

Part I

Aims and Evidence



1

Welcome to Dictionopolis!

— The human word-store —

Before long they saw in the distance the towers and flags of Dictionopolis sparkling in the sunshine, and in a few moments they reached the great wall and stood at the gateway to the city.

‘A-H-H-H-R-R-E-M-M-’, roared the sentry, clearing his throat and snapping smartly to attention. ‘This is Dictionopolis, a happy kingdom, advantageously located in the Foothills of Confusion and caressed by gentle breezes from the Sea of Knowledge . . . Dictionopolis is the place where all the words in the world come from. They’re grown right here in our orchards.’

Norton Juster, *The Phantom Tollbooth*

‘Words glisten. Words irradiate exquisite splendour. Words carry magic and keep us spell-bound . . . Words are like glamorous bricks that constitute the fabric of any language . . . Words are like roses that make the environment fragrant,’ asserts the writer of a textbook urging people to improve their vocabulary.¹

Few people regard words with the awe and reverence of this author. Most of us use them all the time without thinking. Yet words are supremely important. Everyone needs them, and a normal person probably comes into contact with thousands in the course of a normal day. We would be quite lost without them: ‘I wanted to utter a word, but that word I cannot remember; and the bodiless thought will now return to the palace of shadows,’ said the Russian poet Mandelstam.²

The frustration of being without words is vividly expressed in Stevie Smith’s poem ‘In the park’:

‘Pray for the Mute who have no word to say.’
Cried the one old gentleman, ‘Not because they are dumb,
But they are weak. And the weak thoughts beating in the brain
Generate a sort of heat, yet cannot speak.
Thoughts that are bound without sound
In the tomb of the brain’s room, wound. Pray for the Mute.’

On a less poetic level, someone who has had a stroke can illustrate clearly the handicap suffered by those who just cannot think of the words they want. For example, K.C., a highly intelligent solicitor, was quite unable to remember the name of a box of matches: ‘Waitresses. Waitrixies. A backland and another bank. For bandicks er bandiks I think they are, I believe they’re zandicks, I’m sorry, but they’re called flitters landocks.’ He had equal difficulty when shown a telephone: ‘Ooh that, that sir. I can show you then what is a zapricks for the elencom, the elencom, with the pidland thing to the . . . and then each of the pidlands has an eye in, one, two, three, and so on’.³

Most people are convinced that they need to know a lot of words, and become worried if they cannot recall a word they want. Yet most of the time they will have relatively little difficulty in remembering the thousands of words needed for everyday conversation. This is a considerable feat.

However, speakers of a language are unlikely to have given much thought to this remarkable skill. Even those who deal with language professionally, such as speech therapists and teachers, know relatively little about how humans cope with all these words. Their lack of knowledge is not surprising since there is little information readily available about key issues, such as ‘How are words stored in the mind?’, ‘How do people find the words they want when they speak?’, ‘Do children remember words in the same way as adults?’, and so on.

This is the topic of this book. It will primarily consider how we store words in our mind, and how we retrieve them from this store when we need them. The overall aim is to produce outline specifications, as it were, for a working model of the word-store in the human mind. This turns out to be a huge subject. In order to narrow it down somewhat, the book will focus on the spoken words of people whose native language is English. English has been selected because, up till now, more work has been done on it than on any other language. And spoken speech has been chosen because native speakers of English talk it before they learn to read or write it. Reading, writing and other languages will therefore be mentioned only intermittently, when work on them illuminates the topic under discussion. The decision to concentrate on spoken English means that bilingualism and multilingualism are not directly discussed – though hopefully the findings will shed light on how people cope with the vocabulary of more than one language.

Mazes Intricate

Mazes intricate,
Eccentric, intervolved, yet regular
Then most, when most irregular they seem.

Milton’s description of the planets in *Paradise Lost*⁴ could apply equally well to the human word-store. Planets might appear to the untrained observer to wander randomly round the night sky, yet in fact their movements are under the control of

natural laws which are not obvious to the naked eye. Similarly, words are not just stacked higgledy-piggledy in our minds, like leaves on an autumn bonfire. Instead, they are organized into an intricate, interlocking system whose underlying principles can be discovered.

