
Asking questions about syntax

The title of this chapter is deliberately rather ambiguous. I am hopeful that by this point, you will be able to construct a basic syntactic description of a language, either a language that you speak well yourself, or one for which you can find a native speaker consultant. Section 9.1 outlines the kind of questions that you will need to investigate. Section 9.2 provides a short case study of Welsh, illustrating how these questions could be answered. These sections, then, refer to asking questions about syntax in the most literal way. However, I also hope that the discussion in the previous chapters has ignited some curiosity about the human language faculty more generally. In Section 9.3, I briefly outline some issues and questions surrounding our syntactic abilities that are currently widely debated within linguistics. Section 9.4 looks at possible further directions to pursue in your study of syntax.

9.1 SYNTACTIC DESCRIPTION: WHAT QUESTIONS TO INVESTIGATE

This section aims to give you a framework with which to write a basic syntactic description of a language that you know well, or for which you can access data readily. If appropriate, you can ask one or more native speakers to act as language consultant(s). Make sure you give the source(s) of your data, including attributions to the literature (i.e. cite your sources). Acknowledge any help given by language consultants.

- Give the name by which the language is known to its native speakers, plus its English name, if any. State its language family and the principal locations in which it is spoken.
- Your description should include *some or all* of the questions outlined in (1) to (13), depending on what features of the language you consider to be most interesting from a syntactic, morphosyntactic and typological point of view. Give enough information on (and illustration of) any feature to make it comprehensible to someone who has no prior knowledge of the language.
- *All* parts of the discussion must be illustrated with appropriate and sufficient data, glossed and translated. Number each example, following the conventions used in this textbook. If your language uses a writing system other than the Roman alphabet, cite data using whatever standard system of transliteration is used for this language.

- Give a list of abbreviations used in the gloss where these differ from those found in this textbook.
- You won't need to discuss syntactic properties that are not manifested in your language. For instance (question (8) below), not all languages mark morphologically the relationship between a head and its dependents; see Section 4.3.7. If you were discussing Chinese, question (8) wouldn't be relevant. In such cases, you can simply state that your language does not, for instance, display head- or dependent-marking. Similarly, you don't need to mention the antipassive construction unless your language has an ergative alignment (Chapter 7).
- You can collapse questions together where this makes sense for your language. For instance, questions (8) and (10) touch on the same kinds of data, and for some languages it would be appropriate to discuss them together.
- Make sure, when answering each question, that you provide adequate explanations: do not leave the reader to work out for themselves what your data show.

Some basic questions to consider:

- (1) What is the neutral, or unmarked, constituent order (sometimes termed 'word order') in the clause, if there is one? (Chapter 1, Chapter 6). If there is no neutral constituent order, describe the main principles of linearization. You should at least illustrate a transitive clause with two full NP arguments, and an intransitive clause. Are the orders the same in both these clause types?
- (2) What alternative neutral constituent orders are possible, if any? How marked are these?
- (3) What are the main word classes (or syntactic categories) in your language? Discuss any that have especially interesting properties. Focus on the LEXICAL classes N, V and A. You can expect any language to have a distinct class of nouns and verbs. Most languages will also have a distinct class of adjectives. Most will also have at least one or two (and maybe dozens) of adpositions. Justify all word classes that you posit: in other words, give evidence from its morphosyntactic properties and syntactic distribution to demonstrate that each proposed class should be regarded as distinct (Chapter 2, Chapter 4). Include some of the main FUNCTIONAL categories that your language distinguishes.
- (4) Is your language predominantly head-initial or predominantly head-final? Illustrate with data from more than one word class of heads. Are there any difficulties in establishing a predominant linearization? (Chapter 4). Remember that you are looking at the ordering of heads and their complements here, rather than the position of adjuncts with relation to heads.
- (5) How does your language express clausal negation? (Chapter 3).
- (6) Describe the main strategies for joining clauses together that are found in your language. What kinds of complementation occur? For instance, does the language have both finite and non-finite complement clauses? Does it have clausal

