Introduction

This book consists of three parts, each of which is an introduction to a separate discipline: cognitive science, linguistics (a branch of cognitive science) and English grammar (a branch of linguistics).

Part I, called 'How the mind works', is a very modest alternative to Steven Pinker's bestseller of the same name (Pinker 1998a), and is a personal selection of rather commonplace psychological ideas about concepts and mental networks and the activation that flows round them, together with a few novelties such as default inheritance and node building. These ideas are selected so as to provide a foundation for the next part.

In Part II, 'How language works', I make a theoretical point that's exactly the opposite of the one made famous by Pinker, following the mainstream Chomskyan tradition (Pinker 1994). Where Pinker finds a 'language instinct', I find ordinary cognition. Like other 'cognitive linguists', I believe that language is very similar to other kinds of thinking. I also believe that the fine details that we linguists find when looking at language tell us a great deal not only about language, but also about how we think in general. Every single phenomenon that I know about, as a linguist, is just as you'd expect given the way in which (according to Part I) the mind works.

Finally, Part III, 'How English works', gives a brief survey of English grammar. The chapter on syntax summarizes my little 1998 textbook *English Grammar* which supported my first-year undergraduate course on English grammar. The students seemed to enjoy learning to draw dependency arrows and appreciated the idea that this was a skill that they could apply to virtually any English sentence.

I should explain that the book's structure is itself a little like the structure of thought: it's a network. Admittedly, it doesn't look like a network at first sight; if you look at the table of contents you'll see the usual hierarchical structure of parts, chapters and sections. But if you look more carefully, you'll find that most of the chapters and sections correspond across the three parts. For example, Section 2.2 discusses general principles of classification which are then applied in 6.3 to the principles of how we classify words, which in turn lead into the exposition of English word-classes in 10.1.

The structure based on parts and the one indicated by the cross-links between parts correspond to the two structures of the intellectual picture that I want to present. The hierarchical structure follows the academic divisions: Part I is the broad discipline of cognitive science, which includes linguistics (Part II), which

includes English grammar (Part III). Each of these disciplines has its own logical structure, so the chapters and sections try to follow this logic. But the cross-links are the book's main point because they show how various general ideas from cognitive science apply to language and explain its characteristics. It's not just that there are some parts of language that are similar to other parts of thinking. What I'm claiming is that the whole of language can be explained in this way, so I have to justify the claim in detail with a link from every section in Part II to some section in Part I.

Fortunately, the corresponding sections in the three parts follow exactly the same order because they follow the same logic, which means that you can read the book either linearly or laterally. A linear reading takes you through a course in cognitive science, then through a course in linguistics and finally through a course in English grammar, each following its own internal logic. A lateral reading takes you from a section in Part I into its corresponding section in Part II and on into a section in Part III – or, if you prefer, in the opposite direction.

How you cope with this choice is, of course, up to you. One obvious solution is to combine the linear and lateral approaches. If you follow this strategy, you'll start at the beginning of Part I, read the first section, then read the corresponding section in Part II, then the one (if there is one) in Part III, then back to the next section in Part I; and so on. This is how I hope more advanced students will read it, and to encourage them I've added a note at the end of most sections in Parts I and II recommending that they should stray into a section of the next part, where (to increase the temptation) they'll also find a summary of this section. This is what I call the '**advanced route**'. But I accept that some readers will prefer to follow a purely linear route which takes them straight through the book, and don't need sign-posts.

If you're a teacher, you may like to know how I would use this book as a textbook for my undergraduate teaching. I would spread it across two years, with Part III for first-year students and Parts I and II for the second year. First-year undergraduates can certainly cope with the grammatical analyses of Part III, especially if they make use of the material on the website; indeed, these analyses aren't much harder than those that are standardly taught in many countries to primary school children. The practical experience of exploring the 'real language' of texts is an excellent foundation for the more theoretical exploration in the first and second parts, and is probably especially important for students who have come through the more or less grammar-free schools of most English-speaking countries (Hudson and Walmsley 2005). I've mapped out a '**novice route**' through the book which basically takes them through Part III, but with little excursions into the corresponding sections of Part II. The 'advanced route' should suit second-year students, who can obviously use their discretion about revisiting Part III.

If you're a student, then I should explain my policy on bibliographical references. I assume that you're a typical modern student with easy access to the internet and more IT skills than time. I also assume that you'd like to be able to follow

up some of the research that I quote, but without having to cope with the dense technicalities of research literature. With these two thoughts in mind, I decided to make as much use as I could of two wonderful resources: Wikipedia (en.wikipedia.org) and the second edition of the Elsevier *Encyclopedia of Language and Linguistics* (Brown 2006) which your university may well make available to you online.

