British Journal of Psychology (2001), 92, 129–170 *Printed in Great Britain*

The language machine: Psycholinguistics in review

Gerry T. M. Altmann*

Department of Psychology, University of York, UK

Psycholinguistics is the empirical and theoretical study of the mental faculty that underpins our consummate linguistic agility. This review takes a broad look at how the field has developed, from the turn of the 20th century through to the turn of the 21st. Since the linguistic revolution of the mid-1960s, the field has broadened to encompass a wide range of topics and disciplines. A selection of these is reviewed here, starting with a brief overview of the origins of psycholinguistics. More detailed sections describe the language abilities of newborn infants; infants' later abilities as they acquire their first words and develop their first grammatical skills; the representation and access of words (both spoken and written) in the mental lexicon; the representations and processes implicated in sentence processing and discourse comprehension; and finally, the manner in which, as we speak, we produce words and sentences. Psycholinguistics is as much about the study of the human mind itself as it is about the study of that mind's ability to communicate and comprehend.

> By degrees I made a discovery of still greater moment. I found that these people possessed a method of communicating their experience and feelings to one another by articulate sounds. I perceived that the words they spoke sometimes produced pleasure or pain, smiles or sadness, in the minds and countenances of the hearers. This was indeed a godlike science, and I ardently desired to become acquainted with it.
>
> Mary Shelley *Frankenstein, or, the modern Prometheus* (Penguin edition, p. 108)

Through language we each of us cut through the barriers of our own personal existence. In doing so, we use language as an abstraction of the world within and around us. Our ability to interpret that world is extraordinary enough, but our ability to abstract from it just certain key aspects, and to convey that abstraction through the medium of language to another individual, is even more extraordinary. The challenge for psychology has been to reveal, in the face of extraordinary complexity, something of the mental representations and processes that underpin our faculty for language. The purpose of this review is to convey those aspects of psycholinguistic research that have shaped the current state-of-the-art. The reader should bear in mind, however, that the *Handbook of psycholinguistics* (Gernsbacher, 1994) contains in excess of 1100 pages and a subject index with barely fewer words than the number originally suggested for, but subsequently exceeded by, this

* Requests for reprints should be addressed to Dr Gerry Altmann, Department of Psychology, University of York, Heslington, York YO10 5DD, UK (e-mail: g.altmann@psych.york.ac.uk).

review. The full depth, richness and scope of psycholinguistics thus goes far beyond the limits afforded here.

Psycholinguistics boomed (as did the rest of psychology) in the early to mid-1960s. The Chomskian revolution (e.g. Chomsky, 1957, 1965, 1968) promoted language, and specifically its *structures*, as obeying laws and principles in much the same way as, say, chemical structures do. The legacy of the first 50 or so years of the 20th century was the study of language as an entity that could be studied independently of the machinery that produced it, the purpose that it served, or the world within which it was acquired and subsequently used. The philosopher Bertrand Russell (1959) was sensitive to this emerging legacy when he wrote: 'The linguistic philosophy, which cares only about language, and not about the world, is like the boy who preferred the clock without the pendulum because, although it no longer told the time, it went more easily than before and at a more exhilarating pace.' Subsequently, psycholinguistic research has nonetheless recognized the inseparability of language from its underlying mental machinery and the external world.

The review begins with some brief comments on the early days of psycholinguistics (including both early and current British influences on the field). It then moves to a selection of current topics in psycholinguistics, beginning with the language abilities of newborn infants, and moving on from how infants represent the speech they hear to how they acquire a first vocabulary and how later, as adults, they represent and access words in the mental lexicon (both spoken and written). From there, we move on to the acquisition of grammatical skills in children and the processing of sentences by adults and to text and discourse understanding. The article then considers how adults produce, rather than comprehend, language, and ends with a brief overview of some of the topics that are not covered in-depth in this review.

Psycholinguistics: the early days

Psycholinguistics is, as Wilhelm Wundt (1832–1920) noted in *Die Sprache* (1900), as much about the mind as it is about language. All the more paradoxical, then, that perhaps the earliest use of the term 'psycholinguistics' was in J. R. Kantor's *Objective psychology of grammar* (1936), in which Kantor, an ardent behaviourist, attempted to refute the idea that language reflected any form of internal cognition or mind. According to Kantor, the German psycholinguistic tradition was simply wrong. The term became more firmly established with the publication in 1954 of a report of a working group on the relationship between linguistics and psychology entitled *Psycholinguistics: A survey of theory and research problems* (Osgood & Sebeok, 1954/1965); the report was published simultaneously in two journals that, separately, served the linguistics and psychology disciplines. Almost 50 years on, research into the many different aspects of the psychology of language is now published in a vast range of journals, and accounts for around 10% of all publications in psychology,[1] a figure that has remained remarkably constant given the approximately fivefold increase in the annual publication rate across psychology as a whole since the 1950s.

[1] The figure is estimated from a variety of keyword searches through the PsycLIT database (American Psychological Association). It is possibly a generous estimate of the publication output that would fall under the psychology of language rubric.

Psycholinguistics suffered a turbulent history during the first part of the 20th century, not least because of the behaviourist movement. Even William James, who foresaw many psycholinguistic issues in his *The principles of psychology* (1980, 1950), had turned his back on Wundtian psychology at the very end of the 19th century. Blumenthal (1970), in his historical overview of the early years (and on which parts of this section are based), described psycholinguistics in the early to mid-20th century as the study, in the West at least, of verbal learning and verbal behaviour—a reflection of the behaviourist approach to language learning (the more mentalist approach advocated by Wundt still prevailed in German, and to an extent Soviet, psychology during that time). Within linguistics, the Bloomfieldian school was born (with Bloomfield's *Language* published in 1933) which, although acknowledging the behaviourist endeavour within psychology, promoted the study of language independently of psychology, and took to the limits the taxonomic approach to language. Notwithstanding the behaviourist backdrop, a significant number of empirical studies reported phenomena in those early days that still predominate today (mostly on reading or speech perception; e.g. Bagley, 1900; Cattell, 1886; Dodge & Cline, 1915; Huey, 1900, 1901; Pillsbury, 1915; Pringle-Morgan, 1896; Stroop, 1935; Tinker, 1946). Theoretically, the field moved on (or at least, should have done) following Karl Lashley's (1951) article on serial order in behaviour. Despite no reference to Wundt, there were considerable similarities with the Wundtian tradition. Specifically, Lashley sought to show that the sequential form of an utterance is not directly related to the syntax of that utterance (a theme to be found in Wundt's writings, and later taken up by the Chomskian school), and that (partly in consequence) the production of an utterance could not simply be a matter of complex stimulus–response chains as the behaviourist movement would have it. Skinner, in his *Verbal behaviour* (1957), took on-board some of these limitations of behaviourism when, despite advocating that psychology abandon the mind, he argued for a system of internal mediating events to explain some of the phenomena that the conditioning of verbal responses could not explain. The introduction of such mediated events into behaviourist theory led to the emergence of *neo-behaviorism*, most notably associated, within language, with Charles Osgood.

The year 1957 was something of a watershed for psycholinguistics, not because of the publication of *Verbal behaviour*, but because of the publication of Chomsky's *Syntactic structures* (1957)—a monograph devoted to exploring the notion of grammatical rules. Subsequently, in his review of Skinner's *Verbal behaviour*, Chomsky (1959) laid to rest the behaviourist enterprise (at least as it applied to language). Space precludes the breadth of argument, but crudely speaking no amount of conditioned stimulus-to-verbal-response associations could explain the infinite productivity (and systematicity) of language. With Chomsky, out went Bloomfield, and in came mental structures, ripe for theoretical and empirical investigation. Chomsky's influence on psycholinguistics, let alone linguistics, cannot be overstated. Although there have been many critics, specifically with regard to his beliefs regarding the acquisition of grammar (see under 'From words to sentences' below), there is little doubt that Chomsky reintroduced the mind, and specifically mental representation, into theories of language (although his beliefs did not amount to a theory of psychological process, but to an account of linguistic structure). Indeed, this was the sticking point between Chomsky and Skinner: Skinner ostensibly eschewed mental representations, and Chomsky proved that language was founded on precisely such representation. Some commentators (e.g. Elman *et al.*, 1996) take the view, albeit tacitly, that the Chomskian revolution threw out the associationist baby with the behaviourist bathwater. Behaviourism was 'out', and with it associationism also. Symbolic computation was 'in', but with it, uncertainty over how the symbolic system was acquired (see under 'From words to sentences' below). It was not until the mid-1980s that a new kind of revolution took place, in which the associationist baby, now grown up, was brought back into the fold.

In 1986 Rumelhart and McClelland published *Parallel distributed processing* (1986b; see Anderson & Rosenfeld, 1998, for an oral history of the topic, and R. Ellis & Humphreys, 1999, for an explanation and examples of its application within psychology). This edited volume described a range of connectionist, or neural network, models of learning and cognition.[2] 'Knowledge' in connectionist networks is encoded as patterns of connectivity distributed across neural-like units, and 'processing' is manifest as spreading patterns of activation between the units. These networks can learn complex associative relations largely on the basis of simple associative learning principles (e.g. Hebb, 1949). Importantly, and in contrast to the ideals of the behaviourist traditions, they develop internal representations (see under 'From words to sentences' below). The original foundations for this paradigm had been laid by McCulloch and Pitts (1943) and further developed by Rosenblatt (1958). Rumelhart and McClelland's collection marked a 'coming of age' for connectionism, although many papers had already been published within the paradigm. One of the most influential models in this mould was described by Elman (1990; and see M. I. Jordan, 1986, for a precursor), who showed how a particular kind of network could learn the dependencies that constrain the sequential ordering of elements (e.g. phonemes or words) through time; it also developed internal representations that appeared to resemble grammatical knowledge. Not surprisingly, the entire enterprise came under intense critical scrutiny from the linguistics and philosophy communities (see e.g. Marcus, 1998a, 1998b; Pinker & Mehler, 1988), not least because it appeared to reduce language to a system of statistical patterns, was fundamentally associationist, and eschewed the explicit manipulation of symbolic structures: the internal representations that emerged as a result of the learning process were not symbolic in the traditional sense.

Critics notwithstanding, statistical approaches to language (both in respect of its structure and its mental processing) are becoming more prevalent, with application to issues as diverse as the 'discovery' of words through the segmentation of the speech input (e.g. Brent, 1999; Brent & Cartwright, 1996), the emergence of grammatical categories (Elman, 1990), and even the emergence of meaning as a consequence of statistical dependencies between a word and its context (e.g. Burgess & Lund, 1997; Elman, 1990). Empirically also, the statistical approach has led to investigation of issues ranging from infants' abilities to segment speech (Saffran, Aslin, & Newport, 1999) and induce grammar-like rules (Gomez & Gerken 1999, 2000) to adult sentence processing

[2] Connectionist models are computer simulations of interconnecting cells or units which, when activated, pass that activation along to the other units to which they connect. The amount of activation that passes between two units is modulated by the strength of the connection between them, and the net activation of a unit is determined by its net inputs and a sensitivity function that combines those inputs. Various learning algorithms exist to set the strengths automatically so that a given input pattern of activation across some set of units will spread through the network and yield a desired output pattern of activation across some other set of units. Crucially, these algorithms allow multiple input–output pairings to be learned. See Rumelhart and McClelland (1986b) for the 'first wave' of connectionist modelling, and Altmann (1997) for a non-specialist introduction to how such models work.

(MacDonald, 1993, 1994; MacDonald, Pearlmutter, & Seidenberg, 1994a; Trueswell, 1996; Trueswell, Tanenhaus, & Kello, 1993).

This is where we are now. There is no doubt that connectionism has had a profound influence on psycholinguistic research and cognitive psychology more generally. But despite its attractions (for some at least), it would be disingenuous to ignore the insights and historical convergence among the other disciplines within psychology, linguistics and philosophy that have brought us this far, and which will, like connectionism, take us further.

In the 100 years that have passed since the inception of the British Psychological Society, psycholinguistics has developed into a fully fledged scientific discipline. It is appropriate, in the context of this anniversary issue of the *British Journal of Psychology*, to draw attention to the British influence on that development—an influence that continues to pervade the field. Specific examples of how topics within the field owe their development in major part to British researchers include Morton's and subsequently Marslen-Wilson and Tyler's influence on the development of models of lexical process and representation (concerning the access and organization of the mental dictionary); Cutler and Norris's work on prelexical segmentation processes (the breaking down of the spoken input into representational units that are relevant for lexical access); Mitchell's work on, among other things, language-specific constraints on syntactic processing, and Steedman and Altmann's work on contextual constraints on such processing; Johnson-Laird's influence on the development of 'mental models' (representations of text and discourse); Sanford and Garrod's, and Garnham's, work on inferential processing and referential continuity during text processing (the inferences and representations that enable the hearer/reader to interpret the dependence between an expression in one part of the text and earlier parts of the text); Bryant, Goswami and others on reading and its development; Snowling, Oakhill, Frith and Bishop on disorders of reading and of language more generally (including disorders associated with dyslexia, autism and specific language impairment); Marshall, Shallice, Warrington, and A. W. Ellis on the neuropsychology of language breakdown (following brain injury); and other researchers too numerous to mention, but each of whom has played a significant part in the development of the field as it stands today. The following sections review that field. However, given that it is often difficult to disentangle British influences on psycholinguistics from the other international influences that have contributed to its progress, no attempt is made to do so explicitly in the review that follows.

Language and infancy

It is *in utero* that the foundations are most commonly laid for subsequent language learning and adult language use. It was established in the 1980s that perhaps the first linguistic variation to which newborn babies are sensitive is prosody (variation in the pitch, intensity and duration of the sounds of speech—the *melody*, so to speak). Babies appear to learn the prosodic characteristics of 'material' they hear *in utero*. DeCasper and colleagues (e.g. Cooper & Aslin, 1989; DeCasper, Lecanuet, Busnel, Granier-Deferre, & Maugeais, 1994; DeCasper & Spence, 1986) demonstrated that newborns recognize—indeed prefer—the prosodic characteristics of the maternal voice, as well as the characteristics of particular rhymes spoken repeatedly by the mother during the last weeks of pregnancy. Mehler *et al*. (1988) demonstrated that newborn babies recognize, more generally, the prosodic 'signature' of their mother tongue, even though they have yet to learn the segmental characteristics of their maternal language (the specific sounds, and their combinations, that define the words in the language). Thus, aspects of language can be learned *in utero* and without a 'semantics'; it is not necessary for linguistic variation to map onto meaning for that variation to be learned, even though the greater part of language learning is concerned with establishing precisely such a mapping.

