construction of the lexicon, and in the retrieval of information from that lexicon, may itself develop and, in effect, evolve. We can in fact only guess at what the initial organization of the lexicon might be, and how it may change as infant sensitivities themselves change. The goal of much current research into the acquisition and organization of the early lexicon is to map out these changes.

In the following chapter, we leave behind the question of what, if anything, constitutes the access code, and consider instead what ‘recognizing’ a word actually means. In terms of the analogy with searching through a written dictionary, at what point (in time) do we read off the information contained within the lexical entry, and what determines when we reach this point? In addressing this issue, we can remain agnostic with respect to the nature of the access code. At the end of the day, more important than the denomination of the notes is what you can buy with them.

Words, and how we (eventually) find them

The history of science is littered with examples of analogies that do not work. Often, they are simply inappropriate, simply wrong, or simply confusing. But even when inappropriate, they can prove useful. For instance, it is not unnatural to think of our knowledge about the words in our language as residing in some sort of dictionary. The Oxford English Dictionary (OED), all 20 volumes of it, is as good an example as any—its purpose is to provide, for each entry, a spelling, a pronunciation, one or more definitions, general knowledge about the word itself, and perhaps a quotation or two. Getting to this information is relatively efficient. You scan down the page, ignoring the definitions of the words you are uninterested in until you come to the word you want. Of course, if you are lucky enough to possess the dictionary on CD-ROM, you do not even need to scan down the page (let alone pull the appropriate volume off the shelf); just type in the word, and up pops everything you ever wanted to know about it. But so what? Why should anyone care about how the OED works, CD-ROM or no CD-ROM? In the last chapter we saw that the analogy between accessing a
written dictionary and accessing the mental lexicon is at best fragile. So why carry on with it? The answer to this is simply that it provides a useful starting point from which to proceed, using a vocabulary that is easily understood to describe a process (accessing the mental lexicon) that is easily misunderstood.

Describing the mental lexicon as 'something you've never imagined' is probably accurate, but certainly useless. At least a conventional dictionary can be imagined, and is therefore a useful place from which to start our exploration of the mental equivalence. Most importantly of all, the questions one can ask of a dictionary such as the *OED*, and the questions one can ask of the mental lexicon, are remarkably similar. The answers, though, can be surprisingly different.

Before delving into the mental lexicon and looking at how we retrieve words, we need first to address an important, if basic, question: what, exactly, are these things we call 'words'?

**Words and what they contain**

The purpose of language, and communication in general, is to convey meaning. In spoken language, the word is the smallest stand-alone thing that can do this. It is not, however, the smallest verbal gesture capable of expressing meaning. An 's' added onto the end of 'fact' also expresses meaning; namely that we are dealing with more than one fact. So words can generally be broken down into even finer units, called *morphemes*. This last sentence contains 13 words, but 21 morphemes—the word 'units', for instance, consists of the morpheme 'unit' and the morpheme 's', and the word 'morphemes' consists of 'morph'+'eme'+'s'. Figuring out which kinds of morpheme can be stuck onto which other kinds, and how this affects the meaning of the resulting word, has been studied within a branch of linguistics called *morphology*.

There are different kinds of morphemes: *stems* (e.g. 'unit', 'word', or 'speak') and affixes (e.g. '-s', or '-ing'). Languages differ in terms of where they put their affixes—in English, the most common affix is a suffix, coming at the end of the word, but we also have prefixes which come at the beginnings of words (e.g. the 'pre-' in 'preamature'). Some languages (for instance, Tagalog, the language of the Philippines) have infixes too; these are affixes that are inserted in the stem. Mark Aranoff, a morphologist, has an entire section in his book *Word formation in generative grammar* devoted to one of the few English infixes—'fucking' as in 'fan-fucking-tastic'. But affixes are not the only device we can use for modifying the meanings of words; the irregular past tense in English—'run–ran', 'speak–spoke', and so on—is a remnant of a time when the past tense was produced not by adding the suffix '-ed' onto words, but by modifying a vowel in the stem. In a Semitic language like Hebrew, this is the rule, rather than the exception.

