1 Introduction

1.1 Overview

This book is about how language, and languages inhabit the mind. The discussion will deal not only with 'language' in the broadest sense of that term but also touch on issues of cognition in general.

Looking at the extreme conceptual complexity that characterises the ways humans think, it is tempting to attribute this, in part at least, to the creative power that is afforded us by our language ability. In many other respects we are similar to other, higher mammals that do not have a language system like ours: the information coming from the environment impacts on the senses, is perceived and processed such that the results of these processes acquire particular values and meanings. Whether increasing cognitive sophistication and the development of the language processing system are causally linked in some way is an intriguing issue which we shall not go into. The book will certainly devote space to these larger issues of human cognition although the focus will be on language ability itself, which, for the majority of the inhabitants on this planet, means how we use the language systems (in the plural) that most of us, to a greater or lesser extent, possess.

Language is a topic that seems to fascinate everyone, all the more so now that even the most persistently monolingual communities are becoming more aware of other languages and other cultures. Why, people ask, is acquiring a language so much more straightforward when you are a small child? Can you really possess more than one mother tongue? Why are we adults often so frustratingly slow in picking up a new language in comparison to our sons or daughters? How can we come to know facts about the grammar of a language, use this knowledge to spot and correct our own mistakes and yet continue to make them? Why do some adults acquire new languages more readily than others?

If we seek answers to such questions, it is natural that we should first turn to research that has been carried out in the most obviously relevant fields, namely *child language* (or *first language*) *acquisition*, *second language acquisition*, and also the broader field of *bilingualism* (or *multilingualism*). Over the past decades, research in these fields has yielded some impressive insights about

various aspects of the language acquisition process. Inevitably, though, the more knowledge of the processes and mechanisms underlying language development advances, the more gaps in our knowledge are revealed. The other consequence of this is that the fields develop numerous specialisms each with their own research tools and research traditions each tending to work in isolation from the others. Luckily there is movement in the other direction, namely an increasing need to search across disciplines for problems that remain unresolved. This interdisciplinary trend is assisted by the recognition of an overarching family of disciplines under the rubric of *cognitive science*. It is as a potential contribution to understanding the human mind, in other words to cognitive science in general as well as to more specific fields within the study of language that this book is intended. In this spirit, the chapters that follow introduce and discuss a multidisciplinary theoretical approach called the Modular On-line Growth and Use of Language (MOGUL) framework.

MOGUL is not a brand new theory. Rather, it is designed to reconcile the highly specialised work done in different domains with the need to see mental activity and particularly the growth of new mental structures within a larger perspective. For this to be more than a token endeavour, some serious groundwork will be needed. As already implied, apart from human language ability, the discussion will cover crucial topics such as the role of memory, emotion, and general principles of processing some of which have been only briefly discussed in the language acquisition literature.

In the first instance, the book draws for its inspiration on particular strands of contemporary theoretical research in (psycho)linguistics. However, since the explication of the processes by which language systems grow and are accessed in the mind of individuals will, as just mentioned, take the discussion into more general areas of language and cognition, the book will also capitalise on theoretical and empirical research in the broader field of cognitive science. The theme that unites the whole book is the idea that a comprehensive theory of language acquisition must be one that properly integrates explanations of the structural characteristics of the developing linguistic system on the one hand with explanations of how and why those characteristics develop the way they do in real time. These are, traditionally, separate issues which are investigated in separate fields of research.

In this first chapter, we will go on to discuss some general issues and introduce, in very basic terms, the theoretical framework we call MOGUL and, to conclude, provide an outline of the chapters to follow.

1.2 Theories, frameworks, and safety zones

MOGUL (Modular On-line Growth and Use of Language) provides the researcher with a processing-based perspective on language in the individual.

Although the basic architecture of the model we propose is introduced in the earlier chapters of this book still in terms of the *monolingual* language user, its ultimate focus is the *bilingual* or *multilingual* mind, how it grows and how it operates in real time as well as how language interacts with other perceptual and cognitive systems.

Our proposals are best seen as constituting a theoretical *framework* rather than a 'model' or a 'theory' per se. This is a point that needs some elaboration here since we will assume an understanding of this difference in the chapters that follow. The difference between a theory and a theoretical framework, in the sense that we are using these terms, is in the degree of flexibility that either term implies. It also involves a hierarchical relationship with a framework providing a home for different, specialised but compatible accounts concerning specific areas of language and general cognitive ability all of which may be integrated to provide explanations that cover a larger range of phenomena.

