ERG	ergative case
FEM	feminine gender
GEN	genitive case
INDEF	indefinite
INF	infinitive
IO	indirect object or indirect object case marker
M	modality
MSC	masculine gender
NEG	negative
NEU	neuter gender
NOM	nominative case
NP	noun phrase
OBJ	object or object marker
P1	first person plural
P2	second person plural
P3	third person plural
PASS	passive
PERF	perfective
PL	plural
PP	prepositional phrase
PREP	preposition
PST	past tense
Q	question particle
REFL	reflexive
REL	relative clause marker
S1	first person singular
S2	second person singular
S3	third person singular
SBJ	subject case marker
SG	singular
you(r)$_P$	plural 'you(r)'
you(r)$_S$	singular 'you(r)'

Chapter One

What is Syntax?

There are three men on a train. One of them is an economist and one of them is a logician and one of them is a mathematician . . . (T)hey have just crossed the border into Scotland . . . and they see a brown cow standing in the field . . . And the economist says, 'Look, the cows in Scotland are brown.' And the logician says, 'No. There are cows in Scotland of which one at least is brown.' And the mathematician says, 'No. There is at least one cow in Scotland, of which one side appears to be brown.'

(Haddon 2003: 142)

1 Preliminaries

You have just arrived at the airport in Ankara, Turkey, and are standing at the baggage carousel waiting to be re-united with your possessions. Since the belt fails to disgorge your luggage, you go to the service counter to inquire about it. How do you say 'Where are my two suitcases?' in Turkish? Well, this should be easy: first, you look up each word in your English–Turkish pocket dictionary. This is what you find:

(1) 'where'   *nerede*
     'are'   *var*
     'my'   *benim*
     'two'   *iki*
     'suitcase-s'   *bavul-lar*

You then put the words together in some reasonable order, such as what English has, and come up with the following:

(2) *Nerede var benim iki   bavullar?*
     where are my   two suitcases

The chances are good that the clerk at the service counter will understand what you are trying to say; however, from the point of view of grammaticality, the sentence is a disaster. The correct version is this:

(3) *Iki bavul-ım   nerede?*
     two suitcase-my where

In other words, for the meaning of 'Where are my two suitcases?', Turks say the equivalent of 'Two suitcase-my where?'

   What this example shows is what all travellers know: if you want to speak a language, a dictionary is not enough because the meaning you are trying to express does not fully determine word choice and word order. Here is the additional information not generally included in dictionaries that you need to formulate (3):

(4) (a) The word *var* 'are' must not occur after *nerede* 'where?'.
     (b) The word form *bavullar* 'suitcases' must not occur after *iki* 'two'.
     (c) The word *benim* 'my' may or may not occur before the word for 'suitcase' but either way, the form of the word for suitcase must be *bavul-ım*.
     (d) The word *nerede* 'where?' must come after *bavulım* 'my suitcase'.

Instructions of this kind, which specify the choice and ordering of words and word forms, make up the **syntax** of a language.

   This much characterizes the (minimal) content of syntax. However, actual syntactic rules as formulated by linguists radically differ from those in (4) in the manner in which they are stated. The problem with the rules in (4) is that they are too specific. This is for two reasons. First, the rules in (4) mention individual words. But syntactic rules are better formulated in reference to **classes** (also known as categories, or types) **of words**. Here is the categorial re-statement of (4):

(5) (a) Forms of the verb for 'be' must not follow question words that ask about location. (These verb forms are used only to say 'there exist(s)'.)
     (b) Nouns following numerals (e.g. 'two') must occur in singular form. (Plural forms are used only for un-numbered quantities.)
     (c) Possessed nouns (e.g. 'my suitcase') must occur with a suffix showing the person and number of the possessor. (Possessive pronouns occur before nouns only if the pronouns are emphasized, such as in 'my suitcase, not yours'.)
     (d) Question words must occur at the end of the sentence.

The rules of (5) differ from those in (4): where (4) says *bavullar* 'suitcases', (5) says 'nouns'; where (4) says *iki* 'two', (5) says 'numerals'; where (4) says *nerede* 'where', (5) says 'question words'; and so forth. Thus, (5) is a more useful description than (4) because the rules in (4) describe only the Turkish sentences that include the particular words mentioned in the rules, but (5) helps us formulate many other Turkish sentences as well, such as 'Where are your three houses?' *Üç kitab-ı nız nerede?* or 'Where are his five friends?' *Beş arkadas-ı nerede?* These sentences employ different words but follow the same general mould. Rules stated on classes of words highlight the shared template beneath the sentences which, on the face of it, would seem different just because the words are different.

   The same point can be made if we take the English sentences in (6):

(6) (a) *Jack should buy a hat.*
     (b) *Catherine must wash the dress.*

A word-token-based description as (4) above would fail to capture the similarity between (6a) and (6b). These two sentences differ in word tokens but consist of the same word categories: *Jack* and *Catherine* are both nouns, *should* and *must* are both auxiliaries, *buy* and *wash* are both verbs, and *a* and *the* are both articles.

   Descriptions that mention classes of words rather than individual word tokens are useful not only for stating regularities that hold within a language but also for formulating generalizations across languages. Consider the following sentences of French:

(7) FRENCH
     (a) *Jacques faudrait acheter un chapeau.*
        Jack   should to:buy a   hat
        'Jack should buy a hat.'

     (b) *Catherine doit laver   la robe.*
        Catherine must to:wash the dress
        'Catherine must wash the dress.'

The English and French sentences in (6) and (7) contain different words but follow the same pattern in terms of word categories. Word–token-based rules would need to be formulated separately for the two languages – for example, that in English, *must* has to come before *wash* and in French, *doit* 'must' has to come before *laver* 'wash' – but a single category-based rule saying that the auxiliary has to come before the verb defines a uniform blueprint for the two languages. Similarly, the fact that English *a* and *the* and French *un* and *la* come before the noun can be taken care of by a single categorial rule according to which the article comes before the noun in both languages.

However, even categorial rules may miss some similarities among sentences; and this is the second reason why rules like those in (4) are not optimal. Consider translations of the English and French sentences in (6) and (7) into German and Hungarian:

(8) GERMAN

    (a) *Johannes **soll**   einen Hut **kaufen**.*
          Jack   **should** a   hat **to:buy**
          'Jack should buy a hat.'

    (b) *Katherine **muss** das Kleid **auswaschen**.*
          Catherine **must** the dress **to:wash**
          'Catherine must wash the dress.'

(9) HUNGARIAN

    (a) *Jánosnak **vennie kellene** egy kalapot.*
          Jack   to:**buy should** a hat
          'Jack should buy a hat.'

    (b) *Katinak **mosnia kell** a   ruhát.*
          Catherine **to:wash must** the dress
          'Catherine must wash the dress.'