- subjects? If so, can they be both finite and non-finite? Does your language rely largely on subordination, as is the case for typical European languages? Does it have nominalized embeddings? Or does it, for instance, use co-ordination or verb serialization? (Chapter 3)
- (7) Are the constituent orders occurring in subordinate clauses the same as those in root clauses, or different? If different, describe the differences carefully (Chapter 3). Are there any (other) noteworthy differences between root and embedded clauses?
 - (8) How, *if at all*, does your language mark morphosyntactically the relationship between heads and dependents? (Chapter 4). In other words, is your language largely head-marking or largely dependent-marking? Does it display a mix of both strategies? Illustrate at least with reference to the verb and its arguments.
 - (9) Does your language readily identify distinct constituents? (Chapter 5) If so, give at least two tests for constituency, illustrating with contrasting grammatical and ungrammatical data. Are there distinctive DISPLACEMENT processes for constituents in your language? Perhaps alternatively your language has free word order of the type found in Warlpiri (Chapter 6); if so, illustrate.
 - (10) Describe the way(s) in which the grammatical functions A, O and S are identified in your language (Chapter 6). Does this rely predominantly on constituent order, on agreement or cross-referencing, or on case-marking? Does your language exemplify an accusative or an ergative alignment? Make sure you give enough data to illustrate this. If your language has ergative alignment, is this purely morphological, or is it also (a much rarer possibility) syntactic? In other words, is there a clear SUBJECT relation in your language? Is it possible to identify a syntactic pivot?
 - (11) Your language almost certainly has some readily identifiable ways to change the grammatical functions or relations, either increasing or decreasing the valency of a verb (Chapter 7). Does it have a passive? If ergative, an antipassive? An impersonal construction? An applicative? A causative?
 - (12) Describe how *wh*- (i.e. constituent) questions are formulated (Chapter 8). Does the language have *wh*-fronting or *wh*-in-situ, or perhaps both? Discuss the main strategies for forming relative clauses in your language (if appropriate, you could alternatively discuss these under question (6), as part of subordination). Are there other *wh*-constructions in your language; for instance, is focus movement found, and if so, is it similar to *wh*-question formation?
 - (13) Are there any other interesting syntactic constructions that are not covered by these questions? If so, explain and illustrate them.

You may be wondering why it's worthwhile to investigate the grammars of languages. I hope that the preceding chapters have answered this question, but, in case not, you should consider the fact that every week, languages are becoming extinct. Today there are perhaps 6000 or so languages in the world; we don't know the

exact number, and to some extent the answer depends on what counts as a distinct ‘language’ rather than a ‘dialect’. Most of these languages, perhaps as many as 90 per cent of the world’s total, are endangered. A language that has only a handful of speakers, even a few hundred or a few thousand speakers, is unlikely to survive to the end of this century. Many languages will become extinct by the end of this decade (see Crystal 2000; Dixon 1997).

When a language dies out because its speakers have chosen to speak (or been browbeaten into speaking) one of the large ‘global’ languages, much of the culture of that society is likely to die out too (see Nettle and Romaine 2000). Just as biological diversity is endangered by the relentless march of westernized societies, so linguistic diversity is threatened by domination from the world’s major languages, including English. Every time another language becomes extinct, we lose the opportunity to discover something more about the human language faculty; every language investigated to date has fascinating constructions and patterns that we may never know about unless linguists (including native speaker linguists) uncover them. If you decide that you want to undertake linguistic fieldwork, you will need professional training in all its aspects, which includes handling often complex socio-political situations.

(For more information about how to describe the syntax and morphosyntax of a language, see T. Payne (1997, 2006).)

9.2 A CASE STUDY: GRAMMATICAL SKETCH OF COLLOQUIAL WELSH

This section provides a necessarily brief grammatical sketch of Colloquial Welsh, illustrating the kinds of answer that could be given to questions in the previous section. The term ‘Colloquial Welsh’ is used by linguists to indicate, broadly speaking, the modern spoken language.

Colloquial Welsh is spoken in many, though not all, parts of Wales, where it has around half a million native speakers. Welsh speakers are also scattered throughout Britain, and there is a Welsh-speaking community in Argentina. The language is known as *Cymraeg* to its native speakers. Welsh is a member of the Celtic language family, a branch of Indo-European, and is thus related ultimately to English.

Let’s now turn to the syntactic properties. In the examples that follow, I have deliberately left in place the ‘messy’ morphological details that characterize natural languages. See if you can spot some of these. If you’d like further information, please contact me.

In neutral constituent order, a finite element (either a main verb or an auxiliary) is in clause-initial position. The subject immediately follows. With an inflected main verb, this gives VS(O) order, as in (14) and (15); no other neutral word orders occur. Many VSO languages allow an alternative SVO order, but Welsh does not, as (16) shows:

- (14) Gwerthodd Elin y delyn.
 sell.PAST.3SG Elin the harp
 ‘Elin sold the harp.’

(15) Diflannodd y delyn.
 disappear.PAST.3SG the harp
 'The harp disappeared.'

(16) *Elin gwerthodd y delyn.
 Elin sell.PAST.3SG the harp
 ('Elin sold the harp.')