To complicate matters further, different kinds of affix do very different things. Some affixes are called *influenational*; these include the plurals ('-s'), and the various ways of inflecting verbs (e.g. for 'govern': 'governs', 'governing', 'governed'). Inflectional affixes do not change the meaning of the word, but convey additional information relevant to it. Derivationale affixes, on the other hand, do change the meaning; they are used to derive new words, so from 'govern' we can get 'governor', 'government', 'governance', 'governable', 'ungovernable', 'governability', and so on. Although related in meaning, each of these words means something different, and in the case of 'governable' and 'ungovernable', they mean exactly the opposite. But not all derived words are related; 'casual' and 'casualty' are unrelated in meaning, as are 'depress' and 'express' (although a glance at the *OED* will reveal their common historical ancestry). Another complication (there are several more) is that although the majority of stems can be free-standing, there are some inflected words which do not contain a free-standing stem, such as the verbs 'permit' and 'submit'. And whereas the meaning of 'ungovernable' can be deduced by stripping away the affixes and recovering the meaning of the stem 'govern', the meaning of 'permit' and 'submit' cannot be deduced from the meaning of 'mit'.

So words are complicated things. And knowing whether it should be called a word or a morpheme, an affix or a stem, a prefix or a suffix, an inflection or a derivation, matters far less than knowing that words have an internal structure. Somehow we have to strip off the excess (but important) baggage, and reveal the word's core. And sometimes it looks as if we ought to do this, but in fact we should not (as in 'permit', 'report', and so on). Linguistics has told us an enormous amount about how words are structured and how the meaning of a word is dependent on the meanings of the different morphemes it is composed of; it tells us which kinds of morpheme can be combined with which other kinds, and in which order. But that is just the periodic table again—it tells us what the result is, but it does not really tell us how the result comes about. It does not explain how the brain comes to acquire the conventions that tell us that 'un-' on the front of a word results in a
meaning that is the contrary of whatever the word meant in the first place. It does not tell us where this knowledge is stored, or how it is stored, or how the brain takes a complex word like ‘unspeakable’ and breaks it down into its components, or even whether it does break it down at all. All these questions fall under the remit of psycholinguistics. And sadly, there are few definite answers, only hints. But of one thing we can be certain: at the root of each word is a meaning, and recovering this meaning is precisely what a dictionary is for.

Of course, we knew all along that words convey meaning, and that the mental lexicon is a store of word meanings. But so is the OED, and yet physically they could hardly be more different. But what, if any, are the consequences of the physical differences? The fact that they evidently are different does not mean that they are necessarily used any differently—for instance, the OED in book form could hardly be more different from the OED on CD-ROM, and yet there are aspects of their use which are common to both of them. Apart from the fact that the OED and the mental lexicon are physically different, what else is different about them?

**Accessing our dictionary**

We already know that not all dictionaries are the same, and that depending on which dictionary we use, we can access the words (and narrow down the search) on the basis of how they are spelled, how they are pronounced, what they rhyme with, what they look like, how long they are (as in a crossword dictionary), or even how frequently they occur in the language at large (as in some specialized dictionaries used by psycholinguists). But crucially, however we do it (see Chapter 5 for some discussion of this), it is an inevitable consequence of accessing the dictionary that we will encounter, during the search, other words that share certain features with the word we are ultimately interested in finding, whether they share their spelling, pronunciation, rhyme, shape, length, or frequency. It is in this respect that our intuitions about what we do with a written dictionary are quite at odds with what we actually do with our own mental lexicon. For instance, although we do not burden our minds with the definitions of the other words that we pass as we scan down the page of a written dictionary, the same is not true of the process by which we access the mental lexicon. We do burden our minds with the contents of the neighbouring words we encounter as we narrow down the search. Our intuition that we do not is wrong, and our expectation on the basis of what appears to be a similar process (using the OED) is also wrong. The challenge, of course, is to prove that these intuitions are wrong.

It seems somewhat unreasonable to access the meanings of the words ‘ram’ and ‘ramp’ simply because they are encountered during the search for ‘rampart’. It would be equally unreasonable for an Australian listening to the word ‘acoustic’ to access the entry for ‘acabra’ (a traditional Australian hat) just because they start off sounding the same, or for a naturalist to access the meaning of the word ‘pichiciago’ (a kind of armadillo) just because it starts off like ‘pitch’. It would surely make sense only to look up the definition of the word being looked for, as we do with written dictionaries, and not to look up the definitions of all the other words that just happen to overlap in their first few sounds. So why, when searching the mental lexicon, do we access the meanings of neighbouring words? And how can we, as psycholinguists, be so sure that this happens? As we shall see, it is unclear how things could possibly happen any other way.