Are these sentences cut by the same template as their equivalents in English and French in (6) and (7)? The answer is yes and no. On the one hand, all these sentences are very much alike in that they all include a verb and an auxiliary as well as two nouns and an article. But on the other hand, they are also different since, as the use of bold makes it clear, the order of the auxiliary and verb is not the same: in English and French, auxiliaries directly precede verbs, in German, they indirectly precede them, and in the Hungarian sentences, they directly follow them. Here are the three order patterns:

(10) (a) ENGLISH and FRENCH:
          Noun & **Auxiliary & Verb** & Article & Noun
      (b) GERMAN:
          Noun & **Auxiliary** & Article & Noun **& Verb**
      (c) HUNGARIAN:
          Noun & **Verb & Auxiliary** & Article & Noun

(The & sign stands for 'immediately precedes'. For example, (10a) reads: 'The noun must immediately precede the auxiliary, which must immediately precede the verb, which must immediately precede the article, which must immediately precede the noun.')

The problem with (10) is that it specifies three different sentence templates. But, as noted above, the three structures are not entirely different: while they differ in the order of some of the categories selected, they agree in the selection of the categories. This similarity among the four languages can be made explicit if we specify selection separately from order, as in (11). (In (11a), the comma is used to indicate co-occurrence between categories without specifying their linear order: (A, B) stands for 'A and B must co-occur in some yet unspecified order'.)

(11) (a) RULE OF SELECTION:
          – for ENGLISH, FRENCH, GERMAN, AND HUNGARIAN:
          (Noun, **Verb**, **Auxiliary**, Article, Noun)
      (b) RULES OF ORDER
          – for ENGLISH AND FRENCH:
          Noun & **Auxiliary & Verb** & Article & Noun
          – for GERMAN:
          Noun & **Auxiliary** & Article & Noun & **Verb**
          – for HUNGARIAN:
          Noun & **Verb & Auxiliary** & Article & Noun

Thus, (11) breaks down the notion 'combination of word categories' into two components: 'selection of word categories' and 'ordering of word categories'. The selection rule in (11a) abstracts away not only from the specific words but also from the specific order in which they appear. Thus, (11) succeeds in capturing both similarities and differences among sentences that differ in word order but contain the same categories. Their similarity is made explicit by the selection schema in (11a) while their difference is shown by the order schemata in (11b).

That the selection of words and their ordering are distinct components of grammaticality is evident also from the fact that there are ungrammatical sentences that violate constraints on order but not on selection. Suppose somebody wants to say in English 'Peter has found a lizard in the yard' and comes up with *Peter has found a lizard yard the in*. This sentence is selectionally sound – it has all the right kinds of words – but is ungrammatical because it does not observe the right order of the noun, article and adposition. Another ungrammatical alternative is **Peter has found a lizard in yard*. The sentence is flawed in a more basic way: it is ungrammatical by selection because the article *the* is not mis-ordered but simply left out.

In sum: syntactic rules specify the selection and order of words. For the rules to be general, they have to be stated on word classes rather than individual words and with selection information separated from ordering. Just as different words are concrete manifestations of word categories, different

orderings of the same words are concrete manifestations of selection patterns: the ordered strings 'Auxiliary & Verb' and 'Verb & Auxiliary' are tokens of a single selectional structure 'Verb, Auxiliary', just as the individual words *boy* and *book* are tokens of the general category 'Noun'.

Before closing this brief introduction to the concept of syntax, it is worth noting the broad applicability of the term 'syntax': in the sense in which we defined the term, all complex objects have syntactic structure. Here are some syntactic rules for a meal at dinner:

(12) (a)   RULE OF SELECTION:
           Soup, salad, meat course and dessert may make up a dinner.
      (b)   RULE OF ORDER:
           Soup must precede salad, which must precede the meat course, which must precede dessert.

Cooking recipes are also 'syntactic descriptions': American cookbooks even follow the two-part format by first listing the ingredients selected and then giving the order in which they need to be mixed. Just as sentences can be selectionally grammatical but ungrammatical in order or ungrammatical even by selection, a cake can also be ruined because the right ingredients have been mixed in the wrong order – such as adding flour to the egg whites before beating them – or because the very choice of the ingredients is incorrect – such as using salt instead of sugar. Images, too, have their own syntax. For example, the picture of a chicken may be ill-formed either by the choice or by the arrangement of the parts. The picture of a chicken fitted with human legs is ill-formed by selection; one with both eyes on one side of the head is selectionally correct but it suffers from an ill-formed arrangement of the parts.

As we are about to embark on a more detailed discussion of syntactic descriptions, an obvious question that arises is about the purpose of all this. What is the goal of the entire endeavour of describing the syntax of sentences? And why should descriptions be general as was suggested in the preceding discussion? In attempting to answer these questions, we will first look at the big picture: the goals of science in general (Section 2). Following that, we will explore how these goals apply to the science of linguistics (Section 3) and then to syntax in particular (Section 4).

2 Studying the world

2.1 EXPLANATION

2.1.1 *Why-questions*

On a Saturday morning, 1 February 2003, the space shuttle Columbia was on its way back to Earth when it exploded in mid-air killing all of its seven crew members. As the event was replayed on TV over and over again with the world watching in horror, there was one question on most people's mind: why did this happen?

The impact of the disaster was different for people depending on what their expectations had been. Those who did not know that a similar disaster destroyed the space shuttle Challenger 17 years earlier might have thought that the explosion of the Columbia was simply impossible. If in turn you had asked people old enough to remember the Challenger catastrophe, they probably would have said it was possible that the Columbia would explode but certainly not that it was necessary or even likely. And some of the engineers at the National Aeronautics and Space Administration who had raised doubts earlier about the space-worthiness of the Columbia might have considered the crash likely – but not that it would necessarily happen.

During the weeks and months following the Columbia disaster, reports about the subsequent investigation were splashed on front pages of newspapers everywhere. Just as the people around the globe, the investigative board also asked the question: why did it happen? Their task was to show that the crash was not only possible and not even just probable but that there were factors involved that rendered it necessary.

A host of different circumstances were considered and their status as the cause for the crash kept fluctuating between being impossible, possible and likely. As one newspaper article reported, 'The board has emphasized that everything is under consideration, no matter how seemingly irrelevant or obscure or unimaginable' (*The Milwaukee Journal Sentinel*, 20 February 2003: 12A). That the shuttle was hit by space debris, a meteorite, or lightning, that its old age was the problem, that there had been undue budget pressures on the project and that poor management decisions had been made were all considered unlikely but possible. Even sabotage, which was first ruled out as impossible by the investigative board, was later brought back into the realm of possibilities. The idea that a bit of foam insulation peeled off the fuel tank and struck the craft's left wing at take-off allowing hot gases to penetrate the wing upon return was first tentatively suggested; then ruled impossible; then revived as a possibility, until it eventually emerged as the most probable cause for the crash.

The impact of Columbia's crash on ordinary people and scientists alike and the thinking that underlay the subsequent investigation aiming to find the cause illustrate an important universal characteristic of the human mind. Throughout our lives, we notice things around us and wonder why they are the way they are. Although our desire to know may sometimes be fuelled by some further aspiration – wanting to improve things in the world, such as in the case of the Columbia investigation, or seeking money or prestige for ourselves – often there is no such ulterior motif. When people send inquiries to the science columns of newspapers about why leaves turn colour in autumn, this is out of sheer curiosity. Or, if you notice one October morning that the cup of water left on the backyard table froze overnight but the cup of wine standing next to it did not and you ask yourself about the reason for this difference, this may again be without any further goal: you simply want to understand.