A fixed constituent order identifies the grammatical functions, A, S and O. Welsh has accusative alignment, and a clear subject relation. S and A are identical; both immediately follow the finite verb or auxiliary, and, in a VSO clause, O immediately follows S. This is seen in (14). Both the S and A relations trigger subject agreement on a finite verb, under restricted conditions, as is illustrated below.

A finite auxiliary (in bold) occurs in clause-initial position in both (17) and (18); again, the subject immediately follows, and there is also a non-finite lexical verb lower down in the clause:

(17) **Gwnaeth** Elin werthu 'r delyn.
 do.PAST.3SG Elin sell.INFIN the harp
 'Elin sold the harp.'

(18) **Mae** Elin wedi / yn gwerthu 'r delyn.
 be.PRES.3SG Elin PERF / PROG sell.INFIN the harp
 'Elin has sold/is selling the harp.'

The main difference between these two clauses, apart from the different auxiliaries, is that (18) is an overtly aspectual clause. Aspectual particles, including *wedi* (PERFECT) and *yn* (PROGRESSIVE) co-occur with the auxiliary *bod*, 'be', as shown in (18). *Bod* is the only aspectual auxiliary; there is no 'have' auxiliary in Welsh. (The auxiliary *bod* 'be' has suppletive (= unpredictable and irregular) morphology, as you can see from (18).)

Welsh is a strongly head-initial language. A verb precedes its direct object, as in *gwerthu'r delyn*, 'sell the harp', just seen in (18). All adpositions are prepositions, as in *wrth y drws* 'at the door'. Nouns precede their possessors, as in *ci Elin* (dog Elin), 'Elin's dog'. Across all categories, in fact, heads precede complements.

Negation is a particularly complex area of Welsh syntax. Here, I illustrate clausal negation in finite clauses:

(19) Ddiflannodd y delyn **ddim**.
 disappear.PAST.3SG the harp NEG
 'The harp didn't disappear.'

(20) Wnaeth Elin **ddim** gwerthu 'r delyn.
 do.PAST.3SG Elin NEG sell.INFIN the harp
 'Elin didn't sell the harp.'

(21) Dydy Elin **ddim** wedi / yn gwerthu 'r delyn.
 NEG.be.PRES.3SG Elin NEG PERF / PROG sell.INFIN the harp
 'Elin hasn't sold/isn't selling the harp.'

- (22) Werthodd Elin **mo** 'r delyn / *ddim y delyn.
 sell.PAST.3SG Elin NEG the harp NEG the harp
 'Elin didn't sell the harp.'

As the first three examples show, clausal negation involves a negative adverb, *ddim*, which occurs in post-subject position. However, the adverb *ddim* can't appear before a direct object, as (22) shows; instead, a form *mo* is used, which literally means 'nothing of'. There are also changes to the finite verbs and auxiliaries in initial position in (19) through (22), as you'll see if you compare them with the affirmative clauses seen earlier; these changes occur because the clause is negative.

The lexical categories N, V and A are clearly distinct in Welsh, as they are in other European languages. The morphosyntactic categories that these word classes inflect for are very familiar from European languages, so I will not illustrate these specifically; examples occur in the data that follow. The inflectional categories include: for nouns, number and gender (masculine/feminine); for verbs, tense and subject agreement; and for adjectives, comparison. Welsh has a large class of prepositions, and these are more interesting, since most of them inflect to agree with their pronominal objects. Table 9.1 illustrates a characteristic paradigm. The bare citation form of the preposition is *wrth*.

Table 9.1

Inflectional paradigm for the Welsh preposition *wrth*, 'at'

	Singular	Plural
First person	<i>wrth-a i</i> at-1SG me 'at me'	<i>wrth-on ni</i> at-1PL us 'at us'
Second person	<i>wrth-at ti</i> at-2SG you 'at you'	<i>wrth-och chi</i> at-2PL you 'at you (PL)'
Third person	<i>wrth-o fo</i> / <i>wrth-i hi</i> at-3SG.M him / at-3SG.F her 'at him' / 'at her'	<i>wrth-yn nhw</i> at-3PL them 'at them'

As the occurrence of an inflectional paradigm for prepositions indicates, Welsh is head-marking rather than dependent-marking. Unlike in English, there is no case-marking whatever on either pronouns or nouns. Welsh has extensive agreement morphology. Heads agree with a following pronominal argument for person, number and, in the third person singular forms only, gender (excluding finite verbs, which display no gender agreement). In all instances, agreement crucially co-occurs *only* with a following pronominal argument, and *never* with a lexical noun phrase. This is illustrated in (23) through (25). Agreement occurs on six distinct categories of head, of which three are illustrated here: finite verbs agree with a pronominal subject: (23); non-finite verbs agree with their pronominal object: (24); prepositions also agree with their pronominal object: (25). In each case, the agreeing head and the following pronominal that it agrees with are underlined:

- (23) a. Cerddon nhw i 'r dre.
walk.PAST.3PL they to the town
'They walked to town.'
- b. Cerddodd / *Cerddon y genod i 'r dre.
walk.PAST.3SG / walk.PAST.3PL the girls to the town
'The girls walked to town.'
- (24) a. Gwnaeth Meic eu gweld nhw.
do.PAST.3SG Meic 3PL see.INFIN them
'Mike saw them.'
- b. Gwnaeth Meic (*eu) weld y genod.
do.PAST.3SG Meic 3PL see.INFIN the girls
'Meic saw the girls.'
- (25) a. arni hi b. arnyn nhw
on.3SG.F her on.3PL them
'on her' 'on them'
- c. ar yr eneth / *arni 'r eneth
on the girl / on 3SG.F the girl
'on the girl'
- d. ar y genod / *arnyn y genod
on the girls / on.3PL the girls
'on the girls'

Example (23a) shows a verb agreeing with a plural pronominal subject in person and number. However, when the verb has a lexical noun phrase subject, as in (23b), the agreeing form is ungrammatical. For finite verbs, there is no 'bare' citation form; instead, the third person singular is the default form, as (23b) shows.

For non-finite verbs, such as *gweld* in (24a), the agreement element is not a verbal inflection, but rather a preverbal marker (here, *eu*) which agrees with the pronominal object – here, in person and number. As (24b) shows, the agreement marker cannot occur with a lexical noun phrase object.

Examples (25a) and (25b) show an inflecting preposition, *ar* 'on', agreeing with a pronominal object. In (25c) and (25d), we see once again that when the preposition has a lexical noun phrase object, there is no agreement, and instead, the preposition occurs in its 'bare' citation form, *ar*.

Unlike more canonical head-marking languages, Welsh does not have true pronominal affixes; in other words, person and number cannot be reliably identified solely from the verbal or prepositional inflection. Table 9.2 illustrates this with the past tense paradigm for a regular verb, *gweld*, 'see' (giving 'I saw', 'you (sg.) saw' etc.). Note that there are only three distinct forms of the verb, *gwelis*, *gwelodd* and *gwelso*: the pronouns that follow are therefore essential to identify the participant. Colloquial Welsh is, then, not what is termed a 'pro-drop' language: the subject pronouns cannot generally be omitted.

Table 9.2

Inflectional paradigm for the past tense of the Welsh verb *gweld*, 'see'

	Singular		Plural	
First person	<i>gwelis</i> see.PAST.SG	<i>i</i> I	<i>gwelso</i> see.PAST.PL	<i>ni</i> we
Second person	<i>gwelis</i> see.PAST.SG	<i>ti</i> you	<i>gwelso</i> see.PAST.PL	<i>chi</i> you(pl)
Third person	<i>gwelodd</i> see.PAST.3SG	<i>o / hi</i> he/she	<i>gwelso</i> see.PAST.PL	<i>nhw</i> them

There are two remaining inflectional paradigms for lexical verbs in Colloquial Welsh, namely the future tense and the conditional. Both of these have parallel properties to the past tense in terms of verbal agreement.

As is typical for a European language, Welsh makes extensive use of subordination. Constituent order is the same both in finite root and embedded clauses: (26) illustrates a finite embedded VSO clause, bracketed:

- (26) Dywedodd Aled [darllenith Elin y papur].
 say.PAST.3SG Aled read.FUT.3SG Elin the paper
 'Aled said that Elin will read the paper.'

As well as finite complement clauses, as in (26), Welsh has infinitival complement clauses. The syntax of the latter is actually rather complex, since some of these are interpreted as finite, others as non-finite. Examples (27) and (28) illustrate these two types: both are introduced by a small functional element, *i*, which I've glossed as 'to' since it looks identical to the preposition *i* 'to'. The complement clauses in these examples are superficially identical, but have very different meanings and properties:

- (27) Dywedodd Aled [i Elin ddarllen y papur].
 say.PAST.3SG Aled to Elin read.INFIN the paper
 'Aled said that Elin had read the paper.'
- (28) Disgwylodd Aled [i Elin ddarllen y papur].
 expect.PAST.3SG Aled to Elin read.INFIN the paper
 'Aled expected Elin to read the paper.'
 (i.e. *not* 'Aled expected that Elin had read the paper.')

