

Chapter 2

Modelling the Lexicon: Some General Considerations

2.1 Introduction

In order to understand current research on multilingual processing and to appreciate the role early theories on the mental lexicon have had in shaping this new area of inquiry, it is useful to go back in time and examine some of the most influential works published over the years. Notably, the present chapter tackles problems which do not have a specifically multilingual focus; however, in research “relating to the multilingual mental lexicon the same kinds of organizational and operational issues arise as in L1-focused research” (Singleton 1999, p. 83), the difference being that in the case of L2 they are further complicated by questions having to do with precisely the fact that more than one language comes into the picture. Therefore, what is said in the present chapter with respect to L1 lexical processing is also relevant to L2 or, by extension, to Ln.

In the following subsections, an attempt will be made to explore different aspects of the multifaceted concept of the mental lexicon. The chapter begins with the discussion of the hypotheses referring to the internal structure of the lexical entries. Issues to be examined include, inter alia, the type of the stored information as well as the way this information is organized within an entry. Next, the chapter goes on to discuss the wider issue of the domain of the lexicon. It offers a brief presentation of different definitions of the mental lexicon. Presently, it reviews the most influential monolingual models of lexical processing to be discussed within two broader theoretical frames of reference—the modularity theory and connectionism. The discussion of the numerous lexical access models will be supplemented with some research evidence that the models seek to account for.

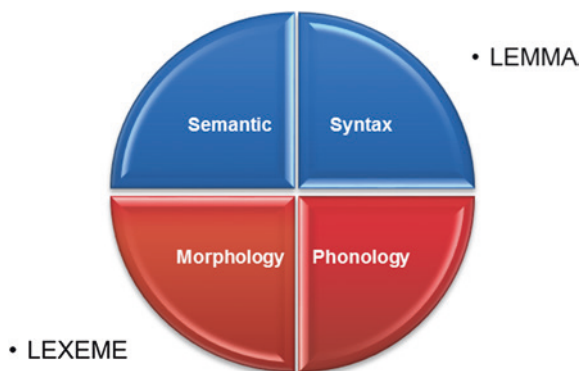
2.2 The Internal Composition of a Lexical Item

The mental lexicon includes a large number of lexical entries containing all the information on individual words. But what precisely are these individual words and what do they consist of? Within the psycholinguistic tradition two propositions concerning the issue of the internal structure of the lexical entry merit attention; the first, whose most fervent supporter is Levelt (1989, 1993) and the other, put forward by Bierwisch and Schreuder (1992).

Many linguists (cf. Aitchison 2003a, 2012; Levelt 1989, 1993) support the view that all the information “behind” a word can be allocated to two separate components: a semantic component called a *lemma*¹ (including the information on the word’s meaning, its connotations, style, and its syntactic pattern) and a formal one, frequently referred to as a *lexeme* (including the word’s morphology, phonology and orthography). According to Levelt, each lemma has a lexical pointer which “indicates an address where the corresponding word-form/information is stored” (Levelt 1989, p. 165). Levelt enumerates four main internal features of a lexical item (cf. Fig. 2.1): meaning, defined as the semantic information which lists “a set of conceptual conditions that must be fulfilled in the message for the item to become selected” (Levelt 1989, p. 165), syntax (including syntactic arguments and other properties), morphology and phonology. He also points to some stylistic, pragmatic and affective attributes of a word.

Another influential perception of the internal structure of a lexical entry was proposed by Bierwisch and Schreuder (1992). There are many similarities between Levelt’s and Bierwisch and Schreuder’s models since both approaches argue in favour of similar internal features of the lexical entry. Bierwisch and Schreuder enumerate the phonetic form, the grammatical form, the argument structure and the semantic form. There is, however, one fundamental difference between the models.

Fig. 2.1 The structure of lexical representations in the mental lexicon (adaptation based on Levelt 1989)



¹ The term *lemma* was first used by Kempen and Huijbers (1983) in reference to the part of the lexical entry which relates to its meaning and syntax.

Namely, the manner in which the meaning is represented. The key question is whether semantic representations of items are identical with general world knowledge or whether it is possible to draw a direct line between word meanings and concepts which represent encyclopedic information. Similarly to other psycholinguists and unlike Levelt, who is a fervent proponent of the holistic approach to meaning representation, Bierwisch and Schreuder (1992) believe that the internal structure of the lexical meaning of the entry is a composition of more primitive units. In contrast, Levelt asserts that the meaning of lexical items is “represented as a whole which cannot be decomposed into separate elements” (Levelt 1993, p. 28).

The findings so far, then, are that although there exists a general agreement as to the constituents of a word, two conflicting approaches concerning the issue of the representation of meaning are still under discussion. The first approach, called the one-level model or the network model (cf. Levelt 1989), considers semantic and conceptual knowledge identical. The other approach, known also as the two-level model (cf. Bierwisch and Schreuder 1992), differentiates between a word’s semantic meaning and the more general conceptual knowledge the item refers to. It needs to be noted that this latter theory relates to the proposition advocating the modularity of mind (cf. Fodor 1983, 1989). In the light of this approach a human linguistic system forms a closed mental module which does not depend on other mental faculties but is to some extent interconnected with them. Hence a lexical item is connected with the more general conceptual domain of world knowledge.²

All in all, the commonly adopted approach to the structure of a lexical entry and to the representation of meaning is the compromise between the two presented options. On the one hand, this approach draws a boundary between semantic and conceptual knowledge and conceives of them as non-identical. On the other hand, it admits that these two types of knowledge are strongly related (cf. Aitchison 2003a, 2012; Pickering and Garrod 2013; Randall 2007).

2.3 Towards the Model of the Mental Lexicon

The following sections pertain to a number of issues concerning the mental lexicon. Firstly, an attempt has been made to review various definitions of the phenomenon in question from a diachronic perspective. Subsequent sections relate to the internal organization of the mental lexicon; to be more precise to the actual number of storage systems. In brief, questions arise how many lexicons are to be found in the brain and how the semantic and formal (morpho-phonological and orthographical) components of lexical entries are stored. Are they stored together in a unitary modality-neutral lexicon or rather separately within two different modality-specific lexicons, and, if so, are there any direct links between the two?

² The modularity theory will be further discussed in Sect. 2.6.1.

2.3.1 *The Mental Lexicon Defined*

The term mental lexicon was introduced by Oldfield in 1966 (Oldfield 1966; in Singleton 1999) and since then it has been the focus of attention of a number of psycholinguists all over the world. It has been researched and re-defined from various perspectives many times. One of the early definitions was proposed by Fay and Cutler who attempted to describe the mental lexicon in terms of the lexicon metaphor as “the listing of words in the head” (1977, p. 509). The evidence they cite to support their claim demonstrates that the majority of words, excluding onomatopoeias, are characterized by arbitrary sound-meaning relations. Fay and Cutler (1977, pp. 508–509) offer the following description of the mental lexicon:

What is this mental dictionary, or lexicon, like? We can conceive of it as similar to a printed dictionary, that is, as consisting of pairings of meanings with sound representations. A printed dictionary has listed at each entry a pronunciation of the word and its definition in terms of other words. In a similar fashion, the mental lexicon must represent at least some aspects of the meaning of the word, although surely not in the same way as does a printed dictionary; likewise, it must include information about the pronunciation of the word although, again, probably not in the same form as an ordinary dictionary.

While some linguists compare the mental lexicon to a written dictionary, others describe it as a network of interconnected nodes similar to bundles of neurons in the brain. Aitchison (2003a, p. 248) rightly argues that “the mental lexicon is (...) concerned above all with links, not locations” and observes that “the lexical connections in the mind are far from what we normally imagine a dictionary or lexicon to be”. When a word is activated, other words of similar form (Stamer and Vitevitch 2012), meaning (Mirman 2011), syntax (Kim and Lai 2012), orthography (Carreiras et al. 2013) or emotional content (Bayer et al. 2012) are also activated, suggesting that the mental lexicon is complex and highly interconnected. Emmorey and Fromkin (1988) propose to view the mental lexicon as

that component of grammar in which information about individual words and/or morphemes is entered, i.e. what a speaker/hearer of a language knows about the form of the entry (its phonology), its structured complexity (its morphology), its meaning (its semantic representation), and its combinatorial properties (its syntactic, categorical properties) (...) also orthographical or spelling representation (Emmorey and Fromkin 1988; in Gabrys-Barker 2005, p. 38).

According to Singleton (1999), the mental lexicon is a module in human long-term memory which contains the speaker’s all knowledge concerning words in his or her language(s). Marslen-Wilson rightly describes the mental lexicon as “the central link in language processing” (1992, p. 9). In Levelt it is argued that the speaker’s mental lexicon is a “repository of declarative knowledge about the words of his language” (1989, p. 182). For the purpose of this work, however, a much more recent definition by Roux (2013, p. 82) seems suitable, which sees the mental lexicon as “the collective representation of words in the mind, which draws together contextual, personal and interpersonal dimensions of meaning, and assists most fundamentally in the acquisition, retention and expression of language.”

Before anything more is said about the structure of the mental lexicon, it is imperative to realize that finding common patterns in language errors is believed to provide valuable information about the nature of the internal lexical storage system. Thus, error analysis constitutes a basis and seems to be a perfect source of data in research on language processing (cf. Fromkin 1973). Admittedly, errors in any language system have an incalculable explanatory value. The evidence from word searches and “slips of the tongue”, selection errors known as malapropisms, but also psycholinguistic experiments and research with aphasic patients, show that lexical items in the mental lexicon are interconnected in a wide variety of ways.

Fay and Cutler (1977) based their model of the mental lexicon on malapropisms (cf. Vitevitch 1997; Goldrick et al. 2010). These are speech or writing errors “in which a word similar in sound to the intended one is uttered as in *The cold is being exasperated by the wind* instead of *The cold is being exacerbated by the wind*” (Aitchison 2003b, p. 71). However, there are three basic conditions an erroneous word needs to meet in order to function as a malapropism. Firstly, the meaning of the error and the target word needs to be unrelated. Secondly, erroneous intrusion should sound similar to the intended word. Thus, using *tattoo* instead of *book* cannot be classified as a malapropism; whereas substituting *tattoo* for *taboo* would be. Lastly, the word becomes a malapropism providing it has the so-called recognized meaning in the user’s language. Consequently, coining a non-existent or ungrammatical word by adding some affixes does not make a word a malapropism. Moreover, Fay and Cutler claim that

(...) the malapropisms, have some interesting properties. First, the target and the error are of the same grammatical category in 99 % of the cases. Second, the target and the error frequently have the same number of syllables (87 % agreement in our list). Third, they almost always have the same stress pattern (98 % agreement) (Fay and Cutler 1977, pp. 507–508).