The kind of curiosity that permeates our everyday life about why things are the way they are is the ultimate basis of scientific endeavour. The difference between what an everyday person does and what the scientist does is only that the latter takes a more persistent, more rigorous and more systematic approach to inquiry than the rest of us.

What exactly do we mean when we are asking for an explanation of something? The need for an explanation arises in a person's mind when there is a mismatch between the way he observes reality and the way he expects it to be. In other words, a why-question emerges from a gap between the actual world as we see it and the contents of our mind. The Columbia disaster is a case in point: as noted above, many people had believed it could not happen, some had imagined it was possible, a few may have thought it was likely to happen, but nobody had thought it necessary. Yet, it did happen, proving everybody wrong and giving rise to the question 'why?'.

An analysis of this example shows that the mismatch between what is experienced and what is expected may vary in depth. First, at times, the way things are seem likely but not necessary. Second, the gap is wider if the actual state of affairs seems possible but not necessary and not even likely. And, third, something may seem impossible – that is, totally beyond the realm of conceivable alternatives – even though it is real.

Let us further illustrate these three scenarios. In winter, retail prices of fruits and vegetables can be expected to go up. It is likely that this would happen but, given the complexity of the factors that influence pricing, it is not necessary.

For the second scenario, take the example of the water and the wine left outdoors overnight. That the water should have frozen is not miraculous provided you know that H_2O freezes at 32 degrees Fahrenheit and that overnight temperatures dipped below 32 degrees. But what about the wine? Why should water freeze sooner than wine? Wine is like water in many ways: it can be boiled, it seems to have similar viscosity, it quenches thirst as water does, and so forth. But wine is also different from water in some ways, such as in colour and taste. Therefore, when it comes to freezing, we have some reason to expect wine to behave like water but also to expect it to behave unlike water. Reality, however, comes down hard on one side: water will freeze before wine will. In this case, both the freezing and the non-freezing of the wine at 32 degrees F seem possible and the problem is why, of two equally likely alternatives, one is true and the other is not.

The example of leaves turning colour exemplifies the third type of gap between thought and reality: for the botanically uninitiated, the way things actually happen may seem not even possible. Since from March to September, oak trees show a bright green colour, one would expect them to remain verdant for the rest of the year just as evergreens do. The colour change of oak leaves thus might appear entirely beyond expectations: reasoning and reality go separate ways, this time even more radically than in the previous examples, and we ask 'why?'.

'Why X?' always means 'why X rather than Y?'. Asking a why-question requires detachment from the here and now and the ability to see alternatives to it, whether from past experience or from imagination. As long as we are prisoners of a particular view of the world, the possibility that things could be otherwise is dimly or not at all perceived: the very notion of an alternative is beyond our grasp. It is only when we are able to distance ourselves from a given situation and approach it from the angle of other conceivable alternatives that we come to realize that it is not necessary that it be the way it is. A person whose musical experience does not extend beyond bluegrass music is not likely to wonder why bluegrass is the way it is; in fact, he may not even be aware of the special conventions of this genre. But somebody coming to bluegrass from a different musical style will be struck by these constraints and is bound to ask why they are the way they are.

The fact that a why-question presupposes the inquirer's ability to see alternatives to the way things are throws light on why it is such an interesting experience to meet new people, to visit foreign countries and to learn new languages. In the course of these experiences, alternatives to known patterns come into vision and prompt us to take an outsider's look at our own ways and ask why they are the way they are. Science fiction offers a similar experience in that it takes us into a different world and thus inspires questions about our own. Reading about human-like creatures with two heads makes us realize that we, too, could have two heads and makes us wonder why we just have one.

By way of a summary, (1) provides a schematic representation of the mental scenarios that prompt why-questions. X stands for a state of affairs, such as leaves turning colour. In (a), (b) and (c) are shown the three kinds of gaps in order of increasing size that may separate reality and expectation: in (a), it is thought that something is probable but not necessary; in (b), it is believed that something is possible but not necessary and not even probable; and in (c), what is observed is thought to be entirely beyond the realm of possibilities.

(1) • OBSERVATION: 'X exists'
     • THREE ALTERNATIVE REACTIONS TO THE
       OBSERVATION:

	'X is possible'	'X is likely'	'X is necessary'
(a)	YES	YES	NO
(b)	YES	NO	NO
(c)	NO	NO	NO

     • THE RESULTING QUESTION: 'Why does X exist?'

2.1.2 Answers to why-questions

Since a why-question sprouts from a gap between what is observed and what is expected, what an explanation must do is close or at least narrow this gap.

Minimally, all that an explanatory attempt achieves is to bring the observed fact into the realm of possibilities, without showing that it is necessary or even likely. In other cases, the explanation goes one step further towards a tighter fit between fact and expectation: it renders the observed fact likely but not necessary. And optimally, an explanation renders the observed fact necessary so that all conceivable alternatives are ruled out.

Here are examples of the three degrees of explanation. Suppose that you have just moved from California to Chicago and on a May morning you wake up to snow covering the ground. Snow in May might seem impossible to you; but if the locals tell you that this has happened before, your expectations change: the statement 'Snow may fall in Chicago in May' renders the observed fact possible, although not likely and certainly not necessary. We will call this kind of explanation **permissive**: it permits something to happen but does not require or even probabilize it.

To illustrate the second kind of explanation, let us assume that the day you see snow on the ground in Chicago is in January rather than May. In this part of the world, snowfalls are frequent in the winter months and this generalization renders the observation likely. We will call this a **probabilistic** explanation.

The third and most satisfying type of explanation renders an observation not only possible and not only probable but necessary. This is so if the scene is the Arctic region rather than Chicago. For that part of the world, there is an exceptionless generalization which says the ground is always packed with snow, and thus the observation about snow on one particular day is an inevitable consequence of this general truth. We will call this kind of explanation law-like, or **nomological** (from the Greek word *nomos* 'law').

Additional examples of nomological explanations come from two of our earlier scenarios: water freezing before wine and leaves turning colour. A nomological explanation for why water freezes before wine would appeal to the chemical composition of wine: while wine contains some water, it also contains alcohol, whose freezing-point is known to be lower than that of water. As a result of this recognition, the fact that appeared mysterious before – that wine does not freeze when its twin substance, water, does – now seems natural: things have 'fallen into place'. Before the explanation, the observation seemed possible but not necessary. After an explanation has been offered and accepted, the real state of affairs remains the only possible option and thus a necessary one, with all counterfactual alternatives ruled out.

Similarly, leaves of oak trees turning colour in autumn can be explained nomologically. The colour change comes about in response to the leaves gradually separating from the branch – a process triggered by the lessening of daylight hours and the coldness of nights. Once the leaves are cut off from the nourishment coming from the trunk, photosynthesis cannot take place any more to produce chlorophyll – the substance responsible for the green of the leaves. Other colours such as yellow and red thus far masked by the green then show up. Given that we understand that the conditions needed for the production of the chlorophyll gradually disappear, the colour change is inevitable: the leaves could not but lose their green. As Hercule Poirot, musing at his lifelong efforts to explain 'The Murder', remarks: 'when the right solution is reached, everything falls into place. You perceive that **in no other way** could things have happened' (Agatha Christie 1986; emphasis original).