During the 1980s, William Marslen-Wilson demonstrated that we can recognize a word even while it is still being heard (before, even, the speaker has finished saying it). We therefore access the lexical entry of a
word well before the corresponding physical stimulation has ceased (that is, before its acoustic offset). In one of the first demonstrations of this, people were asked to repeat aloud as quickly as they could what they heard over headphones (to shadow what they were listening to). Marslen-Wilson found that often they would start to vocalize a word before it had finished playing on the tape. This was not simply some blind repetition of the sounds they heard, because if the words were jumbled up so that they made no sense (‘up words jumbled he they so no sense the made’), people could no longer shadow as fast—so they were clearly interpreting what they were listening to, and were therefore recognizing the individual words before repeating them. In other experiments, he asked people to press a button as soon as they heard a particular word on a tape (word-monitoring, similar in spirit to the syllable-monitoring task mentioned in Chapter 5). He found that once you took into account the time it takes to decide to press a button, and the time it takes to press it, people were responding so fast that they must have been initiating their response well before the end of the word.

Marslen-Wilson found that the time it takes to recognize a word correlates very well with how much of the word has to be heard before it becomes uniquely distinguishable from all the other words in the language that share the same beginning. So ‘slander’ becomes uniquely distinguishable only when the /d/ is encountered. Before then, the input would be compatible with ‘slant’.

An important component of the account of lexical access developed by Marslen-Wilson is that the entries in the mental lexicon are not simply accessed, they are activated. The idea that information is activated has a long established history in psychology, although its application to word recognition became more widespread in the late 1960s following the work of John Morton, now Director of the Medical Research Council’s Child Development Unit in London. One way to think about this is to remember that ultimately, all the information in the mental lexicon is stored within the neural structures of the brain. When a pattern of light enters the eyes, or a sequence of sounds enters the ears, those stimuli do not access anything within the brain, even if they result in the recognition of, for instance, a politician speaking or a baby babbling (or both, if they are indistinguishable). Instead, the stimulation passes through the neural circuitry of the brain, being modified by, and in turn, stimulating (or activating) different parts of the circuit. Only certain kinds of stimulus will provide the right kind of stimulation for some particular part of the neural circuit—the stimulus is a key that can activate a part of the circuit, and depending on which part is activated, we experience ‘seeing a politician’ or ‘hearing a baby’. There will be more of this later, but for now, the important point is that nothing is accessed; it is activated. And although we might just as well continue to refer to lexical entries, we shall return later to the idea that the mental lexicon is in fact a collection of highly complex neural circuits.

So what has this to do with why we access/activate the meaning of anything but the intended word? Why does this suddenly make it reasonable to suppose that we start to activate words and their meanings even before they become uniquely distinguishable from their neighbours? The answer has to do with the quite reasonable assumption that sounds entering the auditory system (i.e. the ear and beyond) stimulate the neural circuitry as they enter the system—a sequence of sounds is much like the combination to a safe; the tumblers in a combination lock fall into place as the correct sequence of rotations is performed, without waiting until the sequence is complete. Similarly, those neural circuits which require a particular sequence of sounds (before a particular word is ‘experienced’) will become activated as that sequence enters the system. So the neural circuits that encode what we think of as lexical entries could quite reasonably become activated on the basis of a developing (but not yet completely developed) sequence of sounds—/slan/ would activate the neural circuits associated with (and hence would activate the meanings of) both ‘slander’ and ‘slant’. But so much for what is possible in principle. What actually happens? Where is the proof?

What we need is a way of establishing which meanings of a word have been activated, and when. The priming task (first mentioned in Chapter 1) does just this. The task here is to decide whether a word that has just appeared on a computer screen is a real word in their language (e.g. ‘broom’), or a nonword (e.g. ‘brom’). How long it takes people to make a response (a lexical decision response) depends on all sorts of things. Nonwords that are similar to real words take longer to say ‘no’ to than nonwords that are very different, and real words that are used infrequently take longer to say ‘yes’ to than words that are used frequently. But the recognition of a real word can also be faster if a related word has been seen beforehand—lexical decision times to ‘broom’ are faster following ‘witch’ than following the unrelated control word ‘pitch’. This effect is called priming; ‘witch’ (the prime) can prime ‘broom’ (the target), ‘doctor’ can prime ‘nurse’, ‘bug’ can prime ‘ant’, and so on. Activating the prime causes the target to be activated faster.
Conversely, if a target word is activated faster (primed), you can be sure that the priming word must have been activated.