The greater flexibility of a framework lies in the fact that it should accommodate, at least to some extent, different combinations of theoretical perspectives in order to increase its coverage and thereby, at least potentially, its explanatory power. The success of a multidisciplinary framework is measured by the extent to which it can provide better and wider-ranging explanations, in the case of language ability, than any of the separate theoretical accounts it draws on. It should facilitate the integration of different strands of research which, to some extent, have their own traditions and terminology. At the same time, there must be limits to the flexibility provided by a framework: there has to be a basic set of principles to which any candidate account for a particular aspect of the phenomena to be explained should conform. You could not, for instance, incorporate an account which relies on a denial of *any* innate properties of the mind if one of the underlying principles is that such properties must exist however they may be specified or constrained. A framework that is too vague and too unconstrained is of little use. The result is likely to be an eclectic mix-and-match one, mixing up essentially contradictory parts taken from diametrically opposed theoretical perspectives. Hence it is important to state in advance that MOGUL, if not a theory, is a notably 'biased' theoretical framework.

With the current expansion of interdisciplinary research within cognitive science, the time is ripe to build more ambitious frameworks. The immense benefits of focusing on very specific areas of concern and thereby excluding a range of external factors should be set against the advantages provided by a wider perspective. Although the constraints that, say, theoretical linguists or psycholinguists or neurolinguists have placed on the scope of their investigations are fully understandable and have enabled each group to make impressive progress in their particular research fields, there is nonetheless a role for those working in those areas to step outside their safety zones. It has to be said, of course, that paying lip service to the principle is one thing, implementing it

is quite another. In any event, it is important to find some way past awkward issues that have remained fuzzy and unresolved within a particular subarea of cognitive science and, more generally, to consider what particular research findings in one area say about the functioning of the mind in general. The reinterpretation, or richer interpretation of research findings taking into account other perspectives may then proceed. Take some very simple examples from second language research. The presence of a systematic non-native feature in a language learner's production, say, a missing inflectional 3rd person singular *-s* or past tense *-ed* in English, or a missing subject pronoun as in the sentence, *[1] 'Is not a good idea', might be interpreted in two ways. It might indicate a lack of knowledge of their native-like equivalents. Alternatively, the learner might be able in principle to produce the native equivalents but just has not managed to: their absence may then be attributed to processing overload. This second interpretation would, however, be given firmer support if a properly worked-out processing theory were resorted to (cf. Hawkins and Chan 1997; Lardiere 1998a). Facilitating such a match-up between research in different domains is the type of contribution provided by a theoretical framework such as the one to be proposed in this book.

A final note on framework relates to falsifiability. We wish hypotheses to be vulnerable, that is falsifiable at least in the sense that they generate falsifiable predictions which can be put to the test. If such predictions are not borne out, this at least should weaken the hypothesis and lead to some reformulation. At a higher level, hypotheses are linked up in the form of a more general explanation, i.e. a theory. Any such hypothesis which generates predictions that are not adequately supported by empirical findings will obviously pose a problem for the theory itself. This may not necessarily falsify the theory in one go but at least it will raise doubts which have to be dealt with. It may also be the case that two different hypotheses emanating from different theories may explain the same data. Again, a simple example is provided by the findings in early investigations of L2 English of native Spanish speakers where learners typically produce utterances with preverbal negation, not evidenced in native English (**no is good* instead of *is not good*). Some theories would expect the early stages of acquisition to be predominantly influenced by Spanish, i.e. by language transfer (e.g. Lado 1957). Dulay and Burt (1974), however, whose theoretical approach led to a minimising of the role of L1 wherever alternative explanations could be found, pointed out that children learning English as an L1 also produce preverbal negation even though they are not exposed to it. Whereas it might *look* like Spanish interference in L2 English, it could also be seen as reflecting a developmental stage that is common to both L1 and

[1] The asterisk * is a conventional way in linguistics of indicating a violation of native-speaker grammatical norms. It may still conform to the learner's current (developmental) grammar.

L2 learners. Or indeed both factors could be at work. Ambiguous results can only be resolved by placing them within wider interpretative frameworks.