In (2) there is a summary of the three types of answers to why-questions as they close the gap between fact and expectation to varying degrees, proceeding from minimal to maximal explanatory force: (a) schematizes permissive explanations, (b) schematizes probabilistic explanations and (c) represents nomological explanations.

(2)	•	WHY-QUESTION: 'Why does X exist?'		
	•	THREE ALTERNATIVE ANSWERS TO THE QUESTION:		
		'X exists because it follows from Principle P that . . .		
		. . . X is possible.'	. . . X is likely.'	. . . X is necessary.'
(a)		**YES**	NO	NO
(b)		**YES**	**YES**	NO
(c)		**YES**	**YES**	**YES**

Exactly how does an explanation effect this crucial transformation in the mind? The key notion is **generalization**. All explanations must include three crucial components: a fact to be explained – called an **explanandum**; at least one generalization to which appeal is made – called the **explanans**, or explanatory principle – and at least one additional statement called the **bridge statement** that shows how the explanans ties in with the explanandum. An explanation renders an observed fact possible, likely or necessary by revealing that the object under study belongs to a class some, most, or all members of which exhibit the observed characteristic(s).

Here is the explanatory schema for the water and wine example:

(3) EXPLANANDUM:
Wine freezes at a lower temperature than water.

EXPLANATORY PRINCIPLE:
Alcohol freezes at a lower temperature than water.

BRIDGE STATEMENT:
Wine is a kind of alcohol.

However, explaining things is a much more complex endeavour than it might seem. First, an explanation may appeal to more than one general principle. Thus, in the chlorophyll example, it is not enough to invoke the principle that photosynthesis requires water: other relevant principles having to do with how trees sense the shortening of days and so on are also at play. Second, things may have several competing explanations. Thus, an

alternative attempt at explaining the colourful autumn foliage may be made by positing little fairies that paint the leaves yellow and brown in the depth of the night. Although in this particular instance it would not be difficult to show that the latter theory lacks independent support and thus is inferior to the former, the task of choosing between alternative explanations may in other cases be much more difficult.

Third, explanations are never ultimate: the explanans itself is also in need of an explanation. To say that wine does not freeze at 32 degrees Fahrenheit because it is a kind of alcohol is a low-level explanation in that it raises the question of why alcohol does not freeze at 32 degrees even though water does? A higher-level explanation would have to do with the molecular composition of alcohol and water and the behaviour of the ingredient atoms and their component particles under varying temperatures. But even these principles would call for further explanations. The existence of complex explanations and of alternative explanations, and the need to search for explanations of explanations render science a forever challenging and, in the final analysis, forever elusive pastime of human beings.

In sum: Section 2.1 discussed why-questions and what it takes to answer them. Explanations are called for when there are mismatches between observation and expectation. If an observed fact is deemed possible but not probable, or is seen as probable but not necessary, the explanation needs to narrow the range of expectations to the one state of affairs actually observed. If the observed fact is not even deemed possible, then the explanation first needs to broaden the spectrum of conceivable alternatives before narrowing it to the one that holds true. In either case, an explanation succeeds if it eliminates the conflict between perceived reality and the human mind: by changing the mind's content it re-aligns the two.

In order for something to be explained, it first has to be observed and described. The next two sections will take a look at the notions of observation and description.

2.2   Observation

A why-question asks about something that has been observed; that is, it presupposes that somebody has looked at the world and has singled out a certain portion of it for attention. Thus, for instance, asking a question about leaves turning colour in the autumn presupposes that we see leaves as separate from the trunk and from the insects that crawl on them, and their colours as separate from their shapes; or asking why Jim is absent in class implies that Jim has been perceived as an entity separate from his classmates and separate also from the chair that he usually sits on in class. Accordingly, the first step in scientific inquiry is observation: delimiting an object as a separate part of the world.

Facts and observations about facts are not the same thing: an observation reflects how facts are perceived. Observations are our images formed of reality rather than reality itself. No matter how much we try to make perceptions be true to reality, they always carry the stamp of the observer. As the Danish physicist Niels Bohr is reputed to have said: 'It is not enough for us to explain what things seem like; we need to explain what things really seem like.' Bohr did not contrast how things 'seem' with how things 'are' but with how things 'really seem'. The tongue-in-cheek expression 'really seem' is self-contradictory and it reflects Bohr's doubts as to whether the true nature of reality can ever be revealed to the observer.

The observer's effect on perception is shown by the fact that things are perceived differently depending on the observer and his goals and means. For example, newborn babies may not see their caretakers as separate from themselves and may not perceive their own hands and feet as parts of their own body. Dogs and bees are colour-blind and thus miss out on some properties of things that humans readily perceive – just as we miss out on nuances of smell that dogs are adept at picking up.

The human capacity to observe things is in fact greatly constrained by our limited abilities: we cannot observe things that are too big, or too small, or too distant, or that are inside objects too dense to be penetrated by our senses. These limits can be stretched somewhat by telescopes, microscopes, X-ray machines, radar, sonar, MRI and other instruments but the fact remains that it is only a part of reality that is open to human observation. Since scientific inquiry pertains only to things that can be observed, the entire endeavour of science is limited to the humanly observable portion of the world.

2.3   Description

Before an observation can become an explanandum, it must be put in verbal form. Verbal renderings of observations are called **descriptions**.

As shown by the three observers' different ways of recording their observation regarding the brown cow in the epigraph at the start of this chapter, descriptions may be couched at different levels of generality. The mathematician's statement – that there is at least one cow in Scotland, of which one side appears to be brown – is the most cautious and most factual one, the economist's – 'The cows in Scotland are brown' – is the most general, and the logician's take on the facts falls between the two.

In its simplest form, the description of a set of objects may be enumerative, simply presenting each object in its entirety. Good descriptions, however, are analytic rather than enumerative: they describe objects by dissecting them into component parts and properties. Analytic descriptions reduce an apparently unique object to a unique combination of non-unique components and thus relate it to various other things.

Suppose, for example, that I am asked to describe lemons. I may offer the following analytic description: 'Lemons are a tropical fruit of the size of a small apple, yellow by colour, their consistency similar to that of an orange and their flavour like grapefruit laced with vinegar.' This description characterizes lemons by six properties: general type, provenance, size, colour,

consistency and flavour. By situating lemons along each of these descriptive parameters, I assign them to classes of things. When I say lemons are a fruit, I am relating them to pears, bananas and strawberries. By saying that they are tropical, I am relating them to crocodiles and rain forests. By saying that the size of lemons is that of a small apple, I am lumping them together with apples, potatoes, eggs and small balls of yarn. By describing them as yellow, I am linking them with bananas and sunflowers. By describing their consistency, I am relating lemons to oranges, grapefruits and limes; and by describing their flavour, I am assigning them to yet another class which consists of acidic substances. This characterization captures both the uniqueness of lemons and also their relatedness to other things in the world. On the one hand, there are many other things that have the same colour as lemons *or* the same size *or* the same flavour *or* the same consistency *or* the same provenance *or* the same fruitiness. But there is nothing else that has the same size *and* the same colour *and* the same flavour *and* the same consistency *and* the same provenance *and* the same fruitiness. Lemons, in other words, are non-unique in their individual properties but are unique in the combination of their properties.

