Syntax IN the lexicon
Ray Jackendoff
Tufts University

Two continuing themes of generative grammar: mentalism and formalism

Of the two, mentalism is more important; the formal technology has to adapt when necessary, in order to better fit the facts of language.

Formal technology should support the integration of the theory of language with theories of other cognitive capacities.

Taking "knowledge of language" very seriously:

(1) a. What linguistic units and linguistic complexes does a speaker of language L store in long-term memory?
  b. What procedures does a speaker use to construct novel composite structures?
  c. What procedures does a language learner use to construct the material in (1a,b)?

Standard answer to (1a): *the lexicon.*
Standard answer to (1b): *the grammar.*
Standard answer to (1c): *Universal Grammar* and/or *the Language Acquisition Device.*

Rules of the grammar have to be stored in long-term memory too! And rules that differ among languages have to be acquired.

Generative grammar has assumed throughout its history that lexicon and grammar are fundamentally different kinds of mental representations.

My argument: There are benefits to be gained by giving up this assumption.

Slippery slope: There are things you need to store in the lexicon that are progressively more and more rule-like, so there seems less and less reason to distinguish them from things that everyone accepts as rules.

Where do you draw the line between words and rules? You don't have to: When you get to the bottom of the slippery slope, you discover it's not so bad down there after all (cf. HPSG, Cognitive Grammar, Construction Grammar).

What's a word?

In anybody's theory, a word is linkage of composite phonology, composite semantics, some syntactic features.

(2)   Phonology: /kæt/
       Syntax: +N, -V
       Semantics: FELINE, PET, etc.

In a mentalist theory, this association is instantiated in long-term memory.

Not all words have features in all three domains.

(3)   Phonology and semantics, no syntax:
        hello, ouch, yes, oops, upsey-daisy, allakazam, feh, uh-oh

Combine only in paratactic constructions (4a) and quotative constructions (4b,c).

(4) a. Hello, Bill.
    b. "Hello," she said. (cf. "Shema Yisroel," she said.)
    c. the word *hello*   (cf. the word *pupille*)

(5)   Phonology and syntax, no semantics:
        epenthetic *it*, *do*-support *do*, *of* in *N of NP*

(6)   Syntax and semantics, no phonology:
        PRO, *pro*

(7)   Phonology, no syntax or semantics
        hey diddle diddle, e-i-e-i-o, brrr-raka-taka

What *else* do you have to store?

For one thing, idioms, which are syntactically composite.

(8)   Phonology: /kɪk#ðə#bʌkət/
        Syntax: [$_{VP}$ V [$_{NP}$ Det N]]
        Semantics: DIE (X)

Idioms can have argument structure, just like verbs.

(9)   take NP for granted
        put NP on ice
        give NP the once-over

Many, perhaps most, idioms have canonical syntactic structure, but some don't:

(10)a. day in day out
        by and large
        give NP the once-over (what's *once-over*?)
        (all) V/N-ed out (*all knitted out*, *all coffeed out*, etc.)

b. How about XP?
   What about XP?
   Prt with NP!   (*off with his head*, *down with the government*, etc.)

The idioms in (10b) in addition are full utterances, but not sentences: they don't embed.

(11)   *I wondered (whether) how about lunch.   (cf. "How about lunch?" I wondered.)
       *Jill asked what about Bill.   (cf. Jill asked, "What about Bill?"
       *The crowd demanded (that) off with his head. (cf. "Off with his head!" they demanded.)

"Constructional idioms" with canonical syntax, but to unusual semantic ends:

(12)a. *Way*-construction:
      [$_{VP}$ V pro's way PP]
      Bill belched his way out of the restaurant. (= 'Bill went out of the restaurant belching')
    b. Time-*away* construction:
      [$_{VP}$ V [$_{NP}$ (*time*)] away]
      Fred drank the afternoon away. (= 'Bill spent/wasted the afternoon drinking')
    c. *Head off* construction:
      [$_{VP}$ V pro's head/tush/butt off]
      Suzie sang her head off.   (= 'Suzie sang a lot/intensely')

Each of these involve stored composite syntax, stored idiosyncratic link to meaning, stored
    phonology of designated elements *way*, *away*, and *off*.