In the mid-1980s, a student of William Marslen-Wilson’s, Pieien Zwitserlood, used a version of the priming task called cross-modal priming to explore when, during the sound sequence, words are activated. In cross-modal priming, the priming word is presented in the auditory modality, and the target is presented visually. Zwitserlood and Marslen-Wilson reasoned that if lexical entries are activated before the end of a word, and if this activation is the only one it takes to get priming to related words, it should be possible to find cross-modal priming effects when the visual target word is presented on the screen part way through the auditory presentation of the priming word. So people would hear only the first part of a word, and at the end of that part, a related word (or unrelated control word) would be flashed up on the screen. The actual experiment was performed in Dutch in The Netherlands, but it translates very easily into English.

Zwitserlood used the Dutch equivalent of ‘captain’, and played people a recording of this word up to and including the /t/. At this point, the sound stopped, and a related word (e.g. ‘ship’) appeared on the screen. Sure enough, she found priming—the word ‘ship’ was responded to faster than the corresponding control word. Of course, this simply shows that the lexical entry for ‘captain’ can be activated before the entire word has been heard. But the clever thing about this experiment was that ‘captain’ was not the only word compatible with the fragment played to people; ‘captive’ is just as good a continuation (the words can only be discriminated between on the basis of the final phoneme). And crucially, Zwitserlood also found priming to words related to these alternative continuations. In other words, the two alternatives that were compatible with the auditory input were both activated. And just to really prove the point, Zwitserlood demonstrated that if the visual targets were flashed up on the screen at the end of (and not part way through) ‘captain’ or ‘captive’, then only the related target words were primed—there would be no priming to words related to ‘captive’ when presented at the end of ‘captain’.

So it looks pretty cut-and-dry; as the acoustic input enters the system, we activate all the lexical entries compatible with the input that we have heard so far. This is exactly what Marslen-Wilson’s theory had predicted. And as the input becomes incompatible with certain alternatives, so those alternative entries begin to de-activate. But there was a further aspect to Zwitserlood’s experiments that was important.

Recall that one of the determinants of response times in lexical decision experiments is the frequency of occurrence of that word in the language at large; the more common words are responded to faster than the less common words. And although the more common words tend also to be the shorter words, it has been shown that this is not simply a length effect—once length is held constant, it is still the case that more frequent words appear to be recognized faster than less frequent words. The priming words used by Zwitserlood (‘captain’ and ‘captive’ being just one pair) did not have the same frequencies; one member of the pair was always more frequent. And what Zwitserlood observed was that there was generally more priming from the more frequent word than from the less frequent word. It is as if the lexical entries corresponding to the more frequent words become more strongly activated on the basis of similar acoustic input than their less frequent neighbours. Again, this had been predicted by the theory.

So the cross-modal priming studies show that we activate the entries for all possible words compatible with the acoustic input. But does this not mean that there is a real danger of system overload? How do we prevent an explosion of lexical possibilities? How do we choose which possibilities are the right ones? And how is it possible that we can activate the meanings of all these words without recognizing that we have done so? Apparently, activation does not imply recognition. But if it does not, what exactly is recognition? What does it mean to say that we recognize (or even hear) a word? And when does this recognition happen? To return to the OED metaphor: it looks as if access is what goes on when we activate the possibilities, and recognition is what happens when we (somehow) determine which of these possibilities is the right one. But how do we do that?

The effects of acoustic mismatch

According to Marslen-Wilson’s theory, lexical access is a kind of race; different lexical entries compete in the race, but there can be only one winner—we recognize a word when it has been identified as the winner. But for there to be a winner, there have to be losers. So what determines whether, and when, a competitor falls by the wayside?

The most obvious factor is compatibility with the acoustic input. There is extensive evidence showing that acoustic mismatch leads to a rapid decline in the activation of a lexical entry. Whereas a word like
describe how a vowel, for instance, can be 'coloured' by the consonants that precede and follow it. The fact that vowels are not perceived categorically allows this colouring to be used in anticipating the identity of the following segment. But something very similar can occur when one consonant is followed by another. And this is where the problems start: if the consonant were to actually change as a result of this process, a mismatch would occur. And this would mean that we would then fail to activate the intended meaning. Just how bad is the problem?