That scientific descriptions are analytic is important since analysis is an indispensable tool of explanation. Here is an example of how the two are related. Consider the case of a patient who reports to his doctor that he has allergic reactions to the following kinds of food: chicken salad, potato salad, eggnog, Hollandaise sauce, custard pie, chicken soup and ham-and-noodles casserole. The doctor's job is to explain why the patient is allergic to just these foods but not, say, to beef bouillon or turnip salad. This requires working out what it is that the seemingly diverse foods have in common. Since they are all distinct as wholes, their similarity must reside in their parts. Analysis of the respective recipes will reveal that two of them – chicken salad and potato salad – contain mayonnaise. Two of the other foods also have an ingredient in common: both chicken soup and ham-and-noodles casserole contain noodles. But breaking down four of the seven kinds of food into their immediate ingredients does not yet solve the problem since mayonnaise is still distinct from noodles and both are distinct from eggnog, Hollandaise sauce and custard pie. Additional analysis reveals, however, that noodles, mayonnaise, eggnog, Hollandaise sauce and custard pie all share a single ingredient: eggs, which must therefore be the cause of the patient's allergy.

This example shows how analytic descriptions are instrumental in explaining things. In addition, note that the doctor's job is not only to explain why the patient gets sick from certain foods, but he also needs to be able to predict what other food items are likely to evoke the same reaction so that he can warn the patient against them. Prediction is a by-product of explanation and, as such, it becomes possible only by analysis. In this instance, the prediction is that any food that contains eggs will cause problems for the patient.

In addition to explanation and prediction, there is a third benefit to analytic descriptions: they are the only descriptive option if the set of things to

be described is infinite in size. There are many such things: the infinite variety of human fingerprints, human faces, human diseases, snowflakes, birch trees, cloud configurations and hair patterns of giraffes. Members of such infinite sets could never be 'enumeratively described' – that is, presented in their entirety – but they can be described in terms of their finite number of properties.

Constructing useful descriptions is difficult. The example of the doctor's task to find the crucial food ingredient that causes the allergy illustrates this. In order for the description to become a tool of explanation and prediction, the analyst must hit upon the right kind of parts and properties of things and find the right level of analysis. First, note that foods could be described in many ways other than in terms of their ingredients – such as their colour, taste and consistency. But none of these analyses would be helpful in finding the common denominator of the offending food items in our example and thus explaining and predicting the patient's allergic reactions. Second, even if the analysis undertaken is on the right track by breaking the food items down into their ingredients, the crucial ingredient responsible for the allergy would be missed if the analysis did not go far enough but ended with identifying mayonnaise as being an ingredient in some of those foods. Thus, finding the right description that will make explanation and prediction possible is not a mechanical task: it takes insight to hit upon the right parts and properties and to zero in on the right level of analytic depth.

Hidden behind the idea of describing an object as the sum of its parts, properties and their relationships lies the assumption that a whole equals the sum of its parts and their relations. This assumption is known as the principle of **compositionality**. Although in many instances, compositionality turns out to be a fruitful assumption, it does not always hold. Is the full nature of a lemon really accounted for by the sum of its properties? Can the behaviour of a crowd of people be reduced to the behavioural features of the individuals that make it up? And does the sum of the characteristics of your best friend do full justice to what he or she is really like? Similarly, how could the idiomatic meanings of expressions like *kick the bucket* or *spill the beans* be accounted for in terms of the meanings of the words they are made up of? Compositionality is an empirical hypothesis about things in the world and it may or may not always work.

Let us summarize Section 2. Scientific inquiry is rooted in two realities. One is outside the human mind – the world; the other is inside it – the mind that perceives reality and wants to explain it. We make observations, describe them and then ask why they are the way they are. Observations are limited by the observer's goals and means. Descriptions must be analytic for three reasons. First, infinitely large sets of phenomena cannot be described individually but analytically. Second and third, descriptions are tools of explanations and predictions.

Asking a why-question is a creative activity in and of itself since it means that the person asking it is capable of visualizing alternatives. Since a why-question arises from a gap between what the observer perceives and what he

expects, an explanation must close the gap by showing, minimally, that the observed fact is possible; or, better, that it is probable, and, optimally, that it is necessary.

3 Studying language

3.1 OVERVIEW

In Section 1, two basic points were made about syntax: syntactic statements specify the choice and ordering of words; and they have to be formulated in a maximally general way. Section 2 served to place the generality requirement in a broader context in that it showed how this requirement was dictated by the goals of scientific inquiry in general. Next, we will return to the first point: the content of syntax, and view it within a larger context of linguistics – the science of natural human languages.

As all other artifacts – objects constructed for a purpose, such as chairs, houses or computers – sentences, too, have three basic types of properties: structure, function and location. Structure refers to composition out of smaller parts; for sentences, these are sounds, morphemes, words and so on. An observation about language structure is that the verbal affix *-ing*, as in *singing*, must follow the verb stem that it goes with rather than precede it. Function means utility: what sentences are used for. An observation about sentence function is that certain sentences, such as *Can you tell me what time it is?*, may be used both as questions and as requests. The third concept, location, refers to the spatial and temporal context in which sentences occur. For example, you can note that double negatives, such as *I did not see nothing*, occur in twenty-first-century non-standard dialects of English as well as in standard versions of Russian and Serbo-Croatian.

The study of these three basic domains of linguistic observation is divided among the various subfields of linguistics. Structure is the concern of descriptive linguistics. Language function is studied in semantics, pragmatics, stylistics, psycholinguistics and sociolinguistics. The textual, spatial, temporal, natural, social and physiological context in which sentences occur is the subject of discourse analysis, dialectology, areal linguistics, historical linguistics, developmental linguistics, sociolinguistics, psycholinguistics and neurolinguistics.

In the next three sections, we will take up the three concepts of linguistic structure, function and location in more detail.

3.2 STRUCTURE

An object has structure if it can be analysed into component parts. As was noted in Section 1, any complex object has structure. By recognizing that sentences have structure – they consist of words – we recognize their fundamental similarity to all other complex phenomena in the world: minerals, galaxies, plants, human bodies, societies and religions.

Sentences are not the largest units of speech, they are themselves components of larger chunks called discourses. A discourse is a coherent sequence of sentences produced either by a single person, such as a diary, lecture, poem or novel, or by a set of interlocutors, such as a conversation or exchange of letters. Although descriptive linguistics may ultimately adopt the discourse as its most fruitful domain, in this book we will follow the dominant tradition and restrict our analysis to sentence-size chunks while delegating the study of discourse context to the study of the location of sentences (see Section 3.4 below).

The description of a language cannot simply be a dictionary-like listing of all the sentences that it consists of; instead, what is needed is an analytic description that describes parts and properties of sentences. This is not only because, as we saw in Section 2.3, analytic descriptions facilitate explanations and predictions but also because the number of sentences that any language consists of is infinite and, as also noted above, infinite sets cannot be enumerated.