Words cannot be heaped up randomly in the mind for two reasons. First, there are so many of them. Second, they can be found so fast. Psychologists have shown that human memory is both flexible and extendable, provided that the information is structured.⁵ Random facts and figures are extremely difficult to remember, but enormous quantities of data can be remembered and utilized, as long as they are well organized.

However, to say that humans know ‘so many’ words and find them ‘so fast’ is somewhat vague. What number are we talking about? And what speed are we referring to? Let us briefly consider these two points.

Native speakers of a language almost certainly know more words than they imagine. Educated adults generally estimate their own vocabulary at only 1 to 10 per cent of the real level, it has been claimed.⁶ Most people behave somewhat like the rustics in Oliver Goldsmith’s poem ‘The Deserted Village’. The villagers gather round to listen in awe to the schoolmaster, whose verbal knowledge amazes them:

Words of learned length and thund’ring sound
Amazed the gazing rustics rang’d around,
And still they gaz’d, and still the wonder grew,
That one small head could carry all he knew.

While admiring the word power of their local schoolteacher, the rustics did not realize that the word-store within each one of their heads was probably almost as great as that of the teacher. Even highly educated people can make ludicrously low guesses. In the middle of the last century Dean Farrar, a respected intellectual, pronounced on the vocabulary of some peasants after eavesdropping on them as they chatted: ‘I once listened for a long time together to the conversation of three peasants who were gathering apples among the boughs of an orchard, and as far as I could conjecture, the whole number of words they used did not exceed a hundred.’⁷ They managed with this small number, he surmised, because ‘the same word was made to serve a multitude of purposes, and the same coarse expletives recurred with a horrible frequency in the place of every single part of speech’.

More recently, the French writer Georges Simenon was reported as saying that he tried to make his style as simple as possible because he had read somewhere that over half the people in France used no more than a total of 600 words.⁸ Simenon’s figure is perhaps as much the product of wishful thinking as his claim to have slept with 10,000 women in his life. At the very least one should probably exchange the numbers of words and women, though 10,000 words is still likely to be an underestimate.

An educated adult might well know more than 150,000 words, and be able to actively use 90 per cent of these, according to one calculation.⁹ This figure is

controversial, because of the problems of defining ‘word’ and the difficulty of finding a reliable procedure for assessing vocabulary knowledge. However, Seashore and Eckerson were pioneers of a method now widely used for measuring vocabulary size. It might be useful, therefore, to consider how they reached their conclusions, even though they are now thought to have overestimated the total, and their techniques have been subsequently modified.

Seashore and Eckerson defined a ‘word’ as an item listed in the 1937 edition of Funk and Wagnall’s *New Standard Dictionary of the English Language*, which contains approximately 450,000 entries. They reduced this to 370,000 by omitting alternative meanings. Of these, they reckoned that just under half, about 166,000, were ‘basic words’ such as *loyal*, and the remaining 204,000 or so were derivatives and compounds, such as *loyalism*, *loyalize*, *loyally* and *Loyal Legion*. Obviously it is impractical to test anyone on all the words in the dictionary, so a representative sample of the total needs to be obtained. The researchers did this by taking the third word down in the first column of every left-hand page. This gave a list of 1,320 words, which they divided into four. Several hundred college students were tested on their ability to define the words on each list and to use them in illustrative sentences.

Seashore and Eckerson found that their subjects were surprisingly knowledgeable. On average, the students knew 35 per cent of the common ‘basic words’ on the list, 1 per cent of the rare ‘basic words’ and 47 per cent of the derivatives and compounds. When these proportions were applied to the overall number of words in the whole dictionary, the average college student turned out to know approximately 58,000 common ‘basic words’, 1,700 rare ‘basic words’ and 96,000 derivatives and compounds. The overall total comes to over 150,000. The highest student score was almost 200,000, while even the lowest was over 100,000.