The newborn baby is armed, however, with more than just an appreciation of the prosodic characteristics of what will probably become its mother tongue. It is armed also with an ability to recognize, in a particular way, the individual sounds of the language (the phonemes) which, combined in different ways, give rise to the words of the language. Liberman, Harris, Hoffman, and Griffith (1957) demonstrated that phonemes are perceived *categorically*—despite an almost infinite range of sounds that could make up the dimension along which the initial phonemes of the words 'buy' and 'pie' vary, we appear to perceive just two phonemes; /b/ and /p/. Eimas, Siqueland, Jusczyk, and Vigorito (1971) demonstrated that this mode of perception is not learned, but is present in young infants, and Bertoncini, Bijeljac-Babic, Blumstein, and Mehler (1987) demonstrated subsequently that it is present even in newborns (and see Nakisa & Plunkett, 1998, for a computational account based on a genetic learning algorithm). And although not all languages use the same categories within a given dimension (Thai, for example, has an extra phoneme where we only have /b/ and /p/), babies appear sensitive to all used categories (e.g. Lasky, Syrdal-Lasky, & Klein, 1975; Streeter, 1976) until around 8–10 months, by which time they have lost their earlier sensitivity to categories that are not relevant within their own language (e.g. Werker & Lalonde, 1988; Werker & Tees, 1984). Our perception of these categories is modulated by a variety of influences: for example, Ganong (1980) demonstrated that if a segment that is ambiguous between /b/ and /p/ replaces the final segment of the word 'clap' it will tend to be perceived as /p/, but the same acoustic token at the end of 'blab' will be perceived as /b/. Also Summerfield (1981) demonstrated that the perceived rate of speech modulates perception—the /p/ uttered in 'pie' (spoken quickly) could be acoustically identical to the /b/ uttered in 'buy' (spoken normally); and yet we would still perceive the first word as 'pie'. Infant perception is also modulated in this way (e.g. Miller & Eimas, 1983). Thus, our interpretation of the acoustic input is determined by our interpretation (at a variety of different levels of analysis) of the surrounding input.

Liberman *et al*.'s (1957) original observation was partly responsible for the idea that the manner in which we perceive speech is uniquely human and quite speech-specific. For a time, it was believed that there existed phoneme detectors that operated in much the same way as motion detectors (e.g. they could be 'fatigued'; Eimas & Corbit, 1973; but see Ades, 1974, for evidence against position-independent phoneme detectors). However, it since transpired that many of these effects are not confined to human perceivers: a range of other species perceive phonemes categorically (e.g. Kuhl & Miller, 1975), with their perception also modulated by speech rate (Stevens, Kuhl, & Padden, 1988). The precise mechanism that brings about the appearance of discontinuous perception is the subject of some considerable controversy: Massaro (1987, 1994) has pointed out that perception could be continuous but that the application of a decision rule (operating preconsciously) would lead naturally to the appearance of discontinuities in the appropriate identification and discrimination functions. Nonetheless, it would appear that the newborn infant brings with it into the world a perceptual mechanism that is neither specific to humans nor to speech, but which endows it with some considerable advantage. A problem for the infant is to know that different instances of the same word *are* the same word; categorical perception may provide the infant with a solution to that problem.

The relevance of these observations on prosodic sensitivity and discontinuous perception of phonemes concerns the nature of the mental representations that are constructed on the basis of the novel input that the newborn encounters. Newborns apparently recognize what they hear in terms of syllabic units, and anything that is not a 'legal' syllable is neither recognized nor distinguished in the same way (e.g. Bertoncini & Mehler, 1981; Mehler, Dupoux, & Segui, 1990). Only legal syllables have the prosodic characteristics that the infant is already familiar with, and the infant therefore recognizes syllables through recognizing familiar prosodic patterns. Presumably, the infant subsequently can categorize these familiar patterns in terms of their phonemic content also.

To conclude: the newborn infant is set up to organize what it hears in linguistically relevant ways, as if it were born to recognize the building blocks of the words it will learn subsequently. This ability need not be based on some innate, *language-specific* mechanism, but need only be based on a mechanism, perhaps statistical in nature, with which to learn the prosodic tunes of the language (a statistical regularity in its environment), and on a mechanism shared with other species with which to identify and discriminate finer segmental information in the face of linguistically irrelevant variation.[3] For the infant, language is not an independent entity divorced from the environment in which it is produced and comprehended; it is a part of that environment, and its processing utilizes mental procedures that may not have evolved solely for linguistic purposes.

Contacting the lexicon I: spoken word recognition

The importance of a syllabic basis to early linguistic representations pervades the literature on *lexical access*—the manner in which the mental representations of the words in the language are accessed. In the early 1980s, research on English and French established syllable-bounded representations as central to the access process (e.g. Cutler, Mehler, Norris, & Segui, 1986; Mehler, Domergues, Frauenfelder, & Segui, 1981); the syllabic structure of the maternal language apparently could influence the nature of the representations that 'contact' the mental lexicon following auditory input. Thus, French has a syllabic structure (and indeed, a prosodic structure) that is different in significant ways from English, and similarly for languages such as Spanish, Catalan or Japanese (cf. Otake, Hatano, Cutler, & Mehler, 1993; Sebastián-Gallés, Dupoux, Segui, & Mehler, 1992). How these representations, as reactions to the speech input, develop from infancy onwards has only recently been explored (see Jusczyk, 1997, for a review). But all the indications are that the prosodic/syllabic attributes of the language being learned have a fundamental influence on the sensitivities of the infant, as do statistical regularities in the

[3] Although other species appear to share with humans some of the mechanisms that have been postulated to underpin the learning of language, they do not share with humans the same capacity (or any capacity, in some cases) for language. In part this may reflect the evolutionary pressures that have accompanied the population by particular species of specific evolutionary niches (they may not have needed, to survive, the social organization that may otherwise facilitate the evolution of language); see Deacon (1997) for further discussion.

language (see Jusczyk, 1999, for a review; and Saffran *et al.*, 1999, for an empirical demonstration of statistical learning in infants). The infant language device is, again, a product of the environment in which it finds itself, and appears to be at the mercy of the statistical regularities within that environment.

Learning words

The task for the infant as it begins to acquire a lexicon, and learn the meanings of words, is by no means simple (see Bloom, 2000, for a recent review on word learning): how are children to know which of the many sounds they hear correspond to which of the infinite range of possibilities before them? For example, children may be able to work out that, among the sounds in the spoken utterance 'look, the dog's playing with a ball', the sounds corresponding to 'dog' are intended to correspond to the animal in front of them (perhaps because they already know that 'ball' refers to the ball, and have a sufficient grasp of syntax to realize that 'dog' is a noun and will hence refer to something). But children must still work out whether 'dog' corresponds to the concept associated with dogs, or with animals more generally; or to things of that shape, or to things of that colour; or to its head, or to all of it. Given the infinite number of hypotheses that children might test (Quine, 1960) how are they to reject all but the correct one? An early suggestion was that the child is armed with certain innate primitive concepts, and that as primitive hypotheses they either undergo continual revision and modification (e.g. Bruner, Oliver, & Greenfield, 1966), or are innately ordered so that the child 'guesses' the basic-level concept before the superordinate or subordinate concept (e.g. J. A. Fodor, 1981; see also J. A. Fodor, 1998). More recently, it was proposed that children are constrained, or biased, to interpret words in certain specific ways (see Markman, 1990, for a review). Thus, children tend to assume that nouns refer to whole objects rather than to their parts or their substance (Gentner, 1982; Markman & Hutchinson, 1984); that nouns are labels for objects of the same shape (e.g. Imai, Gentner, & Uchida, 1994; Landau, Jones, & Smith, 1992; see Smith, 1995, for a review); that nouns are labels for objects of the same kind ('dog' applies to poodles and alsations) rather than for objects that have some relationship ('dog' applies to dogs and bones—Markman & Hutchinson, 1984); and that each object can only have one label (Markman & Wachtel, 1988; cf. E. V. Clark, 1987). However, the evidence for these constraints is based on relatively weak statistical trends, and despite initial optimism there is growing evidence that their explanatory power is limited, and that these constraints may in fact *result* from early lexical development, rather than *guide* it (e.g. Nelson, 1988, and see below).

How children acquire the meanings of verbs has enjoyed greater consensus (but see under 'From words to sentences' below). R. Brown (1957) first demonstrated that children can use their knowledge of syntax (see the next section) to constrain their interpretation of words. Thus, the (non-)word 'sib' is interpreted differently depending on the syntactic context: 'In this picture, you can see sibbing/a sib/sib'. Subsequent studies demonstrated that children as young as 2 years who are watching an action described by a verb can use the syntactic context within which the verb occurs to determine transitivity (whether or not a verb takes a grammatical object): e.g. 'Big Bird is gorping with Cookie Monster' vs. 'Big Bird is gorping Cookie Monster' (see Gleitman, 1990, for a review). Thus, the acquisition of verb meaning requires a basic syntactic competence (to which we return below in 'From words to sentences'). Indeed, a basic syntactic competence is also implicated in the acquisition of noun meaning: R. Brown's (1957) demonstration included 'see a sib' ('sib' is a *count noun*, as is 'dog', for example) and 'see sib' ('sib' here is a *mass noun*, as is 'butter'), and children were sensitive to this syntactically marked distinction. The fact that the acquisition of both nouns and verbs is sensitive to syntactic context suggests a common theme. Smith (1999; Smith, Jones, & Landau, 1996) has argued that biases such as those discussed above in respect of early noun learning may result from general associative learning principles; in particular, that regular association between one perceptual cue (e.g. the syntactic form of a description) and another (whatever is being referred to) causes perception of the first cue to direct attention to the second (cf. 'goal-tracking' in animal learning research; W. James, 1890/1950; Rescorla & Wagner, 1973). For example, the object-shape bias may arise because of an early association between descriptions of the form '. . . a dog' or '. . . the dog' and the statistical regularities that define membership of the class of objects that can be described as 'dog'. Crucially, the first names that children learn are for objects whose names refer to categories of objects of similar shape, and not similar colour, substance or function (and equally crucially, the shape bias emerges only after a certain number of nouns have been learned). Thus, the syntactic configuration ('the/a X') can cue the perceptually relevant cue (e.g. shape) through basic associative learning processes. In principle, an equivalent account should be possible of the acquisition of verb meaning through syntactic cueing (see under 'From words to sentences' below).

More recently, Burgess and Lund (1997) described an approach to the acquisition of meaning which takes further some of the principles embodied in recent connectionist models (e.g. Elman, 1990). They describe a computational model which calculated the co-occurrence statistics for words in a sample of language; words that have similar meanings will tend to co-occur with the same kinds of other words. Using a multi-dimensional scaling technique, they were able to show how the different words in the language grouped together along dimensions of similarity that could be interpreted as semantic—thus, semantic 'categories' emerged as a function of the co-ocurrence patterns of the words in the language. Of course, this demonstration could not take into account the grounding of word meaning in the external world, but the principle (meaning as knowledge of the context in which a word occurs) is the same. This principle pervades contemporary theories of the nature of conceptual structure—theories of what constitutes 'knowing' or 'having' a concept. The early view (e.g. Katz & Fodor, 1963) assumed that a concept was a list of necessary and sufficient features that constituted membership of a category. Given the problems inherent in such a definitional approach (one problem being that of exceptions), alternatives were soon adopted: the 'family resemblance' account (e.g. Rosch & Mervis, 1975) assumes that a concept is an abstraction of the commonalities across different instances; the 'exemplar' account assumes that membership of a category is dependent on similarity to stored exemplars (e.g. Medin & Schaffer, 1978); accounts based on 'schemata' assume the encoding of prototypical attributes of a member of the category and the associated encoding of how these attributes interrelate (see Rumelhart, 1980, for an overview); and the 'explanation-based' approaches (e.g. Johnson-Laird, 1983; Murphy & Medin, 1985) assume that a concept includes information about the interaction between members of the category and other objects in the world, as well as information about the relationships between the different attributes of each of those members. These later approaches tend towards accounts in which concepts are abstractions across multiple experiences of exemplars of a category, with the abstraction encoding both attributes of the exemplars themselves, and the contingent (predictive) relationships between these attributes and attributes of the context (causal or otherwise). Once again, predictive structure in the environment is seen as determining cognitive representation (see McRae, de Sa, and Seidenberg (1997) for discussion of correlational approaches to featural representation and meaning; and Komatsu (1992) for a review of alternative views of conceptual structure).

Accessing words

Somehow, words are learned and their meanings acquired, and the result of this learning process is a mental lexicon in which each of 60 000 to 75 000 words can be distinguished uniquely from each of the others on a variety of dimensions. Research into the factors that influence the manner in which adult lexical access proceeds has a long history. There is a range of phenomena associated with word recognition that has been studied over the course of the last century, although perhaps the most commonly cited phenomena have been that words are recognized faster if they follow a semantically related word than an unrelated word (the 'semantic priming' effect; D. E. Meyer & Schvaneveldt, 1971; see also Moss & Gaskell, 1999), that they are also more easily recognized if embedded in appropriate sentential contexts (Bagley, 1900; Marslen-Wilson, 1973; Marslen-Wilson & Welsh, 1978), that words that are frequent in the language are recognized more quickly than words that are infrequent (Savin, 1963), and that words can be recognized before their acoustic offsets (e.g. Marslen-Wilson, 1973; Marslen-Wilson & Tyler, 1975, 1980). An early insight into the processes of lexical access was that lexical representations are not like dictionary entries to be *accessed*, but are representations to be *activated* (Morton, 1969, 1970). Morton's logogen model was instrumental in its influence on contemporary theories of lexical access, and was quite distinct from models which assumed a process analogous to a serial search through a lexicon in which the entries are ordered in some way (cf. Forster, 1979). Within Morton's model, word detectors, which stored a word's visual, phonological and semantic properties, would become activated as a function of the auditory (or visual) input; once they reached threshold, they would 'fire'. Influences on recognition times, such as word frequency or context, would manifest themselves as changes to the recognition threshold or resting level activation (frequency) or as dynamic changes to the activation level of the logogen (context). Subsequently, Marslen-Wilson, Tyler, and colleagues (Marslen-Wilson, 1987; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978) developed the 'cohort' model of spoken word recognition (see McClelland & Elman, 1986, for an influential computational variant).