On the basis of their findings, Fay and Cutler proposed a model which assumes that lexical storage is phonologically governed. The mental lexicon is conceived of as a network which “lists entries that have similar phonological properties near each other” (Fay and Cutler 1977, p. 512). Recapitulating, words beginning with the same phoneme are listed together, whereas words sharing the same second phoneme are grouped in a subcategory of that class and so on. It needs to be added that Fay and Cutler do not exclude the possibility of arrangement by syntactic category. Nevertheless, they do not provide any further details of such a concept.

Admittedly, there are considerable similarities between a traditional dictionary and the human mental lexicon. They are both organized along some underlying principles based on the characteristics that words share. Clearly, in the case of a written dictionary, the basic criterion of organization is orthography. Words in a book lexicon are always stored in alphabetical order. Consequently, if we want to look up a word we need to identify its initial letter, find words beginning with that letter and, finally, again in alphabetical order, exhaust the possibilities until the right entry has been found. Locating the word enables us to gain access to all the related data hidden “behind”—the semantic, phonetic, and pragmatic information.

Similar to a dictionary, the mental lexicon is comprised of a substantial number of lexical entries with linguistic information “behind” them, the complexity of the storage, however, being far more sophisticated.

In the first place, the lexical entries in a traditional dictionary are static, whereas the mental dictionary is dynamic. Not only do languages evolve constantly, but the individual linguistic knowledge of a language speaker also changes over time (cf. Aitchison 2003a). Consequently, the mental representations change—new meanings are added, while words which are rarely or never used become inaccessible. Another critical difference between a tangible dictionary and the mental one is the accessibility of the information being stored. In a book dictionary we can easily get equal access to any of the chosen entries. By contrast, words stored in our mind have different degrees of accessibility. It is argued that frequency of use, context and imageability³ are the most common factors influencing the accessibility of a given word. A further, but concurrently the most radical, difference is the form of the stored information. A written dictionary is simply an inventory of verbal information. The dictionary in the human brain, on the other hand, includes both verbal—linguistic and non-verbal—conceptual data. Schreuder and Flores d’Arcais (1989) describe this characteristic feature of the human mental lexicon in the following way:

A word in the mental lexicon has, besides its lexical properties, nonverbal percepts, conceptual representations and images that are derived from “real-life experience” and are stored in episodic memory (Schreuder and Flores d’Arcais 1989, p. 422).

As Bakhtin (1981) formulates it, “every word smells of the context (...) in which it has lived its intense social life” (Bakhtin 1981; in Gass and Selinker 1994, p. 276). In communication, language users depend on the contexts in which words appear to a significant degree, inferring word senses on the basis of linguistic as well as non-linguistic data, the latter being frequently even more informative.

However influential the lexicon metaphor can be, many cognitive psychologists and psycholinguists reject it claiming that the mental lexicon is much more than just a repository of lexical items. The advocates of the cognitive approach posit that the mental lexicon consists of concepts and their linguistic realizations, both phonological and orthographic. They conceive of it as a conceptual system. As Gabryś-Barker puts it,

A mental lexicon should be seen more as a conceptual system than a pure inventory of entries, a system which is composed of concepts and their linguistic realisations both phonological and orthographic, and with strong emphasis put on lexical processing (...) that is to say, access and retrieval as evidence of the working structure of the mental lexicon (Gabryś-Barker 2005, p. 39).

Notably, the standard position in language processing is that the mental lexicon is a largely fixed resource, acquired during early development. Although people can of course add new lexical entries during their adult life, this is generally seen

³ According to Aitchison, imageability is “the extent to which something can be visualized” (2003b, p. 57).

as a marginal activity. Studies of processing assume that people already know the language that they use and that there is a clear demarcation between acquisition and processing. (cf. Aitchison 2003a, 2012; Cutler 2005). In addition, the lexicon is treated as a store that principally consists of small units (either words or morphemes) and that knowledge of larger units is largely limited to idioms, which are regarded as fairly peripheral to “core” language processing.

More recently, Pickering and Garrod (2013) proposed an alternative view of the mental lexicon that is consistent with the Dynamic Systems Theory (cf. Briggs and Peat 1989). They based their proposition on the evidence from dialogues which shows that interlocutors make use of fixed or semi-fixed expressions during a particular conversation with meanings that are established through that conversation. They also argued that language users “routinize” (Pickering and Garrod 2005, p. 87) these expressions by storing them in the mental lexicon, normally for that conversation alone. This requires a conception of the lexicon in which complex expressions (of all kinds, not just established idioms) can be stored alongside more traditional lexical units. On this view, the lexicon can be constantly and dynamically updated, and the strict division between acquisition and adult usage is removed.

The final paragraphs of this section attempt to shed some light on the research concerning the much debatable problem of the size of the mental lexicon. It is generally believed that the mental lexicon is comprised of a huge number of lexical entries; however, its exact size remains undefined. In the early research conducted by Seashore and Eckerson in 1940 (in Aitchison 2012) the number of words stored in the mental lexicon of an educated adult was estimated at about 150 thousand receptive words with 90 % available for production. A similar study carried out by Diller in 1978 resulted in an unpredictably high number of about 250 thousand words, whereas the more recent work by Levelt (1989) rated the productive vocabulary of an educated adult at no more than 30 thousand word families. According to Clark (1993), on the other hand, adult speakers of a language have at their disposal between 20 and 50 thousand productive words, the amount of receptive vocabulary being “considerably larger” (1993, p. 13). All things considered, average educated adult language users have at their disposal a production vocabulary of between 20 and 50 thousand words and comprehension vocabulary of between 150 and 250 thousand words.

Why are the research results so diverse? Many linguists postulate that such sharp differences are connected with the failure to distinguish between productive and receptive vocabulary. Consequently, different experiments employ either active vocabulary exclusively or involve both passive and active words. Some researchers concentrate on active vocabulary (thus achieving lower numbers); other experimenters use both passive and active words (those used exclusively for comprehension and those for both comprehension and production). Another typically cited explanation for such a discrepancy in the results are various, not infrequently, incompatible methodologies. Nonetheless, whatever the answer to the question about the amount of mental word-stores, the actual number seems to exert little impact on the way the lexicon functions.

2.3.2 *The Internal Organization of the Lexicon*

Turning to the internal organization of the lexicon, the number of components of the human word store is a complex issue which is far from being settled. There are many models and the number of components of the mental lexicon they distinguish markedly vary. Some scholars (cf. Carroll 1994) apply the term *mental lexicon* to mean only the semantic sub-lexicon. Others (cf. Garman 1990) distinguish between the semantic lexicon and the phonological one. Alternatively, there are models which disregard the word's orthographic representation and instead concentrate on two levels called semantic and phonological sub-lexicons (cf. Levelt 1989; Aitchison 2003a, 2012). On the other hand, many psycholinguists perceive the orthographic representation an inseparable component of a lexical item. Consequently, in their models of the mental word-store they describe two modality-specific phonological and orthographic components within the formal layer of the lexicon (cf. Emmorey and Fromkin 1988; Randall 2007; Fernández and Smith Cairns 2011). The validity of the latter type of models has been proved by experiments involving priming effects of different modalities on word production and recognition (cf. Harley 2004).

It is widely agreed that the semantic and formal components of a lexical item are not stored together. Aitchison (1987, 2003a, 2012), Levelt (1989), Garman (1990), or more recently Randall (2007) and Fernández and Smith Cairns (2011), all agree that the semantic aspects of a word are located in one layer and the information on the formal aspects is kept in a separate part of the word-store. The two levels, however, are assumed to be connected by a wide net of direct links. A common argument supporting this view refers to the tip-of-the-tongue phenomenon in which the meaning of a word and a number of its syntactic properties are available for the speaker but the word's form cannot be retrieved (cf. Ecke 2009; Ecke and Hall 2013). By comparison, in his mental lexicon model Levelt (1989) adopted a twofold lemma vs. lexeme distinction to the entire lexicon, thus creating two separate stores: a lemma lexicon containing lemmas and a form lexicon comprised of morpho-phonological forms. Clearly, this division has only a metaphorical function which is to show that the internal organization of the mental lexicon is twofold, according to the meaning of items, as well as according to their morpho-phonological features.

The still debatable problem of the number of lexicons coincides with the issue of the modality of input and output. Are there two modality-specific lexicons or do we use the same store both while reading and listening? Undoubtedly, the advantage of the former assumption is the economy of storage, its drawback being the expense of retrieval in contrast to the latter proposal characterized by the simpler retrieval at the expense of complex storage. In short, the model which allowed for the maximum storage capacity might, at the same time, invalidate the most efficient retrieval. However, as Aitchison observes,

In dealing with words in the mind (...) we must treat storage and retrieval as interlinked problems (...). 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 (Aitchison 2012, p. 10).

With regard to organization, then, Fay and Cutler (1977) believe that there is one single mental lexicon for both production and comprehension instead of two separate lexicons. This assumption has been based on the analysis of common speech errors such as malapropisms or slips of the tongue. By contrast, Garman's model (1990) accounts for the existence of two separate specialized stores: one for generating and one for identifying words. Here the evidence supporting this view comes, above all, from neuropsychological research which has proved a number of discrepancies between comprehension of spoken and written input and production of spoken and written output. According to Ellis and Young's model (1988, 1996), on the other hand, there is one semantic lexicon incorporating four-modality specific interconnected sub-lexicons.

2.3.3 *The Internal Relations Within the Lexicon*

The structure of the lexicon is not the only debatable issue concerning the human word-store. Equally controversial is the matter of the relations within the mental lexicon. A highly advanced classification of various internal connections occurring in the mental lexicon was proposed by Levelt (1989), who distinguished between intrinsic and associative links. Intrinsic relations occur when items are linked through at least one component of the fourfold information on a word—meaning, morphology, syntactic category or phonology. Associative relations, on the other hand, hold between words which show no direct semantic, phonological or morphological links, but which frequently co-occur in speech or writing.