Later researchers have pointed out a number of flaws in Seashore and Eckerson’s methodology. The students might have been able to guess the meaning and use of derivatives from a knowledge of the ‘basic words’ to which they are related. Also, bright students tend to overestimate their knowledge. Take the word *kneehole*. This is the space under a desk for a person’s knees. Yet someone who was ‘quite sure’ he knew the word, suggested it was a hole worn by a person’s knee through thin fabric trousers. In contrast, less good pupils think they know words which are similar to others. When asked to use the word *burrow* in a sentence, one child wrote: ‘May I *burrow* your pencil?’, confusing it with *borrow*, and another: ‘You take away rubbish in a *wheelburrow*’, instead of *wheelbarrow*.

The ‘big dictionary effect’ is another problem: the bigger the dictionary used, the more words people are found to know, partly because bigger dictionaries include more homonyms (different words with the same form). The word *must* probably elicits the meaning ‘should, is obligated to’ (‘You *must* wash your hands’) in the mind of someone asked about it. Yet a dictionary sample might have picked on *must* ‘the newly pressed juice of grapes’, or even *must* ‘a state of frenzied sexual excitement in the males of large mammals, especially elephants’.

It's also difficult to know what level of knowledge is being tapped. One person claiming to know *aardvark* might think of it only as a strange wild animal, but another might be able to describe it as a nocturnal mammal with long ears and a snout which feeds on termites and inhabits the grasslands of Africa.¹⁰

In spite of these problems, self-assessment of a dictionary sample has turned out to be the best way of estimating vocabulary size, mainly because it allows a large number of words to be reviewed. The method has been refined somewhat since Seashore and Eckerson's pioneering work: non-words are normally included in the sample, in order to detect unreliable respondents. Different levels of list are tested, each controlled for the frequency of occurrence of the words selected. Students are no longer always asked to give a straight 'yes-no' answer to whether they know it, but can also reply 'maybe' if the word sounds vaguely familiar.¹¹

On the basis of this method, some conclusions are possible. An educated adult speaker of English can understand, and potentially use, at least 50,000 words, with a word provisionally defined as a 'dictionary entry'. Modern dictionaries usually include different forms of a word under the same entry, so *sing*, *sings*, *sang*, *sung* would all come under the headword *sing*. However, they normally provide separate entries for derivatives whose meaning cannot be reliably guessed, so *singer* would have an entry to itself, because it does not just mean 'someone who sings', but more usually 'someone who sings for a living'.

This guesstimate of 50,000 is based on informal tests with British English university students. But the total may be on the low side. The reading vocabulary of the average American high school graduate has been assessed at about 40,000 words,¹² with the total rising to 60,000 or perhaps even 80,000 if all the proper names of people and places and all the idiomatic expressions are also included.¹³ Only a few thousand of these words will be routinely used, but many more, such as *anteater*, *barometer*, *crustacean*, *derogatory*, can be understood or actively produced if required.

Compare these totals with the vocabulary of any of the 'talking apes', animals who have been taught a language-like system in which signs stand for words. The chimps Washoe and Nim actively used around 200 signs after several years of training, while Koko the gorilla supposedly used around 400. None of these animals approached the thousand mark, something which is normally achieved by children soon after the age of 2. And animals trained more recently, such as Lana (a female chimp) and Kanzi (a male bonobo) have an even more limited vocabulary, since they have been taught to manipulate pre-set symbols on a keyboard whose number does not exceed 200.¹⁴

In conclusion, the number of words which an educated adult native speaker of English knows, and can potentially use, is unlikely to be less than 50,000, and may be much higher. These high figures suggest that the mental lexicon is arranged on a systematic basis.