In the cohort model, words' representations are activated as a function of the fit with the acoustic input, with mismatch against the input causing a decrease in activation. Like the logogen model, all potential candidate representations are activated (cf. Marslen-Wilson, 1987; Zwitserlood, 1989) but, unlike the logogen model, there is no threshold beyond which they 'fire', so information concerning the word's phonological or semantic properties becomes activated as a function of that acoustic fit (although different semantic properties become available more rapidly than others; Moss, McCormick, & Tyler, 1997; see also McRae, de Sa, & Seidenberg, 1997). Another difference relative to the earlier logogen model concerns the manner in which contextual information influences the selection of lexical hypotheses; in the cohort model, context does not modulate the activation of a word's representation (as it does in the logogen model), but rather modulates the process by which active candidates are subsequently selected for integration with the ongoing syntactic and/or semantic analysis. Finally, word frequency effects are manifest within the cohort model as differences in the sensitivity of the function relating goodness-of-fit to activation, with high frequency words having a faster rise-time than low frequency words (Marslen-Wilson, 1990).

More recently, Marslen-Wilson and Warren (1994) established that the smallest acoustic details can influence the activation (up or down) of candidates, suggesting that the speech input is not encoded as an interpreted sequence of phonemes, or syllables, prior to its match against stored lexical representations (and see Gaskell & Marslen-Wilson, 1997, for a connectionist interpretation). This renders the prior observation regarding sensitivity to syllabic structure mildly paradoxical: on the one hand, it appears as if the language-specifics of syllable structure play an important part in determining the segmentation of the spoken utterance into representational units that subsequently contact the lexicon (cf. Cutler *et al.*, 1986; Cutler & Norris, 1988); on the other hand, refinements to the cohort model suggest that the syllable, despite its ontological significance, is not the unit of lexical access. In fact, there is no paradox here: if 'segmentation' of the spoken utterance reflects the 'cutting up' of the speech input into 'chunks' which then contact the lexicon, the acoustic details which are matched against the lexicon need not correspond to those on which basis the input is segmented. However, 'segmentation' need not reflect any 'cutting up' as such, but may instead reflect constraints on the goodness of fit between acoustic input and lexical representation—statistical properties of the language may render certain lexical hypotheses more likely than certain others, given the surrounding acoustic input, and these statistical properties are likely to include constraints on syllabic structure.

An enduring puzzle for proponents of the cohort model has been how a word-recognition system based on establishing goodness-of-fit against the acoustic input could cope with the range of noise (extraneous and intrinsic) within that input. People often mispronounce words, sometimes in 'lawful' ways: 'hand' might be pronounced as 'ham', and 'thin' as 'thim' in the context of 'hand me the thin book' (uttered as 'hameethethimbu'), and yet it is well-established that even slight mispronunciations cause significant reduction in activation of the intended candidate (Marslen-Wilson, 1993; Marslen-Wilson & Warren, 1994). However, Gaskell subsequently demonstrated that whereas, for example, 'thim' does not ordinarily activate the representation for 'thin', it *does* do so just in those cases where such variability is lawful given the surrounding phonetic context (in this case, a subsequent bilabial)—'thim girl' does not, 'thim boy' does (Gaskell & Marslen-Wilson, 1996, 1998). Moreover, a computational system that is sensitive only to statistical regularities in the input is quite able to learn the occasions on which such activation is or is not appropriate (Gaskell, Hare, & Marslen-Wilson, 1995). Once again, the interpretation of input is determined by a combination of that input and its surrounding context.

A defining feature of the cohort model is that, given an input compatible with more than one alternative, the alternatives are activated in parallel as a function of their goodness-of-fit to the acoustic input and their frequency, with some modulation, at some stage within the process, by surrounding context. There are a number of conditions under which the input may be compatible with more than one alternative lexical candidate. The first is simply that speech input is noisy, and a given stretch of sound may be compatible with a number of alternative candidates (with differing degrees of fit). A second condition obtains when different candidates might be activated by different but overlapping parts of the input. Shillcock (1990) demonstrated that the lexical representations for both ‘wombat’ and ‘bat’ will be activated when hearing ‘put the wombat down’, even though ‘bat’ is neither intended nor compatible with the prior input (there is no word ‘wom’ which could end where ‘bat’ would begin); see Gow and Gordon (1995) and Vroomen and de Gelder (1997) for constraints on such activation, and Norris (1994) for computational issues surrounding such overlap. A third condition under which multiple alternatives will be activated obtains for homophones—words which sound the same (and hence share the same acoustic input) but mean something quite different. Historically, the main theoretical, and empirical, concerns have included whether all meanings are indeed activated in parallel; whether more frequent meanings are activated to a greater extent than less frequent ones; and whether sentential context influences the activation of the relevant/irrelevant meanings in some way (see Simpson, 1984, 1994, for a review). Towards the end of the 1970s, it appeared that alternative meanings *are* activated in parallel (Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979), and constraining sentential context does not prevent the activation of the irrelevant meanings. However, these studies did not determine whether the alternatives were activated to the same extent. In fact, they are not: the dominant, or more frequent, meaning appears to be more accessible (cf. Duffy, Morris, & Rayner, 1988; Tabossi, Colombo, & Job, 1987), with sentential context able to make the non-dominant meaning as accessible as the dominant one, although not, apparently, more accessible (e.g. Duffy *et al.*, 1988). See Lucas (1999) for a meta-analysis of the different studies, and Tabossi and Zardon (1993) for conditions under which only the contextually appropriate meaning *is* activated.

A final issue in this section concerns the fact that many words are *morphologically complex*, and are composed of a root and one or more affixes (e.g. the verb ‘review’ + affix ‘er’ = the noun ‘reviewer’). How are such words represented in the mental lexicon? Taft and Forster (1975) argued that the root word is located (through a process of ‘affix-stripping’), and then a list of variations on the root word is then searched through (see also Taft, 1981). Marslen-Wilson and colleagues (e.g. Marslen-Wilson, Tyler, Waksler, & Older, 1994) have provided extensive evidence to suggest that polymorphemic words are represented in terms of their constituent morphemes (with an entry/representation for ‘review’, and an independent entry/representation for the affix ‘er’). However, the evidence also suggests that morphologically complex words which are *semantically opaque* are represented as if they were monomorphemic (the meaning of ‘casualty’, for example, is not related to ‘causal’, hence the opaqueness). Thus some morphologically complex words are represented in their decomposed form (as distinct and independent morphemes), while others are not. Determinants of whether a word is represented in decomposed or whole-word form include semantic transparency, productivity (whether other inflected forms can also be derived), frequency and language (see Marslen-Wilson, 1999; McQueen & Cutler, 1998, for reviews). In respect of the *access* of these forms, for *phonologically transparent* forms, such as 'reviewer', the system will first activate, on the basis of 'review', the corresponding stem. It will then activate some abstract representation corresponding to the subsequent suffix 'er', and the combination of these two events will cause the activation of the corresponding meaning. For *phonologically opaque* forms, such as 'vanity' (from 'vain'), the phonetically different forms of the same stem would map directly onto (and cause the activation of) that abstract representation of the stem (making the strong prediction, hitherto untested, that the sequence /van/ should prime not only 'lorry', but also 'conceit').

Theories concerning the acquisition, representation and processing of *inflectional* affixes (e.g. 'review' + affix 'ed' = past tense 'reviewed') have been particularly controversial. The controversy has centred on the traditional belief that children's *overregularization* of irregular verbs points incontrovertibly to the acquisition of rules that become over-applied. Much of the debate has focused on the acquisition of past tense verb forms. There are approximately 180 verbs in the English language that do not obey the traditional 'add -ed' rule of past tense formation. Thus, whereas 'walk' becomes 'walked' and 'research' becomes 'researched', 'run' becomes 'ran', 'go' becomes 'went', 'hit' stays as it is, and 'eat' becomes 'ate'. Children initially get both regulars and irregulars right, but then pass through a stage when they regularize the irregulars (saying 'goed', for example) before a final stage when they get the irregulars right again (e.g. Ervin, 1964; see also Marcus *et al.*, 1992). The behavior looks rule-driven, with the first stage indicative of some form of rote learning, the second stage indicative of the acquisition of a productive rule, and the third stage indicative of both rule application and rote memorization of irregulars. The controversy stems from the demonstration that a connectionist model, based on the extraction of statistical regularities in the environment, apparently could exhibit this same staged learning behaviour in the absence of explicit rule-driven processing (Rumelhart & McClelland, 1986a). Pinker and Prince (1988) argued against the particular input representations employed in the model, and against the assumptions embodied in its training schedule concerning the changing ratio of regulars and irregulars in the child's input (as well as arguing against connectionist models of language more generally). Some of these criticisms were addressed in subsequent, and equally (if not more) successful, models of the developmental profile of verb morphology (e.g. Plunkett & Marchman, 1991, 1993; Seidenberg & McClelland, 1989; see also Marcus, 1995, for a dissenting view of the success of such models; and Marslen-Wilson & Tyler, 1998, for review of the neural correlates underlying the processing of regular and irregular forms, and implications for the debate). It is testimony to the progress that controversy engenders that Bloom (1994, p. 770) ends a brief review of this controversy with: 'it might not be unreasonable to expect this very specific issue—Why do children overregularize and why do they stop?—to be resolved within some of our lifetime.' In all likelihood, Bloom is right.

Contacting the lexicon II: the written word

Evolution has only twice brought about the encoding and transmission of information in durable form: the first time through the genetic code, and the second time through the written word.[4,5] Some of the earliest research on reading was concerned with establishing the perceptual unit(s) of word recognition (with the perceptual identification of such units being the precursor, ultimately, to the extraction of meaning). For example, Cattell (1886) first reported the somewhat paradoxical finding that there are occasions when words can be recognized faster than individual letters. Subsequently, Reicher (1969) confirmed this 'word superiority' effect (see also T. R. Jordon & Bevan, 1996), with Baron and Thurston (1973) demonstrating an equivalent effect for letters embedded in pronounceable vs. unpronounceable non-words (see also McClelland & Johnston, 1977). These later data posed a challenge to one of the earliest models of letter recognition (the Pandemonium model; e.g. Selfridge & Neisser, 1960), which had assumed, in effect, that the only input to the letter identification process was a prior stage of featural analysis. The word-superiority effect implied that higher-level information could feed into the letter identification process (although the non-word data implied that it need not be lexical-level information). This finding led subsequently to the development of McClelland and Rumelhart's (1981) interactive activation model of letter perception (a connectionist model), which contained 'units' (cf. detectors) at the featural, letter and word levels, with letter-level units receiving information from both the featural and word levels. The model explained the word superiority effect in terms of feedback from the lexical level to the letter level, and the pronounceable non-word ('pseudoword') superiority effect in terms of analogy to real words (so 'mave' would cause activation of the word units for 'pave', 'cave', 'mate' and so on, which in turn would feed activation back down to the letter level).

The McClelland and Rumelhart model embodied the claim that letters are not recognized one-by-one as if in isolation; instead, their recognition is modulated by their surrounding context. Research by Evett and Humphreys (1981), among others (see also M. Coltheart, 1981; McClelland, 1976; Rayner, McConkie, & Zola, 1980), suggested, moreover, that letters are not recognized as letters *per se*, but are recoded into an abstract orthographic code that is independent of typeface. They found that strings of letters presented briefly in lowercase, whether words or non-words, primed subsequent words presented in uppercase if the second (word) string shared letters with the first (see Forster, 1993, for a discussion of the claim that changing case precludes low-level visual summation in this paradigm). More recently, T. R. Jordan (1990, 1995) has demonstrated that abstract orthographic information (on a letter-by-letter basis) is not the sole determinant of word identification; coarser shape information (spanning more than one letter) can also be recruited to the process of word identification (cf. Cattell, 1886; see Henderson, 1982, for an historical overview).

Although recognition of a word's physical characteristics, at some abstract level of encoding, is a necessary prerequisite to word identification, other factors mediate the recognition process also: word frequency (e.g. Forster & Chambers, 1973); familiarity (e.g. Connine, Mullenix, Shernoff, & Yelens, 1990; Gernsbacher, 1984); concreteness

[4] The 'written word' would of course include words that have never touched paper (but are stored on computer media), and words that are not used for the purposes of communicating with other humans (e.g. computer code written in hexadecimal).
[5] Technically, oral language constitutes the encoding and transmission of information in durable form, to the extent that cultural transmission (cf. oral histories) is durable. In which case, it is noteworthy that on both occasions (encoding in DNA and in oral language) evolution accompanied the transmission of information with the development of mechanisms for the encoding of grammar.

(C. T. James, 1975); and age of acquisition (Carroll & White, 1973; Lyons, Teer, & Rubenstein, 1978). (See also with regard to age of acquisition Gilhooly and Watson (1981) for an early review; Morrison and Ellis (1995) for more recent evidence; and A. W. Ellis and Lambon Ralph (2000) for a connectionist perspective.) With regard to the latter variable, research in Japanese (Yamazaki, Ellis, Morrison, & Lambon-Ralph, 1997) showed that naming of words written with a single Kanji character was influenced by both the age at which the word was acquired and the age at which the character was learned. Age of acquisition (like the other variables) has also been shown to influence reading accuracy in children (V. Coltheart, Laxon, & Keating, 1988; Laxon, Coltheart, & Keating, 1988). The number of meanings of a word also influences recognition: words that have more than one meaning are recognized faster than words with just one meaning. This result is consistent with the more general findings concerning *neighbourhood effects* (cf. M. Coltheart, Davelaar, Jonasson, & Besner, 1977). Here, words with many neighbours, defined in terms of letter overlap, tend to be identified faster than words with fewer neighbours, although the effect is generally more noticeable with low-frequency words (Andrews, 1989). An important factor here is not necessarily the number of neighbours, but their frequencies relative to the target word (Grainger, 1990; Jared, McRae, & Seidenberg, 1990). Such results are easily accommodated within the successors to the original McClelland and Rumelhart (1981) interactive activation model (e.g. Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989; but see Spieler & Balota, 1997).