Lexical items can be intrinsically related through their meaning. A word is linked with its hyperonym (*banana—fruit*), co-hyponyms (*banana—apple*), near-synonyms (*wide—broad*), antonyms (*wide—narrow*) etc. All these interrelated links form a network called a semantic field. Another form of intrinsic links are morphologically-determined relations between derivatives of one item, which simultaneously share some semantic features (e.g., *govern, government, governmental, governor*). Evidence supporting the existence of such types of relations between individual lexical entries, again, comes from the analysis of speech errors. Fay and Cutler (1977) and much later Fikkert (2005, 2007) point to yet another type of intrinsic relation—the one based on phonological features which may be responsible for substitution errors such as the already discussed malapropisms. The authors claim that “words with the same initial or final segments seem to be connected as they cause errors in speech production such as *week* for *work*” (Fay and Cutler 1977, p. 514). Finally, there is some evidence on syntactically conditioned connections between entries coming from research with aphasic patients who have lost access to the entire class of words (cf. Haverkort 2005).

The second type of connections between entries in the mental lexicon are the associative relations. This kind of link occurs between entries which do not share any semantic, phonological or morphological features but which tend to co-occur in language use. The existence of associative relations has been evidenced in a

variety of experiments using different methodologies, the most common of them being priming tests (cf. Carr and Dagenbach 1990; Kroll and Sunderman 2003; Dijkstra 2005; Dijkstra et al. 2010). It is argued that if a word is found to prime another, then the words could be closely connected in the mind (cf. Aitchison 2012). Another group of experiments employed to support the existence of associative links in the mental lexicon encompasses association tests.⁴

2.3.4 *Lexical Storage: The Full Listing Hypothesis Versus the Decompositional Hypothesis*

One of the most hotly disputed controversies connected with the mental lexicon seems to be the issue of whether words are stored as whole units or as roots plus affixes. The following paragraphs address two fundamental questions concerning the lexical storage of polymorphemic words. Prior to the presentation of two influential hypotheses seeking to explain the storage of morphologically complex words, a question addressing the problem of what precisely is stored will be tackled.

The issue of lexical storage is strongly related to the phenomenon of word primitives which are commonly defined as the smallest meaningful elements stored in the mental lexicon. For many decades linguists have tried to determine how words consisting of more than one morpheme (e.g., *government*) are stored within the lexicon. Are they stored as independent units, or, as many linguists have suggested, are complex words decomposed into their constituent elements (e.g., *govern* and *ment*), which would support the morphemic organization of the lexicon? Depending on the perception of word primitives, linguists advocate in favour of one of the following theories.

The Full Listing Hypothesis was first proposed by Butterworth in 1983 and since then it has gained a number of supporters (cf. Henderson et al. 1994). In the light of this theory, derivations are stored, similarly to a written dictionary, as separate, independent entries (e.g., *go* and *goer* are stored as independent units). Consequently, both for comprehension and production they are accessed separately. In the light of more recent studies (cf. Vigliocco and Hartsuiker 2005), the only advantage of this hypothesis seems to be the so-called access efficiency.

The alternative proposition, known also as the *Decompositional Hypothesis*, has gained far more advocates (cf. Levelt 1989; Taft 2004; Frost and Ziegler 2007) and for this reason the idea of the morphemically-governed organization of the lexicon will be elaborated further in what constitutes the final paragraphs of the present section.

In the *Decompositional Theory*, words are seen as bundles of morphemes, and since morphemes are believed to be the smallest meaningful units of

⁴ For a thorough discussion of this methodology see Gabryś-Barker (2005), Fitzpatrick (2007) and Roux (2013).

language, consequently, the smallest element to be stored is no longer a word but a morpheme. Morphemes are typically ascribed to one of two categories: free morphemes (functioning as independent words) and bound morphemes (all sorts of meaningful affixes which do not, however, function independently and which require the accompaniment of a free morpheme, thus changing its meaning and generating a new word). In the light of this hypothesis, to produce a morphologically complex word (also called a polymorphemic word) separate morphemes need to be accessed and subsequently melded into one unit (which may be very elaborate at times; e.g., *anti-dis-establish-ment-arianism* constitutes six morphemes). Similarly, on encountering a polymorphemic word our brain needs to decompose it into separate morphemes to be accessed individually.

Critics of the *Decompositional Theory* (cf. Bozic et al. 2013) point, among others, to the problem of the lengthening of the recognition time that the hypothesis would need to endorse. Undoubtedly, due to the fact that in the case of complex words many more units would have to be accessed, additional processing would be inevitable. As a consequence, the amount of time necessary to access a complex word would be much longer. On the other hand, scholars supporting the *Decompositional Hypothesis* (cf. Levelt 1989; Taft 2004; Frost and Ziegler 2007) postulate that its obvious advantage seems to be the economy of storage. Morphemic organization ensures that there is no redundancy in the representation of related words created by using either derivational (e.g., *trusty*, *distrust*, *untrustworthy*) or inflectional (e.g., *jumps*, *jumped*, *jumping*) morphemes.

It needs to be stressed that there is a fair amount of experimental evidence supporting the hypothesis under discussion. The literature on the topic abounds with data coming from priming tasks, lexical decision tasks, spoken error analysis or experiments with brain-damaged subjects, in particular those suffering from Broca's aphasia. For instance, in priming tasks responses to a simple word (*hunt*) are speeded by a prior presentation of a related word (*hunter*), suggesting that these words share some entries in the mental lexicon (cf. Reichle and Perfetti 2003; Rossell et al. 2001; Dijkstra et al. 2005). Moreover, many experiments (cf. Garrod 2006) have also confirmed that the priming effect accompanying the morphologically related word pairs is stronger than that for word pairs overlapping in exclusively orthographic (*planet* vs. *plan*) or semantic form (*imitate* vs. *copy*).

Another source of evidence in favour of the morphemically-governed organization of the lexicon are the lexical decision tasks in which words are mixed with nonwords (pseudowords). The oft-cited experiments measuring the reaction time (RT) show that the longer the word (i.e. the more morphemes it has), the longer the reaction time; in other words, the more time we need to decompose it to understand the meaning of the constituent parts and to evaluate their validity (cf. Taft 1981; Reid and Marslen-Wilson 2003, 2007; Marslen-Wilson and Tyler 2007).

A further example utilizing the RT paradigm to support the *Decompositional Theory* comes from the research conducted by MacKay concerning morphological processing in language production (cf. MacKay 1978). When a group of participants were asked to derive nouns (such as *government*, *existence*, *decision* etc.) out

of the aurally presented verbs (*govern*, *exist*, *decide*), MacKay noticed that RTs varied significantly depending on the “complexity” of the derivation. And thus, *government* was identified as the fastest item (no phonological changes), *existence* was slower (resyllabification), while *decision* turned out to be the slowest (two phonetic changes). The interpretation put forward by MacKay strongly supports the *Decompositional Theory*. He claims that the results confirm the assumption that subjects are able to make those changes when producing morphologically complex words, which means that such words are not stored simply as independent units. However, the popular criticism of the research was linked to the very form of the experiment. Namely, the participants were explicitly instructed to derive morphologically complex nouns from a presented list of verbs. The task itself required derivational processing which may not occur normally. Thus, many linguists find MacKay’s research unreliable.

A similar lexical decision task paradigm was adopted by Taft and Forster (1975) and Taft (1981), who worked on the “prefix-stripping” strategy in word recognition. They concluded that in lexical decision tasks words with prefixes had greater RTs than words without them. Thus, *remind* (prefix *re-*) was identified faster than *relish* (“pseudoprefix”). The proposed interpretation of the obtained results is as follows:

Morphological processor automatically “strips off” anything that looks like a prefix (e.g., “RE”), then searches for the base in the lexicon. With words like REMIND, it will find MIND (real word), but with words like RELISH, will not find *LISH and will have to restart the search for the whole string (Taft and Forster 1975, pp. 642–643).

It needs to be stressed, however, that the presented results and Taft and Forster’s interpretation were also rejected by many who, as in the previous case, criticized the methodology of the conducted experiment. The opponents claimed that although the participants were not instructed to strip off prefixes, “maybe they were implicitly told this by the kind of word list they got” (cf. Rubin et al. 1979, p. 760).

Aitchison, who is also an ardent advocate of the decompositional approach, to support her theory uses error analysis of a spoken discourse—he example she gives is *She wash upped the dishes* instead of *She washed up the dishes* (Aitchison 2003a, p. 65). In her interpretation, the error may be suggestive of the organization of the internal lexical storage. In other words, the error has been committed since the brain has accessed the preposition *up* instead of the verb *wash*. The researcher believes that such errors verify the *Decompositional Hypothesis* and prove the fact that words are stored as morphemes. To generate this sentence our brain needs to access the verb *wash*, the preposition *up* and the past tense morpheme *-ed*. If derivations were stored as in the *Full Listing Hypothesis*, such an error would not occur—our brain would store the word *washed* as a separate item. We would not have to go into the process of building a word; instead, the already prepared items would wait to be accessed.

Yet another source of evidence supporting the hypothesis under discussion are the results of experiments with patients suffering from Broca’s aphasia (cf. Tyler et al. 1995). On the basis of these and other findings, models of word recognition were created which typically include a processing stage in which complex words are split into their constituent morphemes before meaning-based representations are accessed.

It has been generally agreed upon that some morphologically complex words share their lexical entries with the related forms. Nevertheless, the question of precisely which complex words are stored as morphemic units remains unanswered. It is worth mentioning that many researchers (e.g., Levelt 1989) emphasize the difference between lexical entries and lexical items. Levelt postulates that not all lexical items constitute separate lexical entries. And thus, inflections are items belonging to one single entry (e.g., *going*, *goes*, *gone* are all to be included under *go*). Derivations, on the other hand, are to be treated as separate entries (e.g., *goer*). The presented assumption has been confirmed by some experiments showing that the decomposition into morphemes is typically stronger for words composed of inflectional suffixes than for those formed with derivational endings (cf. Stanners et al. 1979; Chialant and Caramazza 1995; Blevins 2004).

All in all, hypotheses of morpheme processing are classified in relevance to the type of explanation they offer for the identification of polymorphemic words. Proponents of the *Decompositional Theory* claim that the meaning of a complex word is composed of its constituent morphemes (cf. Taft and Forster 1975; MacKay 1978). From this perspective, the meaning of *schoolbooks* would be created by first identifying the word's components (e.g., *school+book+s*) and then accumulating its meaning from these components. Conversely, supporters of the *Full Listing Theory* argue that complex words are stored and represented as independent units (cf. Rubin et al. 1979; Butterworth 1983; Henderson et al. 1994). In the light of the latter theory, the word *schoolbooks* is stored as a single entity, with individual representations for their components: *school* and *books*. Moreover, even the singular form of the word, *schoolbook* has its separate representation.

