any amount of idiosyncrasy in a pattern implies it's "in the lexicon," not the grammar. But then, how does grammar generate *spill the beans*? Generative grammar has been evasive.

Opposite (but soluble) problem for PA: "Lexical redundancy rules" are schemas, indistinguishable in form from productive schemas like (4) – and both are in the lexicon. How to tell them apart?

The status of nonproductive schemas

An example: *hard* and *harden*

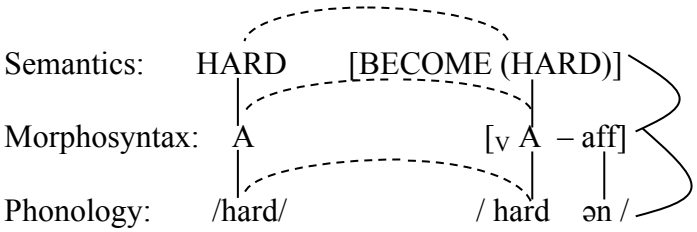
(6)

Semantics:	$HARD_1$	Semantics:	$[BECOME\ (HARD_1)\ ]_2$
Morphosyntax:	A_1	Morphosyntax:	$[_V\ A_1\ aff_3\ ]_2$
Phonology:	$/hard/_1$	Phonology:	$/\ hard_1\ ən_3\ /_2$

- Coindex 1 in *hard* connects the three components: *interface links*
- Coindex 1 also connects interior of *harden* by interface links
- Coindex 1 also connects the word *hard* to interior of *harden*: *relational links*
- Coindex 2 connects the three components of whole word *harden*
- Coindex 3 connects phonology of affix with its morphosyntax

Coindices can be interpretated as ends of association lines:

(7)



Schema for *–en*:

(8) Semantics: $[\text{BECOME } (\underline{X}_x)]_y$
Morphosyntax: $[_V\ \underline{A}_x\ \text{aff}_3\ ]_y$
Phonology: $/\ \underline{\ldots}_x\ \text{ən}_3\ /_y$

Coindices x and y are *variable indices*: linked to any word with same pattern.

(8) is a lexical redundancy rule: it is not expected to apply across the board, and it admits idiosyncratic material. Problem: it looks just like fully productive schemas. We don't want to say (4) is "in the grammar" and (8) is "in the lexicon" – especially since in PA the grammar *is in* the lexicon!

Proposal: Productivity is marked on the variable in the schema. Variable in (4) is productive, variable in (8) is not.

Possibility: a schema could have one productive variable and one nonproductive one. Example: English place names.

(9) a. Arrowhead *Lake*, Loon *Mountain*, Wissahickon *Creek*, Laurel *Hill*, Sugar *Island*
b. *Mount* Everest, *Lake* Michigan*, Cape* Cod
c. the Indian *Ocean*, the Black *Sea*, the Hudson *River*, the White *Mountains*
d. the *Bay* of Fundy, the *Gulf* of St. Lawrence, the *Cape* of Good Hope

Choice of name is productive (*Morris Mountain*); choice of geographical feature is not (*the Halle Mountains* but **the Mountains of Halle*)

The schemas, all in "the same place" (productive variables have double underline; notice that (10a,b) are morphological, and (10c,d) are phrasal):

(10) a. Name N [e.g. *Loon Lake*: compound]
b. N Name [e.g. *Mount Washington*: compound]
c. the Name N [e.g. *the Lehigh River*: phrasal]
d. the N of Name [e.g. *the Gulf of Mexico*: phrasal]

Schemas with nonproductive variables capture patterns among stored lexical items; tolerate idiosyncrasy, treat novel uses as exceptional. Schemas with productive variables allow free online combination.

BUT: Productive schemas actually have two functions: (a) *generative* function, to create new expressions, and (b) *relational* function, to capture patterns among stored items, just like nonproductive schemas. Examples:

- Compounding is productive, but there are thousands of conventionalized instances – same schema, used generatively and relationally
- Stored regular plural in *raining cats and dogs*, *keep tabs on NP*, *odds and ends*: productive plural schema being used relationally

- Experimental evidence: frequent regular plurals are stored, but still fall under productive plural schema
- In syntax: Productive VP schema is used relationally for VP idioms

Can there be schemas that are only used generatively? No, because you can always memorize new pieces of structure, and when you do, they're supported relationally.

