A. Hierarchical structure.
What does it really mean for language not to have ‘hierarchical structure’? For decades, if not centuries, the study of language has revealed that strings of words appear to function as units that constitute parts of larger units. Evidence for this phenomenon comes not just from compositional semantics, but from the challenge of trying to characterize in a general way what constitutes a well-formed expression in a language. E.g., a unit of a certain type, regardless of its internal makeup, appears in a particular sequence with respect to units of other types, and has a certain function with respect to the entire expression of which it is a part. This relationship is what is thought of as ‘hierarchical structure’ or ‘phrase structure’. The ability of speakers to create novel utterances suggests that they have knowledge of the structural and combinatorial possibilities.

If we agree on such facts, then the question must be not whether there is such a thing as hierarchical structure, but how it is represented in the mental representation of language. What form does the knowledge take? Is there a symbolic representation with this structure, or is the structure a description of the sequence of computations of the system that produces and interprets strings? If I must perform actions A and then B in order to perform a complex action C (e.g. I reach out my hand and grab the cup as part of drinking the tea), C has a hierarchical structure in the sequence of actions, and the static symbolic representation C -> AB abstracts away from the actual dynamics. And A, B and C may be abstractions as well. But as humans we seem to categorize action in this way.

If we do not agree on such facts, then the challenge becomes even more interesting: can we adequately describe how language works without reference to categories and hierarchical structure, and can we do so without smuggling such reference in by the back door?

B. Long distance dependencies.
These are, presumably, what linguists call A-bar or filler-gap constructions. They are characterized by two properties:
(1) a constituent of a sentence (see A. above) is in a position that does not in itself determine its function and (2) there is no constituent in a particular position in the sentence to which a function must or may be assigned (depending on the verb and other factors). This empty position, a ‘gap’, determines the function of the constituent in the A-bar position, and the relationship between the two is called a ‘chain’.

There are many A-bar constructions in languages. The wh-interrogatives are the parade case (“Who are you talking to ___?”). The assumption that there is a gap after ‘to’ that forms a chain with the A-bar constituent captures the general property of ‘to’ that it does not appear alone without an NP complement.

There are several varieties of wh-constructions cross-linguistically when multiple wh-questions are considered. There is the English type, where a wh-phrase appears in A-bar position (as in “Who are you talking to ___?”) and if there are more than one wh-phrase, only one appears in this position. The others appear ‘in situ’ and have their scope interpreted with respect to the wh-phrase in A-bar position. E.g. “Who did you say ___ gave what to whom?”.

There is the German (dialect)/Hungarian/Hindi type, where a wh-phrase appears in A-bar position, but may appear in an intermediate A-bar position in a complex sentence, with an expletive wh-word marking the scope of the question (“Was glaubst du wen er ___ gesehen hat”/what think you who.ACC he saw has/ ‘Who do you think he saw?’). And there is the Slavic type, where a wh-phrase appears in initial position, and where all wh-phrases appear in initial position in a multiple wh-question.

Languages without A-bar wh-questions have "wh-in-situ."

An interesting question is, can the German/Hungarian/Hindi type of intermediate wh-question be characterized simply through reference to linear order? Reference to the initial position of the complement clause would appear to require reference to the complement clause, which implicates hierarchical structure and (I believe) cannot be characterized simply in terms of linear order (without smuggling). I attach a paper by Joachim Sabel on such ‘partial wh-movement’. I don’t think that it is necessary for participants to read the paper closely or completely, but a look at some of the examples and the discussion at the beginning of the paper may provide a sense of the complexity of the phenomenon and the challenges for accounting for it, even if we take hierarchical structure for granted.

Beyond this, it is well-known that not all chains are equally acceptable. I refer of course to the famous Ross ‘island constraints’ and subsequent manifestations of the same idea. A question that I have been struggling with for many years is to what extent it is possible to account for judgments of violations of island constraints not in terms of grammar but in terms of processing complexity, that is, the complexity of the computation of the meaning of a sentence with a particular chain. A classic example is “Which of these books did you meet the person who wrote ___?” which is decidedly less than perfect and which is an instance of a violation of the Complex NP Constraint (the complex NP being “the person who wrote which of these books” and the constraint being that the wh-phrase cannot be moved out of such a configuration, or form a chain with a gap in such a configuration).