Space precludes discussion of all the factors influencing word identification, but one final one concerns the regularity of the pronunciation of the word; words with regular pronunciations (e.g. 'mint') appear to be identified in a qualitatively different manner than words with irregular pronunciations (e.g. 'pint'), a distinction embodied in the *dual-route* model of word recognition (M. Coltheart, 1978; see also Humphreys & Evett, 1985). According to this model, regularly spelled/pronounced words are identified by translating the spelling of the word into its sounds and then accessing the word's lexical representation via that phonological encoding, whereas irregular words are mapped directly against their lexical representations. Considerable evidence for such a distinction comes from a double dissociation observed in acquired dyslexia—reading problems that arise following brain damage. Here, *surface dyslexics* are impaired in their reading of irregular words (often pronouncing them as if regular; e.g. Marshall & Newcombe, 1980), implying damage to the direct lexical route, while *phonological dyslexics* have little problem with irregular words but have difficulty pronouncing pronunceable non-words, implying damage to the phonological route (e.g. Shallice & Warrington, 1980). Interestingly, interactive activation models are able to model such data without the need to postulate distinct processing systems (Plaut, 1997; Plaut *et al.*, 1996; Plaut & Shallice, 1994; Seidenberg & McClelland, 1989). They also model successfully the finding that the effects of regularity impact more on low-frequency words than on high-frequency ones (Andrews, 1982; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). This interaction with frequency is also apparent in studies of the confusions that participants make when having to categorize, for example, 'meat', 'meet', or 'melt' as food; van Orden (1987) reported considerable errors for the homophone 'meet' (see Lukatela, Lukatela, & Turvey, 1993, for a priming study), with Jared and Seidenberg (1990) noting that this effect occurred primarily for low-frequency words. This frequency by consistency-of-spelling interaction is also mediated by a semantic variable, imageability (Strain, Patterson, & Seidenberg, 1995), with low-frequency irregularly spelled words named faster if they were more imageable (see Plaut, 1997, for how this three-way interaction can be accommodated within connectionist models of reading). Taken together, the data suggest that high-frequency words tend to be recognized 'directly', and low-frequency words via an element of phonological recoding, with other factors such as the richness of the semantic representation (cf. imageability) helping to overcome the problems inherent in recognizing low-frequency irregularly spelled words.

Learning to read

Contrary to popular belief, just as we are not taught to comprehend spoken language, so we are not taught to read. What we *are* taught, under the guise of learning to read, is remarkably limited; we are taught that certain sounds correspond to certain letters on the page, that (in English at least) the correspondence is often dependent on position and/or the identity of surrounding letters, and that this correspondence is often quite unpredictable. But aside from specific examples of the mapping between printed and spoken word, little else is given explicitly. What children *do* with that information is left largely to the individual child.

Until the early 1990s it was generally agreed that children go through a series of stages as they develop their reading skills (e.g. Frith, 1985; Gough, Juel, & Griffith, 1992; Marsh, Friedman, Welch, & Desberg, 1981; Morton, 1989; Seymour & Elder, 1986). According to such accounts, the first stage involves using idiosyncratic visual cues as a basis for associating a printed word with its spoken form. As these cues cease to differentiate between the growing number of words entering the child's (sight) vocabulary, they gradually become more refined (relying less on course overall word shape and crude letter information). With increasing vocabulary size, and explicit instruction, the child internalizes the relationship between letters and sounds, and uses this relationship to recognize novel words (cf. Share, 1995). To begin with, the relationship may apply only to some letters within each word; only later will it be applied systematically across the word (Ehri, 1992). Finally, a shift occurs whereby the skilled reader bypasses the phonological route and uses a more direct orthographic route for the more frequent words in the language. More recently, an alternative conception of the learning process has arisen (e.g. Goswami & Bryant, 1990; Harm & Seidenberg, 1999; Snowling, Hulme, & Nation, 1997), based on advances in connectionist modelling (e.g. Plaut, 1997; Plaut *et al*., 1996). According to this more recent view, staged-like reading behaviour is an emergent characteristic of a unitary and continuous learning process during which orthographic, semantic and phonological factors each influence recognition. What changes as learning proceeds is the relative balance of these factors as vocabulary size increases and words are learned with different phonological characteristics (e.g. regular vs. irregular spelling), semantic characteristics (e.g. high vs. low imageability) and (among other differences also) frequencies of occurrence.

Eye movements during reading

Many of the effects described above on isolated word recognition can be observed also in the patterns of eye movements during reading (see Rayner, 1998, for a review, as well as an early review of eye movement research by Tinker, 1946). For example, frequent words engender shorter fixation times (Inhoff & Rayner, 1986), whereas lexically ambiguous words such as 'bank' often engender longer fixation times (Rayner & Duffy, 1986), as do syntactically ambiguous words (Frazier & Rayner, 1982). Various cognitive processes also influence fixation durations, including the reanalyses that are required following an initially incorrect choice of grammatical structure in cases of syntactic ambiguity (Frazier & Rayner, 1982—see the next section), the resolution of anaphoric dependencies between a referring expression and its antecedent (e.g. Ehrlich & Rayner, 1983—see under 'Sentences, discourse and meaning' below), and the additional 'wrap-up' processes that occur at the ends of clauses or sentences (Just & Carpenter, 1980). The sentential context also influences fixation times: the reductions in subsequent fixation duration because of parafoveal preview when the previewed word is highly predictable are far greater than when it is less predictable (Ehrlich & Rayner, 1981).

When reading text, information is taken up from more than just the currently fixated word. McConkie and Rayner (1975, 1976) demonstrated that information is taken up from a perceptual window spanning a few characters to the left of the current fixation point and 14–15 characters to the right. This 'perceptual span' varies as a function of orthography, with 'denser' orthographies, such as Japanese Kanji, having smaller spans (Ikeda & Saida, 1978). From within the perceptual span, the currently fixated word will be identified, but words in the parafovea will not be; instead, partial word information based on coarse letter information will aid identification of that parafoveal word when it is subsequently fixated (Rayner, 1975; Rayner, Well, Pollatsek, & Bertera, 1982; Underwood & McConkie, 1985). This effect appears to be mediated by abstract non-letter specific information (Rayner *et al.*, 1980), as well as by phonological information (Pollatsek, Lesch, Morris, & Rayner, 1992). This latter study measured fixation times to a target word when, on the previous fixation (when the target was in parafoveal view), a homophone had appeared in that position (the homophone was then replaced by the target during the saccade to the target position). Fixation times were reduced for homophones, and also (but less so) for orthographically related words (relative to unrelated words). Surprisingly, semantically related words do not provide any such advantage: if the word 'song' is replaced during the saccade from the previous fixation by the target word 'tune', there is no advantage relative to an unrelated word in place of 'song' (Rayner, Balota, & Pollatsek, 1986).

Despite these many factors which influence fixation times (and there are more), the main determinant of fixation times is word length (longer words requiring longer fixations; Just & Carpenter, 1980). Nonetheless, models of eye-movement control (e.g. Reichle, Pollatsek, Fisher, & Rayner, 1998), which attempt to predict fixation times and saccadic movements through the sentence, have to take each of these factors into account.

From words to sentences

The meaning of a sentence goes beyond the meaning of its component words; in English, the ordering of those words can change quite fundamentally the meaning conveyed by them: 'The man ate up all the fish' implies no more fish; 'The fish ate up all the man' implies no more man. The convention in English for taking the elements before the verb as (generally) indicating the person/thing doing the action, and the elements after the verb as the person/thing at which the action was directed, is a convention of *grammar*. 'The man ate up all the fish' means something quite different from 'Yuki stroked the cat', and yet there are commonalities in meaning because of their shared *syntactic structure*—the man and Yuki did the actions (they are the grammatical *subjects*), and the fish and the cat were the things the actions were directed at (they are the grammatical *objects*). Consequently, the *dependency* between 'The man' and 'the fish' is the same as that between 'Yuki' and 'the cat'. The syntactic structure of a sentence reflects, simply, the dependencies, such as these, that exist within a sentence between its component elements.

How children acquire knowledge of the range and significance of such dependencies—the rules of grammar—has been the subject of considerable attention over the last few decades. In part this has been because of an apparent paradox: if children do not know the syntactic categories (noun, verb and so on) of novel words, how can they induce the rules that govern their ordering? But if children do not know these rules, how can they deduce the relevant syntactic categories from the positions of individual words in the sentence? Broadly speaking, three classes of solution have been proposed to break the paradox. The first assumes that children converge on a body of grammatical knowledge through gradual refinement of non-grammatical representations (e.g. Karmiloff-Smith, 1979): they calculate the distributional properties of each word (their positions relative to the other words in each sentence) and cluster words and phrases together that have similar properties until these clusters gradually come to resemble categories such as noun, verb and so on (cf. the Burgess & Lund, 1997, model mentioned earlier). Pinker has argued against such an approach because of the sheer number of distributional facts that would have to be encoded, many of which would have no relevance whatsoever to the correct categorization of words (e.g. Pinker, 1987, 1995). Instead, he argues for a *semantic bootstrapping* procedure by which children determine the semantic category associated with the meaning of a word (these categories are 'given'), and then determine the syntactic category associated with that word on the basis of crude innate knowledge about the mappings between semantic and syntactic categories (Pinker, 1984, 1987). Once children have induced a body of syntactic knowledge in this way, they can determine the distributional characteristics of the categories, and can then use those characteristics to determine the syntactic category of novel words (when, perhaps, the semantic categories they have available are too crude to determine the syntactic category of the novel word). Of course, how those crude mappings between semantic and syntactic categories 'enter' the genome is unclear. The third class of solution to the learnability paradox has been proposed by Gleitman (see Gleitman, 1990; Gleitman & Gillette, 1995, for reviews). Her *syntactic bootstrapping* hypothesis maintains that the structure of an event that a child sees (in terms of causal relationships, numbers of participants and so on) guides the child's interpretation of the corresponding sentence, and conversely, that the child's interpretation of the structure of the sentence guides the child's attention within the scene. If a child knows the meaning of the words 'Daddy' and 'peas' and hears 'Daddy is eating peas' while viewing the corresponding scene, he or she will be able to induce both the meaning of the verb 'eat' and the syntactic rule which determines that, in English at least, the subject (most generally the causal agent) precedes the verb, and the object (referring to the thing that the action is directed at) follows it. Indeed, even if the child only knew the meaning of 'Daddy', but knew also that '-ing' tended to occur at the ends of verbs, not nouns, this same rule could be induced, as well as the meaning of 'peas'. The acquisition of verb meaning is thus inseparably bound to the acquisition of syntactic (and event) structure; the child's task is not to map individual words onto individual objects or actions, but to map sentences onto events (and vice versa).

The semantic bootstrapping hypothesis requires a degree of innate language-specific knowledge that neither of the other hypotheses requires. Gleitman's syntactic bootstrapping hypothesis (a misnomer given that the bootstrapping relationship between syntax and semantics is reciprocal) and the distributional approach are in fact quite similar, and both are compatible with the proposal put forward by Smith in respect of the early acquisition of word meaning (see under 'Contacting the lexicon I' above). Research on the connectionist modelling of grammatical knowledge can also inform the debate (see Elman *et al*., 1996, for a review). Elman (1990) described an influential model in which a connectionist network had to learn a fragment of English. The network was presented with sequences of short sentences, one word after another, and its task was to learn to predict what the next word in its input would be. Although it could not predict the actual next word, it could predict a *range* of words corresponding to the ones that, in its experience, could occur in that subsequent position given the words that had preceded it (i.e. given the context). It predicted classes of words corresponding to nouns and verbs, and to transitive and intransitive verbs (and finer distinctions still). In effect, it induced syntactic categories on the basis of a distributional analysis of its input: it encoded the predictive contingencies between a word and its context in such a way that words which overlapped in respect of their contextual dependencies would overlap in respect of the internal representations that developed within the network. Contrary to Pinker's objections (see above), the model did not encode irrelevant dependencies between words in its input, because the nature of the prediction task meant that only *predictive* dependencies would be encoded (see Altmann, 1997, for a description of how and why the model worked, and how it could be extended to encode 'meaning'). More recently, Altmann and Dienes (1999) and Dienes, Altmann, and Gao (1999) demonstrated how a simple extension to this model could learn to map structure in one domain onto structure within another—precisely the task required if, as in Gleitman's approach, structure in language is to be mapped onto structure in the world, and vice versa. Such *emergentist* approaches to grammar learning, and language learning more generally, are summarized in both Elman *et al*. (1996) and MacWhinney (1999).

The controversy surrounding the emergence of grammatical competence was initiated in part by Chomsky's assertions regarding a language acquisition device akin to a mental organ (e.g. Chomsky, 1968; see Bates & Goodman, 1999, for a concise refutation of the Chomskian argument). However, Chomsky's influence extended further: the early 1960s saw the initiation of a considerable research effort to validate the psychological status of syntactic processing (the construction of representations encoding the dependencies mentioned at the beginning of this section), and to show that perceptual complexity was related to linguistic complexity, as defined by *transformational grammar* (Chomsky, 1957, 1965). However, it soon became apparent (e.g. J. A. Fodor & Garrett, 1966) that whereas the syntactic structures postulated by transformational grammar had some psychological reality (not surprisingly, given that they reflect aspects of meaning also), the devices postulated by linguistics for building those structures (e.g. the transformations that formed a part of the grammatical formalism) did not (see J. A. Fodor, Bever, & Garrett, 1974; and Valian, 1979, for a review). Subsequently, the emphasis shifted, in large part following Bever's lead (Bever, 1970), towards examination of the psychological mechanism (as opposed to the linguists' equivalents) by which syntactic dependencies are determined during sentence processing—*parsing*. Specifically, Bever pointed out that in cases of ambiguity, where more than one dependency (or structure) might be permissible, the human parser exhibits consistent preferences for one reading rather than another; thus, despite the grammaticality of 'the horse raced past the barn fell' (cf. 'the car driven past the garage crashed'), the preference to interpret 'raced' as a main verb (instead of as a past participle equivalent to 'driven') is so overwhelming that the sentence is perceived as ungrammatical (and the preference is said to induce a 'garden path' effect). Other examples of ambiguity lead to less extreme perceptions, but nonetheless demonstrate the parser's preferences: 'he delivered the letter he had promised her last week' (the delivery may have occurred last week); 'he put the ball in the box on the shelf' (the ball may already have been in the box), and 'she watched the man with the binoculars' (the man may have had the binoculars). These examples (and there are many others) all permit more than one interpretation, and yet there is a very strong tendency to adopt the interpretation that is the alternative to the one implied in parentheses. Following Bever, a number of researchers (most notably Frazier) articulated various operating principles that would give rise to such preferences (e.g. J. D. Fodor & Frazier, 1980; Frazier, 1979, 1987; Frazier & Clifton, 1995; Frazier & Fodor, 1978; Kimball, 1973, 1975; Wanner, 1980, 1987; Wanner & Maratsos, 1978). Crucially, these preferences were determined not by the alternative *meanings* that could be derived at the point of ambiguity, but by the alternative *structures*. Frazier's work was particularly influential because it maintained that these preferences arose as an inevitable consequence of the mental machinery and the principles which governed its operation.