There are several sources of infinity for the set of sentences making up a language, many of which have to do with the infinite set of ideas that can be expressed. First, note that it is possible to talk about an infinite number of things. Thus, one can say: *Sammy owns two houses, Sammy owns three houses, Sammy owns twenty million houses*, and so on. If we run out of numerals, we can always name the largest number available and add *plus one, plus two* and so on. Second, we can talk about an infinite number of events or states, as in *I know that Sammy is a jerk, I know that you know that Sammy is a jerk, I know that you know that I know that Sammy is a jerk*, and so on. Third, even a single event can have an infinite number of instances that differ in length: *The cougar ran, The cougar ran and ran, The cougar ran and ran and ran*, and so on. And, fourth, degrees of intensity are also infinite in number, as in *The hunter was tired, The hunter was very tired, The hunter was very, very tired*, and so on. It is the unlimited repeatability of certain meaningful elements in a sentence that enables language to express an infinite number of thoughts by using a finite stock of items.

How exactly do analytic descriptions of linguistic structure work? Consider first an enumerative description – one where sentences are not analysed into parts. In such an account, each statement pertains to one sentence only, as is the case in (1).

(1)   OBSERVED SENTENCES   DESCRIPTIVE STATEMENTS

   (a)   *There is a squirrel* ——————— (A)   *There is a squirrel behind*
          behind the tree.   *the tree.* is a well-formed
                                                          sentence in English.

   (b)   *There is a frog* ——————— (B)   *There is a frog behind the*
          behind the tree.   *tree* is a well-formed
                                                          sentence in English.

(c) *Is there a squirrel* ——————— (C) *Is there a squirrel behind*
   behind the tree?   *the tree?* is a well-formed
                                       sentence in English.

(d) *Is there a frog* ——————— (D) *Is there a frog behind*
   behind the tree?   *the tree?* is a well-formed
                                       sentence in English.

As the lines indicate, there is one 'rule' for every sentence in this account and each rule applies to one sentence only. In an analytic account, on the other hand, there is a many-to-many relationship between descriptive statements and objects to be described: each descriptive statement pertains to only one portion of a sentence, but each applies across sentences.

(2)   OBSERVED SENTENCES   DESCRIPTIVE STATEMENTS
     (a) *There is a squirrel*   (A) The unstressed *there*
         behind the tree.   immediately precedes the verb
                                       in declarative sentences.

     (b) *There is a frog*   (B) The unstressed *there* must
         behind the tree.   immediately follow the verb
                                       in questions.

     (c) *Is there a squirrel*   (C) The verb must agree with its
         behind the tree?   subject in number.

     (d) *Is there a frog*   (D) Nouns must occur with
         behind the tree?   an article.

                                   (E) Articles must precede their
                                       nouns.

                                   (F) The meaning 'squirrel' may be
                                       expressed by the sound form
                                       squirrel.

                                   (G) Vowels must be nasal
                                       preceding nasal consonants.

Additional rules stating that certain words or phrases are repeatable any number of times will allow for an infinite number of sentences.

The rules in (2) make reference to various kinds of structural characteristics: the presence of certain words (as in (D)), the presence of an ending in a word (as in (C)), the sound form that goes with a meaning (as in (F)), and details of pronunciation (as in (G)). The branches of descriptive linguistics analysing sound form are called **phonetics** and **phonology**. (While phonetics and phonology have distinct domains – phonetics analyses individual speech sounds and phonology analyses their sequences – the term phonology may be used broadly so as to include both.) The analysis of meanings is called **semantics**. Meaning-form linkages are formulated as rules of the **lexicon**, **morphology** and **syntax**. How these components fit together will be discussed in Section 4.

3.3 FUNCTION

All complex objects have structure but not all of them have function. Artifacts, such as spoons, books and trains, are functional; natural objects, such as rocks, mountains or thunderstorms, are not – unless somebody presses them into service, such as using a rock to crack open a nut.

Language is a human artifact and it does have functions. The most obvious purpose that language serves is communication – that is, for rendering thoughts public and addressing them to somebody.

Communication includes three components: an idea that the speaker has, a form in which he expresses the idea and an addressee that the idea is communicated to. While all three of these factors may be prominent in an act of communication, there are uses of language where only one or two of them stand out, with the others receding or not present at all. Thus, in puns and in poetry, form may acquire supreme significance that may outweigh content. In small talk, both literal content and form are underplayed – the very fact of exchanging a few words with somebody on an inconsequential topic such as the weather can communicate goodwill and camaraderie. Also, simply expressing thoughts may be de-coupled from wanting to convey them to somebody else, such as when writing a diary. And in thinking, neither addressing another person nor even trying to express one's thoughts is present.

While communication is the most obvious function of language, human language is neither necessary nor sufficient for human communication. Postures, gestures, facial expressions and various kinds of bodily contact may not only complement – or cancel – a verbal message but may also entirely replace it. Besides, language may not even be the best means for communicating some messages. The process of fitting ideas into linguistic form is often frustrating: a stream of thoughts is somewhat like a waterfall which, when reaching the foot of the mountain, loses its power by being forced to split into rivulets that follow already existing ruts in the ground. Linguistic expression forces the speaker to utilize publicly agreed-upon means for bringing forth what is on his mind and thus possibly to compromise some of the intended content.

Identifying the functions of artifacts is important if we want to explain structure: the structural characteristics of artifacts must support or at least be compatible with their function. For example, it would be hard to understand why bicycles were built the way they are if we did not know what they were used for. Does language function similarly leave its marks on language structure? This question will be taken up in Section 4 of Chapter 7.

3.4 LOCATION

Everything – both simple and complex objects and objects with or without function – occurs somewhere. In other words, everything has a location. The location of an object can be specified in terms of space, time, and

natural and social context. For example, boats are normally found in water; and two legs as a means of locomotion are located in humans and birds. The location of a set of objects relative to their surrounding conditions is called their **distribution**. For example, one can describe the distribution of oak trees in Wisconsin by referring to the areas where they grow and the distribution of different hair colours among people by specifying the body type and age factors that condition them.

The location, or distribution, of sentences is definable in reference to linguistic and extralinguistic context. What is the **linguistic context** in which sentences are located? First, sentences are generally not used in isolation: they occur as parts of sentence sequences forming coherent discourses, such as conversations, lectures, poems or novels. Second, sentences are part of a language that includes many other sentences that are all similar: they share grammatical characteristics with each other. We will call the first type of context **syntagmatic** and the second, **paradigmatic**. Here are non-linguistic examples for the clarification of these terms. Paprika and whipped cream share the paradigmatic context of Western cuisine since they are both available in that style of cooking even though they may never be used together in any one dish; while whipped cream and powdered sugar share not only paradigmatic but also syntagmatic context since whipped cream is generally sweetened with powdered sugar.

Apart from the linguistic context, the other factor defining the location of sentences is the **extralinguistic context**: the particular speech situation in which a particular sentence is uttered and, more generally, the natural and social environment in which the language is used.