The answer is that it is as bad as having to recognize 'Hameetha-thimboo' as meaning 'Hand me that thin book'. Word-final consonants such as the /d/ in 'hand', the /t/ in 'that' and the /k/ in 'book' are often dropped completely. And instead of articulating the /n/ in 'thin' by closing off the mouth with the tip of the tongue against the back of the upper teeth (and allowing air through the nasal passage), the speaker might anticipate the following /b/ and instead close off the mouth at the lips (still allowing air through the nasal passage). This would result in 'thin book' being articulated as 'thim book'. And because the /d/ had been dropped from 'hand me', the preceding /n/ may combine with the /m/ to produce 'hamee'. These kinds of changes, generally at the ends of words, are surprisingly common, although the extent to which they occur, and how they occur, can depend on the language being spoken. But if acoustic mismatch leads to the deactivation of lexical candidates, what hope is there of recognizing the intended words after these changes have occurred? If these kinds of effects are more common than not, how could we ever recognize a sentence in its entirety?

The answer, once again, is tolerance. In this case, the tolerance is context-sensitive. The nonword 'thin' will activate the meaning associated with 'thin', but only in the context of a following word. But it cannot be just any old word, it has to be a word in the context of which it would have made sense for what was originally an /n/ to become an /m/. Whereas the 'thin' in 'thin book' would activate the lexical entry for 'thin', the 'thin' in 'thim slice' would not. This was demonstrated by another student of William Marslen-Wilson's, Gareth Gaskell, in a series of experiments using variations on the priming theme. This naturally begs the question of how the system 'knows' to do this.

Linguists have produced a whole range of rules which describe the range of circumstances in which these different kinds of word-final changes can occur. The rules are complex—the Cambridge encyclopedia of language writes one such rule as: 'an alveolar nasal becomes bilabial

'book' might prime 'page', the nonword 'boog' (pronounced to rhyme with 'book') would not—changing the voice onset time (see Chapter 3) of the final phoneme from a /k/ to a /g/ would be enough to cause rapid deactivation of the lexical entry for 'book'. But if the smallest deviation can lead to a decline in activation (and see Chapter 5 for further examples), what is going to happen each time we hear a word pronounced slightly differently, or each time a bit (or worse still, a lot) of background noise changes the acoustic signal? There has to be some tolerance in the system.

In fact, it turns out that there is; a slight deviation does not cause a lexical entry to self-destruct, it merely causes a decline in the activation, which means that the activation can pick up again if subsequent input is still compatible with that entry. Of course, if that deviation occurs at the start of a word, it may prevent the intended word from being activated in the first place. But it is not just any small deviation that leads to this; it is the smallest acoustic deviation that could in principle distinguish between one word in the language and another—in other words, the smallest detail that would cause one phoneme to be perceived as another. Indeed, the categorical perception of phonemes discussed in Chapter 3 is an example of how variation in the acoustic signal associated with a particular phoneme is tolerated up to a certain degree, beyond which any further variation causes the sound to be perceived quite differently.

In general, then, a word can be recognized when there has been sufficient mismatch between the acoustic input and that word's competitors. Often this will be before the word's acoustic offset, but sometimes it may be after. 'Ram' could continue as 'ramp' or 'rampart'. But if the sequence being heard was something like 'The ram roamed around', the lexical entries for 'ramp' and 'rampart' would become deactivated when 'roamed' was encountered, resulting in the eventual recognition of 'ram'.

So far so good. But there is one further detail that needs to be considered. Words are rarely spoken in isolation, but are spoken in the (seamless) context of other words coming before and after. And this is important for a number of reasons, not least because people are generally quite lazy in their articulation, and the position and shape of the articulators at any one moment reflects not simply the sound to be produced at that moment, but also the sound that will be produced next. We encountered a version of this phenomenon in Chapter 5 under the guise of co-articulation. Generally, the term is used to
before a following bilabial consonant. Yet despite their complexity, there has been a temptation to believe that (or at least to talk as if) the human mind runs these rules in reverse in order to recover what was originally meant. Do we really do this?