The second reason why words are likely to be well organized in the mind is that they can be located so fast, literally in a split second. This is apparent above all

from the speed of normal speech, in which six syllables a second, making three or more words, is fairly standard.¹⁵ And experiments have confirmed this figure, showing that native speakers can recognize a word of their language in 200 ms (milliseconds) or less from its onset, that is, approximately one-fifth of a second from its beginning.¹⁶ In many cases this is well before all the word has been heard. Indeed, the average duration of words used in the experiments was around 375 ms – almost twice as long as the recognition time. One way in which the researchers demonstrated this was by pointing to the behaviour of subjects in a ‘speech shadowing’ task. Shadowing is a fairly common technique in psycholinguistic experiments, and is reminiscent of simultaneous interpretation. The experimenter asks the subjects to wear headphones into which a stream of speech is played. Subjects are then asked to repeat what they hear as they hear it. People who are good at shadowing can repeat back speech with a delay of little more than 250–275 ms – around one-quarter of a second. If we assume that 50–75 ms is taken up with the actual response, and deduct this from the overall time taken, then we get the figure of 200 ms (one-fifth of a second) quoted above. These good shadowers are not just parroting back what they hear. They are genuinely ‘processing’ the words, since they correct mistakes, such as changing *tomorrance* to ‘tomorrow’.

The detection of non-words provides further evidence of fast and efficient word-searching ability. Subjects are able to reject a sound sequence which is a non-word in around half a second. This has been shown by means of a lexical decision task, an experiment in which subjects are asked to decide whether a sequence of sounds is a word of the language or not.¹⁷ Some of the sequences presented were real words, others non-words, such as *vleesidence*, *grankiment*, *smollite*. Subjects were asked to press a button as soon as they heard a non-word. They did this surprisingly fast, in just under half a second (450 ms) from the point at which the sound sequence diverged from being a possible real word. Once again, this suggests that speakers are able to conduct an orderly search through their mental word-store in a surprisingly short length of time.

Of course, the fact that speakers are usually able to distinguish fast between real words and non-words is something which we can also sometimes see happening for ourselves, as in the following extract from a short story, ‘De Bilbow’ by Brigid Brophy. Barney is questioned by his foreign girlfriend about the meaning of a word:

‘There is an English word I am not knowing. I am not finding it in the dictionary . . . “Bilbow”.’

‘Bilbow?’

‘Yes.’

‘There’s no such word. It’s a surname, not an ordinary word.’

‘Please? You are not knowing this English word?’

‘I AM knowing,’ Barney said. ‘I’m knowing damn well the word doesn’t exist.’

Note that Barney responded without hesitation. This is quite a feat. Suppose he knew 60,000 words. If he had checked through these one by one at the rate of

100 per second, it would have taken him ten minutes to discover that *bilbow* didn't exist. The problem sequence *bilbow*, incidentally, came from Shakespeare's *Henry V*,¹⁸ in a passage in which the French-speaking Katherine mispronounces the English word *elbow*.

Native speakers, then, seem able to carry out a thorough search of their word-store in well under a second, when they need to recognize a real word or reject a non-word. These figures relate to words that are clearly words and non-words that are unlike actual words, since most of us have a grey area of sequences such as *concision* which sound as if they might be 'real' words, but we are not quite sure.

Most humans are also impressively fast at finding the words they need when they produce speech. Unfortunately, we cannot time the production process as easily as we can measure recognition speed. Some researchers have made attempts in this direction by arguing that pauses in speech, which are measurable, often occur before major lexical items. They may therefore have been caused by word searching.¹⁹ However, the pauses vary in length, and their interpretation is controversial: we cannot easily tell whether a speaker is pausing to choose the words themselves or the order in which they will occur. So we cannot produce convincing figures for selection times, especially as some words seem to be easier to find than others.

Indeed, some words seem to be particularly hard to seek out. Almost everybody has had the annoying experience of not being able to think of the particular word they want, even though they are sure they know it. Yet such problems probably seem more frequent than they really are. Even when struggling to find a particular word, normal speakers have plenty of others at their disposal in order to carry on a reasonable conversation. This can be illustrated by a fictional but not unrealistic dialogue from Douglas Adams's science-fiction satire *Life, the Universe and Everything*.

Arthur shook his head in a sudden access of emotion and bewilderment.

'I haven't seen anyone for years,' he said, 'not anyone. I can hardly even remember how to speak. I keep forgetting words. I practise you see. I practise by talking to . . . talking to . . . what are those things people think you're mad if you talk to? Like George the Third.'

'Kings?' suggested Ford.