Constructional idioms that lack distinguishing phonological content (and therefore are not very
    wordlike):

(13)a. Sound+motion construction
      [$_{VP}$ V PP]
      The trolley rumbled around the corner. (= 'The trolley went around the corner,
      rumbling')
    b. Inverted NP construction:
      [$_{NP}$ a/this/that N of an N]
      that travesty of a theory (= 'that theory, which is a travesty') (related to P. Postal's
      "imposter" construction)
    c. Light verb construction:
      [$_{VP}$ V NP NP/PP]
      Pat gave Kim a hug. (= 'Pat hugged Kim')
    d. Paratactic conditional:
      [S, S]
      You take one more step, I bust you one.

Knowledge of these constructions has to be stored as an association between a syntactic complex and a semantic complex. But the basic formalism is the same as that for words, except that the syntax, like the phonology and semantics, is composite.

Is this "just periphery"? *My* question: What do you store in longterm memory? These phenomena are all part of the answer.

Question: What is the distinction between idioms and constructions that are syntactically canonical, like (9) and (12), and those that are not, like (10)? Canonical ones ought to "cost less" in some sense. How should the theory reflect that difference?

Standard way to create syntactically composite expressions: build them from smaller units, using rules of grammar that apply to lexical items to build phrases that *aren't* lexical items. Problem: Syntactically composite expressions *within* the lexicon create a sort of ordering paradox: generalization *in* the lexicon can be stated only in terms of rules that apply *outside* or "after" the lexicon.

A related problem: the treatment of morphology

Earliest generative grammar (esp. Lees): morphological structure derived by transformations. Difficulty: Transformations are supposed to be productive, but much morphology is not productive, in three ways.

1. In many cases, instances of a rule have to be listed individually (*recital*, *recitation* vs. **incital* and **incitation*; *confusion*, *refusal* but **refusion* and **confusal*).
2. Existence of "derived" items without a "root" (*ambition*, *ablution*, *retribution*.
3. Idiosyncratic "derived" meanings that can't be generated by a general meaning-blind transformation (*recitation* and *recital* again).

Rules with some combination of these properties: "semiproductive."

One solution: Lakoff's system of exception features. Solved first and second difficulties (and paved the way for generative semantics), but did not solve the third.

Chomsky's solution: Semiproductive morphology is "in the lexicon," prior to lexical insertion. In the air at the time: Regularities in the lexicon are accounted for by "lexical redundancy rules," distinct from rules of phrasal grammar. LFG and HPSG take this distinction as foundational.

Some formulations: Lexical redundancy rules are derivational. BUT doesn't solve the problem of semantic idiosyncrasy.
Other formulations: Lexical redundancy rules measure redundancy among fully stored items.

Later: Head-raising as a way of syntactically deriving morphologically complex items (e.g. Hale and Keyser). BUT this solution is subject to all three difficulties above. And is head-movement in the syntax? (Then ordering paradox) Or in the lexicon? (Then lexicon is not just long-term memory storage.)

Alternative: Lexical rules as nodes in an inheritance hierarchy – structural schemas with variables in them, from which items in the lexicon can inherit their own structure.
Recital inherits structure both from *recite* and from schema (14a) – plus idiosyncratic semantic structure of its own.
Ablution inherits structure from schema (14b); since there is no root *ablute*, it has to "pay" for its pseudo-root.
Both schemas inherit structure from the more general nominalization schema (14c).

(14)a. $[_N$ V + al]
    b. $[_N$ V + tion]
    c. $[_N$ V + *aff*]

(14a-c) are listed in the lexicon too: pieces of morphophonological and morphosyntactic structure that must be learned as part of one's knowledge of English.
Problem: The very same format works for *productive* morphology, e.g. (15) for English present participles.