Nonproductive schemas in phrasal syntax too ("syntactic nuts"):

- Geographic terms in (10c,d)
- NPN construction: *theory after theory*, *face to face* but **boy beside boy*, **day before day*
- Determiner construction: *what a theory*, *such a theory*, **who a linguist*, **where a city*

Therefore, two kinds of schemas:
- Productive (open) schemas: both generate new expressions and motivate stored items
- Nonproductive (closed) schemas: motivate stored items

Change of perspective: ALL schemas can be used relationally; a subset can ALSO be used generatively. Productive schemas are therefore ordinary schemas that have "gone viral"!

Conclusion: The generative property of language, Humboldt's "infinite use of finite means," emerges from and rides on top of the system of lexical relations!

Therefore, linguistic theory also has to account for the much more widespread relational role of rules/schemas – how to capture the relations among stored items.

Do we really need nonproductive schemas? They might be just analogy/association (Pinker's "dual-route" theory).

Cases that don't fall under any more general pattern:

(11) laugh/laughter
bomb/bombard
hate/hatred
humble/humility

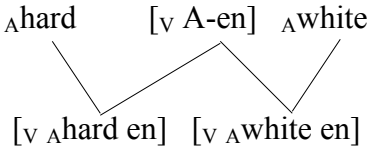
But: For 50+ instances like *harden*, ~1000 instances like *construction*, schemas seem more attractive.

An argument from acquisition: Learners form hypotheses for possible rules, based on input. What is the form of these hypotheses? In PA: tentative schemas! Learners have no idea in advance whether tentative schema will be productive, so they'll end up with lots of failed (nonproductive) schemas. Why throw them away? They may help learning of new words.

How do schemas help relate lexical items?

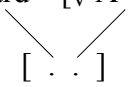
Not by *deriving* one word from another! Popular alternative: domain-general *inheritance hierarchy* (cf. HPSG, Cognitive Grammar, Construction Grammar)

(12)



One interpretation: inheritance as maximizing economy. Don't duplicate information ("impoverished entry theory").

(13)



One of many problems with impoverished entires: complex item should take longer to access than its base. False!

Alternative: Brain embraces redundancy, to stabilize memory and processing (e.g. thematic role marking, depth perception). Full-entry theory: lexical items are encoded in their entirety, even when redundant.

Schema codifies or *motivates* generalizations among lexical entries. Saussure: signs are less arbitrary when motivated by related words.

Structure-sharing – a broader notion than inheritance

Inheritance assumes asymmetry: one word inherits structure from another or from a schema. But with "sister" pairs like *assassin/assassinate*, *linguist/linguistics*, second word is built phonologically on first; but the reverse in semantics.

(14) Semantics: [PERSON WHO [MURDERS POLITICIAN]$_4$]$_5$
Morphosyntax: N_5
Phonology: /əsæsən/$_5$

Semantics: [MURDER POLITICIAN]$_4$
Morphosyntax: [$_V$ N – aff$_6$]$_4$
Phonology: / əsæsən$_5$ ejt$_6$ /$_4$

Coindex 4 links semantics of *assassinate* to part of semantics of *assassin*.
Coindex 5 links phonology of *assassin* to part of semantics of *assassinate*.

Complex words without lexical bases: base has no morphosyntactic category. There are *lots* of them!

(15) Semantics: $BEAUTIFUL_7$
Morphosyntax: $[_A - aff_8]_7$
Phonology: / gɔrdʒ əs/$_8$ /$_7$

The schema:

(16) Semantics: $\underline{PROPERTY}_w$
Morphosyntax: $[_A - aff_8]_w$
Phonology: / … əs$_8$ /$_w$

More sister relations: *ambitious/ambition*, *contagion/contagious*, *cognition/cognitive*
No way to say which is derived from which. (And we don't want an "abstract root")

(17) Semantics: $DESIRE_{11}$ Semantics: $[HAVING\ (DESIRE_{11})]_{12}$
Morphosyntax: $[_N - aff_{13}]_{11}$ Morphosyntax: $[_A - aff_{14}]_{12}$
Phonology: / æmbɪ$_{15}$ ʃən$_{13}$ /$_{11}$ Phonology: /æmbɪ$_{15}$ ʃəs$_{14}$ /$_{12}$

- Defective morphosyntax
- Linked to different affixes
- Share part of semantics (coindex 11) and part of phonology (coindex 15)

Two items motivate each other to the extent that they share structure, as encoded in terms of relational links.