The difference between the two clause types stems from the kind of verb that occurs in the matrix clause. When the 'upstairs' predicate is a verb like *dweud*, 'say', as in (27), or *meddwl*, 'think', the infinitival clause is interpreted as finite. In fact, there is good evidence that these clauses really are finite, including the fact that they have the same interpretation as ordinary tensed clauses. Conversely, when the 'upstairs' predicate is a verb like *disgwyl* 'expect' or *dymuno*, 'wish/want', the infinitival clause is not interpreted as finite; instead, very like its English translation, it tends to refer to

future events that have not yet happened. Syntactically, the embedded clause in (28) does not behave like a finite clause, either. For instance, it is negated with a distinct negator that occurs in non-finite clauses, here shown in bold:

- (29) Disgwyliodd Aled [i Elin **beidio** â darllen y papur].
 expect.PAST.3SG Aled to Elin NEG with read.INFIN the paper
 ‘Aled expected Elin not to read the paper.’

Clausal subjects are another form of subordination in Welsh, but these can only be non-finite, as in the bracketed clause in (30), and not finite, as in (31):

- (30) Mae [mynd i'r cyfandir] yn gyffrous.
 be.PRES.3SG go.INFIN to the continent PRED exciting
 ‘Going to the continent is exciting.’
- (31) *Mae [bydd Aled yn mynd] yn gyffrous.
 be.PRES.3SG be.FUT.3SG Aled PROG go.INFIN PRED exciting
 (‘That Aled will be going is exciting.’)

Various valency-changing operations occur in Welsh. The main valency-reducing process is the passive, illustrated in (33), which is formed from the active construction in (32):

- (32) Mae'r plismon wedi dal y lladron.
 be.PRES.3SG the policeman PERF catch.INFIN the thieves
 ‘The policeman has caught the thieves.’
- (33) Mae'r lladron wedi cael eu dal (gan y plismon).
 be.PRES.3SG the thieves PERF get.INFIN 3PL catch.INFIN by the policeman
 ‘The thieves have been caught (by the policeman).’

As in passives generally, the subject of the active sentence in (32) – *y plismon*, ‘the policeman’ – is demoted or deleted in the passive; in (33), there is an optional *gan* ‘by’ phrase, containing the agent. And the direct object of the active construction (*y lladron* ‘the thieves’) is promoted to subject position in the passive. These changes can be seen from the constituent order: as always, the subject immediately follows the finite verbal element in Welsh. Like many other languages, Welsh has an auxiliary-plus-main-verb passive construction: the added auxiliary is *cael* ‘get’ in (33). The lexical verb, *dal* ‘catch’, also has an agreement marker in the passive, which agrees with the promoted subject in (33): here, it is *eu*, third person plural, agreeing with the plural subject (*y lladron* ‘the thieves’). Literally, the passive reads ‘The thieves have got their catching by the policeman’. There is no morphological passive in Welsh. (Note also that in keeping with the usual restrictions on agreement in Welsh, the finite auxiliary does not agree with a lexical subject NP in (33), and is therefore singular rather than plural.)

Welsh also has a causative construction, which uses a causative verb such as

gwneud, ‘make’, or *peri*, ‘cause’; there is no morphological causative. This is a valency-increasing construction.

There is also an impersonal construction, indicated by a morphological change in the verbal inflection:

- (34) **Torrwyd** y ffenest (gan y bachgen).
 break.PAST.IMPERSONAL the window by the boy
 ‘The window was broken by the boy.’

Unlike the passive, the impersonal construction does not involve the promotion of an object to subject position: *y ffenest* ‘the window’ is the object of the finite verb, and there is no subject at all in (34). The impersonal verb is marked for tense only, and never agrees with the post-verbal argument, even if this is a pronoun. This confirms that the post-verbal element is indeed the object, and not the subject.

Turning finally to *wh*-constructions, Welsh has a *wh*-fronting construction as shown in (35); more examples can be seen in Chapter 8:

- (35) [Pa ferch] welaist ti ___ neithiwr?
 which girl see.PAST.2SG you last.night
 ‘Which girl did you see last night?’