'No, no,' said Arthur. 'The things he used to talk to. We're surrounded by them for heaven's sake. I've planted hundreds myself. They all died. Trees! I practise by talking to trees.'

Arthur cannot remember the word *trees*. Yet while he struggles to retrieve it he uses approximately 50 other different words seemingly effortlessly, with no conscious searching. Such fast and efficient retrieval must be based on a structured system, not on random rummages around the mind.

Our conclusions so far, then, are as follows: the large number of words known by humans and the speed with which they can be located point to the existence of a highly organized mental lexicon.

However, the requirements of massive storage capacity and fast retrieval are not necessarily the same. This can be illustrated by an analogy. Suppose the words in the mental lexicon were like books. If we wanted to store thousands of books, how would we do this? The simplest method would be to find a large room and to stack them up in heaps which go from floor to ceiling. We would start at the side of the room opposite the door and carry on heaping them up until the room was quite full. Then we would shut the door. In this way we could store the maximum possible number of books. But suppose we then needed to consult one of them. How would we find it? We might never locate the book we wanted, unless it happened to be one of the few stored near the door.

In brief, the system which allowed the greatest storage capacity might not be compatible with efficient retrieval. And there might be further discrepancies between storage requirements and speedy retrieval. To continue with the book analogy, libraries often keep all really big and heavy books near the floor. But this means that they cannot be kept in strict sequence. Similarly, in the human mind, extra long words might need a specialized storage system which could separate them from shorter words, and which might cause some delay when it came to retrieving them.

In dealing with words in the mind, therefore, we must treat storage and retrieval as interlinked problems but not identical ones. Although common sense suggests that the human word-store is primarily organized to ensure fast and accurate retrieval, we cannot assume that this is inevitable. Humans might have adopted a compromise solution which is ideal neither for storage nor for retrieval.

Words in the Mind and Words in Books

The human word-store is often referred to as the 'mental dictionary' or, perhaps more commonly, as the *mental lexicon*, to use the Greek word for 'dictionary'. There is, however, relatively little similarity between the words in our minds and words in book dictionaries, even though the information will sometimes overlap. Let us therefore look at some of the differences between a human's mental dictionary and a book dictionary. The dissimilarities involve both organization and content.

With regard to organization, book dictionaries standardly list words in alphabetical order. As a first guess, one might suggest that the mental lexicon of someone who can read and write could also be organized in this way. After all, many of us spend a considerable amount of time looking things up alphabetically in telephone directories and indexes. So, one might assume that educated English speakers had set up their mental lexicons to fit in with their alphabetical expectations.

This is an easy hypothesis to test. People occasionally make mistakes when they speak, selecting one word in error for another. If the mental lexicon was organized in alphabetical order, one might expect speakers to accidentally pick an adjacent entry when making errors of this type. So, in place of the musical instrument 'zither' one would predict, perhaps, the wrong selection of *zit* 'a spot on the skin',

or *ziti* ‘pasta in the form of tubes resembling large macaroni’ which precede and follow *zither* in *The New Oxford Dictionary of English (NODE)*. Similarly, in error for the word ‘guitar’ one might expect someone to accidentally pick *guinea* or *guipure* or *guise*, or perhaps *guiver*, *Gujerati*, *gulch*, *gulden*, *gules*, *gulf*, all words which are near neighbours in *NODE*.

But mistakes of this type are quite unlikely, as becomes clear when we look at a few ‘slips of the tongue’, such as ‘He told a funny antidote’, with *antidote* instead of ‘anecdote’, or ‘The doctor listened to her chest with his periscope’, with *periscope* replacing ‘stethoscope’. These errors suggest that even if the mental lexicon turns out to be partially organized in terms of initial sounds, the order will certainly not be straightforwardly alphabetical. Other aspects of the word’s sound structure, such as its ending, its stress pattern and the stressed vowel, are all likely to play a role in the arrangement of words in the mind.