(15) $[_{Vprespart}$ V + ing]

For productive morphology, we don't think you have to store every form. You automatically produce present participle for new verbs (e.g. *skyping*). Looks more like syntax.
What's the difference between (14) and (15)? (cf. words vs. rules controversy)

Proposal: The difference is not "where" the schema is in the grammar – "lexicon" vs. "syntax." (14) and (15) have exactly the same format, so they should both be in the same component. Since (14) can't be in the syntax, (15) must be in the lexicon.
Proposal: Distinction between (14) and (15) is in character of the variable.
- Variable in (15) is marked *productive*, can be used to create new forms online.
- Variable in (14) is marked *semiproductive*, captures generalizations among listed items.
- Productive schemas can also serve as ancestors of stored items (e.g. *scissors*, *trousers*).

Evidence: Principles for naming geographic features in English.

(16)a. [$_N$ Name$_{prod}$ N$_{semiprod}$]
     Arrowhead <u>Lake</u>, Biscayne <u>Bay</u>, Loon <u>Mountain</u>, Laurel <u>Hill</u>, Claypit <u>Pond</u>, etc.
  b. [$_N$ N$_{semiprod}$ Name$_{prod}$]
     <u>Lake</u> Michigan, <u>Mount</u> Everest
  c. [$_{NP}$ the Name$_{prod}$ N$_{semiprod}$]
     the Indian <u>Ocean</u>, the Black <u>Sea</u>, the Hudson <u>River</u>, the Ventura <u>Freeway</u>, the San
     Andreas <u>Fault</u>
  d. [$_{NP}$ the N$_{semiprod}$ of Name$_{prod}$]
     the <u>Bay</u> of Fundy, the <u>Gulf</u> of Aqaba, the <u>Isle</u> of Man

Any name is possible in these schemas: this variable is productive.
Which schema is appropriate depends on which type of feature (*lake*, *ocean*, etc.), which must
    be learned: this variable is semiproductive.
These schemas have one variable of each type, so semiproductivity can't be "in the lexicon"
    and productivity "in the syntax." They both have to be in the same component, as in (16).
Conclusion: Distinction between productivity and semiproductivity is localized not in the type
    of rule, but in the type of variable.

Returning to idioms and constructions

Phrase structure rules can be stated as syntactic schemas without any association to phonology
    or semantics.

(17) [$_{VP}$ V$_{prod}$ – (NP$_{prod}$) …]

No problem situating this in the lexicon: a piece of syntactic structure with productive
variables.
- Licenses novel, freely combined VPs in syntax.
- Acts as ancestor in the inheritance hierarchy of all VP idioms and constructions (e.g. (8),
(9), (12) – accounts for difference between them and the idioms and constructions with
noncanonical syntax.
- Avoids ordering paradox that arises from keeping syntax and lexicon distinct.

(17) is itself a descendant of the even more general schema (18): setting of head parameter in
    X-bar theory.

(18) [$_{XP}$ X$_{semiprod}$ …]

Since variable in (18) is semiproductive, it allows for other, competing schemas.

Consequence of this treatment of phrase-structure in the lexicon: syntax must be stated in terms of licensing conditions (cf. HPSG, LFG), rather than derivations.

Notion of movement has to be eliminated. How to deal with long-distance dependencies and the like? See treatments in e.g. HPSG, LFG, and Parallel Architecture/Simpler Syntax.

Overall hypothesis within Parallel Architecture: Syntax is in the lexicon. Rules of grammar are schemas with productive or semiproductive variables, stored in the brain in the very same format as words: pieces of stored structure.

Online composition

What's the mechanism that accomplishes online composition to build novel sentences from lexical entries?

The formal device is Unification. Unification can combine partially overlapping structures (cf. my recent article in *Language*), which makes it more flexible than Merge.

Unification can be stated declaratively: competence theory that says what sentences are licensed by the grammar.

Or it can be stated procedurally: description of language processing.

Hence grammar of this sort can serve either as a model of competence or of performance – important desideratum for explaining how language is incorporated into the brain.

Chomsky's conjecture: All differences among languages are encoded in the lexicon, and there is only one, universal, combinatorial rule.

Present theory of grammar has this property:
- Everything previously called a rule of grammar is now a lexical item.
- There is only one combinatorial rule left, only it's not Merge, it's Unification.