This construction leaves a gap in the position from which the *wh*-phrase has moved, which in this case is the direct object position. A relative clause formed on the direct object position of a finite verb is exactly parallel:

- (36) y ferch welaist ti ___ neithiwr
 the girl see.PAST.2SG you last.night
 ‘the girl you saw last night’

For both *wh*-fronting and relative clause formation, the subject and object of finite verbs in Welsh behave in a similar way: in both cases, there is a gap in the clause. Further down the Accessibility Hierarchy (see Chapter 8), a resumptive strategy either may or must be used, under rather complex conditions. Here, I will simply illustrate one such construction, a *wh*-question formed on the object of a preposition. Note that the preposition inflects and that a resumptive pronoun is optional:

- (37) Pwy gest ti ’r anrheg ganddo (fo)?
 who get.PAST.2SG you the present with.3SG.M him
 ‘Who did you get the present from ___?’

Compare here the English translation, which has a gap in the position of prepositional object.

Focus constructions in Welsh are structurally parallel to *wh*-questions, and also involve the fronting of a constituent; some examples can be seen in exercise 4 in Chapter 5.

There are many more fascinating features of Welsh syntax (see Borsley *et al* 2007); the case study provides a brief sketch and illustrates the major typological properties of the language.

9.3 SOME QUESTIONS CONCERNING SYNTAX

In this section, I briefly address some questions and controversies that have been widely debated in recent linguistic research. Some central readings are provided to whet your appetite.

- Why is it so hard for adults to learn a new language?

Many of us have *attempted to learn* at least one language, in other words by making a conscious effort. And we mostly find it very difficult, even if we're immersed in the new culture. As adults, we can only envy the effortless, naturalistic way in which children learn the ambient language(s) of their culture without any instruction. Moreover, across the world, it's totally normal for children to learn – natively – more than one language; it's English-speaking cultures that are abnormal in being so overwhelmingly monolingual. So children have a head-start in language learning, and adult language learners are always at a grave disadvantage compared to children. As we approach puberty, our language-learning ability declines; beyond puberty, it largely atrophies. After that we may, with persistence, become fluent speakers of a new language, but we won't become native speakers – and we won't have the same intuitions about grammaticality as someone who learnt the language as a child. Language, in common with many other acquired skills, such as musical ability, has what is called a critical period (or a sensitive period) for learning. If learning takes place beyond that period, it is no longer effortless, and acquisition will probably be less than native-like.

In biological terms, none of this is terribly surprising. Many other animals exhibit sensitive periods for various systems, including motor systems (involving movement), sensory systems, and behaviour. An example often cited is birdsong. Some (though not all) species of songbirds have to learn their songs by hearing an adult model, inevitably inviting comparisons with language. If they don't have an appropriate adult model (for instance, if they are reared without a singing adult male bird), their song fails to develop properly. Some aspects of song are therefore genetically determined, but input from the environment is crucial. What is important in both birdsong and language, then (and indeed, in many other biological systems) is the interaction between genes and environment. Humans have a language faculty which is genetically specified. This does not mean that there is *in any sense* a single 'language gene', or, most likely, even a dedicated group of language genes. It means that in normal situations, we all acquire at least one language as children: that is hardwired in our species. The interaction of many genes is almost certainly involved. And input from the environment is needed before the child's brain can get to work building a language. No other

species has a language faculty, and no other species can acquire a human language, even under intensive instruction. But for the genetic predisposition to learn language to be triggered, cultural input is required. Language learning in children relies on normal human interaction, including exposure to language data: this is the environment.

How much of our linguistic ability is pre-specified by our genes, and how much is down to the influence of our environment, is currently a hot topic of debate within linguistics. If all languages have some feature in common, is it inbuilt, part of the genetic recipe for being human? Or can it be attributed to the fact that we are all exposed to human cultures, which present children with certain uniform experiences of the world? Frankly, we don't really know.

When adults do try to learn another language, we typically find it easier to learn languages that are closely related to our own, or that are similar typologically – in other words, languages which have similar characteristics, such as sharing the same word order. As in other spheres of life, the familiar is easier to grasp than the radically new. But this brings us to the next question.

- Are all languages equally complex? And are some languages harder to learn than others?

Impressionistically, few people (including professional linguists) are in any doubt that some languages are harder for each of us to learn as adults. But can languages be intrinsically hard, or intrinsically easy? It used to be generally considered that all languages were, essentially, equally complex, and that complexity in one area of the grammar would be balanced out by simplicity in another area. Recently, that view – which was more ideological than evidence-based – has been challenged from many quarters, and it now seems indefensible. The collection of papers in Sampson *et al.* (2009) provides much interesting discussion. One of the authors (Guy Deutscher) calls the claim that 'all languages are equally complex' nothing more than an urban legend!