To conclude, the understanding of how an adult native speaker/hearer processes inflected word forms has increased considerably over the last decade. Experimental studies using a range of different psycholinguistic methods and techniques, e.g., lexical decision or priming, have led to a number of consistent and replicable results, e.g., frequency effects for inflected word forms in lexical decision tasks or priming effects for inflected word forms in different kinds of priming experiments. To account for the theoretical interpretation of these and other results on morphological processing in adult native speakers, many hybrid theories, including both separated and compositional representations, have been proposed (cf. Marslen-Wilson and Tyler 1980; Caramazza et al. 1988; Taft 1994; Clahsen et al. 2003, Marslen-Wilson 2007). In these theories, “complex words are identified via a ‘race’ between compositional and whole-word lookup processes” (Reichle and Perfetti 2003, p. 227). A good example of the hybrid theories are *dual-mechanism models* which hold that morphologically complex word forms can be processed both associatively, i.e. through stored full-form representations and by rules that decompose or parse inflected word forms into morphological constituents (Chialant and Caramazza 1995; Clahsen 2006; Bozic and Marslen-Wilson 2010; Bozic et al. 2013). In brief, hybrid theories, highlight that some words are more prone to decomposition than others.

2.4 Theories of Semantic Representation

In this section the problem of conceptual representation of meaning will be discussed in relation to the storage of conceptual features and their retrieval from memory. The key question here is whether semantic representations of words are identical with general world knowledge or whether it is possible to draw a line between word meanings and concepts which represent encyclopedic information. In brief, this section addresses the problem of the representation of meaning in our mind.

Conceptual representations are assumed to build an independent network that is frequently referred to as semantic or conceptual memory (Levelt 1993). It needs to be noted that semantic memory is not the same as the mental lexicon, which is often compared to a dictionary. Rather, it is a mental encyclopedia independent of the formal linguistic representations of the lexical items (cf. Levelt 1993). Clark and Clark (1978) explain the distinction between the two by saying that not all concepts stored in semantic memory have names in the mental lexicon. A typical way of presenting conceptual representations of lexical items is a rich network of sense relations. Semantic information is given meaning only by the way it relates to other information. Putting it bluntly, “words are organized in an interconnected system linked by logical relationships” (Aitchison 2003a, p. 103). And thus, a definition of a word (concept) is always created in relation to other words (concepts).

Initially, it was believed that it is possible to measure the distance between words in the network, thus defining their mutual relations. Some later studies, however, postulated that the network is far more complicated and much less stable than was once assumed. It is now generally agreed upon that concepts are represented in a network of interconnected nodes and that the distance between the nodes represents similarity between the items. And thus, a typical mode of describing conceptual representations is as an associative network. Originally, associative links among lexical items were believed to be fixed and stable and to reflect the internal organization of words in semantic memory. The major research tool seeking to describe this static model were free association tests (cf. Deese 1962, 1965). In this model the meaning of a word was believed to be the sum of all its associations. Additionally, the model attempted to classify various relations among words such as syntagmatic and paradigmatic relations.

As for the syntagmatic-paradigmatic shift, two prominent scholars dealing with this intriguing phenomenon were Melčuk and Zholkovsky. They argued that, in contrast to adults, children have words organized differently in their mind (Melčuk and Zholkovsky 1988). They found that in word association tasks adults give associations within the same category; i.e. the word *sun* typically evokes words such as *moon* or *star*, whereas children tend to associate words paradigmatically; i.e. the word *sun* triggers *yellow*, *hot* or *shines*. Altogether, it was concluded that the relations change with age from syntagmatic to paradigmatic.

Below basic models of semantic representations are presented; namely, the hierarchical network model, the semantic feature model, and the spreading activation model.

2.4.1 The Hierarchical Network Model

As has been indicated earlier, the storage of conceptual representations can be depicted as a system of interconnected elements. Hierarchical network models posit that a word's meaning depends on its relation (a network of relations) to other words and that semantic information is arranged in a network. However, a new notion introduced here is hierarchy. Collins and Quillian, the major proponents of this model, argue that semantic representations of words belonging to one category create a hierarchical system (cf. Collins and Quillian 1969, 1970). And thus, as illustrated in Fig. 2.2, words with more general meanings are placed higher in a network, whereas more specified words tend to be positioned lower in the hierarchy; e.g., the word *animal* is located over *fish*, which in turn is superordinate to *salmon* or *shark*.

Another significant assumption of the presented model is the cognitive economy according to which semantic information referring to more than one word is stored at the highest possible node and is accessible to all the subordinate nodes through the network of internal relations; e.g., the information that *A salmon can swim* or that *A salmon has fins* is stored at the *fish node* which is superordinate to the *salmon node* and is true of all fish. Essentially, word properties are stored at the most general (i.e. the highest) level possible (cf. Fig. 2.2).

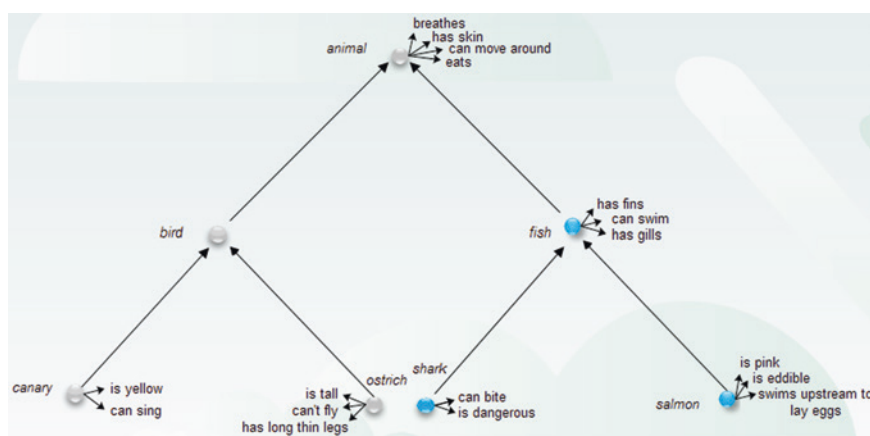


Fig. 2.2 A hierarchical network model of semantic information related to animals (adapted from Collins and Quillian 1969)

Table 2.1 Predictions of the hierarchical network model theory that proved to be wrong (adaptation based on Harley 2004)

Problem	Sample sentences	Model predicts	Finding
Familiarity effect	A. <i>A bear is an animal</i> B. <i>A bear is a mammal</i>	B faster than A	A faster than B
Typicality effect	C. <i>A robin is a bird</i> D. <i>An ostrich is a bird</i>	C = D	C faster than D
Concept property associations	E. <i>An animal breathes</i> F. <i>A bird breathes</i>	E faster than F	E = F

To check their model Collins and Quillian employed sentence verification tasks (cf. Collins and Quillian 1969, 1970). They assumed that it takes longer to verify a sentence containing information from the most remote nodes in the hierarchy, e.g., *A bear is a mammal*, than a sentence using information from closer nodes, such as *A bear is an animal*, since the lower levels inherit the information from the higher levels. This kind of familiarity effect has not, however, been confirmed in empirical research (cf. Table 2.1 below). The model has also been criticized for the invalidity of accommodating the typicality effect which posits that all the words from the same level of a given hierarchy, e.g., *robin*, *ostrich*, *canary* etc. are to be considered equal. Hence the relation between *robin* and *bird* and *ostrich* and *bird* should be perceived as equal. Nevertheless, the postulate has not been evidenced in sentence verification tasks. Conversely, the research carried out proved the predictions of the hierarchical framework inaccurate. Table 2.1 below compiles the basic problems the model does not account for. For instance, in contrast to Collins and Quillian’s assumptions rejecting the familiarity effect, familiar words are indeed recognized faster than unfamiliar words irrespective of their position in the hierarchy.

2.4.2 The Spreading Activation Model

In view of the criticism validated by the results of the numerous experiments (cf. Table 2.1) an improved version of the hierarchical network model was presented and until now seems to be the most satisfying model of the semantic memory. The basic change concerns the notion of hierarchy. Collins and Loftus, the major advocates of the spreading activation theory, postulate that the meanings of words form a network of semantic relations. The network, however, is not hierarchical any more (cf. Collins and Loftus 1975). No longer are the links within the network organized along the superordinate and subordinate principles. Instead, it is argued that the relations between semantic representations are not of equal importance. In brief, some nodes are more accessible than others and the degree of accessibility depends on the frequency of use and the word’s typicality (Collins and Loftus 1975). Additionally, the authors claim that the distance between nodes is determined by structural characteristics, e.g., taxonomic relations (cf. Rosenman and Sudweeks 1995) or the already-mentioned typicality (cf. Fig. 2.3).

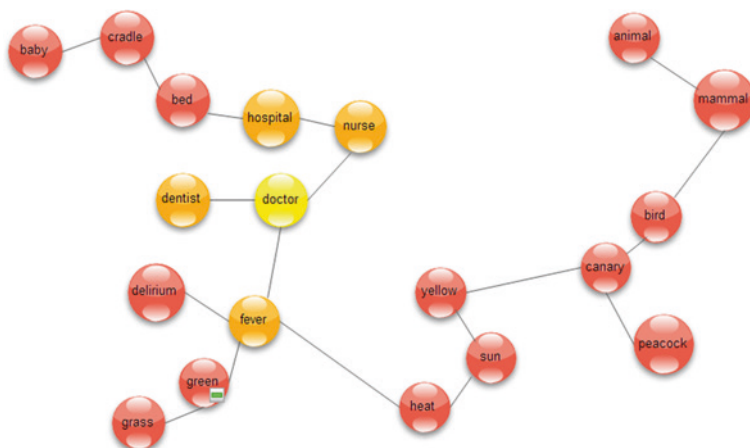


Fig. 2.3 A diagrammatical representation of a possible semantic network for DOCTOR (adapted from Collins and Loftus 1975)

More importantly, the model also seeks to account for the problem of semantic economy. Whereas the hierarchical model assumed that the word's semantic properties were stored, for the reason of economy, at the highest possible nodes thus eliminating the redundancy, the revised theory has it that certain features which are typically associated with a given word are stored with the semantic representation of that word, against cognitive economy, quite redundantly. Collins and Loftus' model also encompasses the typicality effect as developed by the prototype theory (cf. Sect. 2.4.3). Hence the distance between two nodes is conditioned by the typicality of these words and not by the hierarchy of organization; e.g., the connection between *bird* and *penguin* is weaker than between *bird* and *pigeon*. To test the efficacy of their model, Collins and Loftus employed the semantic priming paradigm. The obtained results appeared to support the assumption of the automatic spreading activation mechanism to be found in the processing of semantic representations.