MOGUL provides related examples of hypotheses that are inadequate if taken by themselves without any broader explanatory context. One of the claims to be put forward in the following chapters will be that a language grows or shrinks within an individual as a direct result of that individual's processing history: using a language causes it to grow or be maintained whereas not using it leads to changes in ability that are often viewed as loss or decline. This processing-based account of development obviates the need to posit some special language acquisition device, but there is a problem: several approaches that otherwise differ sharply from one another offer the same kind of explanation ranging from behaviourism, classical connectionism, emergentism through to MOGUL, only the last-mentioned approach assuming a domain-specific language ability. The claim that, irrespective of its precise nature, there exists a general principle that controls how all development proceeds cannot of itself distinguish between the various competing accounts. However, MOGUL embeds this principle, which we call 'acquisition by processing', within a theoretical framework whereby the general developmental mechanism is nevertheless constrained in very different ways according to the domain in which development is taking place. This makes it different from, for example, an emergentist account which sees a similar processing principle working in the same way across all areas of cognition (O'Grady 2003).

1.3 Terminological and conceptual traps

Some attention really needs to be paid at the outset to a few common terms that admit of different interpretations. As always, the academic terminology in use can be quite slippery. To start with, the term 'bilingual' is understood here in its broadest, most flexible sense to include two-year olds simultaneously acquiring two or more languages but may also refer to, say, adolescents or adults learning a foreign language at home or in a formal classroom after they have fully acquired their mother tongue. The essential point here is that human language ability involves, where external circumstances permit, and to different degrees, the acquisition and use of different language systems. This holds whether these systems are registers, accents, and dialects within what society perceives as a single language or whether they are associated with different languages. The question of whether, say, Mandarin, Hakka, and Cantonese are different languages sharing one writing system or rather dialects or varieties of a single language called Chinese, is dictated by society and not by the individual mind or brain. Secondly, the term 'bilingualism' will also be treated as equivalent to 'multilingualism' where the differences between 'two' and 'three' are not currently relevant to the discussion simply because

'bilingualism' has conventionally been used in this way for so long. Perhaps 'multilingualism' will soon become the default term. In any case, we shall use them both here as synonymous expressions. By the same token, the term 'second' in 'second language (L2) acquisition' will remain, as is standard in the literature, a convenient term standing for 'second or other' languages in whatever context they are acquired although L1a and L1b more accurately describe two languages that are acquired side by side from the start by the very young child. An L2 speaker/learner is a type of bilingual who has already acquired one language, typically called a native language or mother tongue and has acquired a new language (to some extent) either in the community which speaks that language or otherwise abroad, typically in a formal school setting. Again, a 'heritage language' speaker is another type of bilingual who has started acquiring one language at home as a child (the heritage language) but grows up in a community that speaks another language, and this second language has, for the child, actually become the dominant one. The heritage language is likely to differ somewhat from that of an equivalent monolingual speaker of that language (Montrul 2009). Again, the intended implication behind these generous uses of the terms 'bilingual' and 'second' is that it is natural and normal for a human being to operate more than one language system (cf. Cook 1992, 2007; Roeper 1999).

So much for terms directly related to language. There are also a small number of general terms referring to different approaches to understanding how the mind works, ones that need clarification because of their standard and, we argue, overrestrictive use. The first one is *connectionism*. The second one is *emergentism* and there are a couple of other terms deserving a brief comment here, terms associated with the general debate about how best to characterise human language ability, namely *innatist* and *nativist*. It is important to establish the meanings of these terms as used in the book right at the outset as their use evokes certain specific theoretical approaches which have to be distinguished from the actual concepts themselves.

Take connectionism, for example. Connectionists use the model of a neural network to explain how the mind works. At a general level, the idea of the mind as a network, or perhaps 'network of networks' must be uncontroversial. There is an ongoing debate between connectionists, in the common use of that term, who see symbolic representations as unnecessary for explaining how the mind works and classicists who insist that the mind is a representational system. Here, as in many places, by 'connectionists' is meant 'radical' or 'eliminative' connectionists. Here the claim is that the networks are composed of simple units. Each unit has a particular activation potential and when it is activated it triggers activation in units that are connected to it within the network. This is called spreading activation. Change in the system can be accounted for by changes in activation patterns or connection strength and without recourse to symbolic

representations. However there are also ways of integrating the two otherwise diametrically opposed approaches and so there are connectionists who see neural networks as a way of implementing a representational system (see Rey 1991; Gregg 2003b; Marcus 2003). Few people would deny at some general level that the mind is connected up and consists of networks within networks and most would also agree that experience has the effect of adjusting or adding to these networks. Frequently encountered patterns in the environment should somehow and to a greater or less extent weight the network differently than rarely encountered patterns. So, at that very general and arguably trivial level of environmental impact on the developing mind, all accounts are connectionist. The devil, of course, is in the detail.