It also now seems that the demographic properties of a language – including the number of speakers it has and the extent of its spread around the world – directly correlate with the linguistic complexity of the language (Lupyan and Dale 2010). At least in terms of their morphosyntax, large global languages which have many millions of speakers, such as English, have been found to be massively simpler than languages with small populations (fewer than 100,000 people) which are spoken only in one region of the world. One of the main factors seems to be that the large global languages are under pressure to become simpler over time because they are learned by adult learners – who, as we saw earlier, are not very good at learning the complexities of language when compared with child learners. Conversely, 'esoteric' languages – the small languages of remote communities – may maintain their linguistic complexity exactly because it facilitates learning by infants; complex morphosyntax seems to provide cues to language structure, and since children are so good at learning such complexities, there is no pressure for it to decrease within a small, closed community. For instance, Levinson (2006) discusses a language called Yélí Dnye, spoken by fewer than 4000 people on a remote island (Rossel

Island) several hundred kilometres off the coast of New Guinea. This language, Levinson suggests, is so complex that it lies at 'the boundaries of learnability': adult incomers to the community cannot learn it, and children seem to need an entire community of speakers to learn it successfully, so that if their parents migrate, the offspring may fail to acquire the language fully. It seems, then, that some languages genuinely are more complex than others.

Having reached the end of this book, you should be clear that complexity lies in differing areas of the grammar from language to language. But in standard circumstances (living within a normal linguistic community) children seem to learn each language *as a system* with equal ease, as far as is known. Certainly, there are no languages so hard that their speakers don't become fluent until they're 18 years old. So we can say that whatever complexities a language throws at children, they can cope. Does that mean that all languages are 'the same' in terms of their inherent difficulty for children? I leave this as an open question.

- Do all languages manifest broadly the same syntactic properties?

From reading this book, you will know that, on the face of things, there is a great deal of syntactic and morphosyntactic diversity between languages. Does that mean that languages can vary from each other at random, differing in essentially any way? Recently, some eminent linguists have suggested that this is the case (Evans and Levinson 2009). These authors reject the idea that languages are built to a universal pattern, citing many examples of 'esoteric' data that are not common to all languages. They claim that languages can differ in fundamental ways, resulting in a 'jungle' of linguistic complexity. Equally, there is, in their view, no language faculty – no innate template for language learning that is shared by all members of our species. But this seems to hugely overstate the case. Since all normal children are able to learn a language or several languages very quickly indeed, and without any instruction, and since no other species can achieve anything remotely similar, it seems wrong to deny that we are biologically pre-programmed for language learning.

A very different view to that of Evans and Levinson is outlined by linguists Ray Jackendoff and Peter Culicover (Jackendoff 2002; Culicover and Jackendoff 2005). These linguists suggest that what is often termed 'universal grammar' – the biological endowment for language learning in our species – provides a 'toolkit': a set of basic principles for building languages, which each language customizes in its own unique ways. There is no reason to expect that everything the toolkit can build will be found in all languages, and this is clearly correct. But the toolkit constrains what *can* be built; when properly investigated, languages do *not* vary from each other at random, but rather, look extremely similar. For instance, Morcom (2009) investigated whether or not languages all have a distinct lexical class of nouns and verbs, something which has been denied by certain linguists. She looked at the most controversial languages, and discovered that in each case, there was indeed a distinct noun word class and a distinct verb word class. Careful investigation of this nature by trained linguists often uncovers patterns that are not obvious on the surface.

Having finished revising this book for the third edition, I am left with the impression that despite the very evident cross-linguistic diversity in syntax and morphology, the languages of the world are similar in many crucial ways. All distinguish several word classes (Chapter 2) and, among these word classes, it seems that all languages have a class of nouns and a separate class of verbs. All languages have predicates and all have participants in the event denoted by the predication. All languages have ways of negating clauses, of asking questions, of giving commands. Most languages (perhaps all) distinguish between simple sentences and complex sentences, although not all languages make use of the kind of subordination that is familiar from European languages, as we saw in Chapter 3. All languages have heads which, together with their dependents, form phrases – the constituents of sentences; these were the topics of Chapters 4, 5 and 6. Although languages do not share the same set of constituents, the same sorts of test for constituent structure can be applied in all languages. These tests rely on harnessing the native-speaker intuitions which we all have about our native language(s), by calling on us to make grammaticality judgements. Hierarchical constituent structure is a universal linguistic feature, though it is clearly exploited more in some languages than others; free word order languages, such as those discussed in Chapter 6, make less use of constituency. All languages exhibit dependencies between elements in a sentence, such as those examined in Chapters 4, 5, and 8. All languages have at least one method of encoding grammatical relations – constituent order, morphological case or verbal agreement – as we saw in Chapter 6. All languages appear to exploit variations in constituent order or word order to foreground or background elements, to add focus and emphasis, or to show the topic of a sentence. The vast majority of languages have at least some valency-changing processes, and processes of promotion and demotion which change the basic grammatical relations borne by noun phrases (Chapter 7). The remarkable unity among totally unrelated languages is nicely illustrated in Chapter 7 by the applicative construction, which turns up again and again across the world, and which has much the same grammatical effects in each case. And though not all languages have *wh*-movement, all languages have a way of forming *wh*-questions, and probably relative clauses too (Chapter 8).