2.4.3 The Componential Approach

The semantic feature view stands in contrast to the hierarchical network model (cf. Smith et al. 1974). This kind of approach, also termed the componential approach, proposes that words can be decomposed into a bundle of primitive semantic elements. As a result, words similar in meaning share some of their semantic features known as *defining features*, but they also incorporate some *characteristic features* specific only to them. This theory is connected with two influential categorisation theories to be discussed below.

Two contrasting standpoints concerning the phenomenon of the nonverbal, conceptual representations are: the classical view and the prototype theory derived from cognitivism. The classical theory is based in ancient Greece and it prevailed in psychology, philosophy, and linguistics until the 1950s. It is based on objectivism and essentialism which in turn constitute the very core of the Aristotelian model of categorization. In the light of essentialism,

all reality is made up of objectively existing entities with properties and relations among them. Some properties are essential and others are not. Classical categorization links categories to properties. Objectivist cognition assumes that people reason in terms of abstract symbols and that those symbols get their meaning via a correspondence between those symbols on the one hand and entities and categories in the world on the other. (Lakoff 1987, p.173)

Aristotle enumerates two aspects of a thing: essence described as “the parts which are present in such things, limiting them and marking them as individuals, and by whose destruction the whole is destroyed (...)” (Metaphysics 5.8.3.) and accidents referred to as “that which attach to something and can be truly asserted, but neither of necessity nor usually, e.g., if someone in digging a hole for a plant has found treasure” (Metaphysics 5.30.1.). The above-mentioned aspects can be explained on the example of the word *flower*. The essence of a *flower* is that it is a plant; its colour or smell is just accidental and does not influence the judgment of whether the entity is a flower or not.

However influential the classical theory of categories may be, it has been “reappraised in all of the cognitive sciences” (Lakoff 1982, p. 3). In the 1970s a competing theory of natural categorisation was proposed by Eleanor Rosch (1975).⁵ Since the theory centered on the so-called prototypical members of the group of possible referents of a given word, it was labeled the prototype theory. It can be briefly described as “a hypothesis that people understand the meaning of words by reference to a highly typical example” (Aitchison 2003a, p. 94). In a short time the theory gained a wide group of supporters, including Bolinger (1977), Lakoff (1982, 1987), Wierzbicka (1985) or Langacker (1987) and more recently Smith and Minda (2002) or Taylor (2003). Unlike the purely theoretical argumentation of the objectivist metaphysics and psychology, the prototype theory is based on empirical evidence, “experimental results and the interpretation of these results” (Lakoff 1982, p. 8).

The subsequent paragraphs provide a short presentation of the number of differences concerning these two highly influential theories. The first difference to be discussed is the so-called componential analysis. In the classical view categories are defined in terms of a conjunction of necessary and sufficient conditions. Entities can be described in terms of smaller parts—components or features of binary structure (present [+] or absent [–]). All members of the category have to share the same necessary and sufficient features (cf. Taylor 1990). It needs to be noted that categories are homogenous; i.e. that all members have equal status and

⁵ The theory found its philosophical grounds in the works of Ludwig Wittgenstein (1953).

they need to share the same features. Thus, there are no worse or better examples. In other words, no single *cat* is more cat-like than others (cf. Lakoff 1987). In the prototype theory, on the other hand, entities belonging to one category do not have to, and rarely do, possess the same inventory of features. Wittgenstein (1953) presents it as follows:

Consider for example the proceedings that we call: 'games'. I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all? (...) For if you look at them you will not see something that is common to all, but similarities, relationships, and a whole series of them at that (Wittgenstein 1953, p. 31).

Thus, it can be concluded that the underlying principle behind the categorisation is family resemblance (Wittgenstein 1953). However, not all members possess the same inventory of features still forming one common category. Rosch and Mervis refer to a category as to "a set of items of the form AB, BC, CD, DE. That is, each item has at least one, and probably several, elements in common with one or more other items, but no, or few, elements in common to all items" (Rosch and Mervis 1975, p. 66). An item classified as belonging to a given category shares features with a few others, but not necessarily with all, members of the same category. Categories are not homogenous, which means that e.g., some birds are more *birdy*, like the robin, while some are less *birdy* like the penguin (cf. Rosch 1975). It could be concluded that the theory accounts for worse and better members of one category. The most representative entities for the entire category are called prototypes. Prototypes have a privileged place in memory as they occupy the central role in the category and, consequently, are retrieved quicker.

Another discrepancy between the two theories concerns category boundaries. In the classical view category boundaries are clear-cut and stable. And thus, the decision whether an entity belongs to the category or not is based on objective features. Moreover, no factors can influence those categories. As Lakoff points out, "category boundaries do not vary. Human purposes, features of context, etc. do not change the category boundaries" (Lakoff 1982, p. 15). Thus, they would often demand redefinition or the creation of new categories. Internal definition is the only factor affecting the category, the structure of a category is context independent, no subjective factors can affect the category and thus, psychological factors seem to be unimportant. No matter how humans perceive a given item, it is categorized regardless of the subjective interpretation. In contrast, the prototype theory shows that there are no clear-cut boundaries. Instead, any boundaries are described as flexible, susceptible to subjective factors such as human purposes. Many experiments proved (cf. Black 1949; cited in Ungerer-Schmidt 1996; Labov 1973) that prototype-based categories merge into each other, and their boundaries instead of being clear can be described as fuzzy. In his publication Labov (1973) elaborates on an experiment in which the subjects were asked to name various containers (e.g., *cup*, *bowl*). The results showed that the labels provided by the participants varied substantially. Furthermore, the same participants were not consistent in their responses. Labov later concluded that a word

has its core meaning which is most central and invariant as well as its peripheral meanings.⁶ As a consequence, advocates of this theory emphasize that the meaning of a word should be analyzed on a continuum.

To conclude, the objective of the present section was to discuss the two most influential theories related to lexical meaning. As has been shown, the theories are very different. Many scholars emphasize the fact that the prototype theory being based on empirical evidence seems to be far more convincing than the classical one based on non-empirical speculations (cf. Lakoff 1987). In the following this section various models of lexical access will be examined.

2.5 Models of Lexical Access in the Mental Lexicon

Having discussed the issues concerning the structure of mental representation of words and their meaning in the human mind, the chapter will now proceed to elaborate on the selection of the most influential models of lexical access and retrieval. Obviously, it would be almost impossible, and for the sake of the present work unnecessary, to discuss and compare all the models of lexical access that have been proposed. Thus, this section has been limited exclusively to the most influential language processing models that can be found in psycholinguistics.

Lexical production and recognition are very quick processes. In his research endeavouring to analyze word recognition patterns Marslen-Wilson (1989) found that a word is recognized usually about 200 ms after its onset; this means even before the speaker managed to finish uttering that word. Not only is the mechanism of lexical access rapid, but it is also highly sophisticated and complex. Word recognition involves receiving a perceptual signal, rendering it into the phonological or orthographic representation and then accessing its meaning. The opposite process of producing a word requires first choosing the meaning for the intended concept, then recovering its phonological or orthographic representation, and finally converting it into a series of motor actions.

To date, many methods have been used, many paradigms followed to analyze lexical access in speech production and comprehension. A typical methodology adopted to search for the key to the lexical access enigma has been the analysis of malfunctions (e.g. different types of selection errors, slips of the tongue or the tip-of-the-tongue phenomenon; see Aitchison 2012 for a detailed discussion). Other methods used the already-mentioned picture naming, lexical decision tasks and priming. Yet another source of research data derives from speech pathologies such as aphasia. Aphasic patients who have lost parts or all of their linguistic abilities have provided linguists with a substantial amount of data concerning the processes of lexical access and retrieval (cf. Dell et al. 1997; Biran and Friedmann 2012).

⁶ Cf. Kellerman's famous study of core and peripheral meanings of the German verb *brechen* (*break*) (Kellerman 1978).

However complex and demanding the research on the mental lexicon might be, psycholinguistic literature abounds with models of lexical access. There are many properties according to which the models can be grouped. Firstly, some models focus on word recognition, others on production. There are also models which try to combine these two processes. Another distinctive property is the type of search involved in lexical processing: here serial (*indirect*) or parallel (*direct*) models are distinguished. The serial models have it that words are accessed individually, one by one, at the phonological, orthographic and semantic levels. The parallel models, on the other hand, postulate that words are searched simultaneously. A further property is interactivity—the question of whether lexical information can travel backwards and forwards between different levels of lexical representation and affect their retrieval.

What differentiates the models to be examined in the paragraphs to follow is the sequence of interaction; a property which divides the models into direct and indirect ones (cf. Garman 1990, p. 260). The first category of models distinguished by Garman are indirect access models, which depict lexical processing as “looking up a word in a dictionary” or “finding a word in a library” (Singleton 2000, p. 170). The indirect access models, also known as multi-step models, are predicated on two-stage access “via a search procedure and then a retrieval procedure” (Singleton 1999, p. 84). A much-debated representative of the indirect kind of model to be described below is Forster’s serial search model functioning within a modular system paradigm.

When referring to direct access models Garman uses the metaphor of a “word-processing package which allows items stored by name to be accessed simply by the typing in of as many letters as are sufficient to distinguish the relevant name from all other stored names” (Singleton 2000, p. 170). In other words, direct models view lexical processing as a one-stage phenomenon. Two oft-cited representatives of the direct type of model to be discussed below are the logogen model and the cohort model.