The mid-1980s saw the beginnings of a shift in the theory underlying ambiguity resolution. Crain and Steedman (1985), and then Altmann and Steedman (1988), proposed that what really mattered was the context within which a sentence was being understood. They argued that the preferences observed previously were an artefact of the manner in which sentence processing had hitherto been studied: most studies investigated the processing of single sentences divorced from the natural contexts in which they might normally occur (there were notable exceptions, including perhaps the first demonstration of contextual influences on parsing, Tyler and Marslen-Wilson (1977)). They, and subsequently others, demonstrated that these preferences could be changed if the sentences being studied were embedded in appropriate contexts (e.g. Altmann, Garnham, & Dennis, 1992; Altmann, Garnham, & Henstra, 1994; Altmann, Garnham, van Nice, & Henstra, 1998; Altmann & Steedman, 1988; Liversedge, Pickering, Branigan, & van Gompel, 1998; Spivey-Knowlton & Sedivy, 1995; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993; Trueswell & Tanenhaus, 1991). Thus, decisions regarding which structure to pursue do after all appear to be informed by the meaning(s) associated with the alternatives.

At about the same time, the focus of research into parsing turned to languages other than English, following Cuetos and Mitchell's (1988) finding that the preferences described by Frazier (1987) were not universal across languages; Spanish, for example, appeared to exhibit the opposite of a preference observed in English. This finding challenged not only the purely structural accounts of parsing preferences (if the structures are equivalent across the languages, why the differences?) but also the accounts based on contextual influences (insofar as these accounts made claims also about what should happen when sentences are processed in isolation, cf. Altmann & Steedman, 1988). Evidently, parsing was guided by a complex interplay of factors. Indeed, the 1990s saw a further shift: an alternative to the structure-based theories, already apparent in earlier research (e.g. Altmann & Steedman, 1988; Bates & MacWhinney, 1987; Ford, Bresnan, & Kaplan, 1982; MacWhinney, 1987), began to predominate parsing research. This alternative views parsing as a process of constraint-satisfaction (e.g. MacDonald *et al.*, 1994a; Trueswell & Tanenhaus, 1994), in which sentence processing consists of the application of probabilistic constraints, in parallel, as a sentence unfolds, with no single constraint being more or less privileged than any other except in respect of its probabilistic strength. This latter approach is predicated not simply on those prior demonstrations of contextual influence, but also on demonstrations that other factors such as lexical frequency, plausibility and so on can also influence the resolution of syntactic ambiguity (e.g. MacDonald, 1993, 1994; MacDonald *et al.*, 1994a; MacDonald, Pearlmutter, & Seidenberg, 1994b; Pearlmutter & MacDonald, 1995; Spivey-Knowlton & Sedivy, 1995; Spivey-Knowlton *et al.*, 1993; Trueswell, 1996; Trueswell & Tanenhaus, 1994; Trueswell, Tanenhaus, & Garnsey, 1994; Trueswell *et al.*, 1993).

In parallel with concerns over the human parser's resolution of ambiguity, there developed a concern over the manner in which aspects of the *meaning* of a sentence are derived as the sentence unfolds through time, and specifically that aspect of meaning associated with the assignment of *thematic roles*. These roles are, crudely speaking, the roles that the participants play in the event being described by the sentence: in 'the man ate the fish', the man is the agent of the eating, and the fish the patient of the eating (the thing being eaten). The verb defines the appropriate roles given the event, and the grammar determines where (in English) the participants filling particular roles will be referred to within the sentence. It is this relationship, between aspects of meaning and knowledge of grammar, that places thematic role assignment at the interface between syntax and semantics (cf. Carlson & Tanenhaus, 1988; Mauner, Tanenhaus, & Carlson, 1995; Tanenhaus, Boland, Mauner, & Carlson, 1993; Tanenhaus, Carlson, & Trueswell, 1989; Tanenhaus, Garnsey, & Boland, 1990). An influential account of parsing in which aspects of the role-assignment process govern the parsing process was developed by Pritchett (1988, 1992). Subsequently, a number of studies investigated the possibility that verb-based information (contained within a verb's lexical entry), as opposed to grammatical information more generally, can 'drive' the parsing process (e.g. Boland, Tanenhaus, & Garnsey, 1990; Boland, Tanenhaus, Garnsey, & Carlson, 1995; Ford *et al.*, 1982; McRae, Ferretti, & Amyote, 1997; Mitchell, 1987, 1989; Mitchell & Holmes, 1985; Trueswell *et al.*, 1993). Indeed, there was a corresponding shift in linguistic theory also, with the advent of lexicalized grammars (cf. Ades & Steedman, 1982; Bresnan, 1982; Joshi, 1985; Steedman, 1987, 1990). This research led, most recently, to an account of sentence processing in which the human parser uses verb-based information to actively *predict*, at the verb, what kinds of linguistic expression will come next and which things in the context these expressions might refer to (Altmann, 1999; Altmann &

Kamide, 1999). Thus, in a context in which a boy takes a chocolate bar out of his pocket, a subsequent sentence fragment such as 'he ate . . .' appears to be interpreted, at 'ate', to mean that the thing that was eaten was the previously mentioned chocolate, even though the grammatical position associated with this patient role (the post-verbal grammatical object) has not yet been encountered (and even though the boy could eat some other, hitherto unmentioned, food). In effect, thematic role assignments can precede, in sufficiently constrained contexts, the point in the sentence at which grammatical information would ordinarily license the assignment.

Research on the importance of verb-based information led naturally to consideration of parsing in languages whose grammars dictate that the verb appears at the end of each sentence (as is the case in, for example, Japanese and, in certain circumstances, German). For example, Kamide and Mitchell (1999) recently described data suggesting that, in Japanese, the parsing process is not driven by verb-based information. They proposed than an initial sequence of nouns and their associated role-markers allows the parser to predict properties of the verb that must follow. In this case, the theory is similar to that described above in connection with parsing as a predictive process (Altmann, 1999): a sequence of nouns can constrain what will follow (and can allow representations to be activated which reflect the anticipation of what will follow) in much the same way as a verb, in English, can constrain what will follow it. It is thus conceivable that essentially the same processing account may be applicable to languages with such diverse grammars as English and Japanese.

Sentences, discourse, and meaning

Establishing the roles played out in an event, and using grammatical information to determine which of these roles is associated with which particular referring expressions within the sentence, is just one aspect of the derivation of meaning; those participants have to be identified, and the meaning of the sentence integrated with the meaning of (at least some part of) what has come before. Much research over the last 30 or so years has been concerned with these two processes (identification and integration), as well as with the nature of the dynamically changing mental representations that encode integrated meanings both within and *across* individual sentences.

It has been known for many years that we do not maintain an accurate record of the precise words that make up the sentences in a text or discourse. Instead, as soon as the propositional content of a sentence (in effect, the message to be conveyed) has been integrated within the discourse representation, the sentence's surface form (the precise ordering of words and associated grammatical structure that realizes the message) is lost, and only the propositional content remains (e.g. Bransford & Franks, 1971; Sachs, 1967; see also Bartlett, 1932). Moreover, these and other studies (e.g. Garnham, 1981; Glenberg, Meyer, & Lindem, 1987) suggested that it is not even the propositional content of the individual sentences that is maintained, but rather some representation of the situation described or elaborated on in each sentence (reflecting in effect the state of the world and how it has changed). Thus, what is available for subsequent processing is not the semantic content of each sentence, but rather the content that results from integrating that sentence (or its propositional content) within the discourse. Its specific propositional content is then, in effect, forgotten. This distinction between surface form (words and their ordering), propositional content (the specific message conveyed by the sentence) and situation (the state of the world) pervades contemporary theories of discourse representation and process (e.g. Kintsch, 1988; Sanford & Garrod, 1981). Much of the work on the representation of situation was inspired by Johnson-Laird and colleagues' work on *mental models* (e.g. Garnham, 1981; Johnson-Laird, 1983), although work within the formal traditions of linguistics and philosophy was also influential (e.g. Barwise & Perry, 1981). The mental model approach to discourse and text representation assumed that the end-product of comprehension is, in effect, a mental analogue of the situation described (see Altmann, 1997, for a more complete description of this analogue).

Various elaborations of the mental models approach have taken place, with greater emphasis on the processes by which the model is constructed and the factors that influence the construction process (e.g. Kintsch, 1988; Sanford & Garrod, 1981). Much of the work on the latter has focused on the processes of *cohesion* and *coherence* (cf. G. Brown & Yule, 1983; Garnham, Oakhill, & Johnson-Laird, 1982). Cohesion refers to the way in which the interpretation of an expression in one sentence depends on the interpretation of expressions in a previous sentence. The most common example of this is referential continuity—the manner in which the *antecedents* of referring expressions such as 'he', 'it', 'the fish', 'the fish the man ate' will generally have been introduced prior to the referring expression. Coherence refers to the way in which one sentence may be related to another through various steps of *inference*, even in the absence of any cohesion, as in the sequence 'Richard was very hungry. The fish soon disappeared'; a different inference would have been made had the first sentence been 'Richard accidentally poisoned the river', with the meaning of 'disappeared' being interpreted quite differently. As first noted by Haviland and Clark (1974), inferences are often required to establish cohesion; in 'Mary unpacked some picnic supplies. The beer was warm', the beer must be inferred on the basis of the previously mentioned picnic supplies. Haviland and Clark observed longer reading times to the second sentence in this case than when it followed 'Mary unpacked some beer', presumably because of the additional inference required. However, Garrod and Sanford (1982) found that it took no longer to read 'The car kept overheating' after 'Keith drove to London' than after 'Keith took his car to London'. They argued that the mental representation constructed in response to 'the car' must contain information about the role that the car could play in the event just described (Keith driving to London). Given the meaning of 'drive', which requires something to be driven, a role is immediately available in a way that it is not in the beer/picnic case. Unlike full referring expressions (e.g. 'the car'), pronouns require explicit antecedents—hence the infelicity of 'Keith drove to London. It kept overheating'—and one function of pronouns is to keep their referents in *explicit focus* (Sanford & Garrod, 1981). This notion of focus, or, from the linguistic perspective, foregrounding (Chafe, 1976), has proved central to theories of discourse representation and process, not least because theories of how focus is maintained, or shifted, are required to explain not simply the form that language can take in certain circumstances (specifically, the form of the referring expressions, as full referring expressions or as pronouns, as definites or as indefinites), but also the ease and immediacy (or otherwise) with which cohesive and inferential linkages can be established (see Marslen-Wilson & Tyler, 1987, for a review of early on-line studies of discourse comprehension).

The interpretation of referring expressions (or *anaphors*) is dependent on both the form of the expression and the state of the discourse representation against which it is being interpreted. The ease with which a full referring expression (e.g. 'the car') can be resolved, and its referent identified, depends on various factors including the degree of coherence between the sentence and the prior discourse or text. The ease with which a pronoun (e.g. 'it') can be resolved depends on the extent to which its antecedent is in focus. Research on the immediacy with which such resolution takes place led Sanford and Garrod (1989) to propose a two-stage process in which the processing system first locates where within the discourse representation the relevant information is located (the *bonding* stage), and then commits itself to a particular interpretation on the basis of that information (the *resolution* stage). It appears that the bonding stage is, under certain circumstances, immediate, but the resolution stage less so—only in very constrained cases is resolution equally immediate (generally a pronoun that bonds to a focused antecedent); in other cases, there is reason to believe the processor delays commitments lest interpretations involving shifts in focus turn out to be required (Vonk, Hustinx, & Simons, 1992).

Discourse and text understanding rely heavily on inferential processes. Some of these are required for successful comprehension (as in the earlier example of Richard as hungry or accident-prone). Others are not required for successful comprehension, but are more 'elaborative' and provide causal (explanatory) coherence (as in 'Bill was rich. He gave away most of his money', where the inference is that it was *because* he was rich that he gave it away). Considerable research effort has focused on what kinds of inference are made and when (see Broek, 1994; Sanford, 1990; Singer, 1994, for reviews). Most of this research has assumed, however, a 'transactional' approach to language (cf. Kintsch, 1994) in which the comprehender is a passive participant in a transaction that involves transmission of information from the speaker/writer to the comprehender. Relatively little research has focused on the 'interactional' approach, more usual of dialogue and other cooperative tasks, in which language is interactive and mediates a cooperative relationship between the conversational parties. Research in this area has largely been pioneered by H. H. Clark and colleagues and by Garrod and colleagues (see H. H. Clark, 1994; and Garrod, 1999, for a review). One important aspect of the interaction concerns the identification of 'common ground' between speaker and hearer (H. H. Clark & Marshall, 1981; Stalnaker, 1978), requiring speaker and hearer to have some representation of what is in the *other*'s discourse representation. A further aspect concerns inferences at a more social level, regarding the speaker's intentions and the hearer's requirements. Indeed, to fully capture and understand the meaning of discourse requires faculties that go well beyond the linguistic.

From meaning to speaking

Most psycholinguistic research has investigated some component or other of the mental machinery that transforms sound (or print) into meaning. Research on the machinery for transforming meaning back into sound (spoken language *production*) had been considered, until the mid- to late 1980s, the 'poor cousin' of psycholinguistic research. One reason for this belief was that it was only in the 1980s that paradigms were developed that enabled researchers to map out the time-course of (aspects of) the production process (much of which was carried out by Levelt and colleagues; see Levelt (1999) and Levelt, Roelofs, and

Meyer (1999) for reviews). Until then, the major input to psycholinguistic theory was based on analysis of the errors that people make when speaking, a tradition originating with the work of Meringer and Mayer at the turn of the 19th century (Meringer & Mayer, 1895). Two fundamental observations were made: that intended words may be substituted by erroneous ones related in meaning, related in sound, or both; and that parts of words could be involved in these errors, with exchanges ('blushing crow' instead of 'crushing blow'), perseverations ('beef needle' instead of 'beef noodle'), anticipations ('pirst part' instead of 'first part'), and blends ('a lot of brothel' instead of 'a lot of bother' or 'a lot of trouble'). Notwithstanding Freud's views on the origins of speech errors as revealing something of the psychodynamic aspects of the individual (e.g. Freud, 1914), the nature of these errors indicates something quite fundamental about the machinery of production. A resurgence of interest in speech errors in the 1970s (e.g. Fromkin, 1973) led to two theories of speech production that still underlie all contemporary accounts of the production process. The first of these (Garrett, 1975) is based on the observations that word exchanges or blends preserve grammatical category (the word is the wrong one, but it is the right part of speech and in the right part of the sentence) and that phonemic exchanges tend to be short-ranged and across words from different grammatical categories. Thus, syntactic function (and grammatical ordering in a language such as English) is determined separately from, and independently of, the ordering of phonemes (and indeed, morphemes—'order' + 'ing'). The second (Shattuck-Hufnagel, 1979) is based on the observation that errors involving phonemes tend to preserve syllable position (e.g. Boomer & Laver, 1968), suggesting that when a word's phonemes are retrieved from the mental lexicon, they are somehow coded for their position within the syllable.