My overall impression, then, is that the syntactic diversity among languages from different families and different regions of the world is not trivial, but that the overwhelming homogeneity which exists between languages is far more impressive. In particular, when we look at language isolates (languages with no known relatives) and find that they too utilize the same ‘toolkit’, it seems safe to say that languages are, unmistakably, amazingly similar in design.

9.4 LAST WORDS: MORE SYNTAX AHEAD

My feeling when I started writing the first edition of this book (published by Arnold/OUP in 1998) was that there is an awful lot of syntax out there in the world, much of it rather daunting. This is a view that students of syntax often appear to share! I hope that by now you are familiar with many of the basic concepts needed in order to

understand the ways in which syntax operates in the natural languages of the world, and feel less daunted by its complexities. There is certainly much more syntax out there than a short book can cover, but my intention has been to introduce you to the major syntactic constructions found in the world's languages, and to the main ideas, terms and concepts used by linguists to discuss syntax.

Of course, the syntax part of the grammar of a language doesn't exist in isolation. We have seen in nearly every chapter how it interacts both with form (morphology) and meaning (semantics). Not all languages have much morphology, as I've often noted, in the sense of having variations in the form of words. However, many languages use morphology to signal the kinds of syntactic processes that I've talked about in this book, such as the formation of passives or of causatives. Many languages use morphology – case-marking, verbal agreement, or both – to distinguish between the core participants in a clause, although some languages rely almost entirely on word order to do this. Matters morphological have arisen over and over throughout this book, and if these have interested you, you may wish to move on to a specialized book on the topic, such as Lieber (2010).

Although the topics of semantics, pragmatics, and discourse are beyond the scope of this book, all of these areas also critically interact with syntax in all languages. For instance, there are important discourse factors involved in the selection of syntactic constructions; as an example, we saw in Chapter 7 how the need to allow noun phrases to be coreferential with previous noun phrases within the sentence can give rise to constructions such as the passive and the antipassive.

Having completed this introduction, you are now ready to further your study of syntax. There are various (overlapping) paths your study might take. One is to look at descriptions of languages, studying grammars written by linguists. In a good grammar, the chapters on syntax and morphology should be very prominent, and should ideally cover (at least) all the areas we've seen in this book: word classes, grammatical categories (Chapter 2); simple sentences and complex sentences (Chapter 3); heads and their dependents, head-initial or head-final syntax, head-marking or dependent-marking morphology (Chapter 4); constituent structure (Chapter 5); case, agreement, constituent order and grammatical relations (Chapter 6); syntactic processes which change grammatical relations, such as passives and/or antipassives, causatives and applicatives (Chapter 7); and *wh*-questions, relative clauses, and focus constructions (Chapter 8).

You might also take a course in linguistic fieldwork, which will build on the knowledge gained throughout this text, and might ultimately lead you to investigate the grammar of languages as yet undescribed (of which there are many).

A further way your study might proceed is by looking at syntactic theory. In order to explain the syntactic differences and similarities between languages, linguists need first to know how alike (and unlike) the world's languages are. This requires good descriptions of the sort mentioned earlier. Most linguists want not merely to describe languages in isolation, however, but to discover the ways in which their structures are related, even when there are no genetic relationships between the languages. For instance, the morphology and syntax in the majority of languages operates on the basis of either the nominative/accusative system or the ergative/

absolute system, with the former predominating cross-linguistically, as we saw in Chapter 6. Why do languages generally 'choose' one system or the other as their major system, in spite of the existence of several other logical possibilities? The likelihood is, as I suggested in Chapter 6, that the most economical way of, say, distinguishing between grammatical relations is to use one or other of these major systems.

Such economies in the grammar of a language are of interest to theoretical linguists, in part because we hope they will ultimately tell us something about how children can learn their native languages so quickly, regardless of all the complexities that exist. Linguists are also interested in language as a faculty unique to one species, *Homo sapiens*, and in addition, through the study of the human language faculty, we seek to discover more about the remarkable properties of the brain and of human cognition.