2.5.1 *The Serial Search Model*

The best known and one of the most influential indirect models is Forster’s autonomous search model (Forster 1976; Murray and Forster 2004). The processes of access and retrieval described in this model resemble looking up a word in a written dictionary or looking for a book in a library, the only difference being the organizing principle, which in the case of a written dictionary is alphabetically-governed, while in the mental lexicon it is claimed to be frequency-dependent. This is how Garman summarizes Forster’s suggestion:

We enter, looking for a particular book; we do not go straight to the main shelves where the books are located, since there are simply too many of them to permit efficient search of them in this direct fashion. So we go instead to the catalogue. Searching through the catalogue, we find something that matches what we are looking for; but we retrieve from

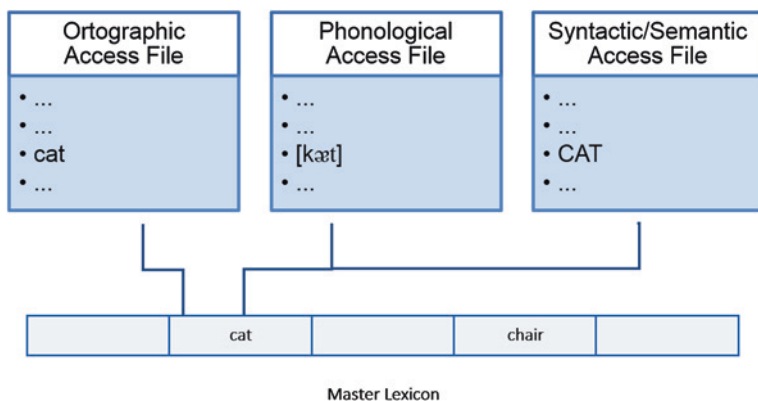


Fig. 2.4 Forster's serial search model of lexical access (adaptation based on Foster 1976)

this stage of the process, not the book itself, but an abstract location marker, telling us where to find the book on the shelves. Armed with this, we implement the second stage of the process, by using the marker to guide us to the right book on the shelves (Garman 1990, pp. 266–267).

A similar metaphor has been adopted by Singleton, who compares the first stage to finding the right page in a dictionary, the major difference between these two processes being the aforementioned principle behind the organization of the entries. The moment the “page” has been found (on the basis of the initial properties of the signal), the search goes on through the “page” governed by word frequency. Once the abstract location marker has been found, the second stage starts. Forster assumes that lexical entries are searched sequentially until the appropriate word is selected and believes that the mental lexicon consists of two levels: one containing access files and the other the master file (*the lexicon proper*; cf. Forster 1976; Fig. 2.4). There are two stages of word processing. In the first stage, following perceptual processing, the serial search of access files starts. The only information on a word available in the access files is its address in the master file. To put it differentially, access files comprise the stimulus features of a word, i.e. its *access code* and the *pointer* to the matching entry in the master file. The master file includes all the information on a given word—phonological, orthographic, morphological, semantic and syntactic data. It needs to be noted that not only is the master file a complete representation of each individual lexical item, but it also includes cross-references between all the items stored in the master file, thus providing for the semantic priming effect. In order to cater for different modalities, through which lexical items are perceived and generated, and two directions of lexical access Forster proposed three separate access files which organize words either by orthographic, phonological or syntactico-semantic properties and are linked with the master file by pointers. These discrete operating subsystems process lexical information independently of one another. A schematic visualization of the model has been presented in Fig. 2.4.

Garman (1990) notes that Forster's search model incorporates two key features a good model of the mental lexicon should have. It is characterized by the diversity of access and, at the same time, the unity of storage. Whatever the channel of communication, it is always the same entry in the master file. Access to each word depending on the channel leads always through a suitable access file:

If one is listening to speech, one processes each spoken word by going first to the phonological access file; if one is reading written language one goes first to the orthographic access file; and if one is producing language on the basis of particular meaning intentions, one goes first to the syntactic/semantic access file. The access file in question then facilitates access to the master files (Singleton 2000, p. 174).

For convenience, the access files are divided into separate bins based on the initial sound or letter. The words in a bin are arranged in a descending order of frequency so that more frequent words can be searched faster and matched with the acoustic string prior to low frequency words. In this way, Forster eventually managed to accommodate the frequency effect within his model.⁷ The effect has been confirmed by a substantial body of evidence from lexical decision tasks where high frequency words were identified faster than low frequency items. The model also includes the lexicality effect assuming that there occurs exhaustive search for nonwords and terminating one for actual words. Before rejecting a nonword the brain needs to search through the entire master file only to find an empty entry. This has been confirmed by numerous lexical decision tasks which show that it takes about 150 ms longer to identify nonwords than actual words.⁸

Not only does the model account for the frequency effect and the lexicality effect, but it is capable of accommodating the priming effect as well. Forster's model is not interactive in that it does not allow for the cross-referencing of access files and the master file. Words in the master file are accessed only through one file at a time. However, once an entry in the master file has been accessed, cross-references are observed. Thus, the model is able to accommodate the effect of semantic priming. If an individual sees the target word for *doctor* and then subsequently is shown the word *nurse*, the response time for the latter word is rightly expected to decrease.

However detailed the model seems to be, there is still a substantial number of controversies it has been unable to end. Firstly, the model faces the problem of capacity limitations. The evidence from lexical decision tasks supports the idea of empty entries for nonwords, which, if really there, would occupy a lot of space redundantly. Secondly, speech seems to be much too rapid to accept the idea that words are searched sequentially; the model allows for only one entry to be searched and matched with the input at a time. Another repeatedly criticized issue is the fact that the model does not allow for the influence of context on the process of recognition. It also does not give an account of form-based priming

⁷ The original version of the model which presented it as a direct access model (cf. decision trees; Forster 1976, p. 258) failed to incorporate the results supporting both the lexicality and the frequency effect and proved quite to the contrary.

⁸ Again this effect was impossible to implement within the original version of the model.

effect and it cannot explain the role of similarity neighbourhood. Finally, the model does not explain the influence of context on speech production (a phenomenon confirmed by the repetition priming effect). Due to the above limitations the early version of the model underwent extensive changes (cf. Forster 1989). For instance, in the revised version, Forster introduced a discrete comparator for each lexical entry, thus solving the problem of limited capacity (Murray and Foster 2004). He also proposed various models of activity among lexical entries. However, it seems that the presented changes have actually transformed the model in the direction of spreading activation models.

2.5.2 The Logogen Model

The logogen model, in contrast to its serial search equivalent, assumes one-stage parallel processing. Initially, the model was proposed by the British psychologist John Morton in 1969 to account for visual word recognition. Only later was it developed and revised to incorporate both written word recognition and word selection in speech production. The model comprises three elements: the logogen system, the cognitive system and the response buffer. However, its key feature is the logogen system which is defined to be a part of the nervous system responsible for lexical processing (in the initial version of the model it was described as a *neural unit*, to be later changed to the more technical term *logogen*⁹). According to Singleton, the logogen system is “a set of mechanisms (...) specialized for collecting perceptual information and semantic evidence concerning the presence of words to which the logogens correspond” (Singleton 1999, p. 86), whereas Coltheart et al. (2001) describe it as an “information-gathering” device.

Initially, Morton postulated that there was a unitary logogen system, but due to some empirical evidence he revised the idea and divided the system into three parts. He distinguished two specialized input logogen systems: visual and auditory and the output logogen system (cf. Morton and Paterson 1980). It has to be stressed that none of these units includes semantic information about words. This information is stored in the cognitive system, which includes “a collection of semantic information of various kinds” (Singleton 1999, p. 86). The system may, but does not have to, be incorporated in the lexicon itself. What merits special consideration, however, is that meaning is not stored as a single unit for each word; instead, it is computed when required.

In the logogen system every single item is represented by the corresponding logogen which comprises the word’s features (phonological and orthographic characteristics). The moment the acoustic or visual input reaches the logogen, it is changed into appropriate phonological or orthographic representation which

⁹ *logogen*, from Greek *logos*, “word” and Greek and Latin *gen*—“birth”; “to bring to life” (after Singleton 2000, p. 171).

launches the process of word finding. The next step is passing the information to the cognitive system which specifies its semantic and conceptual aspects, and finally to the logogen output system. It needs to be stressed that the links between the elements of the model are bidirectional.

It has to be noted that one of the key features of the model is the so-called threshold level. Each logogen has a “resting” threshold level. Once sufficient evidence has been introduced into the system, the threshold level is reached and the word is activated. This means that, e.g., in the case of a written word, even before all the letters are identified, the word can already be recognized and its code sent to the cognitive system. Clearly, threshold levels differ in value depending on the words’ frequency. Thus a high frequency word has a much lower threshold than a low frequency word and less activation will be needed to “fire” and thus access this word (cf. Harley 2008). In this way the model accounts for the frequency effects of words in a different way, by using the activation and raising of potentials within different words.

Summing up, in this model word recognition is seen as a process of accumulating sufficient information to ultimately access a given word. Once enough information has been gathered, the logogen’s threshold is exceeded, the code is passed to the cognitive system and to the suitable output logogen. The key features of the model are: directionality of access—every word has its logogen, interactivity—it allows for the interaction of semantic and perceptual aspects, and parallel processing—incoming information is checked against logogens when it reaches threshold. The model is versatile in that it accounts for both visual and auditory processing. Nevertheless, it is perceived as very complex and difficult to test experimentally. It also leaves many empirical findings unaccounted for, one of them being the effect of neighbourhood size.

2.5.3 *The Cohort Model*

The cohort model is another variant of the direct access model. It was first proposed by the British psycholinguist William Marslen-Wilson (1973, 1975) and later revised many times to incorporate new findings of psycholinguistic research (Marslen-Wilson and Welsh 1978; Marslen-Wilson and Warren 1994; Marslen-Wilson 2007). The model focuses on explaining the process of spoken word recognition and does not account for other aspects of lexical access, namely visual word recognition and word production. The model is based on the assumption that words are recognized by their onsets in the left-to-right fashion of processing. Once the initial segments of the word are uttered and received, all potential lexical candidates commencing with that very sound(s) are activated and form an initial cohort. This assumption is supported by the tip-of-the-tongue phenomenon (TOT), according to which lexical access is possible when the beginning sounds of a word become accessible (Biedermann et al. 2008). Spoken word recognition is assumed to constitute three stages: access, selection and integration. In the access stage the

perceptual representation of the word triggers the activation of a set or, as Marslen-Wilson (1992) suggests, a cohort of lexical items which share the same acoustic features. On the basis of empirical findings¹⁰ Marslen-Wilson (1992) postulates that a cohort is activated even before the word has been pronounced to the end. In fact, the very first sounds trigger the activation of a group of words beginning with that particular sequence of phonemes referred to as the *word-initial cohort*. As the subsequent sounds are pronounced, more information is given and the cohort is narrowed down up to the word's *uniqueness point*; the point at which only one word candidate is left in the cohort. What proves the existence of the uniqueness point is the finding that a word that has not been fully pronounced can still be guessed. Moreover, the model also defines the point at which nonwords are recognized; namely the point at which the sequence of pronounced sounds fails to correspond to any word of the language. For instance, the nonword recognition point for a potentially English word *daffodip* would be the very last sound /p/ since "only this final sound rejects the possibility of a match" (Singleton 2000, p. 173).