This last account is concerned with the retrieval of the *phonological form* of a word, which is stored separately from a word's conceptual specification. The evidence for this stretches back to the 19th century again, and was elegantly summarized by William James (1890/1950, p. 252): 'The rhythm of a lost word may be there without a sound to clothe it; or the evanescent sense of something which is the initial vowel or consonant may mock us fitfully, without growing more distinct.' The *tip-of-the-tongue* state occurs when we select the conceptual specification of a word (its *lemma*) but somehow fail to retrieve its associated phonological specification (the *lexeme*).

Research into spoken word production accelerated with the development of techniques to induce errors (Baars, Motley, & MacKay, 1975) and tip-of-the-tongue states (R. Brown & McNeill, 1966; see also A. S. Brown, 1991). It was paradoxical, nonetheless, that most of what was known about the production of fluent speech (distinct stages in production, distinctions between lemma and lexeme, planning units, and so on) was derived from the observation of *dysfluencies* in speech production (and not just speech errors, but hesitations and pauses also: e.g. Ford, 1982; Gee & Grosjean, 1983; Holmes, 1988). In the 1980s, a new paradigm emerged, however, that was based in part on the Stroop interference paradigm (Stroop, 1935). Here, a distractor word is seen to interfere with naming a colour or object. Schriefers, Meyer, and Levelt (1990) modified variants of the picture–word interference paradigm (variants originated by Glaser & Düngelhoff, 1984; Lupker, 1979) in the first study to track the time-course with which semantic and phonological information (the lemma and lexeme respectively) are activated. Schriefers *et al*. manipulated the stimulus onset asynchrony (SOA) between the onset of a picture (which the participant had to name) and a spoken word that could be phonologically related, semantically related or unrelated to the name of the picture: both semantically and phonologically related words influenced naming times, but only when semantically related words were presented earlier (relative to picture onset) and phonologically related words were presented later—confirmation that activation of the lemma precedes activation of its phonological form.

This last study marked a new age for research into spoken language production, with this paradigm and subsequently others allowing fine temporal distinctions to be made between different processes underlying spoken word production (e.g. Levelt *et al.*, 1991; Meyer, 1990, 1991; Wheeldon & Levelt, 1995). Most recently, electrophysiological studies and also imaging techniques have been recruited to the armoury of on-line techniques for studying the time-course of lemma selection, morpho-syntactic retrieval and phonological encoding (e.g. Levelt, Oraamstra, Meyer, Helenius, & Salmelin, 1998; van Turennout, Hagoort, & Brown, 1997, 1998; see also Price, 1998, for a review of brain imaging studies). Equally important has been the development of computational models able to explain both the speech error data (Dell, 1986, 1997) and the picture-interference reaction time data (Roelofs, 1992, 1997a, 1997b). The two classes of model (both connectionist) encode similar characteristics, although they differ quite significantly in respect of the flow of information between the different levels (phonological, lemma and so on), with Dell and colleagues' model allowing bidirectional information flow, and Roelof's model allowing bidirectional flow between all but the phonological layers. Crucially, the two models have generated empirically testable predictions which have in turn added to the empirical knowledge base on which contemporary theories are developed (e.g. Jescheniak & Schriefers, 1998; Levelt *et al.*, 1999; Peterson & Savoy, 1998; Roelofs, 1993, 1997a).

The advances that have pervaded research into spoken *word* production have been accompanied also by advances in the study of *sentence* production, most notably with the priming paradigms developed by J. K. Bock and colleagues (e.g. Bock, 1986, 1987; Bock, Loebell, & Morey, 1992). For example, Bock (1986) showed participants a picture of an event that they had to describe (e.g. lightning striking a church) but first presented them with a word that primed either one component ('thunder' for the lightning) or the other ('worship' for the church). Depending on which prime was given, the form of the subsequent sentence that participants uttered would change: 'thunder' might lead to 'lightning is striking the church' while 'worship' might lead to 'the church is being struck by lightning'. Bock (1987) used phonologically related primes, but this time found that the prime caused the related word in the subsequent description to occur towards the end of the sentence, not the beginning. She concluded that the semantic prime makes the appropriate concept more available (or activated), while the phonological prime inhibits the associated concept (in some models, inhibition is required to 'deactivate' a just-uttered concept). These and related findings have suggested that the more activated a concept, the more likely it will be mentioned first—a consequence of a processing system that starts encoding for production as soon as material becomes available. Having encoded the initial portion of the utterance, the system is then constrained in respect of the grammatical encoding of the remainder of the utterance (for reviews, see K. Bock & Levelt, 1994; Ferreira, 2000; Levelt, 1989).

Research into language production is no longer a 'poor relation'. The last two decades have seen methodological advances that enable empirical investigation of questions that are as sophisticated and subtle as any that are posed by researchers of language comprehension.

Psycholinguistics: conclusions and prospects

It is inevitable that a review this of size can capture neither the breadth nor depth of the psycholinguistic research that has contributed to the current state of the art. For some, the most serious omission, given that over 50% of the world's population speak more than one language, will be the wealth of research on bilingualism and second language learning (for reviews, see de Groot & Kroll, 1997; Schreuder & Weltens, 1993). Prominent issues here have concerned the nature of lexical representation and grammatical encoding: questions of lexical organization in the bilingual are concerned with the nature of the representations within the lexicon that are, or are not, shared across the languages, and the determinants of that sharing; and questions of grammatical encoding are similar in respect of the manner in which grammatical knowledge is represented across the languages. A second major omission concerns the considerable body of work on disorders of language following brain trauma (although there was some reference to a small part of this work under 'Contacting the lexicon II' above, but see also Caplan, 1994; A. W. Ellis & Young, 1988). For example, some aphasics have impaired semantic knowledge of living things (e.g. Warrington & Shallice, 1984), others of man-made artefacts (e.g. Warrington & McCarthy, 1983), and these category-specific semantic deficits have provided important insights into the manner in which meaning is encoded across different parts of the cortex (auditory, visual, motor and so on; see Farah and McClelland (1991) for a connectionist perspective; and McRae, de Sa, and Seidenberg (1997) for a recent account of the dissociation based on correlational differences among the features underlying word meaning). Similarly, patterns of semantic breakdown in patients suffering semantic dementia, Alzheimer's dementia or herpes encephalitis also inform models of semantic organization (see Hodges & Patterson, 1997, for a review). But not all deficits are purely semantic, as evidenced by the comprehension deficits in cases of agrammatism (e.g. Zurif & Swinney, 1994). The study of language breakdown has proved critical in constraining theories of normal language function (see also McCarthy, this issue).

A third omission thus far concerns disorders of language development (some of which co-occur with other cognitive developmental disorders; see Goswami, this issue) and the development of language in atypical circumstances. There are many disorders of language development: disorders of reading (e.g. 'dyslexia', which appears to be due most often to an underlying *phonological* deficit; see Snowling, 2000, for a review) and disorders of language understanding and production which go beyond disorders of reading (see Bishop, 1997, for a review). Certain of these disorders affect populations of children that can hardly be described as 'atypical': for example, between 10% and 15% of children (in the UK) are poor readers that can be characterized as having good decoding skills (good application of spelling-to-sound correspondences), age-appropriate word recognition skills, but poor comprehension (Oakhill, 1982, 1984, 1993). Nation and Snowling (1998) showed that their listening comprehension is also poor, with weak word knowledge and poor semantic processing skills. In essence, these children have intact decoding skills but both word-level problems (with distinct patterns of lexical activation as evidenced in priming studies; Nation & Snowling, 1999) and higher-level problems involving the integration of what they hear with the context in which it is heard. Similar but generally more severe comprehension impairments have also been described in the autistic population and in non-autistic hyperlexic children (Snowling & Frith, 1986). The relevance of these disorders (no matter their severity or incidence) is twofold: studying the disorder informs theories of normal development (e.g. Bates & Goodman, 1997, 1999), and understanding the nature of the deficit relative to normal development can help in the construction of intervention techniques that are best suited to the individual child given the nature of his or her deficit (see, for example, Oakhill (1994) for intervention studies aimed at improving the comprehension skills of poor comprehenders, and Olson and Wise (1992) and Snowling (2000) for interventions aimed at improving the decoding skills of dyslexic children).

Examples of *atypical* language development (as distinct from disordered language development) include the acquisition of sign language in deaf communities (see Klima & Bellugi, 1979) and the development of Creole languages in what are in effect multilingual (albeit pidgin) communities (Bickerton, 1983, 1984). The former demonstrates that when the auditory modality is unavailable for language, the visual modality (and motor modality for production) can 'take over'; and interestingly, the same areas of the brain that are implicated in spoken language processing are implicated in signed language processing (Corina, 1999; Nishimura *et al.*, 1999), which is perhaps less surprising when one considers the parallels that exist between the structures, and acquisition, of spoken and signed languages (see, for reviews, Bonvillian, 1999; Poizner, Klima, & Bellugi, 1987; and for an analysis of babbling and its significance in sign language, Petitto & Marentette, 1991). The latter cases, of creolization, illustrate that language acquisition is neither simple imitation nor simple associative processing; if it were, the result would be a straightforward reflection of the linguistic input received during the acquisition process. What makes creolization remarkable is that the first-generation creole speakers in effect generate a language that they have not heard spoken (although whether that language is an abstraction across the input they have been exposed to, or is innately determined, is a controversial issue; see Bickerton, 1984, and associated commentaries).

A final omission (there are still others), concerns the wealth of research on the neuroscience of brain and language (see Gazzaniga, Ivry, & Mangun, 1998, for an accessible introduction; Hickok & Poeppel, 2000; and Price, 1998, for summaries of recent data on word comprehension and production; and Marslen-Wilson & Tyler, 1998, for an example of how such work can inform current psycholinguistic debate). There exists an increasing range of techniques for exploring the neural dynamics of cognitive processing and for making the functional neuroanatomy of the brain, as it goes about its 'daily business', ever more accessible to the cognitive neuroscientist (see Posner & Raichle, 1997, for an accessible review). Our understanding of the neural structure and functioning of the brain and its subparts will inevitably contribute to our understanding of the range, type and properties of neural computation that underlie different mental processes. This in turn will feed into models of how such computations might bring about these distinct mental activities.

While it is both appropriate and necessary to study the adult language faculty independently both of its neural underpinnings and of the manner in which that faculty develops from infancy onwards, one of the lasting lessons from the 20th century is that the adult faculty is an emergent characteristic of a biological system that, in its initial state at least, is as much a device for acquiring language as it is a device for using language (although whether it is a device for acquiring language *per se* is a moot point). The child language system does not suddenly switch off at puberty to be replaced by a system that is the next size up. And although there are discontinuities in the learning curve (e.g. in vocabulary growth and the acquisition of grammar; e.g. Bates, Dale, & Thal, 1995), connectionist modelling has demonstrated how such discontinuities can arise through the operation of a non-linear, but unitary, learning mechanism (e.g. Elman *et al*., 1996). Perhaps the greatest challenge for the next century will be to foster this relationship between theoretical models of learning, the operating principles of which are becoming increasingly well understood, and the empirical study of the biological mechanisms, the workings of which are becoming increasingly accessible, that enable human learning and emergent cognition.

Acknowledgements

I should like to thank Anne Anderson, Dorothy Bishop, Vicki Bruce, Gareth Gaskell, Peter Hatcher, Don Mitchell, Kate Nation and Lorraine Tyler for their comments on an earlier version of this review. They are not responsible for errors of omission, which are my own, and are sadly inevitable given the size of the field, the large number of contributors to the literature, and inevitable constraints on article length.

References

Ades, A. E. (1974). How phonetic is selective adaptation? Experiments on syllable position and vowel environment. *Perception and Psychophysics, 16*(1), 61–66.

Ades, A., & Steedman, M. J. (1982). On the order of words. *Linguistics and Philosophy, 4,* 517–528.

Altmann, G. T. M. (1997). *The ascent of Babel: An exploration of language, mind, and understanding.* Oxford: Oxford University Press.

Altmann, G. T. M. (1999). Thematic role assignment in context. *Journal of Memory and Language, 41,* 124–145.

Altmann, G. T. M., & Dienes, Z. (1999). Rule learning by seven-month-old infants and neural networks. *Science, 284,* 875.

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31,* 685–712.

Altmann, G. T. M., Garnham, A., & Henstra, J. A. (1994). Effects of syntax in human sentence parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory & Cognition, 20*(1), 209–216.

Altmann, G. T. M., Garnham, A., van Nice, K., & Henstra, J. A. (1998). Late Closure in context. *Journal of Memory and Language, 38*(4), 459–484.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264.

Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition, 30*(3), 191–238.