The simplest answer is 'not necessarily'. And one way to imagine what we might do instead is to recall that the task of the infant is to associate sounds with meaning. The infant must therefore associate not just /thin/ with the meaning of 'thin', but also /thim/ with 'thin', and even /thing/ with 'thin' (as in 'The thin carpet was worn through', where 'thin' would be pronounced as /thing/). But what is actually being associated with the meaning of 'thin' is not just the sound that has been heard, but rather the sound that has been heard within a particular context. This context necessarily includes the surrounding sounds. The infant might therefore associate with the meaning of 'thin' all the following: /thin/ in combination with a following /t/ (e.g. 'thin tree'), /thim/ in combination with a following /b/ ('thin book'), or /thing/ in combination with a following /k/ ('thin carpet', where /k/ is the first phoneme of 'carpet'). As an adult, it is then just a matter of recovering whatever meaning was associated with a particular combination of sounds.

Not surprisingly, many linguists have objected to this last possibility—it would require the infant/adult to store all possible pronunciations of each word in all possible contexts. There would be enormous numbers of combinations possible, and it would surely be much easier to simply acquire a very much smaller number of rules to do the same job, each rule applying across a whole range of words (for instance, one rule could apply to all words ending in /n/ when followed by a /b/, /p/, or /m/). Of course, in order to learn the rule, the infant would still have to be exposed to all the different pronunciations in all the different contexts, but at least it would not have to remember each combination of pronunciation and context. And if it did remember each such combination, how much context would be required? The following phoneme? The following word? The entire utterance?

On what basis can anyone reasonably claim that rules are not run in reverse, but that the infant/adult has knowledge of all possible pronunciations in all possible contexts? Surely this would require the most enormous memory space. The fact that the entire OED, which we would also expect to take up a lot of space, fits into something that is less than a millimetre thick and just a few centimetres across does suggest that unimaginably huge volumes of information are none the less manageable. But so what? The brain is hardly a CD-ROM, and its own memory capacity may well be limited, especially given the huge amounts of memory that would be required to store information about all the different pronunciations of all the different words in all their different contexts. But even if all this information could be stored, could it feasibly be learned and feasibly be deployed? And is acquiring, storing, and deploying this information more feasible than acquiring, storing, and deploying what by comparison would be a very small number of rules?

Currently, there is no empirical way to establish conclusively which of these two possibilities we actually use. A rule-based approach appears to have all the advantages stacked up in its favour—it is low on memory, and easily deployed. The alternative approach, based on some representation of the individual pronunciations and their associated contexts, would place an excessive burden on memory, and for all sorts of reasons it has a somewhat implausible ring to it. But then again, so did the idea that we activate the meanings of all the neighbouring words we encounter during the search for the actual word that was spoken. So plausibility is not necessarily a good criterion by which to choose between the possibilities. A more useful criterion concerns the feasibility of the alternatives. This is where, in the absence of data on what we actually do, computational modelling can shed some light on the puzzle. Unfortunately, this means waiting until Chapter 13, and in the meantime we must accept that, one way or another, we can overcome the problems associated with the mispronunciation of words uttered in the context of continuous speech.

Getting at the meaning

We know something about how, and when, lexical entries are activated, and how, and when, they may become deactivated. But what information is contained within a lexical entry? How do we square a question like this with the idea that a lexical entry is simply a kind of neural circuit? Returning to the analogy of a combination lock, we can ask the same kind of question: given the arrangement of its tumblers, what information does a combination lock contain? On the one hand, there is a sense in which a combination lock itself contains no information at all. It is simply a physical arrangement of potentially moveable objects. On the other hand, the precise arrangement of the tumblers determines which
exact sequence will open the lock—the appropriate sequence has meaning by virtue of causing an effect to occur that is specific to that sequence, and to no other. In this sense, the combination lock does contain information, and a skilled locksmith would be able to examine the arrangement of the tumblers, and figure out, on the basis of this information, the sequence required to open the lock. Similarly, even if a lexical entry is nothing more than the neural equivalent of a combination lock, it contains information by virtue of the effect that an input sequence can have (and in Chapter 9 we shall discuss further the nature of meaning, and the nature of the effects that a word may cause). And just as we can still refer to lexical entries when what we are really talking about is some complex neural circuitry, so we can refer to meaning when what we are really talking about is the result of this circuitry becoming activated.