Furthermore, consider a speech error such as ‘The inhabitants of the car were unhurt’, where the speaker presumably meant to say *occupants* rather than ‘inhabitants’. Such mistakes show that, unlike book dictionaries, human mental dictionaries cannot be organized solely on the basis of sounds or spelling. Meaning must be taken into consideration as well, since humans fairly often confuse words with similar meanings, as in ‘Please hand me the tin-opener’ when the speaker wanted to crack a nut, so must have meant ‘nut-crackers’.

Arrangement in terms of meaning is found in some collections of synonyms, such as *Roget’s Thesaurus*, but not generally in book dictionaries, where a desire to be neat and tidy in an alphabetical fashion may outweigh other considerations. For example, the word *horsehair* occurs soon after *horse* in the *Encarta World English Dictionary (EWED)*, but there is no mention of it near the entry *hair*. Similarly, *workhorse* occurs soon after the entry for *work*, but does not appear with *horse*. In brief, the organization of the mental lexicon is likely to be considerably more complex than that of book dictionaries, for whom orderliness is a prime requirement.

As for content, a book dictionary contains a fixed number of words which can be counted. Book dictionaries are therefore inescapably outdated, because language is constantly changing, and vocabulary fastest of all. As the eighteenth-century lexicographer Samuel Johnson pointed out in the preface to his famous *Dictionary of the English Language* (1755): ‘No dictionary of a living tongue can ever be perfect, since while it is hastening to publication, some words are budding, and some fading away.’ Everyone must at times have been frustrated to find occasions when a book dictionary concentrates on an archaic meaning of a word or omits a moderately common item. *COD* (7th edn, 1982), for example, defined *buzz* only in terms of sound. It did not mention its more recent and perhaps equally frequent meaning in the 1980s of ‘a thrill, a euphoric sensation’ until almost a decade later (8th edn, 1990). Or take the word *wimp*, meaning ‘a weak ineffectual person’. This was a vogue word in the early 1980s, as in the ‘lonely hearts’ ad ‘Wimp needs bossy lady’ (*Time Out*, July 1984), or the comment by a singing group that ‘the trying-hard wimps’ were an easy target for humour (*Guardian*, July 1984), or the magazine column which noted that ‘your cad, pale-faced wimp, Byron with

malnutrition, Little Boy Lost . . . have a great appeal for women since they are vulnerable' (*Cosmopolitan*, July 1984). Its adjectives were also widespread: a Sunday newspaper referred to 'the wimpish young schoolmaster' (*Mail on Sunday*, May 1982), and a women's magazine called attention to a calendar featuring 'six most decidedly wimpy males in varying states of undress' (*Over 21*, August 1984). Yet *wimp*-words were slow to find their way into British book dictionaries. The *Oxford English Dictionary* Supplement (1987) finally included them, and showed that they had been around for decades: *wimp* (first occurrence 1920), *wimpish* (1925), *wimpy* (1967), and *wimpishness* (1978). Meanwhile, the teenage vogue greeting *whassup*, from 'what's up?', is still absent from many dictionaries.

The way in which written dictionaries dodder along behind language is amusingly satirized in Douglas Adams's *Life, the Universe and Everything*:

The mattress globbered. This is the noise made by a live, swamp-dwelling mattress that is deeply moved by a story of human tragedy. The word can also, according to 'The Ultra-Complete Maximegalon Dictionary of Every Language Ever', mean the noise that is made by the Lord High Sanvalvwag of Hollop on discovering that he has forgotten his wife's birthday for the second year running. Since there was only ever one Lord High Sanvalvwag of Hollop, and he never married, the word is only ever used in a negative or speculative sense, and there is an ever-increasing body of opinion which holds that 'The Ultra-Complete Maximegalon Dictionary' is not worth the fleet of lorries it takes to cart its microstored edition around in. Strangely enough, the dictionary omits the word 'floopyly', which simply means 'in the manner of something which is floopy'.

– though this judgment on dictionaries is now somewhat unfair. In the last decade dictionaries have narrowed the gap between the first occurrence of a word and its appearance in print, due to the development of computerized databases, which themselves are based on the electronic scanning of recent material. A new word can now make it into a dictionary within months.