Anderson, J. A., & Rosenfeld, E. (Ed.) (1998). *Talking nets: An oral history of neural networks.* Cambridge, MA: MIT Press.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition, 10,* 565–575.
Andrews, S. (1989). Frequency and neighbourhood size effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(5), 802–814.
Baars, B. J., Motley, M. T., & MacKay, D. (1975). Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behaviour, 14,* 382–391.
Bagley, W. C. (1900). The apperception of the spoken sentence: A study in the psychology of language. *American Journal of Psychology, 12*(1), 80–130.
Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology, 4,* 207–228.
Bartlett, F. C. (1932). *Remembering: An experimental and social study.* Cambridge: Cambridge University Press.
Barwise, J., & Perry, J. (1981). Situations and attitudes. *Journal of Philosophy, 78*(11), 668–691.
Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (pp. 96–151). Oxford: Basil Blackwell.
Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes, 12,* 507–584.
Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29–80). Mahwah, NJ: Erlbaum.
Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–193). Hillsdale, NJ: Erlbaum.
Bertoncini, J., Bijeljac-Babic, R., Blumstein, S. E., & Mehler, J. (1987). Discrimination in neonates of very short CV's. *Journal of the Acoustical Society of America, 82,* 31–37.
Bertoncini, J., & Mehler, J. (1981). Syllables as units in infant speech perception. *Infant Behaviour and Development, 4,* 247–260.
Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 270–362). New York: Wiley.
Bickerton, D. (1983). Creole languages. *Scientific American, 249,* 116–122.
Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences, 7,* 173–221.
Bishop, D. V. M. (1997). *Uncommon understanding.* Hove: Psychology Press.
Bloom, P. (1994). Controversies in language acquisition. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 741–779). San Diego, CA: Academic Press.
Bloom, P. (2000). *How children learn the meanings of words.* Cambridge, MA: MIT Press.
Bloomfield, L. (1933). *Language.* Chicago: University of Chicago Press.
Blumenthal, A. L. (1970). *Language and psychology: Historical aspects of psycholinguistics.* New York: Wiley.
Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 575–586.
Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structure. *Journal of Memory and Language, 26,* 119–137.
Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99,* 150–171.
Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.
Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language, 29,* 413–432.
Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language, 34*(6), 774–806.
Bonvillian, J. D. (1999). Sign language development. In M. Barrett (Eds.), *The development of language* (pp. 277–310). Hove: Psychology Press.
Boomer, D. S., & Laver, J. D. M. (1968). Slips of the tongue. *British Journal of Disorders of Communication, 3,* 1–12.
Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2,* 331–350.
Brent, M. R. (1999). Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Sciences, 3,* 294–301.
Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and photactic constraints are useful for segmentation. *Cognition, 61,* 93–125.

Bresnan, J. (Ed.) (1982). *The mental representation of grammatical relations.* Cambridge, MA: MIT Press.
Broek, V. D. (1994). Comprehension and memory of narrative texts. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539–588). San Diego, CA: Academic Press.
Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109,* 204–223.
Brown, G., & Yule, G. (1983). *Discourse analysis.* Cambridge: Cambridge University Press.
Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology, 55,* 1–5.
Brown, R., & McNeill, D. (1966). The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behaviour, 5,* 325–337.
Bruner, J. S., Oliver, R., & Greenfield, P. M. (1966). *Studies in cognitive growth.* New York: Wiley.
Burgess, C., & Lund, K. (1997). Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Processes, 12,* 177–210.
Caplan, D. (1994). Disorders of language. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 1023–1053). San Diego, CA: Academic Press.
Carlson, G. N., & Tanenhaus, M. N. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), *Thematic relations* (pp. 263–288). New York: Academic Press.
Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology, 25*(1), 85–95.
Cattell, J. McK. (1886). The time it takes to see and name objects. *Mind, 11,* 63–65.
Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. N. Li (Ed.), *Language comprehension and the acquisition of knowledge* (pp. 25–56). Washington, DC: Winston.
Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.
Chomsky, N. (1959). Review of Skinner's *Verbal behaviour. Language, 35,* 26–58.
Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.
Chomsky, N. (1968). *Language and mind.* New York: Harcourt Brace.
Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Erlbaum.
Clark, H. H. (1994). Discourse in production. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 985–1021). San Diego, CA: Academic Press.
Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.
Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). London: Academic Press.
Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language, 3,* 245–286.
Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance* (Vol. VI, pp. 534–555). Hillsdale, NJ: Erlbaum.
Coltheart, V., Laxon, V. J., & Keating, K. (1988). Effects of word imageability and age of acquisition on children's reading. *British Journal of Psychology, 79*(1), 1–12.
Connine, C., Mullenix, J., Shernoff, E., & Yelens, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1084–1096.
Cooper, R. P., & Aslin, R. N. (1989). The language environment of the young infant: Implications for early perceptual development. *Canadian Journal of Psychology, 43*(2), 247–265.
Corina, D. (1999). On the nature of left hemisphere specialization for signed language. *Brain and Language, 69*(2), 230–240.
Crain, S., & Steedman, M. J. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 320–358). Cambridge: Cambridge University Press.
Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition, 30*(1), 73–105.
Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25,* 385–400.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 113–121.
de Groot, A. M. B., & Kroll, J. F. (Eds.) (1997). *Tutorials in bilingualism: Psycholinguistic perspectives.* Mahwah, NJ: Erlbaum.
Deacon, T. (1997). *The symbolic species: The co-evolution of language and the human brain.* London: Penguin Books.
DeCasper, A. J., Lecanuet, J.-P., Busnel, M. C., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behaviour and Development, 17*(2), 159–164.
DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behaviour and Development, 9,* 133–150.
Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.
Dell, G. S. (1997). Lexical access in aphasic and non-aphasic speech. *Psychological Review, 104,* 801–837.
Dienes, Z., Altmann, G. T. M., & Gao, S.-J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science, 23*(1), 53–82.
Dodge, R., & Cline, T. S. (1915). The angle velocity of eye-movements. *Psychological Review, 8,* 145–157.
Duffy, S., Morris, R., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429–446.
Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recording. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–143). Hillsdale, NJ: Erlbaum.
Ehrlich, K., & Rayner, K. (1981). Context effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behaviour, 20,* 641–655.
Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behaviour, 22,* 75–87.
Eimas, P. D., & Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Perception and Psychophysics, 4,* 99–109.
Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171,* 303–306.
Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(5), 1103–1123.
Ellis, A. W., & Young, A. W. (1988). *Human cognitive neuropsychology.* Hove: Erlbaum.
Ellis, R., & Humphreys, G. (1999). *Connectionist psychology.* Hove: Psychology Press.
Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14,* 179–211.
Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press/Bradford Books.
Ervin, S. (1964). Imitation and structural change in children's language. In E. H. Lenneberg (Ed.), *New directions in the study of language* (pp. 163–189). Cambridge, MA: MIT Press.
Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology, 33A,* 325–350.
Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General, 120,* 339–357.
Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 291–330). Hove: Psychology Press.
Fodor, J. A. (1981). The present status of the innateness controversy. In J. A. Fodor (Ed.), *Representations* (pp. 257–316). Cambridge, MA: MIT Press.
Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong.* Oxford: Oxford University Press.
Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar.* New York: McGraw-Hill.
Fodor, J. A., & Garrett, M. F. (1966). Some reflections on competence and performance. In J. Lyons & R. J. Wales (Eds.), *Psycholinguistic papers* (pp. 133–179). Edinburgh: Edinburgh University Press.
Fodor, J. D., & Frazier, L. (1980). Is the human sentence parsing mechanism an ATN? *Cognition, 8,* 417–459.
Ford, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful

relations underlying sentences. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 797–827). Cambridge, MA: MIT Press.

Ford, M., Bresnan, J., & Kaplan, R. M. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 727–796). Cambridge, MA: MIT Press.

Forster, K. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 27–85). Hillsdale, NJ: Erlbaum.

Forster, K. I. (1993). Form-priming and temporal integration in word recognition. In G. T. M. Altmann & R. C. Shillcock (Eds.), *Cognitive models of speech processing: The 2nd Sperlonga meeting* (pp. 467–495). Hove: Erlbaum.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behaviour, 12,* 627–635.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies.* Bloomington, IN: Indiana University Linguistics Club.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance: Vol. XII. The psychology of reading* (pp. 559–586). Hove: Erlbaum.

Frazier, L., & Clifton, C. (1995). *Construal.* Cambridge, MA: MIT Press.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6,* 291–325.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178–210.

Freud, S. (1914). *The psychopathology of everyday life.* London: Ernest Benn.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 119–145). Hillsdale, NJ: Erlbaum.

Fromkin, V. A. (1973). *Speech errors as linguistic evidence.* The Hague: Mouton.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*(1), 110–125.

Garnham, A. (1981). Mental models as representations of text. *Memory & Cognition, 9*(6), 560–565.

Garnham, A., Oakhill, J., & Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. *Cognition, 11,* 29–46.

Garrett, M. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and memory* (pp. 133–177). New York: Academic Press.

Garrod, S. (1999). The challenge of dialogue for theories of language processing. In S. Garrod & M. Pickering (Eds.), *Language processing* (pp. 389–416). Hove: Psychology Press.

Garrod, S., & Sanford, A. J. (1982). Bridging inferences in the extended domain of reference. In A. Baddeley & J. Long (Eds.), *Attention and performance* (Vol. IX, pp. 331–346). Hillsdale, NJ: Erlbaum.

Gaskell, G., Hare, M., & Marslen-Wilson, W. D. (1995). A connectionist model of phonological representation in speech perception. *Cognitive Science, 19*(4), 407–439.

Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*(1), 144–158.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Discriminating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12,* 613–656.

Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 380–396.

Gazzaniga, M., Ivry, R., & Mangun, G. (1998). *Cognitive neuroscience: The biology of the mind.* London: Norton.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology, 15,* 411–458.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relatively versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol. II. Language, thought, and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113*(2), 256–281.

Gernsbacher, M. A. (Ed.) (1994). *Handbook of psycholinguistics.* San Diego, CA: Academic Press.

Gilhooly, K. J., & Watson, F. L. (1981). Word age-of-acquisition effects: A review. *Current Psychological Reviews, 1*(3), 269–286.
Glaser, M. O., & Düngelhoff, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Human Perception and Performance, 7,* 1247–1257.
Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1,* 3–55.
Gleitman, L. R., & Gillette, J. (1995). The role of syntax in verb learning. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language.* Oxford: Blackwell Publishers Ltd.
Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26,* 69–83.
Gomez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109–136.
Gomez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4*(5), 178–186.
Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hove: Erlbaum.
Gough, P. B., Juel, C., & Griffith, P. L. (1992). Reading, spelling, and the orthographic cipher. In P. B. Glough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition.* Hillsdale, NJ: Erlbaum.
Gow, D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance, 21*(2), 344–359.
Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29*(2), 228–244.
Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106*(3), 491–528.
Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and verbal Behaviour, 13,* 512–521.
Hebb, D. O. (1949). *The organization of behavior.* New York: Wiley.
Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic Press.
Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences, 4*(4), 131–138.
Hodges, J. R., & Patterson, K. (1997). Semantic memory disorders. *Trends in Cognitive Sciences, 1,* 68–72.
Holmes, V. M. (1988). Hesitations and sentence planning. *Language and Cognitive Processes, 3,* 323–361.
Huey, E. B. (1900). On the psychology and physiology of reading: I. *American Journal of Psychology, 11*(3), 283–302.
Huey, E. B. (1901). On the psychology and physiology of reading: II. *American Journal of Psychology, 12*(3), 292–312.
Humphreys, G. W., & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual route theory of reading. *Brain and Behavioural Sciences, 8,* 689–740.
Ikeda, M., & Saida, S. (1978). Span of recognition in reading. *Vision Research, 18,* 83–88.
Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development, 9*(1), 45–75.
Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40,* 431–439.
James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1*(2), 130–136.
James, W. (1800/1950). *The principles of psychology.* New York: Dover.
Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29*(6), 687–715.
Jared, D., & Seidenberg, M. S. (1990). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120,* 358–394.
Jescheniak, J. D., & Schriefers, H. (1998). Discrete versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory & Cognition, 5,* 1256–1274.
Johnson-Laird, P. N. (1983). *Mental models.* Cambridge: Cambridge University Press.
Jordon, M. I. (1986). *Serial order: A parallel distributed processing approach* (No. 8604). Institute of Cognitive Science, University of California, San Diego.

Jordon, T. R. (1990). Presenting words without interior letters: Superiority over single letters and influence of postmark boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 16*(4), 893–909.

Jordan, T. R. (1995). Perceiving exterior letters of words: Differential influences of letter-fragment and non-letter-fragment masks. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 512–530.

Jordan, T. R., & Bevan, K. M. (1996). Position-specific masking and the word–letter phenomenon: Reexamining the evidence from the Reicher–Wheeler paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 22*(5), 1416–1433.

Joshi, A. K. (1985). How much context-sensitivity is necessary for characterizing structural descriptions: Tree adjoining grammars. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language processing: Theoretical, computational and psychological perspectives.* New York: Cambridge University Press.

Jusczyk, P. W. (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press/Bradford Books.

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences, 3*(9), 323–328.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 57,* 329-354.

Kamide, Y., & Mitchell, D. (1999). Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes, 14*(5/6), 631–662.

Kantor, J. R. (1936). *An objective psychology of grammar.* Indiana University Publications, Science Series No. 1.

Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference.* Cambridge: Cambridge University Press.

Katz, J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language, 39,* 170–210.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition, 2,* 15–47.

Kimball, J. (1975). Predictive analysis and over-the-top parsing. In J. Kimball (Ed.), *Syntax and semantics* (Vol. 4). New York: Academic Press.

Kintsch, W. (1988). The use of knowledge in discourse-processing: A construction-integration model. *Psychological Review, 95,* 163–182.

Kintsch, W. (1994). The psychology of discourse processing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 721–740). San Diego, CA: Academic Press.

Klima, E., & Bellugi, U. (1979). *The signs of language.* Cambridge, MA: Harvard University Press.

Komatsu, L. K. (1992). Recent views of conceptual structure. *Psychological Bulletin, 112*(3), 500–526.

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced–voiceless distinction in alveolar plosive consonants. *Science, 190, 69–72.*

Landau, B., Jones, S., & Smith, L. (1992). Perception, ontology, and naming in young children: Commentary on Soja, Carey, and Spelke. *Cognition, 43*(1), 85–91.

Lashley, K. S. (1951). The problem of serial order in behaviour. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.

Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology, 20,* 215–225.

Laxon, V. J., Coltheart, V., & Keating, C. (1988). Children find friendly words friendly too: Words with many orthographic neighbours are easier to read and spell. *British Journal of Educational Psychology, 58*(1), 103–119.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press/Bradford Books.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences, 3*(6), 223–232.

Levelt, W. J. M., Oraamstra, P., Meyer, A. S., Helenius, P. I., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience, 10*(5), 553–567.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–45.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98*(1), 122–142.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54,* 358–368.

Liversedge, S. P., Pickering, M. J., Branigan, H. P., & van Gompel, R. P. G. (1998). Processing arguments and adjuncts in isolation and context: The case of by-phrase ambiguities in passives. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*(2), 461–475.

Lucas, M. (1999). Context effects in lexical access: A meta-analysis. *Memory and Cognition, 27*(3), 385–398.