So lexical entries are where the meaning of a word resides. But one of the first things one notices when opening up a large written dictionary is that most words have more than one meaning. The word 'pitch', for example, was introduced in Chapter 1 without any explicit definition. And yet it would be almost impossible to look up the word and not discover that it has several distinct senses or meanings: to pitch a ball; to pitch a tent; the pitch of a roof; the pitch of a musical sound; the pitch you get from distilling tar; the sales pitch; the football pitch, and so on. Presumably the mental lexicon must also reflect this multiplicity of meaning. But what are the implications for how we retrieve a single meaning? Do we activate all possible meanings of a word that is ambiguous and has more than one meaning? Do we somehow scan all the meanings (to return, momentarily, to the dictionary metaphor) until we get to the first one that is appropriate given the context in which the word occurs, ignoring any others that we have yet to get reach? Or do we somehow activate only the contextually appropriate meaning, so avoiding a cluttering of our minds with all those other, inappropriate, meanings?

In the late 1970s, David Swinney, then at Tufts University, published a paper that was to prove extremely influential. Not only did it demonstrate (after many years of bitter argument and counter-argument) that the alternative meanings of ambiguous words are activated, but it was also the first demonstration of cross-modal priming, which we encountered earlier. The specific question that Swinney considered was whether or not we activate the alternative meanings of words even when those words are heard in the context of sentences which are compatible with only one of the meanings of the word. For instance, in

'He swam across to the far side of the river and scrambled up the bank before running off', it would hardly be appropriate to interpret 'bank' as a financial institution. Similarly in 'He walked across to the far side of the street and held up the bank before running off', it would hardly be appropriate to interpret 'bank' as a river bank (or to interpret 'hold up' as 'support'). In order to explore what actually happens when we hear sentences such as these, Swinney played people sentences similar in principle to these ones and immediately after they heard the word 'bank', he flashed up on a screen either 'money' or 'river'. The people knew to make a lexical decision as soon as they saw a word appear on the screen. Swinney found that, irrespective of which of the two sentences had been used, both 'money' and 'river' were primed. This showed that both meanings of 'bank' must have been activated.

Of course, at some stage, the inappropriate meaning of 'bank' must be suppressed, and, sure enough, Swinney found that if he presented the target words two or three syllables later (that is, downstream from the ambiguous word), only the target related to the contextually appropriate sense of the word was primed.

These findings aroused an enormous amount of interest, not least because some subsequent studies failed to show the same results. These studies found that in context, only the appropriate meaning was activated. Of course, certain meanings of a word will be produced more (or less) often than certain others, and the evidence suggests that the more frequent the meaning, the greater its activation. This is entirely consistent with the idea, discussed earlier in connection with Zwitserlood's experiment, that the more frequent a word, the greater the activation of its lexical entry. If the institution meaning of 'bank' is more frequent than the river meaning of 'bank', the institution meaning will be the more active. And if, in these studies, the inappropriate meaning is sufficiently infrequent, it might look as if the contextually inappropriate meaning has not been activated, but what has in fact happened is that its activation was so low that it was very quickly deactivated by the context. So despite some initial controversy surrounding Swinney's results, the general consensus is that they are right—we do activate all meanings of an ambiguous word.

At around the same time that Swinney performed his cross-modal priming experiments, Michael Tanenhaus and colleagues at Wayne State University in Detroit performed a similar experiment, using words that were ambiguous between a noun (e.g. 'watch' as a time-piece) and a verb (e.g. 'watch' as a kind of looking). In a sentence like 'John began
to watch the game’, only a verb can follow the fragment ‘John began to . . .’. Armed with this information, we could scan a written dictionary and look only at the entry for ‘watch’ as verb, ignoring the entry for ‘watch’ as noun, and hence ignoring the time-piece meaning of ‘watch’. But does the same thing happen when we search the mental lexicon? Can we eliminate from the lexical search all the words whose syntactic categories are inappropriate given the preceding words in the sentence? Apparently not. Tanenhaus found that the alternative meanings of ‘watch’, related to time-piece and looking, were activated when people listened to sequences such as ‘John began to’. So knowledge of the type of word that must follow (that is, knowledge of its syntactic category) is not used to help constrain the possibilities. But why not? Is this some arbitrary property of the workings of the mental lexicon? Or is there some reason behind this?

Once again, perhaps we are too used to the convenient abbreviations that our written dictionaries provide us with; just because written dictionaries include the syntactic categories corresponding to the alternative meanings of each word does not mean that the mental lexicon does likewise. After all, the OED also includes the approximate date at which each word entered the language—but just because the OED includes this information, and can search for words on the basis of this information, does not mean that the mental lexicon does likewise. So perhaps the simplest interpretation of Tanenhaus’s ‘watch’ experiment is to suppose that syntactic categories are not listed separately as in a dictionary. Why should they be? If the syntactic category of a word is nothing more than a reflection of its meaning (and there will be more about this in Chapters 9 and 13), they will not be listed separately. The meaning would have to be activated before the syntactic inappropriateness could be judged. And when we hear an ambiguous word like ‘bank’, how could we judge which meanings were inappropriate if we did not activate them all?