Sister relation between *schemas*:

(18) pacifism/pacifist
altruism/altruist
solipsism/solipsist
impressionism/impressionist
etc.

The schemas:

(19) Semantics: $\underline{IDEOLOGY}_\alpha$ Semantics: $[ADHERENT\ (\underline{IDEOLOGY}_\alpha)\ ]_x$
Morphosyntax: $[_N - aff_{16}]_\alpha$ Morphosyntax: $[_N - aff_{17}]_x$
Phonology: /$\underline{\ldots}_\beta$ ɪzəm$_{16}$ /$_\alpha$ Phonology: /$\underline{\ldots}_\beta$ ɪst$_{17}$ /$_x$

Greek subscripts α and β denote *variables* that are bound across multiple schemas.
Interpretation: if you find a noun that denotes an ideology (coindex α) that ends in *–ism*, it's a good bet there will also be a noun that denotes an adherent of that ideology, that ends in *–ist*, and that, aside from the suffix, has the same phonology (coindex β).
Sister schemas are important for paradigmatic relations, including ablaut, stem allomorphy, truncations.
Formal relation among sister *schemas* is exactly the same as relation between sister *words*. Continuity between words and rules again!
Is there a counterpart of sister schemas in phrasal syntax? Yes! (Guess whose!) An alternative notion of transformations: an equivalence class among sentence patterns – a sister relation

between constructions. Three candidates: dative alternation, symmetrical predicates, and German V2 vs. V-final.

Conclusions so far (many of which are shared with CxG and CxM)

The lexicon is not simply a list of unstructured exceptions. Lexical items can have internal structure – morphological structure in the case of morphologically complex words, and syntactic structure in the case of idioms and other fixed expressions.

Morphology in PA = grammar of word-sized pieces of language.
The lexicon includes a rich network of relational links among its items.

We have eliminated the strict distinction between lexicon and grammar. Rules of grammar are in the lexicon.

Within the lexicon, we have eliminated a hard distinction between words and rules. They are both encoded in the same format: pieces of linguistic structure connected by interface and relational links. Rules of grammar are encoded as schemas: pieces of structure containing variables, otherwise the same as words.

We have eliminated the distinction between grammatical rules and lexical rules. Both kinds of rules are in the lexicon, encoded in a common format. We have replaced this distinction with the distinction between the generative and relational functions of schemas.

The only difference between productive and nonproductive schemas is a diacritic on the variable.

Nonproductive schemas motivate listed lexical items. Productive schemas both motivate listed lexical items (their relational function) AND allow free generation of novel forms (their generative function).

All these characteristics of morphology are also found in phrasal syntax.

The generative function of schemas is only half of the picture – or less than half – since there are plenty of schemas that lack the generative function. This dethrones the generativity of language from its dominating position in linguistic theory (though we must stress that this issue doesn't go away either). Lexical relations call for at least equal attention.

Beyond language!

"Knowledge of language" = what is stored in long-term memory

The lexicon of PA/Conceptual Semantics/Simpler Syntax/Relational Morphology is a theory of one department of long-term memory: huge lexicon, complete with structured lexical entries, interface links (including words), relational links, and schemas.

What makes language different from other cognitive capacities? (Speaking to Goal from the beginning of the week)

Hypothesis: **Memory is memory**: Much the same organization is found in memory for every cognitive domain.

Differences among domains: (1) units in terms of which the domain is constructed; (2) the way these units can be combined; (3) their interfaces to other domains.

Language is different because it's about syntactic, morphosyntactic, and phonological units, and it interfaces with conceptual structure on one end and auditory (or for sign, visual) perception and motor control on the other end. And it's different because it's used for conventionalized communication.

Are there ways that language is the *same* as other domains? Abstract away from content and function. Some general properties:

- A vast lexicon, with tens of thousands of items
- Lexical items that involve multiple levels of representation, coordinated by interface links
- Hierarchical constituent structure within items, on all three levels
- Relational links among items that pick out shared structure
- Regularities across items that are picked out by schemas
- Schemas that both motivate structure within stored items (their relational role) and assign structure to novel items (their generative role)
- Both free items (e.g. *sheep*), which can occur independently, and bound items (e.g. *-ish*), which occur only attached to other material

Can we find these properties in other domains?