Wikipedia is especially good for Part I as it gives easy access to the rather elementary research ideas that I discuss, but please remember to take it with a pinch of salt. As far as I can tell, the articles I recommend are, by and large, sensible and scholarly, but some of the claims are inevitably controversial, and occasional silliness is hard to avoid in a work that anyone can edit. If in doubt about something you find in Wikipedia, try searching in Google, and especially in Google Scholar and Google Books. For Part II, of course, the *Encyclopedia* is the main point of reference. The articles in both sources are written by experts with whom I can't compete; my main contribution is simply to have put their ideas together in an unusual combination.

More material is available on the book's **website** (www.phon.ucl.ac.uk/home/dick/izwg/index.htm) for those who want it, and especially for those who want to hone the skills that Part III tries to develop; it includes an encyclopaedia of English grammar and Word Grammar, but much more besides.

And of course, for those who want to know more about Word Grammar, there are plenty of publications, not least my most recent (2007) monograph, *Language Networks: the New Word Grammar*. There's no better test for ideas than writing a book about them, whether it's a monograph or a textbook, and this textbook is no exception. Consequently I have to report a number of points where I've changed my mind even since writing *Language Networks*: choice sets (3.3), best landmarks (3.4.3), the notation for coordination and dependencies (7.5) and the mechanism for resolving word-order conflicts (7.6). This is as you'd expect. After all, Word Grammar is a network of ideas in my mind, and as I explain in Part I, any cognitive network is forever changing as it tries to adjust to reality.

Where next?

Advanced: Part I, Chapter 1: Introduction to cognitive science
Novice: Part III, Chapter 9: Introduction to English linguistics

PART I

How the mind works

1 Introduction to cognitive science

Although this book is about language, the first part is not about language as such at all, but about general **COGNITION –** i.e. 'knowledge'. Its aim is to provide a general background to the discussion of language in the second part.

Cognition includes everything you might think of as knowledge – knowledge of people, things, events – and may be as general as so-called 'general knowledge' or as specific as what you know about the room you're sitting in at the moment. If we want to understand cognition, we must answer questions such as the following:

- How is it organized in our minds?
- How do we learn it?
- How do we use it in understanding our experiences, in solving problems and in planning actions?
- How is it related to things that we wouldn't call 'knowledge', such as feelings, actions and perceptions?

The main point of this book is to show how answers to these questions throw light on language; or to put it more negatively, how unlikely we are to understand language if we ignore what's already known about cognition.

Cognition is very complex and diverse, so it's hardly surprising that a range of methods have been used for studying it. The term **COGNITIVE SCIENCE** is often used as a cover term for the various different disciplines that explore cognition, including psychology, neuroscience, artificial intelligence, philosophy and (of course) linguistics. (Wikipedia: 'Cognitive science'.) Nor is it surprising that there's a great deal of controversy about findings and theories, so I can't claim that the theory which I present here is the agreed view of every cognitive scientist. Nor, indeed, can I claim to be an expert on cognitive science (in contrast with linguistics, where I do claim some expertise). What I can claim, though, is that the ideas I present in this part are compatible with elementary cognitive science. Most of the things in these chapters can be found in introductory textbooks, though no other textbook presents this particular combination of ideas and theories.

The main differences between the various disciplines that study cognition lie in their research methods, so it will be helpful to outline here the main methods that underpin the research findings, and especially the methods that are used in the research that I present below. **Psychology** uses many methods, but the most

relevant for us are experiments which measure the time ('response time') taken by people ('subjects') to perform very specific tasks when sitting in front of a computer in a psychological laboratory. **Neuroscience** uses brain scans which reveal the structure of the brain and which parts of the brain are particularly active at any given moment. **Artificial intelligence** uses computer programs that try to simulate human behaviour. **Philosophy** uses logical argument about how knowledge and thought must 'work'. And **linguistics** uses a variety of methods, including the famous 'grammaticality judgement' and other kinds of self-report made by a native speaker. (Wikipedia: 'Psychology', 'Neuroscience', 'Artificial intelligence' and 'Linguistics'.)

As you can see, these disciplines collectively offer an impressive range of methods for studying human cognition, and the ideal situation will be one in which they all support each other – for instance, where the results of laboratory experiments converge with those of observed behaviour and of brain scans. At present this ideal still lies in the future, but the major debates and disputes lie within the disciplines rather than between them. It would be strange indeed if, say, psychologists all accepted a view of the mind which all neuroscientists rejected. Instead, there is enough diversity within each discipline to allow a synthesis, such as the one that I offer here, which combines at least the greater part of the research findings of all of them.

