A rough definition of syntax:

(1) The laws that govern how words combine to form sentences.

This presupposes that we have definitions of “words” and “sentences,” and we don’t. Over a large range of cases, though, speakers seem to have clear judgements about what words and sentences are, and so we can start with those judgments. We might have occasion to get closer to a definition as we go along.

How do we find out about these laws? What is our evidence? We could simply observe what people say, of course, and record the orderings of words that they use. But this has a few large flaws. People make mistakes — they say things that they feel aren’t actually “good.” And it’s incredibly cumbersome and slow to collect examples in so haphazard a fashion. (Web searches, however, have helped somewhat with this.) And finally, as we’ll see, it’s often necessary to know what the meaning of a string of words is before we know what sentence it is. For these reasons, we don’t simply collect examples of spoken sentences.

Instead we ask speakers for their grammaticality judgments. We ask them to explicitly tell us whether a particular string is something they judge to be a well-formed sentence of their language. Grammaticality judgements have their own perils. And we should acknowledge them at the outset so that we do a better job of dodging them. When you ask someone whether a particular string is a grammatical sentence, they are going to tell you is whether or not it is a good sentence. Grammatical and good are not the same. Many overly-educated people have an idea about what a good sounding English sentence is that derives from their schooling. And those with a literary bent will also have an idea about the sound of a sentence, perhaps coupled with an aesthetic that favors some over others. We’re not interested in any of these judgments. They aren’t part of syntax; they are part of sociology.

Let me give you a few famous examples of this sort. I learned in school that the sentences in (2) are bad and that their good counterparts are (3).

(2) a. Who did you visit?
   b. What did you talk about?
   c. I wanted boldly to go where no one has gone before

(3) a. Whom did you visit?
   b. About what did you talk?
   c. I wanted boldly to go where no one has gone before.

(2b) and (2c) involve word orders that aren’t possible in Latin, and this seems to have had an influence on decisions about what are good sentences. As it happens, in my English all of the sentences in (2) and (3) are grammatical, though the ones in (3) employ a register that I practically never use. I suspect that (3) are part of my grammar only because of my schooling.

The cognitive process that grammaticality judgments tap isn’t something about just recognizing things we’ve heard before. We can see this by appreciating that there is a difference in grammaticality between the sentences of (4), even though both are equally “new.”

(4) a. Whenever the earth revolves around its equator, the moon begins to rotate about its axis.
   b. *Whenever the earth revolves around its equator, the moon begins itself to rotate about its axis.

The first of these is grammatical, I judge, whereas the second isn’t. This is what the asterisk “*” means. We seem to have the ability to recognize an arrangement of English words which is completely new as either a good arrangement, i.e., grammatical, or not. This ability, then, isn’t familiarity. Our syntactic knowledge is not something about matching sentences we’ve heard previously. We know something about sentences that we’ve never heard before.

Ungrammatical and difficult to understand are easy to confuse. To see this, consider (5).

(5) I decided to marry on Tuesday.

There are really two sentences here. Each has the same series of words, but they are grouped in such a way that they yield different meanings (as we shall see). The two meanings can be paraphrased as:

(6) a. I decided on Tuesday to marry.
   b. I decided that the day I should marry would be Tuesday.
I can bring out these two meanings with prosody. Let a "," indicate a slight pause, and you can hear the differences in meaning in:

(7) a. I decided to marry, on Tuesday.
   b. I decided, to marry on Tuesday.

Did you hear both meanings in (5) right away?

Now, technically, there are two sentences in (5). So when we ask for a grammaticality judgment, we are asking whether it is possible for a particular arrangement of words to form a certain meaning. When we get around to defining "sentence," in other words, it will have to be based on meaning. In the case of (5), we are finding that there are two grammaticality judgments involved. One pairs that string with the meaning in which on Tuesday describes the time of the decision, and the other pairs that string with the meaning in which on Tuesday describes the time of the (planned) marriage.

Now consider (8).

(8) I decided that my daughter should marry on Tuesday.

It is harder, studies show, for us to understand this with the meaning in which on Tuesday describes the time of the decision. Does that mean that there is only one sentence here? There is some evidence that there are two sentences here, but that it is just more difficult to hear them both. It is still, perhaps, possible to bring out the other meaning with prosody:

(9) I decided that my daughter should marry, on Tuesday.

Unless we are careful, then, we might mistakenly judge a grammatical sentence ungrammatical just because the meaning were pairing that string of words with is difficult to hear.

This problem arises in cases where it's hard to distinguish ungrammatical from requiring a context in which the relevant meaning is easy to hear. Let me illustrate this problem with an example that involves what's called "Ellipsis." Consider:

(10) Jerry will annoy everyone that [\_ Sean will \triangle].

This has the same meaning that (11) has.

(11) Jerry will annoy everyone that [\_ Sean will annoy].

There is understood to be the verb annoy in the sentence in brackets. One way of thinking about why this is so is to imagine that there is a rule of English syntax that allows verbs to go unpronounced when they are spoken somewhere else. This is "Ellipsis," and where there is an unpronounced thing is indicated with "\triangle."

Now consider (12).

(12) a. Whomever she did \triangle got better.
   b. Every assignment to \triangle was hard.

I am tempted to judge these both ungrammatical. If that is correct, then what we need to do is formulate the rule of Ellipsis so that it can apply in (10) but not (12).

But now consider what happens if we put these same sentences within larger ones that supply a context.

(13) a. Whomever Sally didn't tutor got worse but whomever she did \triangle got better.
   b. * Every assignment to do was easy but every paper to was hard.

The first improves, but the second does not. The problem with the sentences in (10) is that they didn't obey the requirement of Ellipsis that the verb that goes unpronounced be spoken somewhere in the context. Once that requirement is met, we can see that there is a difference between the sentences that responds to something else about the rule of Ellipsis. It appears that not all verbs can be elided, even when they match something spoken elsewhere.

We will use grammaticality judgments to get at our knowledge of syntax, but these problems will bedevil us. We will sometimes have to use caution. The following Laws can be useful.

(14) 

Two Laws of Elicitation

a. The sentences for which you elicit a grammaticality judgement should be embedded in a discourse that makes their meaning salient.
   b. Every suspected ungrammatical sentence should be part of a minimal pair, the other member of which is grammatical.

A lot about how speakers distinguish a grammatical from ungrammatical sentence can be expressed on the basis of the classes that words belong to. These are known as morpho-syntactic categories (or just categories). We can see this in two ways.

First, knowing just where a word can stand in a grammatical sentence confers knowledge about where it can stand in other sentences. Suppose, for instance, that I tell you:

(15) Many bloresnicks are grey.

and you decide that I have used a grammatical sentence, albeit one with a word you don't know: bloresnicks. Knowing only that, you are probably able to make grammaticality judgments about sentences in (16).
Your knowledge of English syntax allows you to know, given the position a word has in some grammatical sentence, what its likely positions in any grammatical sentence can be. Knowing where a word in a grammatical sentence is tells you something about the “kind” of word it is, and knowing the kind of a word is enough to allow you to make good guesses about where it can be positioned in any sentence.

Knowing just what a word means also seems to confer knowledge about where it can stand in grammatical sentences. To see this, let me define a word for you:

\( \text{pondel}: ([\text{pand}]]) \) unwanted facial hair.

Knowing this probably allows you to have a feeling about which of the following sentences is grammatical and which isn’t.

(17) a. Many pondels are grey.
    b. He ran pondel the tree.
    c. He made his face pondel.
    d. He removed the long pondel.
    e. She finds Sammy pondel.
    f. He made his face pondel.