Music

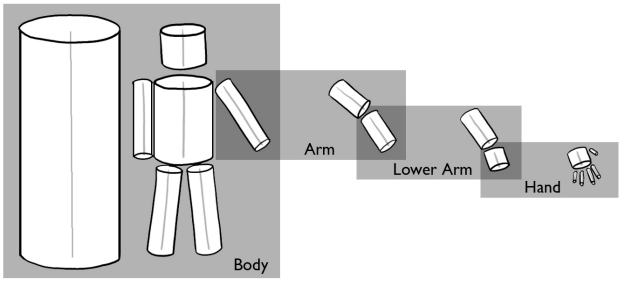
- Vast lexicon: Yes! You can recognize hundreds if not thousands of popular songs, folk songs, nursery rhymes, and, for some people, 45-minute symphonies and even longer operas. You can identify many of them immediately on hearing just a few selected seconds of music (say, upon turning on the radio).
- Lexical items that involve multiple levels of representation: Yes! Lerdahl and Jackendoff 1983: Music cognition involves independent levels of grouping structure, metrical structure, and tonal hierarchy (or "prolongational reduction"), linked by a rich system of interface principles
- Hierarchical constituent structure within items: Yes! Music stored in memory has hierarchical structure on all three levels.
- Relational links among items that pick out shared structure. Possibly a bit different from language: there can be relational links *inside* a melody, picking out repetitions and variations of motives, e.g. first two phrases of *Happy Birthday* are recognized as nearly identical. This sort of relational link does not appear in morphology or syntax, though it

might occur in rhetorical speech registers with heightened affect such as poetry and preaching.

- Regularities across items that are picked out by schemas: Yes! The well-formedness rules and preference rules of Lerdahl and Jackendoff can be recast as schemas. Schemas are also an attractive way to characterize conventionalized forms such as 32-bar popular song form, 12-bar blues, and classical sonata and minuet form.
- Schemas both motivate structure within stored items and assign structure to novel items: Yes! There is no reason to believe that the structures of known music and newly experienced music are different in character. Well-known music may be represented more richly in memory, but according to the same principles.
- Free vs. bound items: Bound items: appoggiaturas and suspensions – fleeting dissonances that have to resolve to consonances and make no sense on their own. Another possibility: the dominant harmony in a dominant-tonic cadence, whose function is to relax into the tonic.

Knowledge of objects

- Vast lexicon: Yes! One can recognize thousands of objects and object types – without the necessary use of linguistic labels.
- Multiple linked levels of representation: Yes! Knowledge of objects involves more than just how it looks: how it feels (haptic representation), what sounds it makes, and how to use it (action representations).
- Hierarchical constituent structure within items: Yes! Following Marr 1982, objects can be understood as having a hierarchical decomposition. A chair has legs, a seat, and a back; the back may have decorations; the decorations may be segmented into various parts, and so on.



- Relational links among items that pick out shared structure: Yes? You can appreciate the similarity of structure between the seats of armchairs and wheelchairs, despite considerable superficial difference. You can appreciate the similarity in function between radically different kinds of bottle openers and between radically different kinds of lamps or faucets.
- Regularities among items picked out by schemas: Yes: Any sort of prototype representation (e.g. a Marr 3D model) is in effect a schema. Schemas can pick out generalizations about the layout of parts: e.g. windows are normally placed in walls, not floors. Rumelhart 1980 uses the term *schema* in precisely this sense, speaking for instance of a schema for a face, with subschemas for noses and eyebrows and so on.
- Schemas both motivate structure within stored items and assign structure to novel items: Again, there is no reason to think that when a novel object is committed to memory, the

principles determining its structure should be any different. (Linguistics has tended to miss this generalization!)

- Free items and bound items: A stripe is bound: there can't be a stripe without a surface. Holes, cracks, and dents: there can't be a hole without a volume in which it is situated.A handle may be physically free (it can be bought in the hardware store) but it is functionally bound (it has to be a handle of something, used to pick that something up).