This is the justification for the first part of my book, in which I try to present a unified view of those areas of cognition that are most directly relevant to language. Having laid this foundation, I shall then be able to show how we apply this general-purpose cognition to language, and I hope to persuade you by the end of the book that language, in spite of its apparent peculiarities, is actually just an ordinary example of human knowledge applied to the particular task of communicating.

There are two reasons for celebrating this result. The first is that it gives us the best possible explanation for the known characteristics of language: they're exactly what we would expect given the kinds of mind we have.

The second reason for celebration is that linguistics acquires a very special role in cognitive science. Language has a far clearer and more intricate structure than any other part of cognition, and only linguistics can explore this structure in detail. Consequently, the window into the human mind that language provides is unusually clear and throws light on areas of thought that other disciplines can't reach. I hope that by the end of the book you'll feel that you have a better understanding not only of how you use language, but also of how you think.

Where next?

Advanced: Part I, Chapter 2: Categorization

2 Categorization

2.1 Concepts, categories and exemplars

One of the most important areas of work in psychology is the study of categorization, which explains how we **CATEGORIZE** or classify items of everyday experience. (Wikipedia: 'Categorization'.) The examples discussed are very familiar and mundane – things like birds, furniture and birthday parties – and the question is how we all manage to find our way through this familiar world so efficiently. How do we know what things are, and what good is this information? The answers are fairly obvious, and make good sense in terms of everyday experience.

2.1.1 Concepts and properties

The main point is that we have a vast stock of **CONCEPTS**, each of which has a set of things we know about it called **PROPERTIES**. (Wikipedia: 'Concept'.) For example, we have the concept 'bird', with the following typical properties:

- It flies.
- It has wings.
- It has feathers.
- It lays eggs.
- It has two legs.
- It has a beak.

If you explore your knowledge of birds, no doubt you can extend this list of properties.

These properties define the general concept 'bird' and distinguish it from other concepts – 'fish', 'aeroplane' and so on. They all emerge in answer to the question: 'what is a bird?' You can explore the properties of any other concept in your mind by asking yourself what you know about it. What is furniture like? What is a birthday party like? In each case, what you're looking for is a list of things that you know – or at least think – about a typical example of the concept in question, including the things that distinguish it from other concepts.

A concept is simply an idea, so it's very different from whatever it represents in the real world (the thing it's a concept 'of'); thus the concept of a bird does

not itself fly or have wings, though a bird does and the concept's properties are designed to fit those of the bird. Similarly, birthday parties and accidents have a time and a place, but the concept of a birthday party or an accident doesn't.

A concept, on the other hand, may be real or imaginary, according to whether or not it matches something in the world, whereas the thing it represents either doesn't exist (as in the case of the concepts 'unicorn', 'Father Christmas' and Peter Pan's flight to Neverland) or does exist. A concept exists, but only as an element of someone's mind; in technical terminology it's part of **CONCEPTUAL STRUCTURE**, a term which means much the same as our everyday word *knowledge* but with the extra idea that this knowledge has a structure.

The main point is that the concept is different from the bit of the world that it represents. On the other hand, the only reason for having concepts is to guide us through the world, so the better the fit with the world, the better they guide us.

2.1.2 Inheritance

How does a concept guide us? Imagine life without any general concepts. The problem you face is total novelty – everything you see and hear is new, so you never benefit from experience. When you put something in your mouth, you have no idea what taste or texture to expect – or even whether it's food. You have no concept for 'potato' or 'apple', or even for 'food'. When you want to open a door, you won't know how to do it because you have no concept for 'opening a door', or even for 'door'.

Now return to reality, complete with the millions of concepts that you've learned during your life so far. When you see a potato, you can predict its taste, its texture and the effect it will have on your hunger. How do you know that it's a potato? By looking at it. But how do you know that it will satisfy your hunger? You can't see this or even taste it, and yet you know it for sure. In lay terms, you 'guess' it: you know that the typical potato stops hunger, you guess (from its appearance) that this thing is a potato, so you can also guess that this thing will stop hunger.

This everyday guessing is something cognitive scientists know quite a bit about, and it's generally considered to involve a process called **INHERITANCE** that will be one of the main themes of later sections (starting with 2.3).

The examples concerning doors and potatoes show how concepts guide us through life. To summarize, when we meet a bit of experience (whether a thing, an event or a person), we take two steps:

- On the basis of what we know already, including its perceived properties – e.g. what it looks or sounds like – we classify it as an example of some concept that we know already.
- Then we infer more information about it by inheriting further properties from that concept.

In other words, it's concepts that allow us to build on past experience by linking perceived properties to those that we can't perceive.