Another term in common use is *emergentism,* a term describing the development of complex systems: new patterns emerge from old ones. The new ones are different than/greater than the sum of the patterns from which they have emerged. As a noted emergentist in the language acquisition literature, William O'Grady, has pointed out, there are also different kinds of emergentism. O'Grady calls his own approach 'general nativism' and works with symbolic representations where other emergentists do not. O'Grady's version of an acquisition device involves the interaction of (innate) general cognitive principles interacting with the linguistic data supplied by the environment so that mental grammars emerge without the need to posit special grammatical principles unique to language acquisition (O'Grady 2008). It will be seen in the following chapter that the MOGUL framework certainly reflects some characteristics in common with emergentism. Indeed, MacWhinney, an emergentist of a rather different kind to O'Grady, commented on the first MOGUL publication (Truscott and Sharwood Smith 2004) calling MOGUL a hybrid model. He concluded that since it appeared to adopt core emergentist assumptions familiar from his own approach (statistical tracking, working memory constraints, item-based grammar, and competition), it thereby became 'largely equivalent to the more fully elaborated Competition Model account' (MacWhinney 2006: 736). In the course of the discussion we hope to demonstrate convincingly that this is an overhasty conclusion and that a different view of the MOGUL framework will emerge. It is unsurprising that there is a tendency to recategorise new ideas in terms of older more established ones which is a fit test applied to new approaches but one that can be quite misleading. MOGUL is neither an old model dressed in new clothing nor is it a hybrid of two or more old models lumped together.

The fact of the matter is that any processing-based account, MOGUL being one of them, has to incorporate a coherent account of memory and memory change and of structures 'emerging'. This will necessarily involve basic notions such as interacting networks, weighting, and competition all of which are accepted features of cognitive processing and development in pretty much any

account of mental architecture whatever the theoretical allegiance. It would therefore be wrong to attribute the use of networks, competition, emergence etc. as a sign of allegiance to one or other of the particular approaches currently available that also use these notions in their own particular ways. So, while we would ally ourselves with what one might call the connectionist 'insight' and the emergentist 'insight' about the nature of mental operations and mental development, MOGUL capitalises on these insights in one particular way.

Finally, there is the term 'innate', that is what is biologically predetermined, and 'innatist' which, like 'nativist' is sometimes used pejoratively. All approaches to language acquisition that we are aware of assume some degree of innateness or nativeness. Even behaviourists would accept that the formation of stimulus–response associations is biologically determined. In this sense, all connectionists are 'innatist': emergentists are innatists, generative linguists are innatist and so on and so forth but this is not the way in which 'innatist' and 'nativist' are normally understood in this context. Therefore, by signing up to a family of theories that assume the existence, *in some form*, of a domain-specific language faculty, MOGUL can indeed be said to be unashamedly innatist, or nativist. Although the mechanisms that drive language use and language acquisition are at some fundamental level generic, the principles which a crucial subset of those mechanisms adhere to, having to do with the construction of grammars, are very much domain-specific, in other words they are not shared with other cognitive systems and are also unique to human beings. This will be discussed in detail, from the next chapter onwards. The basic point here is that most approaches assume competition as an essential feature of growth and are, albeit to differing degrees, connectionist, emergentist and nativist.

1.4 The pros and cons of compartmentalisation: SLA as a case in point

Explaining language and how it is acquired, lost, processed, and represented in the mind obviously involves a multitude of research domains. As we have acknowledged, to make exploration feasible and achieve appropriate levels of theoretical and experimental rigour, it has been necessary for investigators to develop models that isolate particular problem areas and select particular perspectives based on particular research cultures. This is, of course, a standard methodological strategy in science. The specialisation and compartmentalisation of language research has certainly helped to develop new fields of investigation. Our own specialisation, second language acquisition (SLA), provides a classic example of this.