In its early version the model was presented as fully interactive. Marslen-Wilson postulated that a word can be recognized and selected even before it reaches its uniqueness point due to contextual information. It was also believed that a word can be eliminated from the cohort due to context. Indeed, many findings obtained in the course of psycholinguistic research supported the assumption that contextual information has a facilitatory impact on lexical processing. Much evidence in favour of this proposal came from speech shadowing experiments (cf. Marslen-Wilson and Welsh 1978) where the subjects were asked to retell a story they heard. What merits special attention is the fact that words which in the story were incorrect (e.g., mispronounced or misused) were successfully repaired in the process of retelling the story. Moreover, retelling was not accompanied by any hesitation pauses. Marslen-Wilson used this empirical data to support the assumption of the importance of contextual information. He claimed that fluent restoration proved the influence of context since the mispronounced words could have only been corrected on contextual grounds. Another pool of evidence in favour of the role of context came from word-monitoring (Marslen-Wilson and Welsh 1978) and rhyme-monitoring studies (Marslen-Wilson and Tyler 1980).

Understandably, as every controversial assumption, the issue of context effects had many opponents. Many critics emphasized the fact that context cannot cause the elimination of words from a cohort. Ultimately, the mounting criticism and the growing evidence against the validity of the context effect forced Marslen-Wilson to reject the dominant role of context in auditory word recognition. As he pointed out himself, the problem with the context-driven pre-selection lies mainly in the fact that it does not account for the open-endedness and unpredictability of language use (cf. Marslen-Wilson and Tyler 1980). It needs to be added, however, that the more recent versions of the model assume that although "(...) contextual

¹⁰ To recognize a monosyllabic word it takes about 300 ms from the word onset and about 100 ms before its coda (cf. Marslen-Wilson and Welsh 1978).

information has no impact on the selection of the word-initial cohort, (...) once the cohort has been established, word candidates which are inconsistent with the context can begin to be deactivated” (Singleton 1999, p. 94).

What is interesting to note is that whereas the previously described logogen model allowed for various levels of activation, the early version of the cohort model, on the contrary, postulated the existence of the binary membership. It asserted that an item is either active (on) or not (off). The first alternative refers to the situation when a word still belongs to a cohort of word-candidates, the latter describing the situation when the word has already been eliminated from the cohort. Ultimately, later versions of the model rejected the binary membership and postulated the gradual membership. It was suggested that words which do not receive further verification from the incoming acoustic representation have their level of activation gradually lowered; they cannot, however, be eliminated from the cohort. Conversely, they can be activated again if an appropriate signal occurs. Consequently, in the more recent versions of the model (1990, 1993) the importance of input goes beyond the uniqueness point and the deactivation of word candidates is reversible; the alterations which have made the model maximally efficient.

Finally, it needs to be stressed that however controversial the model might be, there is still a fair amount of experimental evidence to be found in favour of its main assumptions. Singleton (2000), for instance, refers to the recognition of non-words. The recognition of nonwords is shorter, he says, in those cases where recognition points come early in words. As Singleton reports, “the more contextually predictable a word is, the shorter the sequence of sounds required to reduce the cohort to a sole candidate” (Singleton 1999, p. 95). The recognition time is much longer when recognition points appear later within a word. On the other hand, the model is still criticized, the basic criticism being connected with the fact that the model accounts for only one type of modality and fails to explain effects of frequency or neighbourhood density. Additionally, some researchers hold it as highly unlikely that we recognize words on the basis of “the noisy and ambiguous acoustic signal which is speech” (Marcus and Frauenfelder 1985, p. 164; in Garman 1990).

2.5.4 Computational Models

The traditional “box-and-arrow” type models discussed so far were determined by “the high-level theoretical principles themselves” (Norris 2013, p. 518) but were unable to explain what processes were going on in the boxes. The situation changed with the development of computational models of reading in the early 1980s. Recent computational models are able to handle realistic lexicons and simulate data from a range of different tasks (e.g. masked priming, lexical decision or eye-movement control). Moreover, current models of word recognition can now perform large-scale simulations using many thousands of words. Finally, they can successfully simulate an interaction between the theoretical predictions and the contents of the lexicon. They make clear assumptions about what is supposed to

Table 2.2 Major computational models of visual word recognition (adapted from Norris013)

Computational models of visual word recognition	Author(s)	The main phenomena the model simulates and the tasks used in simulation
The interactive activation model (IA)	McClelland and Rumelhart (1981) Rumelhart and McClelland (1982)	Word-superiority effect/perceptual identification task
The spatial coding model (SCM)	Davis (2010)	Letter order/lexical decision task; masked priming task
The dual-route cascaded model (DRC)	Coltheart et al. (2001)	Reading aloud/lexical decision task
The letters in time and retinotopic space (LTRS)	Adelman (2011)	Letter order/masked priming task; perceptual identification task
The Bayesian reader (BR)	Norris (2006) Norris (2009) Norris and Kinoshita (2012)	Word frequency, letter order, RT distribution/lexical decision task; masked priming task
Diffusion model	Ratcliff (1978) Ratcliff et al. (2004) Gomez et al. (2013)	Word frequency, letter order/lexical decision task
SERIOl	Whitney (2008) Whitney (2011) Whitney and Cornelissen (2008)	Letter order/lexical decision task; masked priming task

be going on in the boxes and later successfully work out differential predictions of the models (cf. Norris 2005). For the reasons indicated above, there is a common agreement among psycholinguists that computational models are to be preferred over traditional “box-and-arrow” models. However, some obvious limitations all the current models share, the main being their focus on a single domain of behaviour, should not be left unnoticed. Indeed, Norris raises a valid point when he remarks that there is still a need for more integrated theories of word recognition (2013, p. 523).

Table 2.2 presents a selection of the most influential computational models of visual word recognition and points to the basic phenomena the models have been developed to explicate. As for the modelling style of framework within which the models have been created, the most influential style of computational models are connectionist models, the earliest example of which is the Interactive Activation model (IA),¹¹ first proposed by McClelland and Rumelhart in 1981. This type of modelling has for many years been favoured among researchers mainly due to the fact that it is relatively “brain like” (Clark 1993) and relatively easy to understand. An alternative style of modelling—mathematical or computational one—exploits computational procedures or mathematical formulae. It needs to be noted,

¹¹ McClelland and Rumelhart’s model will be discussed in details in Sect. 2.6.2 below.

however, that the interactive activation model (McClelland and Rumelhart 1981), the Spatial Coding Model (Davis 2010) or the dual-route cascaded model (Coltheart et al. 2001) which are typically visualized as connectionist models can also be expressed mathematically (cf. Norris 2013).

2.6 Views on Language Processing

This section provides a contrastive description of the two hypotheses on the linguistic storage (the modularity hypothesis and connectionism), with particular emphasis on their strengths and weakness.

The modular view has it that the mind is “divided into separate compartments, separate modules, each responsible for some aspect of mental life” (Cook and Newson 1996, p. 31). The modularists claim that linguistic meaning is clearly separated from other varieties of meaning and is represented and processed within the language module (cf. Emmorey and Fromkin 1988). The proposed processing is sequential (i.e. one thing at a time—an assumption which makes the processing slow), symbolic (i.e. one token equals one concept) and procedural (linguistic behaviour is governed by rules). The basic problem with this theory, however, is its inflexibility.

Cognitive theories, which are commonly taken to be antipathetic to the modular view, adopt the analogy of brain-style neuronal interactions and depict the mind as a single system—an interactive network. They describe linguistic processing in terms of connection strength rather than rules or patterns. It needs to be noted, however, that in the recent decades the most current models have sought to combine both the modular computational and the connectionist theory (cf. Dell 1988).

2.6.1 *The Modularity Theory*

Customarily, the origins of the modularity theory can be traced as early as in the 18th century, when a German anatomist Franz Josef Gall “developed the view that each intellectual and behavioural attribute was controlled by a specific location in the human brain” (Singleton 1999, p. 111). The current version of the hypothesis became one of the most influential cognitive perspectives of the late 1960s. The major proponents of this modular view of the mind are theoretical linguist Noam Chomsky (1988) and psycholinguist Jerry Fodor (1983, 1989). Whereas Chomsky’s interest in modularity is related exclusively to language acquisition processes, Fodor’s work concentrates on aspects that are processing-oriented. Since the focus of the present chapter is centered around issues relating to language processing, only the Fodorian perspective is discussed.

The modularity hypothesis, according to Fodor (1983), postulates that “the entire language faculty is a fully autonomous module [comprising] a number of

distinct, specialized, structurally idiosyncratic modules that communicate with other cognitive structures in only a very limited way”¹² (Singleton 2000, p. 176). In the light of the Fodorian theory, modules are independently functioning cognitive systems located within the language system. They can be defined in terms of nine characteristic features. Five of the features refer to the way in which modules process information, and as Fodor himself points out (Fodor 1989), are also characteristic of acquired skills. These include: *informational encapsulation* (i.e. the notion that it is impossible to interfere with the inner workings of a module), *unconsciousness* (i.e. the assumption that it is difficult or impossible to think about or reflect upon the operations of a module), *speed* (i.e. the idea that modules are very fast), *shallow outputs* (i.e. the view that modules provide limited output, without information about the intervening steps that led to that output), and *obligatory firing* (i.e. the claim that modules operate reflexively, providing pre-determined outputs for pre-determined inputs regardless of the context).¹³ Another three features, namely *ontogenetic universals* (i.e. the postulate that modules develop in a characteristic sequence), *localization* (i.e. the idea that modules are mediated by dedicated neural systems), and *pathological universals* (i.e. the suggestion that modules break down in a characteristic fashion following some damage to the system), characterize the biological status of modules and play a crucial role in differentiating the behavioural systems from learned habits.¹⁴ The final and concurrently the most controversial feature is *domain specificity*, i.e. the assumption that modules deal exclusively with a single information type.