Turning to the mental lexicon, its content is by no means fixed. People add new words all the time, as well as altering the pronunciation and meaning of existing ones. Humans, however, do not just add on words from time to time, in between utterances. They often create new words and new meanings for words from moment to moment, while speech is in progress. A caller asking an American telephone operator about long-distance charges was told: 'You'll have to ask a zero.' The caller had no difficulty in interpreting this as 'a person you can reach on the telephone by dialling zero'. Similarly, it was not difficult for native speakers to guess that 'The newsboy porched the newspaper yesterday' meant 'The newsboy left the newspaper in the porch', or that the instruction 'Please do a Napoleon for the camera' meant posing with one hand tucked inside the jacket, as in most pictures of Napoleon, even though they had probably never come across these usages before.²⁰

In the examples above, the speakers and hearers were already familiar with other uses of the word *zero* and *porch* and with the characteristics of a famous character

such as Napoleon. They simply reapplied this knowledge in a new way. But human creativity goes beyond this. Quite often, totally new lexical items can be created and interpreted on the spur of the moment. This skill has been tested experimentally.²¹ The researchers gave a short description of a somewhat eccentric imaginary character to a number of students: ‘Imagine that a friend of yours has told you about his neighbour, Elvis Edmunds. Elvis loves to entertain his children in the evenings with several magic tricks that he knows. He often surprises them by pulling dollar bills out of his ear. During the day, Elvis is employed as a professional skywriter. He likes to work best on days when there is not a cloud in the sky. To supplement his income, Elvis carves fruit into exotic shapes for the delicatessen down the road.’ The students were then quizzed about the meaning of the phrase ‘doing an Elvis’ in various contexts, a task they found easy. They were confident, for example, that a sentence they could not possibly have heard before, such as ‘I have often thought about doing an Elvis Edmunds to some apples I bought’, meant ‘carving apples into exotic shapes’. The fluidity and flexibility of the mental lexicon, then, contrasts strongly with the fixed vocabulary of any book, or even an electronic dictionary.

But the biggest difference between a book dictionary and the mental lexicon is that the latter contains far, far more information about each entry. All book dictionaries are inevitably limited in the amount they contain, just because it would be quite impracticable to include all possible data about each word. In any case it is unlikely that anyone has ever assembled the total range of knowledge which could be brought together about any one dictionary entry. As one linguist notes: ‘There is no known limit to the amount of detailed information . . . which may be associated with a lexical item. Existing dictionaries, even large ones, specify lexical items only incompletely.’²²

For example, one popular dictionary suggests that the verb *paint* means ‘cover surface of (object) with paint’. But ‘If you knock over the paint bucket, thereby covering the surface of the floor with paint, you have not thereby painted the floor’.²³ Nor can one patch up the dictionary definition by suggesting that one must intentionally cover something with paint: ‘For consider that when Michelangelo dipped his brush into Cerulian Blue, he thereby covered the surface of his brush with paint and did so with the primary intention that his brush should be covered with paint in consequence of his having so dipped it. But MICHELANGELO WAS NOT, FOR ALL THAT, PAINTING HIS PAINTBRUSH.’²⁴ All this suggests that people have a much more detailed knowledge of the meaning of words than any book dictionary would have the space to specify.

Furthermore, why don’t people wear rancid socks? Or find fetid milk? Nothing suggests that this is abnormal in *EWED*. It defines the word *rancid* as ‘having the strong disagreeable smell or taste of decomposing fats or oils’ and *fetid* as ‘with a rotten or offensive smell’. This suggests that one ought to be able to attach both words to bad eggs or cow dung or even dirty socks. Yet it would sound very odd to say ‘Alphonse was ashamed of his rancid socks’ or ‘Mary’s egg was fetid.’ Written dictionaries list only a small selection of the range of words with which a lexical

item can occur. As one lexicographer comments: ‘The world’s largest data bank of examples in context is dwarfed by the collection we all carry around subconsciously in our heads.’²⁵

Moreover, in book dictionaries, words are mostly dealt with in isolation. A *child* is defined by *NODE* as ‘a young human being below the age of full physical development’. But this fails to inform us how the word *child* relates to all the other words for young human beings, such as *baby*, *infant*, *toddler*, *youngster*. Similarly, *NODE* tells us that *warm* means ‘of or at a fairly or comfortably high temperature’. Yet in order to fully understand *warm*, one needs to know how it slots into the range of temperature words such as *cold*, *tepid*, *hot*. This type of information seems to be an intrinsic part of one’s mental lexicon.