A question that to me is of considerable interest is, Is it possible to account for judgments of such sentences (and there are many other configurations to consider as well) in terms of a computational model of sentence processing that takes into account such factors as hierarchical complexity, surprisal, expectation, frequency, memory, integration cost, etc. — whatever computational and psycholinguistic theory might consider to be relevant to processing complexity? A related question would be, is it possible to do this *without* reference to hierarchical structure, and would that make it easier?

Finally, I note that the filler-gap relationship itself appears to presume hierarchical structure, in that the constituent in A-bar position, the filler, must be a sister of the part of the sentence that contains the gap. So consider
a. She revealed [who you stole the money from __] to the police.
b. *She revealed [who you stole the money from the children] to __
c. Which people did she reveal [who you stole the money from __] to __?

In (1a) “who” forms a chain with the gap, which is in the same clause. In (1b) it does not, because the gap is in a higher clause. In (1c) there are two phrases, each of which is associated with the gap in its own clause — not the other way around. (Although there is some complexity owing to the nesting of the chains.) So it appears that reference to the hierarchical structure is required.

This being said, it is not clear *how much* hierarchical structure is required. Culicover and Jackendoff (2005) [Simpler Syntax, Oxford University Press] argue for minimal hierarchical structure, i.e. no more than is necessary to account for the interpretation. It may well be that within phrases, e.g. VP, the structure is only apparent or virtual, in the sense that it is implicit in the sequence of steps (see A) mapping the string into conceptual structure. And it is possible (and has been generally assumed) that conceptual structure is itself hierarchical, and that imposes this virtual structure on the strings of words that express it. If this is so, we would want to know, exactly how does this work?

C. (Co)reference.

Another place where hierarchical structure has played a central role in linguistic theorizing is the management of referential dependency. The Binding Theory accounts for a vast number of cases involving the possible relationships between dependents (pronouns like “she” and anaphors like “herself”) and their antecedents in terms of ‘binding’, defined over hierarchical structure. (A binds B if A and B are coindexed and A c-commands B = A is the sister of B or a constituent that contains B). An anaphor must be locally bound (A), a pronoun cannot be (B), and a pronoun cannot bind its antecedent (C).

A reasonable question to ask about the Binding Theory is, To what extent are the binding conditions the product of the computation of referential dependencies time as the discourse representation is constructed? In order to take this possibility seriously, we have to understand what is going on in the construction of discourse representations independently. To get an explanation for a binding condition out of this, we cannot “smuggle” hierarchical structure into the discourse structure. For example, condition A rules out (2a) but not (2b) (where the pronoun is assumed to be coreferential with Otto).

(2) a. *He loves Otto.
b. His mother loves Otto.

We could structure the discourse representation so that the subject of a sentence, e.g. “he” or “his mother” is a unit and gets its reference fixed immediately when processed. But we would have to stipulate that a part of the subject, e.g. “his” does not, so that when “Otto” is encountered, “he” is not available for coreference, but “his” is. But what we have done then is shift the structure into the discourse representation wholesale, and reformulate the binding condition in terms of this structure. If this is all we’ve done, then we have not gained any new insight, just reformulated the problem using a different terminology. But if we can motivate a difference in the construction of discourse representation for “he” (or “him”) and “his mother” independently, regardless of where it appears in the sequence, and if this difference explains the coreference possibilities, we may have an explanation.

More complex examples pose similar puzzles.

(3) a. *He believes that Otto will win.
b. *What he believes is that Otto will win.
c. That Otto will win is what he believes.

(4) a. *He just went out and bought the books that Otto wanted to read.
b. The books that Otto wanted to read, he just went out and bought.

These examples have the same hierarchical structure, for the most part. They seem to suggest that linear order interacts with and may preempt hierarchical structure. But whether it is possible to completely dispense with hierarchical structure in the explanation of binding is unclear.

Other cases involve so-called “weak crossover”. (1b) is classically claimed to be unacceptable, but manipulation of the lexical items and context improves the judgment.

(5) a. Who loves his mother?
b. (?)Who does his mother love?
c. Which well-known linguist did his university try to get rid of?
d. Otto, his mother really loves.
e. The person that his mother loves is Otto.

(5b-e) are structurally identical as far as the relationship between the antecedent and the dependent are concerned, but only (5b), with an unspecified unrestricted interrogative, is at all problematic.