SLA in the 1970s had to cut itself off from its applied linguistic, language teaching roots in order to delimit, and impose some order on its scope of enquiry. Attempting to solve a myriad of practical learning/teaching issues on

the basis of existing SLA research, and being constantly expected to do so, seemed highly premature and hence a counterproductive way of conceptualising the task of the SLA researcher. In order to assist the learning process, it was first necessary to strive towards a reasonable depth of understanding. Fledgling theories, although they might have interesting and challenging things to say about what teachers may have long taken for granted were nevertheless a rocky basis for establishing new teaching methods. Hatch's warning – 'apply with caution' – was an appropriate one (Hatch 1978) and language teaching practitioners, especially in those days, might have felt quite justified in following their intuitions and their experience until the field had matured and had more to offer. Second language researchers in the 1980s certainly felt justified in developing their field further without the requirement to continually explore implications of their research for language teaching.

As new, specialised fields of research become more established, it seems natural and important to become less inward-looking and look across the borders to neighbouring areas of inquiry. In generative linguistics, in line with Chomsky's principled insistence on the distinction between competence and performance, and latterly between I-language and E-language, the object of enquiry became manageable by virtue of the fact that a whole range of real-time (performance) phenomena could be excluded from consideration (Chomsky 1986). As generative linguistics developed however, it was possible, without sacrificing the original distinction and theoretical model with which it was associated, to begin to take some account of how children actually acquire languages and how people process language on-line. In other words, the real-time factor, both in terms of millisecond-by-millisecond processing and in terms of changes that take place over weeks and months, is becoming increasingly less peripheral. All this has happened against a background of widening perspectives. That is to say, there seems to be a renewed interest in making research more open to neighbouring disciplines and seeing particular fields like theoretical linguistics as part of a wider scientific enterprise, namely that commonly referred to as 'cognitive science'.

Despite the fact that SLA researchers have tried to impose boundaries on what they need to consider, in particular, excluding issues that relate directly to language teaching methodology, it has always been impossible to treat this domain as anything else but interdisciplinary. To investigate language learning phenomena you clearly have to have and apply a reasonably mature theory of learning and you do need to apply a good theory of language as well, and these two theories must inevitably be taken from different domains of enquiry, some branch of learning theory and some particular approach to the explanation and description of linguistic structure. The challenge for the emerging field of SLA at the end of the 1970s was to contain the object of their investigation within manageable limits, excluding, at least for the time being, all practical, 'applied'

issues, and then, in order to make it immediately even more manageable, contain it still further in practice by concentrating on one or two aspects of the object of enquiry, primarily the acquisition of morphosyntax. This further containment was influenced in great measure by the academic background of the researchers. In the eighties, SLA, for the first time, attracted an influx of researchers whose academic training was not in education but rather in (theoretical) generative linguistics. The immediate effect of this was to considerably refine the way in which the linguistic structure of learner utterances was described and analysed and the application of this particular theoretical framework, with all its variations and reformulations, has, for some considerable time, been yielding many insights into the nature of learner language. In a way, it has formed, to some people's frustration, the core of active research in SLA over the last two decades or more.

Despite considerable theoretical advances in the way the structure of L2 systems is described and analysed, development within SLA has not produced any general consensus position concerning learning theory. Psychological notions such as proceduralisation, automatisation, attention, and consciousness have certainly figured in the literature (see, for example, Robinson 2001) but where theoretical linguistic constructs have tended to play a dominating role, there has been little progress in explaining development over time. Describing states of learner language at different stages in a learner's career is a good start but it is also necessary to explain what the mechanisms are that actually move the learner on from one assumed stage to another (Gregg, 1996; Sharwood Smith, Truscott and Hawkins 2013). Theoretical linguistics, insofar as it abstracts away from real-time issues, cannot claim to be able to supply such an account. One can argue about whether or not there as yet exists an adequate, widely accepted theory of human learning in psychology and an adequate, widely accepted theory of cognitive processing. One can also consider whether language by its very nature might require separate, domain-specific learning and processing accounts. In any case, although interdisciplinary in principle, the generative linguistic core of SLA research rendered it relatively inward-looking and linguistics-oriented at least up until the early nineties and hence transitions from one state of knowledge (and/or ability) to a later state were not dealt with, only the properties of the separate states and how they differed, which is not the same thing. Since SLA has not come up with much in the way of a theory of development that actually specifies the precise properties of developmental mechanisms, the only option available to those interested in adopting a coherent developmental theory of language learning has long been, at least until recently, to drop the idea of language learning as a special case requiring its own theory together with any language acquisition research that relies on that assumption, and instead adopt a general learning theory that purports to explain all kinds of human learning, such as behaviourism, connectionism,