As we shall see below, the system isn't perfect, but on the whole it works well; and most importantly of all, it works fast. The two steps don't need reflection, but happen almost instantly. For instance, if you're driving on a fast road, you recognize other cars and react to them in a split second. The price you pay for this fast thinking is the possibility of error.

Concepts don't have to represent simple concrete objects, but can represent events as well. Take 'birthday party', in the typical sense of a child's birthday party complete with balloons on the front door, presents, party clothes, games, cake and so on. We all know how the party is organized into sub-events starting with the guests' arrival and ending with their departure; and both parents and children have detailed ideas of what the guests' and hosts' roles involve. If you don't know these things, or if your ideas conflict with those of the birthday girl or boy, then disaster threatens.

The concepts that we've considered so far have been concepts for general notions such as potatoes, birds or birthday parties. As you've probably noticed, all these concepts happen to be the meanings that we would expect to find in a dictionary, so they'll be important for the analysis of meaning (8.7); but word meanings are just the tip of a gigantic iceberg of concepts. The meanings that we give to single words are the concepts that are so general, useful and widely held that society at large gives them a label; but most concepts are far too specific and ephemeral for this – concepts such as 'an undercooked potato' or 'the dust on top of my computer'. Concepts like this don't appear in a dictionary, but they can always be put into words.

2.1.3 Categories and exemplars

Even more specific and ephemeral are what psychologists call **EXEMPLARS** – individual examples of experience – in contrast with the more or less general **CATEGORIES** which we use in categorizing the exemplars. This distinction is so important that some psychologists divide memory into two separate areas which they call (not very helpfully) 'semantic memory' and 'episodic memory', with semantic memory for categories ('semantic' because the categories are meaningful) and episodic memory for remembered exemplars ('episodic' because these exemplars occur in 'episodes' of experience). (Wikipedia: 'Semantic memory' and 'Episodic memory'.)

However, this fundamental distinction is controversial and other psychologists believe that a single memory system includes both kinds of information (Barsalou 1992: 129). This is my view too, and Section 4.3 will suggest that exemplars can turn into categories without in any sense having to move from one system to another. The difference between the two kinds of memory is simply a matter of degree – how specific they are, how much detail about times, places and so on they include, and how long we keep them in our minds.

In these terms, then, the aim of categorization is to put exemplars into categories. The next two sections explain the logical structures that this process creates and how these structures help us to understand our experiences.

Another popular distinction separates 'procedural knowledge', knowing how to do something, from 'declarative knowledge', knowing that something is true. (Wikipedia: 'Procedural memory' and 'Declarative memory'.) Procedural knowledge includes skills such as riding a bike or driving a car which we can't report in words, whereas knowing the properties of birds is an example of declarative knowledge which we can put into words; and crucially for this book, language is often classified as procedural knowledge.

The main reason for making this distinction is that procedural knowledge tends to be so automatic that we can't put it into words, but this is merely a tendency; for example, a driving instructor can put the skills of driving into words, and they are anything but automatic for a learner. Moreover, if we can have declarative knowledge about events such as birthday parties, it's hard to see why we can't use declarative knowledge about driving a car while driving.

A reasonable conclusion seems to be that it's better not to separate procedural and declarative knowledge, especially if we distinguish concepts for procedures from the 'motor skills' (as we shall call them) that are discussed in Section 3.1. Rather obviously, rejecting the distinction between declarative and procedural knowledge undermines the claim that language is procedural. This is a triumph for common sense because we obviously can talk about language – as witness Parts II and III of this book.

Where next?

Advanced: Part II, Chapter 6.1: Types and tokens

2.2 Taxonomies and the isA relation

Another matter of common sense is that our concepts are organized, rather than just a disorganized heap. If you think of 'bird', it's obvious that it must be related to 'feather', 'wing', 'fly' and so on. We explore these relations in Section 3.5, but first we start with the most familiar kind of organization, which is called a **TAXONOMY**.

A simple example of a taxonomy includes 'creature', 'bird' and 'robin', in that order, with 'creature' at the top of the hierarchy and 'robin' at the bottom. As we all know, a bird is a creature – at least that was the assumption behind the dictionary definition quoted above in which a bird was a creature with feathers and wings – and a robin is a bird. But of course a creature is not a bird, nor is a bird a robin, so the expression *is a* defines an unequal relation.

This relation is the basis for any taxonomy, and is so important that cognitive scientists have invented the term **ISA** as the name for the relation, so the concept 'robin' isA 'bird', which isA 'creature'. (Wikipedia: 'Is-a'.)

What distinguishes a taxonomy from a mere list of members is that it allows classification at more than one level. For example, creatures can be classified as