

What is syntax?

1.1 SOME CONCEPTS AND MISCONCEPTIONS

1.1.1 What is the study of syntax about?

This book is about the property of human language known as syntax. ‘Syntax’ means ‘sentence construction’: how words group together to make phrases and sentences. Some people also use the term *GRAMMAR* to mean the same as syntax, although most linguists follow the more recent practice whereby the grammar of a language includes all of its organizing principles: information about the sound system, about the form of words, how we adjust language according to context, and so on; syntax is only one part of this grammar.

The term ‘syntax’ is also used to mean the *study* of the syntactic properties of languages. In this sense it’s used in the same way as we use ‘stylistics’ to mean the study of literary style. We’re going to be studying how languages organize their syntax, so the scope of our study includes the classification of words, the order of words in phrases and sentences, the structure of phrases and sentences, and the different sentence constructions that languages use. We’ll be looking at examples of sentence structure from many different languages in this book, some related to English and others not. All languages have syntax, though that syntax may look radically different from that of English. My aim is to help you understand the way syntax works in languages, and to introduce the most important syntactic concepts and technical terms which you’ll need in order to see how syntax works in the world’s languages. We’ll encounter many grammatical terms, including ‘noun’, ‘verb’, ‘preposition’, ‘relative clause’, ‘subject’, ‘nominative’, ‘agreement’ and ‘passive’. I don’t expect you to know the meanings of any of these in advance. Often, terms are not formally defined when they are used for the first time, but they are illustrated so you can understand the concept, in preparation for a fuller discussion later on. More complex terms and concepts (such as ‘case’ and ‘agreement’) are discussed more than once, and a picture of their meaning is built up over several chapters. A glossary at the end of the book provides definitions of important grammatical terms.

To help you understand what the study of syntax is about, we first need to discuss some things it isn’t about. When you read that ‘syntax’ is part of ‘grammar’, you may have certain impressions which differ from the aims of this book. So first,

although we will be talking about grammar, this is not a DESCRIPTIVE GRAMMAR of English or any other language. Such books are certainly available, but they usually aim to catalogue the regularities and peculiarities of one language rather than looking at the organizing principles of language in general. Second, I won't be trying to improve your 'grammar' of English. A PRESCRIPTIVE GRAMMAR (one that prescribes how the author thinks you should speak) might aim to teach you where to use *who* and *whom*; or when to say *me and Kim* and when to say *Kim and I*; it might tell you not to say *different than* or *different to*, or tell you to avoid split infinitives such as *to boldly go*. These things aren't on our agenda, because they're essentially a matter of taste – they are social, not linguistic matters.

In fact, as a linguist, my view is that if you're a native speaker of English, no matter what your dialect, then you already know English grammar perfectly. And if you're a native speaker of a different language, then you know the grammar of that language perfectly. By this, I don't mean that you know (consciously) a few prescriptive rules, such as those mentioned in the last paragraph, but that you know (unconsciously) the much more impressive mental grammar of your own language – as do all its native speakers. Although we've all learnt this grammar, we can think of it as knowledge that we've never been taught, and it's also knowledge that we can't take out and examine. By the age of around seven, children have a fairly complete knowledge of the grammar of their native languages, and much of what happens after that age is learning more vocabulary. We can think of this as parallel to 'learning' how to walk. Children can't be taught to walk; we all do it naturally when we're ready, and we can't say how we do it. Even if we come to understand exactly what muscle movements are required, and what brain circuitry is involved, we still don't 'know' how we walk. Learning our native language is just the same: it happens without outside intervention and the resulting knowledge is inaccessible to us.

Here, you may object that you *were* taught the grammar of your native language. Perhaps you think that your parents set about teaching you it, or that you learnt it at school. But this is a misconception. All normally developing children in every culture learn their native language or languages to perfection without any formal teaching. Nothing more is required than the simple exposure to ordinary, live, human language within a society. To test whether this is true, we just need to ask if all cultures teach their children 'grammar'. Since the answer is a resounding 'no', we can be sure that all children must be capable of constructing a mental grammar of their native languages without any formal instruction. Most linguists now believe that, in order to do this, human infants are born pre-programmed to learn language, in much the same way as we are pre-programmed to walk upright. All that's needed for language to emerge is appropriate input data – hearing language (or seeing it; sign languages are full languages too) and taking part in interactions within the home and the wider society.

So if you weren't taught the grammar of your native language, what was it you were being taught when your parents tried to get you not to say things like *I ain't done nowt wrong*, or *He's more happier than what I am*, or when your school teachers tried to stop you from using a preposition to end a sentence with? (Like the sentence

I just wrote.) Again, consider learning to walk. Although children learn to do this perfectly without any parental instruction, their parents might not like the way the child slouches along, or scuffs the toes of their shoes on the ground. They may tell the child to stand up straight, or to stop wearing out their shoes. It's not that the child's way doesn't function properly, it just doesn't conform to someone's idea of what is aesthetic, or classy. In just the same way, some people have the idea that certain forms of language are more beautiful, or classier, or are simply 'correct'. But the belief that some forms of language are better than others has no linguistic basis. Since we often make social judgements about people based on their accent or dialect, we tend to transfer these judgements to their form of language. We may then think that some forms are undesirable, that some are 'good' and some 'bad'. For a linguist, though, dialectal forms of a language don't equate to 'bad grammar'.

Again, you may object here that examples of NON-STANDARD English, such as those italicized in the last paragraph, or things like *We done it well good*, are sloppy speech, or perhaps illogical. This appeal to logic and precision makes prescriptive grammar seem to be on a higher plane than if it's all down to social prejudice. So let's examine the logic argument more closely, and see if it bears scrutiny. Many speakers of English are taught that 'two negatives make a positive', so that forms like (1) 'really' mean *I did something wrong*:

- (1) I didn't do nothing wrong.