To continue this list of differences between words in the mind and words in books, a book dictionary tends to give information that is spuriously cut and dried. It is likely to tell you that *pelicans*, *sparrows*, *parrots* and *flamingos* are all birds, but will not rank them in any way. Humans seems able to judge that a sparrow is a more ‘birdy’ bird than a pelican or a flamingo. Or more likely, a human would say that a pelican is a ‘funny’ kind of bird. In addition, book dictionaries do not often spare the space to comment on frequency of usage. There is no indication in *NODE* that *abode* is less usual than *house*, or that *coney* is uncommon beside *rabbit*. People, on the other hand, seem well aware of which words are rare and which not. Recently, however, a number of dictionaries have started to include this kind of information.²⁶

Or, to take another facet of words, book dictionaries contain only a very small amount of data about the syntactic patterns into which each word can slot. *Wide* and *main* are both classified as adjectives in *NODE*. But it does not tell us that you can say ‘The road is wide’ but not ‘The road is main.’ Both *eat* and *resemble* are classified as verbs which take an object, but it does not tell us that whereas ‘A cow was eaten by my aunt’ is possible, ‘A cow was resembled by my aunt’ is not.

If we move on to consider how book dictionaries cope with sounds, we note that they normally specify only one pronunciation for each word. Yet native speakers of a language are likely to be able to understand quite different pronunciations by different speakers. In addition, they are likely to have more than one pronunciation in their own repertoire, depending on the formality of the occasion and how fast they are speaking. Sometimes, for example, one might pronounce a word such as *handbag* with all the sounds found in the conventional spelling, and at other times it might sound like ‘hambag’.

The examples listed above could be multiplied. They show that the mental lexicon is indeed a mammoth structure.²⁷ The relationship between a book dictionary and the human mental lexicon may be somewhat like the link between a tourist pamphlet advertising a seaside resort and the resort itself. A tourist pamphlet gives us a small, partial glimpse of a place as it was at some point in the past, with no real idea of how the different parts of the resort fit together to form a whole, living town. Similarly, a book dictionary gives us a spuriously neat, static and incomplete view of the mental lexicon.²⁸

The differences between words in the mind and words in books are therefore profound. The point of this section, however, was not to point out the shortcomings of book dictionaries, which serve a useful, though limited, purpose, enabling us to check on the conventional spelling of a word and to find out its approximate meaning. Indeed, if we did expect a book dictionary to include the same information as the mental lexicon, then we would undoubtedly require a fleet of lorries to cart its microstored edition around in, like the ‘Ultra-Complete Maximegalon Dictionary’ mentioned earlier. The comparison between mental dictionaries and book dictionaries was made in order to show that we cannot deduce much about our mental lexicons from studying the way words are dealt with in books. In the next two chapters, therefore, we shall consider how we can build up a more adequate picture of the human word-store.

Summary

In this chapter we have noted that humans know tens of thousands of words, most of which they can locate in a fraction of a second. Such huge numbers, and such efficiency in finding those required, suggest that these words are carefully organized, not just stacked in random heaps. This book will discuss the storage and retrieval of words by both adults and children, bearing in mind that a system which is ideal for storage might not necessarily be the best for fast retrieval. Its overall aim is to provide outline specifications for a working model of the human word-store.

We further noted that most ordinary dictionaries are limited in scope in comparison with the mental lexicon. Their organization is oversimple, their content is fixed and outdated, and they contain only a relatively small amount of information about each item – even though the situation is improving, as new sophisticated electronic databases are increasingly being used as the basis for recent dictionaries.

The mental lexicon, then, is both large and complex. In the next two chapters we shall consider how we should set about studying this mammoth structure.