Geography and spatial layout

Involved in two different tasks that require similar sorts of knowledge: finding one's way from one place to another, and knowing where to find particular objects.

- Vast lexicon: Yes! How many places, streets, routes, do you know (finding your way around Barcelona)? How many associations of objects with places do you know? Where do you find the milk in the market? Where do you keep your envelopes? Where did you park your car this morning? (And possibly: In which publication did Chomsky say such-and-such?) It's hard to know how to count, but we have huge amounts of such knowledge.
- Multiple linked levels of representation: Hard to know.
- Hierarchical constituent structure within items: Yes. Where is the supermarket? Where is the milk within the supermarket? My brother's house is on a particular street in a particular neighborhood of a particular city, in a particular part of a particular state.
- Relational links among items that pick out shared structure: Maybe. Similarities among particular airports?
- Regularities among items, picked out by schemas: Yes. What one is likely to find in an airport vs. in a bank vs. in a restaurant?
- Schemas both motivate structure within stored items and assign structure to novel items: Yes. One may use what one knows about supermarkets, based on supermarkets one has experienced, to make informed guesses about how a newly encountered supermarket is going to be organized.
- Free items and bound items: Hard to know.

Social knowledge

Involved in many different issues: People you know and what you know about them; conventionalized social actions; and issues of moral value (Jackendoff 2007).

- Vast lexicon: Yes! You know thousands of people to some degree or another, including not only how they look, but their personality and their social ties such as kin, spouses, ethnicity, religion, and occupation. You know lots of rules of proper behavior, though it's hard to know how to count them: the fork goes to the left of the plate, people go to church on Sundays, you should give kids birthday presents….
- Multiple linked levels of representation: Yes! A person is conceptualized as a linked physical instantiation (body) and a social presence (soul) (Bloom 2004, Jackendoff 2007). Social actions have a physical instantiation linked to a social or moral value (e.g. shaking hands is a physical action fulfilling the social function of greeting.

- Hierarchical constituent structure within items: Yes! Family structure, group structure (groups within groups within groups), authority (or rank) hierarchy.
- Relational links among items picking out shared structure: Yes? Shared rank (all sergeants), shared occupation (all psycholinguists), shared intellectual ancestry (all former students of Chomsky). Similarities among games (what soccer and hockey have in common; what pingpong and tennis have in common)
- Regularities picked out by schemas: Yes! All customs, Schank's scripts (how a trip to a restaurant works), Minsky's (1975) frames (how a birthday party works), Goffman's (1974) frames (how a theater performance works), Rumelhart's (1980) schemas (how a buying/selling transaction works), and at a very general level, Fiske's (1991) four elementary forms of human relations.
- Schemas both motivate structure within stored items and assign structure to novel items: The stored items here would be particular occasions stored in episodic memory. They are presumably structured according to the same principles as one's assessment of ongoing events.
- Free items and bound items: Moral values are not free-floating: They must be attached to some action.

Conclusion: at least at this impressionistic level, many general properties of linguistic knowledge are replicated in other domains of knowledge.

People such as Lakoff and (early) Rumelhart might say this shows there's nothing special about language.

But that's too coarse a judgment. Domains differ in what their structures are made of: phonological, morphological, and syntactic units in language; sequences of pitches with duration in music; visual/spatial units in object knowledge and geographical knowledge; persons, units of interpersonal behavior and social/moral value in social knowledge.

In order to develop this line of inquiry further, we need theories of representation in these other domains – theories that are comparable in sophistication with linguistic theory. Linguistics should be leading the way.

Overall conclusions

Parallel Architecture/Conceptual Semantics/Simpler Syntax/Relational Morphology make possible rigorous description of linguistic structure:

- From most general to most idiosyncratic aspects of structure
- Integrated and highly constrained theory of syntax, semantics, morphology (and phonology, though not treated here) – no underlying forms, no movement, no "inputs" and "outputs"
- Takes account of data and arguments from language processing and language acquisition (only touched on very lightly here)
- Leads to better articulated hypotheses of evolution of language capacity
- Makes connection between language and visual/spatial understanding

- Makes connection between lexicon and knowledge of other domains

These properties reflect well on Parallel Architecture as a whole. To the extent that it invites integration within the language faculty and integration of language with the rest of the mind, I'm encouraged us to think we're on the right track.