It is beyond the scope of the present chapter to elaborate on all the aforementioned modular aspects, thus it will suffice to present Fodor’s two major and most controversial postulates describing the processing of linguistic information as *domain-specific* and *informationally encapsulated*.

Domain specificity asserts that each module is capable of processing only certain linguistic information. Fodor emphasizes the fact that this feature of the language module was confirmed in a number of experiments where both linguistic and non-linguistic context of one and the same signal influenced the way it was perceived by the subjects (cf. Liberman et al. 1967). *Informational encapsulation* means that intramodular processing is unrelated to other operating systems and nonlinguistic cognitive processes and that modules do not make use of other information available in the cognitive system as a whole. In other words, Fodor postulates that language module is immune to non-linguistic operations carried out outside the module such as general knowledge

¹² This opinion does not equal with the claim that the language module has absolutely no connection with other cognitive operations (cf. Aitchison 2003a).

¹³ The detailed description of the features has been based on Fodor (1983, 1985).

¹⁴ It is assumed that learned systems do not display these particular regularities (cf. Singleton 2000).

or the influence of context (cf. Singleton 1999). He sees language processing as a system limited to “a formal processor with no semantic role” (Fodor 1983, p. 178). He also clearly distinguishes the linguistic processing from processing of non-linguistic data.

The postulate that the language module is informationally encapsulated, and thus context-independent, constitutes one of the most controversial and widely disputed aspects of the Fodorian theory. Essentially, in the light of the abundant evidence coming from psychological and psycholinguistic research this postulate is difficult to accept. There is a substantial body of research confirming the facilitative role of general knowledge and context in language task performance. Singleton (2000, p. 177) argues that cases have been reported when multilingual speakers fail to understand or even recognize language which they are fluent in if they do not expect to be exposed to that language. Another source of counterevidence highlighting the importance of context in speech production and comprehension derives from experiments carried out with hypnotized subjects who are able to interact. Furthermore, a pool of arguments against encapsulation derives from empirical findings of experiments that involve reduced-redundancy procedures such as cloze tests. In this type of lexical tasks participants are to fill in missing words removed from a cohesive text. In order to do this they need to read the whole text. Results show that the more predictable the target elements, due to some contextual clues, the more successful the performance of the participants endeavouring to guess the missing words (cf. Weir 1988). These findings lend support for the proposal that participants actually utilize all aspects of contextual information (e.g., semantic or syntactic clues) at the same time. Singleton claims that these results account for evidence for effects of *cognitive penetration* (Singleton 1999, pp. 115–116) during processing. Fodor, however, strongly refutes such an interpretation and suggests that what might appear as contextual effect, might also be viewed as a matter of *interlexical excitation*¹⁵ (Fodor 1983, p. 80). He presents his claim in the following manner:

We can think of accessing the item in the lexicon as (...) exciting the corresponding node; and we can assume that one of the consequences of accessing a node is that excitation spreads along the pathways that lead from it. Assume, finally, that when excitation spreads through a portion of the lexical network, response thresholds for the excited nodes are correspondingly lowered (Fodor 1983, p. 80).

To conclude, from the presented evidence it can be inferred that in the Fodorian model lexical knowledge is represented in the network of interconnected nodes. It is perceived as a central part of a larger system which operates independently from other systems. The underlying assumption posits that the mind is modular and comprises special-purpose perceptual processors called *modules*.

¹⁵ It seems justified to emphasize that Fodor’s description of interlexical excitation shows direct similarities with some claims of spreading activation theory (cf. Sect. 2.3.2).

2.6.2 Connectionism

The connectionist theory dates back to the works of McCulloch and Pitts, who in the 1940s presented the first mathematical model describing the functioning of a neuron (McCulloch and Pitts 1943; after Singleton 2000). However, the first influential models of the lexical processing within the connectionist paradigm were proposed only after a long period of silence in the 1970s and 1980s. To explain lexical processing, connectionism adopts the “brain metaphor” (Rumelhart and McClelland 1986, p. 75), which is based on neurophysiological activity in the brain. And thus, the basic feature of all connectionist models, known also as interactive network models, is an analogy between the human brain and the connection of neurons. They all depict the mental lexicon as a network of nodes which have various degrees of activation and perceive lexical processing as activation spreading along a network of interconnected units. In fact, one of the major interests of connectionism is to compute the algorithm that reflects how activation spreads around the network and triggers individual nodes.

It is worth noting that the connectionist approach to lexical processing belongs in a much broader parallel processing perspective which stands in stark contrast to the aforementioned modular theory deriving from the serial processing tradition (cf. Sect. 2.4.1). Firstly, whereas the parallel processing perspective advocates the independence of processing operations (i.e. that many activated items can be handled simultaneously), the serial perspective describes lexical processing as organized in stages (i.e. that activated items can be handled in the one-at-a-time order). Secondly, the connectionist models call into question the Chomskyan/Fodorian perception of language and the mind by rejecting the so-called *symbolic paradigm* which posits that “mental operations involve the manipulation of symbols” (Singleton 2000, p. 179). Instead, the connectionist paradigm seeks to describe information processing in terms of the strength of connections between units in a network rather than in terms of rules. As Singleton puts it, “it is not patterns that are stored (...) but rather the connection strengths between elements at a much lower level that allow these patterns to be recreated” (Singleton 2000, p. 180).

Proponents of parallel processing models stress that an obvious advantage of these models over serial search models is that they can explain the enormous complexity of the information processing in the brain. On the other hand, connectionist models are frequently criticized for their inability to account for syntactic and semantic aspects of language processing. Indeed, the current versions concentrate mainly on the lexical level.

In the subsequent paragraphs two approaches representative of the connectionist tradition will be briefly outlined: localist connectionism and distributed connectionism, also referred to as parallel distributed processing (PDP).¹⁶ In the

¹⁶ Both localist connectionism and distributed connectionism approach will be further discussed in the multilingual context in Chap. 4.

localist models each item is represented by a single unit (node) which is symbolic in nature and has a functional value (cf. McClelland and Rumelhart 1981; Stemberger 1992; Dell 1988; Roelofs 1992, 1999). In contrast, distributed connectionism models assume the existence of distributed representations which are processed in parallel and where units do not bear any functional value. The most representative example of the latter type is the parallel distributed processing model designed by Seidenberg and McClelland (1989). The basic difference between the localist and distributed connectionism models lies in the representation of words. In contrast to the localist connectionism models which assume one-to-one correspondence of lexical units and their mental representations, in the PDP models “knowledge of words is embedded in a set of weights on connections between processing units encoding orthographic, phonological, and semantic properties of words, and the correlations between those properties” (Seidenberg and McClelland 1989, p. 560). In these models there are no entries, bins or logogens. The models do not accommodate the mental lexicon in the traditional sense, nor do they account for the traditional lexical access. As Singleton observes, “different portions of information are simultaneously processed independently of one another (‘in parallel’) on different levels (‘distributed’)” (Singleton 2000, p. 179).

One of the first parallel processing models (pre-connectionist model) is the Interactive Activation Model which was put forward by McClelland and Rumelhart (1981). The model postulates that perceptual processing takes place simultaneously at a more than one level (*parallel processing*). McClelland and Rumelhart (1981) distinguished the feature level, the letter level, the word level and higher levels responsible for top-down input to the word level. The model is not only parallel but it also accounts for interactive processing, which means that in the process of word comprehension there are two co-occurring factors, namely lexical knowledge and the incoming information from the perceived stimulus. Thus, the processing is both top-down (conceptually driven) and bottom-up (data-driven) at the same time. As for the representation of words, the model posits that lexical units have their corresponding nodes which are stored in levels (localist tradition) and are linked with other nodes. It needs to be stressed that nodes are connected bidirectionally with other nodes at different levels of the network.

The model also accommodates the frequency effect. Nodes have their activation levels which are modified by the amount of activation they receive from other nodes (*neighbours*). Nodes corresponding to frequently or recently used lexical items have a lower level of activation and thus are selected faster than nodes which represent words of lower frequency. Communication between nodes is possible due to the spreading activation mechanism. McClelland and Rumelhart (1981) posit that there are two types of connections within the system of nodes: excitatory and inhibitory ones. The former is responsible for increasing the activation level of connected nodes, the latter for decreasing the level.

2.7 Conclusion

This chapter has been meant to serve as a background for a more thorough consideration of various theoretical issues and empirical investigations concerning the multilingual context which will be discussed in the subsequent chapter of the present work. The underlying assumption has been that presenting and explicating the most significant concepts pertinent to the modelling of the mental lexicon would enable the reader to interpret and evaluate the research projects whose design and findings are presented in Chaps. 4 and 5. In order to do so, an attempt has been made to review the classical theories and models concerning the organization of the monolingual mental lexicon. The chapter started with a short discussion on the internal structure of a lexical entry. Subsequently, it provided a brief overview of various definitions of the mental lexicon as an entity, from those depicting it as a dictionary to the ones which view it as a network. Apart from tackling terminological issues, the present chapter dealt with the phenomenon of the representation of meaning in the human mind. The discussion revolved around the most influential models concerning the internal structure of the monolingual lexicon, as well as the numerous models of lexical processing. With regard to lexical processing, the chapter summarized and assessed the better-known psycholinguistic models concerning the organization and functioning of the mental lexicon (Forster's lexical search model, Morton's logogen model, Marslen-Wilson's cohort model), and gave a brief account of most recent computational models of visual word recognition. Finally, special consideration was given to modular and connectionist perspectives on lexical processing. It needs to be noted, however, that despite its substantial size, the section on lexical processing has not exhausted the vast scope of visual word recognition study. In particular, a lot more can be said about the achievements of computational modelling (cf. Norris 2005, 2013). However, as stated above, the chapter has been meant to serve as a background for a discussion of the multilingual mental lexicon, which constitutes the main concern of the present work.

Now that the main issues concerning the concept of the mental lexicon have been delineated, it is time to outline the most prominent hypotheses and models of language storage, processing and retrieval in relation to the mental lexicon of multilingual speakers. More precisely, the following chapter will be devoted to the presentation and discussion of the issues of single vs. multiple lexicons and language selective vs. language nonselective lexical access.

Multilingual Lexical Recognition in the Mental Lexicon of
Third Language Users

Szubko-Sitarek, W.

2015, XVI, 211 p. 22 illus., 12 illus. in color., Hardcover

ISBN: 978-3-642-32193-1