Of course, this isn't true. First, a speaker who uses a sentence like (1) doesn't *intend* it to mean *I did something wrong*. Neither would any of their addressees, however much they despise the double negative, understand (1) to mean *I did something wrong*. Second, there are languages such as French and Breton which use a double negative as STANDARD, not a dialectal form, as (2) illustrates:¹

- | | | | | | | | |
|-----|---------------------|-----------|-------|---------------|----|---------|----------|
| (2) | Je | ne | mange | jamais | de | viande. | (French) |
| | I | NEGATIVE | eat | never | of | meat | |
| | 'I never eat meat.' | | | | | | |

Example (2) shows that in standard French the negative has two parts: in addition to the little negative word *ne* there's another negative word *jamais*, 'never'. Middle English (the English of roughly 1100 to 1500) also had a double negative. Ironically for the 'logic' argument, the variety of French that has the double negative is the most formal and prestigious variety, whereas colloquial French typically drops the initial negative word.

Another non-standard feature of certain English dialects which doesn't conform to prescriptive notions is illustrated in (3), from a northern (British) English dialect:

- (3) I aren't going with you.

Here, the logic argument runs like this: you can't say **I are not* (the star or asterisk is a convention used in linguistics to indicate an impossible sentence), so the

contracted form *I aren't* must be wrong too. It's true that speakers who accept (3) don't ever say *I are not*. But the argument is flawed: standard English is just as illogical. Look how the statement in (4a) is turned into a question in (4b):

- (4) a. **I'm not** going with you.
b. **Aren't I** going with you?

Example (4) does not conform to the usual rules of English grammar, which form questions by inverting the word order in *I can't* to give *can't I*, and *I should* to give *should I*, and so on. Given these rules, the 'logically' expected form in (4b) would be *amn't I* (and in fact this form is found in some dialects). If the standard English in (4) fails to follow the usual rules, then we can hardly criticize (3) for lack of logic. And since *aren't I* is OK, there's no logical reason for dismissing *I aren't*. The dialects that allow either *I aren't* or *amn't I* could actually be considered more logical than standard English, since they follow the general rule, whilst the standard dialect, in (4), has an irregularity.

It's clear, then, that socially stigmatized forms of language are potentially just as 'logical' as standard English. Speakers of non-standard dialects are, of course, following a set of mental rules, in just the same way that speakers of the most prestigious dialects are. The various dialects of a language in fact share the majority of their rules, and diverge in very few areas, but the extent of the differences tends to be exaggerated because they arouse such strong feelings. In sum, speakers of prestige dialects may feel that only their variety of English is 'grammatically correct', but these views cannot be defended on either logical or linguistic grounds.

If, on the other hand, a speaker of English produced examples like (5), then we could justifiably claim that they were speaking ungrammatically:

- (5) *I do didn't wrong anything.
*Do wrong didn't anything I.

Such examples completely contravene the mental rules of all dialects of English. We all agree on this, yet speakers of English haven't been taught that the sentences in (5) are bad. Our judgements must therefore be part of the shared mental grammar of English.

Most of the rules of this mental grammar are never dealt with by prescriptive or teaching grammars. So no grammar of English would ever explain that although we can say both (6a) and (6b), we can't have questions like (7) (the gap ___ indicates an understood but 'missing' element, represented by the question word *what*):

- (6) a. They're eating eggs and chips.
b. What are they eating ___?

- (7) *What are they eating eggs and ___?

The rules that make (7) impossible are so immutable and fundamental that they hardly seem to count as a subject for discussion: native speakers never stop to wonder why (7) is not possible. Not only are examples like (7) ungrammatical in English (i.e. they sound impossible to native speakers), they are ungrammatical in Welsh, as in (8):

- (8) *Beth maen nhw yn bwyta wyau a ____? (Welsh)
 what are they in eat eggs and
 *‘What are they eating eggs and ____?’

In fact, the equivalents to (7) and (8) are generally ungrammatical in the world’s languages. It seems likely, then, that many of the unconsciously ‘known’ rules of individual languages like English or Welsh are actually UNIVERSAL – common to all languages.



Before reading further, note that English does have a way of expressing what (7) would mean if it were grammatical – in other words, a way of expressing the question you would ask if you wanted to know what it was that they were eating with their eggs. How is this question formed?



You could ask: *They are eating eggs and **what**?* (with heavy emphasis on the *what*).

The fact that certain organizing rules and principles in language are universal leads many linguists to conclude that human beings have an INNATE LANGUAGE FACULTY – that is, one we are born with. We can’t examine this directly, and we still know relatively little about what brain circuitry is involved, but we do know that there must be something unique to humans in this regard. All normal children learn at least one language, but no other animals have anything like language as a natural communication system, nor are they able to learn a human language, even under intense instruction. To try and understand the language faculty, we examine its output – namely the structures of natural languages. So by looking at syntax we hope to discover the common properties between languages, and maybe even ultimately to discover something about the workings of the human brain.

As well as looking for absolutely universal principles, linguists are interested in discovering what types of construction are possible (and impossible) in the world’s languages. We look for recurring patterns, and often find that amazingly similar constructions appear in unrelated languages. In the next paragraph I give an example of this type which compares Indonesian and English. You don’t have to know anything about Indonesian to get the point being made, but if the idea of looking closely at exotic languages seems too daunting at this stage, come back to the examples after you’ve read Section 1.2. The notation $\text{>}\text{---}\text{>}\text{---}\text{>}$ marks the start of a section of the text in which the reader is invited to work something out, as in the