Just like function, the location of an object or phenomenon can also serve to explain aspects of structure: the structure of an object or phenomenon may in part be determined by the context in which it occurs. For example, the kinds of trees that grow in some areas are limited and thus predictable by the climatic and soil conditions available. Similarly, languages reflect the general natural and social conditions under which they are used; and the structure of sentences is influenced by the particular speech situation where they are spoken.

Within a given speech situation, the identity of the speaker – whether male or female, whether a child, a young adult or an old person, whether college-educated or not, what region of a given country a person is from, and their current physical and psychological state – will influence how a person speaks. In addition, an individual's particular style of speech will also depend on the identity of the addressee and other attendant conditions.

The aspect of language structure that most clearly reflects surrounding natural and social conditions is vocabulary. For example, it is well known that Eskimo languages have an extensive set of terms for various kinds of snow. Another example: the Shona people of Zimbabwe and Mozambique, forming an agricultural society, have at least 20 different words for walking, such as for 'walk with a squelching noise through a muddy place', 'walk while making a noise of breaking sticks', 'walk naked or almost naked' and so on.

4 Why do languages have syntax?

Let us now return to the first-mentioned property of sentences: structure. In Section 3.2, five types of rules were mentioned as instrumental in describing sentence structure: phonetic-phonological, semantic, lexical, morphological and syntactic. What exactly is the contribution of each of these rule types? Are all of them necessary?

The description of any symbolic object – such as body gestures, traffic signs or religious rituals – must, by definition, include information about meaning, form and the correspondence between the two. Sentences are symbolic objects: every sentence has a form and a meaning. It thus follows that any description of sentence structure must include three kinds of information: about meanings, about sound forms and about what meanings go with what sound forms. As noted above, these components of grammar are called phonology, semantics and the lexicon, respectively. These three components are therefore indispensable, logically necessary parts of grammar. Let us consider each in turn.

Both phonological and semantic descriptions include in some form an inventory of basic components, rules of selection and rules of arrangement. For example, the **phonology** of English will say that English has both /s/ and /p/, that they can be selected to introduce a syllable, and if they are so selected, the /s/ must precede the /p/ (cf. words like /spɪl/ (spelled *spill*) but none like */psɪl/). Similarly, the **semantic** description of English will characterize meanings in terms of the selection and arrangement of components. The principal components include 'predicate', 'argument' and 'tense'. What is meant by 'predicate' is the action, event or state that the sentence is about; 'arguments' are the participants in the action, event or state; 'tense' refers to time. For example in *Joe wrote a letter*, 'wrote' is the predicate, 'Joe' and 'a letter' are the arguments, and the tense is past. Relying on these concepts, semantic descriptions will say that 'predicate', 'argument' and 'tense' may be selected together to form a sentence meaning, and that tense must go with predicates and not with arguments.

The third logically necessary component of grammar is the **lexicon**. Lexical rules show symbolic correspondence relations between small units of meanings and sounds: they assign the right sound forms to the right meanings.

How small should the chunks of form and chunks of meaning be that lexical rules relate to each other? It is easy to imagine symbolic systems where lexical rules apply to the entire meanings and entire forms of symbols. For example, in the system of traffic lights, the colours – red, green and amber – cannot be broken down into components that would separately symbolize parts of the meanings 'wait', 'go' and 'prepare to stop'. In other symbol systems, however, bits of form are separately associated with bits of meaning. A ritual meal is an example: while the whole meal has a meaning, each food item has its own symbolic significance.

Sentences are like ritual meals in that they consist of recurrent parts each of which carries meaning by itself. Thus, lexical rules are best stated not on

entire sentences but on the smallest recurrent meaningful parts of sentences called morphemes. For example, the meaning-form relation underlying the sentence *The squirrels nibbled at the acorns* does not have to be stated by a single rule for the entire sentence: a separate rule can be formulated for each morpheme, such as DEFINITE = *the*, SQUIRREL = *squirrel*, PLURAL = *s*, and so on (all capitals stand for meaning representations, orthography stands for sound form). The lexicon is thus a set of symbolic correspondence rules stating the meanings and forms of morphemes.

At this point in our argument, we have three components of grammar: **phonology** describing the sound forms of sentences, **semantics** describing the meanings of sentences, and the **lexicon** relating sound form and meaning for the smallest bits of form that are meaningful – morphemes. Do these three types of rules suffice to describe what well-formed sentences are like?

The answer is 'no' for two reasons. First, well-formed sentences must include well-formed words; but simply having a list of morphemes for a language does not tell us how to put words together out of them. Whether a word is grammatical or not depends on whether the morphemes that make it up have been properly selected and properly ordered. Thus – unless in a language every word consists of a single morpheme only (such as English *pumpkin* or *turn*), or unless both the selection and the order of words are completely free – an additional rule type is needed. The rule type that specifies the selection and order of morphemes in words is called **morphology**.

Morphology tells us how words are formed out of morphemes. If we now consider how sentences in turn are formed out of words, we find a similar picture. First, just as many words consist of more than one morpheme, most sentences consist of more than one word. There are, to be sure, one-word sentences – such as English *Sit!* – but they are not typical of English or any other language. Second, as we already saw in Section 1, the selection and ordering of words is not free: *Johnny fast* is not well-formed because of the wrong selection of words and *Johnny fast runs* is not well-formed because of the wrong ordering of words. These facts point at the second reason why phonology, semantics and the lexicon are not enough to account for the well-formedness of sentences. This is where **syntax** comes in: as morphology does for morphemes, syntax specifies the proper selection and order of words.

The five components of grammar thus stack up as follows:

- logically necessary for all languages: semantics
  - phonology
  - lexicon
- not logically necessary for all languages: morphology
  - syntax
  - morphology:
    - impossible in languages that have only monomorphemic words

- possible in languages with at least some polymorphemic words
- necessary if the selection and order of morphemes in words is not free
- syntax:
  - impossible in languages that have only one-word sentences
  - possible in languages with multi-word sentences
  - necessary if the selection and order of words in sentences is not free

In addition to describing well-formed sentences, syntax also needs to account for what the sentences mean. The lexicon does provide the meanings of the morphemes contained in sentences; but, as we saw in connection with the Turkish sentence in Section 1, the meaning of the entire sentence is not always the simple sum of the meanings of the words. Similarly, the sum of the phonological forms of words does not always equal the phonological forms of the entire sentence: that the word *cafeteria* may have rising intonation in the sentence *Where is the cafeteria?* cannot be part of the lexical entry of that word since it does not always carry this intonation pattern. Thus, in addition to rules of syntactic structure that describe the selection and order of words, we also need syntactic correspondence rules that relate syntactic structures to meanings, on the one hand, and to phonological form, on the other.

Syntax, therefore, complements the other components of grammar – semantics, phonology, the lexicon and morphology – in two ways. On the one hand, it describes the selection and order of words for producing well-formed sentences; on the other hand, it describes the correspondence between sentence structure and sentence meaning, and sentence structure and sentential sound form.

Here are some examples of syntactic rules for English:

(1) (a) syntactic structure rules:
  - word selection: Every concrete common noun in the singular must occur with an article.
  - word order: The article must precede the noun.
  (b) syntactic correspondence rules:
  (i) between syntactic structure and meaning:
    - In sentences where the predicate has a negative meaning, if they contain no auxiliary, a form of the verb *do* has to be included.
    - In sentences whose meaning is a basic weather statement, there must be an *it*-subject.
  (ii) between syntactic structure and sound form:
    - Yes-no questions carry rising intonation.
    - The two syntactic constituents *is* and *not* can make a single phonological word *isn't*.