It looks, again, as if all that psycholinguists have done is come up with the obvious (although it is probably true to say that hindsight makes the results of any scientific endeavour seem obvious). But, however obvious it may be, the view that has developed of how we access the mental lexicon is substantially different from the view we might originally have had on the basis of how we access a dictionary like the OED. Accessing the mental lexicon is far less restrained. We activate all the lexical entries compatible with the developing sequence of sounds entering the ear. If the same sequence of sounds has more than one meaning, we activate all the meanings compatible with that sequence and only subsequently are the contextually inappropriate meanings suppressed. Of course, an inevitable by-product of all this is that we must activate all manner of spurious, unintended meanings. But given that the mental lexicon must reside within the neural circuitry of the brain, this is in fact a natural (if initially counter-intuitive) way for the system to work—the neural circuits are like so many combination locks, and as the speech input unfolds through time, so do the tumblers of the different combination locks move around, until eventually, just those combination locks whose sequences are completed spring open.

Unlocking the combination

How far should the combination lock analogy be pushed? To return to the opening theme of this chapter, might this not be just another example of an analogy that is simply inappropriate, simply wrong, or simply confusing? Maybe. But the continued use of the analogy helps explain one further fact concerning the manner in which we access the mental lexicon. The fact itself concerns a prediction that can be made on the basis of the way a combination lock operates, although when the prediction was originally tested, combination locks were probably the furthest things from the mind of the tester. The prediction, quite simply, is that wherever in the speech input a sequence is found that could correspond to a word, the lexical entry for that word should be activated. After all, rotate the dial of an old-fashioned mechanical combination lock and so long as the sequence of rotations contains, somewhere within it, the correct sequence for that lock, the lock will open.

In the late 1980s, Richard Shillcock carried out an experiment to determine whether, for example, listeners activate the lexical entry corresponding to ‘bone’ if they hear the sequence ‘He carefully placed the trombone on the table’. He used the cross-modal priming paradigm described earlier in which a word related to ‘bone’ (e.g. ‘rib’) would be flashed up on a screen at the offset of ‘trombone’. Crucially, he also included sentences like ‘He carefully placed the bone on the table’. Shillcock found that the amount of priming he got from ‘trombone’ to ‘rib’ was the same as that from ‘bone’ to ‘rib’. In other words, the lexical entry corresponding to the word ‘bone’ is activated even when ‘bone’ is heard simply as part of the word ‘trombone’.
Subsequent studies (in Italian) have shown similar effects with sentences equivalent to ‘He broke all records for the new distance’. Here, the lexical entry corresponding to the word ‘nudist’ is activated (embedded in the sequence ‘new distance’). And although no one has tested for which lexical entries are activated on hearing ‘rampart’, it would be surprising if the lexical entries for ‘ram’, ‘amt’, ‘amp’, ‘ramp’, ‘part’, and ‘art’ were not activated in addition to that for ‘rampart’.

So it looks increasingly as if the process of lexical access is in fact rather simple: we consider all possible hypotheses about what could be present at any point in the incoming speech signal. But with this simplicity comes an obvious worry; surely there would be an explosion of possibilities? How do we determine which are the right ones? How do we know that we should be reading about ‘a new discovery’ and that nudists have nothing to do with it (even though the meaning of ‘nudist’ is activated)? Presumably, the answers are to be found in the manner by which we string the meanings of words together to give a coherent meaning to the entire sentence. Why we are not even conscious of all these spurious words is another matter. But, perhaps unfortunately, consciousness is beyond the remit of our ascent of Babel.

Psycholinguists have only brushed the tip of a theoretical iceberg. Many questions remain, and controversies abound. For instance, we are still unsure as to the nature of the bilingual lexicon. Only now are we developing theories concerning the processing of morphologically complex words—recall that many of the questions asked in that section of this chapter were never answered. And only now are we developing a better understanding of the tools we use, and the effects that the tools themselves can have on the phenomena we try to observe. But one thing is for certain: whatever the limits of our understanding, we now know not to trust whatever intuitions we may have had on the basis of the dictionaries on our bookshelves.