The two tasks of syntax are presented in diagram (2):

(2)   SOUND FORMS
      |
      | ← syntactic correspondence rules
      |
      WORD SEQUENCES   ← syntactic structure rules
      |
      | ← syntactic correspondence rules
      |
      MEANINGS

In sum, why do languages have syntactic rules? Because

- most sentences consist of more than one word
- the selection and order of the words in sentences are not free
- the sum of the word meanings does not always equal the meaning of the entire sentence and the sum of the word forms does not always equal the phonological form of the entire sentence

5 Conclusions

This chapter discussed syntax in the context of linguistics and of science in general. Here are the major points that have emerged.

ON SCIENCE

'Only when we are able to find life-as-we-know-it in the larger context of life-as-it-could-be will we be able to understand the nature of the beast' (Langton 1995: 2). The basic message of this statement applies to any branch of science including linguistics: what we want to know is how things are as opposed to how they could be. There are thus two cardinal factors that define the space within which any inquiry – whether in science or in everyday life – takes place: one is external reality and the other is the human mind that observes things, describes the observations, compares what it sees with other logical possibilities and, puzzled when there is no match, asks 'why?'. What an explanation does is to bring about a change in our expectations so that what seems impossible is rendered possible, what seems possible but unlikely is rendered likely and, optimally, what seems possible and perhaps likely but not necessary is revealed as the only option and thus a necessary one. Asking a why-question is asking for a reconciliation between reality and the mind. A true explanation succeeds by resolving the conflict through altering the content of the mind so that harmony between the two is restored.

ON DESCRIPTIVE LINGUISTICS

Linguistics is a science dedicated to the study of the structure, function and location of sentences. Descriptive linguists focus on language structure and

aim at providing maximally general analytic descriptions called grammars. A grammar has five basic components, three of which are logically necessary in descriptions of a language just as they are in the descriptions of any other symbols systems. These three are phonology (or, for symbolic systems that are not sound-based, such as sign language, the description of the particular medium they utilize), semantics and the lexicon. Phonetics and phonology describe the proper selection and order of phonetic form elements. Semantics sets limits to the proper selection and arrangement of meaning elements. The lexicon spells out the minimal pairings of meanings and forms. If words include more than one morpheme and the selection and order of morphemes is not free, rules of morphology are needed to specify how morphemes can be selected to make a word, how they should be ordered, and what the exact meanings and pronunciations of the resulting words are. And given that sentences include more than one word with selection and order constrained, an additional rule component, called syntax, is called for.

ON SYNTAX

Minimally, syntax describes the selection and order of words that make well-formed sentences and it does so in as general a manner as possible so as to bring out similarities among different sentences of the same language and of different languages and render them explainable. Since sentences are not just complex objects but symbolic objects that convey meanings, syntactic rules also need to account for the relationship between strings of word meanings and the entire sentence meaning, on the one hand, and the relationship between strings of word forms and the entire sentential phonetic form, on the other.

Structural syntax – describing word selection and ordering – will be detailed in Chapters 2–4; correspondence syntax – describing the relationship between syntax, meaning and phonetic form – will be taken up in Chapter 5.

Notes

Section 1   Preliminaries
The term 'syntax' is used in two ways in the literature: in reference to a particular aspect of grammatical structure and in reference to the subfield of descriptive linguistics that describes this aspect of grammar.

Section 2.1   Explanation
The concept of explanation known as the deductive-nomological, or covering-law, model associated primarily with Carl Hempel's name (see Hempel and Oppenheim 1948) has been adopted in this book because of its intuitive appeal and for lack of a clear alternative. Nonetheless, it should be noted that this is only one of the several concepts of explanation that have been

entertained by philosophers of science and it has been shown to be unsatisfactory in some ways. A brief and clear summary of the issues surrounding the goals and means of scientific explanations is Newton-Smith (2000). More will be said about linguistic explanations in Chapter 7.

Section 2.3   Description

On the infinite number of hair patterns of giraffes, see *National Geographic Magazine*, September 1977: 403–4.

Section 3.1   Overview

For concise summaries of the various branches of linguistics, consult the web site of the Linguistic Society of America (www.lsadc.org).

Section 3.2   Structure

For arguments for discourse, rather than sentence, being the natural domain of linguistics, see Sanders 1970a. On discourse analysis, see for example Beaugrande and Dressler 1981 and Givón 1997.

Section 3.4   Location

For the Shona walking verbs, see Comrie *et al.* 1996: 89.

Section 4   Why do languages have syntax?

The definition of the concept 'word' is actually more complex than given in this chapter; see Dixon and Aikhenvald 2002.

Exercises

1. Choose a foreign language unfamiliar to you. Translate a couple of English sentences into the language by looking up the words in a bilingual dictionary. Show your sentences to somebody who knows the language and ask for the actual translation. Do your translations and the real ones differ? If so, do they differ in the selection of words or in their ordering or both?

2. Consider the following ungrammatical sentences and determine whether they are ungrammatical because of word order, because of word selection, or both.

   (a) *The rabbits have ran back to the hole.*
   (b) *The rabbits run have back to the hole.*
   (c) *They not have enough to eat.*
   (d) *Squirrels hide nuts and underground.*
   (e) *Nuts and acorns are hide underground.*
   (f) *Are the nuts where?*
   (g) *Acorns grow on.*
   (h) *Who do you think who ate the nuts?*
   (i) *What you think squirrels eat nuts and?*

3. Consider the following common reactions of people to an event:

   (a) 'If you had told me before it would happen, I would not have believed it!'
   (b) 'Well, it has happened before!'
   (c) 'I could not believe my eyes!'
   (d) 'Don't tell me this!'
   (d) 'No-o-o-o-o!!!'
   (e) 'I just KNEW this would happen.'
   (f) 'Didn't I tell you???'

What are the initial expectations that can be reconstructed from these reactions? Did the person think that it was impossible for the event to happen, or possible, or probable or necessary?

4. Find an explanation in everyday life or in a science book and spell out its structure by identifying the explanandum, the explanatory principle(s) and the bridge statement(s).

5. Under certain conditions, drivers honk their horn. Describe the distribution of horn-honking in your town. Factors that may determine the distribution may be specific parts of the town, certain stretches of streets, the kind of vehicle, the place where the driver is heading and the driver's state of mind.

6. At the world-famous Niagara Falls, a small sightseeing boat takes visitors deep inside the enormous horseshoe-shaped waterfall. In the midst of the boiling waters, with the towering backside of the falls already in sight through the mist and with people on board scrambling to put on their rain gear to protect themselves against the deluge of water droplets blown at them by the winds, the boat comes to a precarious bobbing halt and the captain's solemn voice is heard announcing: 'Ladies and gentlemen, this is Niagara Falls.' Needless to say, the announcement has no information value – yet, it has an emotional impact on people. What do you think is the function of this announcement and why does it have the effect that it has?