Lukatela, G., Lukatela, K., & Turvey, M. T. (1993). Further evidence for phonological constraints on visual lexical access: TOWED primes FROG. *Perception and Psychophysics, 53*(5), 461–466.

Lupker, S. J. (1979). The role of phonetic and orthographic similarity in picture–word interference. *Canadian Journal of Psychology, 36,* 349–367.

Lyons, A. W., Teer, P., & Rubenstein, H. (1978). Age-at-acquisition and word recognition. *Journal of Psycholinguistic Research, 7*(3), 179–187.

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language, 32*(5), 692–715.

MacDonald, M. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9*(2), 157–202.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994a). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676–703.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994b). Syntactic ambiguity resolution as lexical ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 123–153). Hillsdale, NJ: Erlbaum.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Erlbaum.

MacWhinney, B. (Ed.) (1999). *The emergence of language.* Mahwah, NJ: Erlbaum.

Marcus, G. F. (1995). The acquisition of inflection in children and multi-layered connectionist networks. *Cognition, 56,* 271–279.

Marcus, G. F. (1998a). Can connectionism save constructivism? *Cognition, 66*(2), 153–182.

Marcus, G. F. (1998b). Rethinking eliminative connectionism. *Cognitive Psychology, 37*(3), 243–282.

Marcus, G. F., Ullman, M., Pinker, S., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularisation in language acquisition. *Monographs of the Society for Research in Child Development, 57*(4), 1–165.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science, 14,* 57–77.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints in word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16,* 1–27.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology, 20,* 121–157.

Marsh, G., Friedman, M., Welch, V., & Desberg, P. (1981). A cognitive-developmental approach to reading acquisition. In T. G. Waller & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (pp. 199–221). New York: Academic Press.

Marshall, J. C., & Newcombe, F. (1980). The conceptual status of deep dyslexia: An historical perspective. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 1–21). London: Routledge & Kegan Paul.

Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature, 244,* 522–523.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition, 25,* 71–102.

Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148–172). Cambridge, MA: MIT Press.

Marslen-Wilson, W. D. (1993). Issues of process and representation in lexical access. In G. T. M. Altmann & R. C. Shillcock (Eds.), *Cognitive models of speech processing: The Second Sperlonga Meeting.* Hove: Erlbaum.

Marslen-Wilson, W. (1999). Abstractness and combination: The morphemic lexicon. In S. Garrod & M. Pickering (Eds.), *Language processing* (pp. 101–119). Hove: Psychology Press.

Marslen-Wilson, W., & Tyler, L. K. (1975). Processing structure of sentence perception. *Nature, 257,* 784–785.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1–71.

Marslen-Wilson, W. D., & Tyler, L. K. (1987). Against modularity. In J. Garfield (Ed.), *Modularity in* *knowledge representation and natural language processing* (pp. 57–104). Cambridge, MA: MIT Press/Bradford Books.

Marslen-Wilson, W. D., & Tyler, L. K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Sciences, 2*(11), 428–435.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101*(1), 3–33.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101*(4), 653–675.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29–63.

Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological enquiry.* Hillsdale, NJ: Erlbaum.

Massaro, D. W. (1994). Psychological aspects of speech perception: Implications for research and theory. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 219–263). San Diego, CA: Academic Press.

Mauner, G., Tanenhaus, M. K., & Carlson, G. N. (1995). Implicit arguments in sentence processing. *Journal of Memory and Language, 34*(3), 357–382.

McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 80–91.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.

McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception and Psychophysics, 22,* 249–261.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of the basic findings. *Psychological Review, 88,* 375–407.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578–586.

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8,* 365–368.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics, 5,* 115–133.

McQueen, J., & Cutler, A. (1998). Morphology in word recognition. In A. Spencer & A. M. Zwicky (Eds.), *The handbook of morphology* (pp. 406–427). Oxford: Blackwell.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126*(2), 99–130.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes, 12,* 137–176.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207–238.

Mehler, J., Domergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behaviour, 20,* 298–305.

Mehler, J., Dupoux, E., & Segui, J. (1990). Constraining models of lexical access: The onset of word recognition. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 236–262). Cambridge, MA: MIT Press.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29,* 143–178.

Meringer, R., & Mayer, K. (1895). *Verspechen und Verlesen: Eine psychologischlinguistische Studie. Classics in Psycholinguistics 2.* Amsterdam: John Benjamins Publishers.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29,* 524–545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30,* 69–89.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing words: Evidence of a dependence upon retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.

Miller, J. L., & Eimas, P. D. (1983). Studies on the categorization of speech by infants. *Cognition, 13,* 135–165.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and performance* (Vol. XII, pp. 601–618). Hove: Erlbaum.
Mitchell, D. C. (1989). Verb-guidance and other lexical effects in parsing. *Language and Cognitive Processes, 4*(3/4), SI 123–154.
Mitchell, D. C., & Holmes, V. M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language, 24,* 542–559.
Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 116–133.
Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.
Morton, J. (1970). Word recognition. In J. Morton & J. D. Marshall (Eds.), *Psycholinguistics 2: Structure and processes* (pp. 107–156). Cambridge, MA: MIT Press.
Morton, J. (1989). An information-processing account of reading acquisition. In A. M. Galaburda (Eds.), *From reading to neurons. Issues in the biology of language and cognition* (pp. 43–66). Cambridge, MA: MIT Press.
Moss, H. E., & Gaskell, M. G. (1999). Lexical semantic processing during speech. In S. Garrod & M. Pickering (Eds.), *Language processing* (pp. 59–100). Hove: Psychology Press.
Moss, H. E., McCormick, S. F., & Tyler, L. K. (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes, 12*(5/6), 695–731.
Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316.
Nakisa, R. C., & Plunkett, K. (1998). Evolution of a rapidly learned representation for speech. *Language and Cognitive Processes, 13*(2/3), 105–127.
Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39*(1), 85–101.
Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70,* B1–B13.
Nelson, K. (1988). Constraints on word meaning? *Cognitive Development, 3,* 221–246.
Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., Nishimura, T., & Kubo, T. (1999). Sign language 'heard' in the auditory cortex. *Nature, 397*(6715), 116.
Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*(3), 189–234.
Oakhill, J. V. (1982). Constructive processes in skilled and less skilled comprehenders' memory for sentences. *British Journal of Psychology, 73,* 13–20.
Oakhill, J. V. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology, 54,* 31–39.
Oakhill, J. V. (1993). Children's difficulties in text comprehension. *Educational Psychology Review, 5,* 223–237.
Oakhill, J. (1994). *Individual differences in children's text comprehension.* In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 821–848). San Diego, CA: Academic Press.
Olson, R. K., & Wise, B. W. (1992). Reading on the computer with orthographic and speech feedback. *Reading and Writing, 4,* 107–144.
Osgood, C. E., & Sebeok, T. A. (1954/1965). *Psycholinguistics: A survey of theory and research problems.* Bloomington, IN: Indiana University Press.
Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable: Speech segmentation in Japanese. *Journal of Memory and Language, 32*(2), 258–278.
Pearlmutter, N. J., & MacDonald, M. C. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language, 34,* 521–542.
Peterson, R. P., & Savoy, P. (1998). Lexical selection and honological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 539–557.
Petitto, L. A., & Marentette, P. F. (1991). Babbling in the manual mode: Evidence for the ontogeny of language. *Science, 251,* 1397–1536.
Pillsbury, W. B. (1915). The mental antecedents of speech. *Journal of Philosophy and Scientific Methods, 12,* 116–127.
Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 399–441). Hillsdale, NJ: Erlbaum.
Pinker, S. (1995). Language acquisition. In L. R. Gleitman & M. Liberman (Eds.), *Language: An invitation to cognitive science*. Cambridge, MA: MIT Press/Bradford Books.
Pinker, S., & Mehler, J. (Eds.) (1988). *Connections and symbols*. Cambridge, MA: MIT Press.
Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28,* 73–193.
Plaut, D. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes, 12*(5/6), 765–805.
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computation principles in quasi-regular domains. *Psychological Review, 52*(1), 25–82.
Plaut, D., & Shallice, T. (1994). *Connectionist modelling in cognitive neuropsychology: A case study*. Hove: Erlbaum.
Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition, 38,* 43–102.
Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition, 48,* 21–69.
Poizner, H., Klima, E. S., & Bellugi, U. (1987). *What the hands reveal about the brain*. Cambridge, MA: MIT Press.
Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 148–162.
Posner, M. I., & Raichle, M. E. (1997). *Images of mind*. New York: Scientific American Library.
Price, C. J. (1998). The functional anatomy of word comprehension and production. *Trends in Cognitive Sciences, 2,* 281–287.
Pringle-Morgan, W. (1896). A case of congenital word blindness. *British Medical Journal,* 178.
Pritchett, B. L. (1988). Garden path phenomena and the grammatical basis of language processing. *Language, 64*(3), 539–576.
Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago: University of Chicago Press.
Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
Rayner, K. (1975). Parafoveal identification during a fixation in reading. *Acta Psychologica, 39,* 271–282.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.
Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology, 40,* 473–483.
Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition, 14,* 191–201.
Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12,* 206–226.
Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31,* 537–550.
Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81,* 275–280.
Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.
Rescorla, R. A., & Wagner, A. R. (1973). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning* (Vol. II, pp. 35–51). New York: Appleton-Century-Crofts.
Roelofs, A. (1992). A spreading activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.
Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition, 47,* 59–87.
Roelofs, A. (1997a). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes, 12*(5/6), 657–693.
Roelofs, A. (1997b). The WEAVER model of word-form encoding in speech production. *Cognition, 64,* 249–284.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.
Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review, 65,* 386–408.
Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.
Rumelhart, D. E., & McClelland, J. L. (1986a). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & A. T. P. R. Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 216–271). Cambridge, MA: MIT Press.
Rumelhart, D. E., & McClelland, J. L. (Ed.) (1986b). *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge, MA: MIT Press.
Russell, B. (1959). Foreword to Ernest Gellner's *Words and Things.* London: Gollancz.
Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics, 2,* 437–442.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*(1), 27–52.
Sanford, A. J. (1990). On the nature of text-driven inference. In D. A. Balota, G. B. F. D'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 515–533). Hillsdale, NJ: Erlbaum.
Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language.* Chichester: Wiley.
Sanford, A. J., & Garrod, S. C. (1989). What, when and how: Questions of immediacy in anaphoric reference resolution. *Language and Cognitive Processes, 4*(3/4), SI 235–262.
Savin, H. B. (1963). Word-frequency effects and errors in the perception of speech. *Journal of the Acoustical Society of America, 35,* 200–206.
Schreuder, R., & Weltens, B. (Ed.) (1993). *The bilingual lexicon.* Amsterdam: John Benjamins Publishing Co.
Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language, 29,* 86–102.
Sebastián-Gallés, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language, 31,* 18–32.
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.
Seidenberg, M. S., Walters, G. S., Barnes, M. A., & Tanenhaus, M. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behaviour, 23,* 383–404.
Selfridge, O. G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American, 203,* 60–68.
Seymour, P. H. K., & Elder, L. (1986). Beginning reading without phonology. *Cognitive Neuropsychology, 3,* 1–36.
Shallice, T., & Warrington, E. K. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 119–145). London: Routledge & Kegan Paul.
Share, D. L. (1995). Phonological decoding and self-teaching: *sine qua non* of reading acquisition. *Cognition, 55*(2), 151–218.
Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Studies dedicated to Merrill Garrett* (pp. 295–342). Hillsdale, NJ: Erlbaum.
Shillcock, R. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 24–49). Cambridge, MA: MIT Press/Bradford Books.
Simpson, G. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96,* 316–340.
Simpson, G. B. (1994). Context and the processing of ambiguous words. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 359–374). San Diego, CA: Academic Press.
Singer, M. (1994). Discourse inference processes. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 479–515). San Diego, CA: Academic Press.
Skinner, B. F. (1957). *Verbal behaviour.* NY: Appleton Century Crofts.
Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not induction. In

C. A. Nelson (Ed.), *Basic and applied perspectives on learning, cognition, and development. The Minnesota Symposia on child psychology* (Vol. 28, pp. 1–32). Mahwah, NJ: Erlbaum.

Smith, L. B. (1999). Children's noun learning: How general learning processes make specialized learning mechanisms. In B. MacWhinney (Ed.), *The emergence of language* (pp. 277–304). Mahwah, NJ: Erlbaum.

Smith, L. B., Jones, S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition, 60,* 143–171.

Snowling, M. (2000). *Dyslexia: A cognitive developmental perspective* (2nd ed.). Oxford: Blackwell.

Snowling, M. J., & Frith, U. (1986). Comprehension in 'hyperlexic' children. *Journal of Experimental Child Psychology, 42,* 392–415.

Snowling, M., Hulme, C., & Nation, K. (1997). A connectionist perspective on the development of reading skills in children. *Trends in Cognitive Sciences, 1*(3), 88–91.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science, 8*(6), 411–416.

Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition, 55*(3), 227–267.

Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology, 47*(2), 276–309.

Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics: Vol. 9. Pragmatics* (pp. 315–332). New York: Academic Press.

Steedman, M. J. (1987). Combinatory grammars and human sentence processing. In J. Garfield (Ed.), *Modularity in knowledge representation and natural language processing* (pp. 428–444). Cambridge, MA: MIT Press/Bradford Books.

Steedman, M. J. (1990). Grammar, interpretation, and processing from the lexicon. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 463–504). Cambridge, MA: MIT Press.

Stevens, E. B., Kuhl, P. K., & Padden, D. M. (1988). Macaques show context effects in speech perception. *Journal of the Acoustical Society of America, 84,* 577–578.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1140–1154.

Streeter, L. A. (1976). Language perception of 2-month old infants shows effects of both innate mechanisms and experience. *Nature, 259,* 39–41.

Stroop, J. R. (1935). Studies of interference in serial interactions. *Journal of Experimental Psychology, 18,* 643–662.

Summerfield, Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance,* 7(5), 1074–1095.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)considerations of context effects. *Journal of Verbal Learning and Verbal Behaviour, 18,* 645–659.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research, 49,* 161–167.

Tabossi, P., & Zardon, F. (1993). Processing ambiguous words in context. *Journal of Memory and Language, 32*(3), 359–372.

Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning and Verbal Behaviour, 20,* 289–297.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behaviour, 14,* 638–647.

Tanenhaus, M. K., Boland, J., Mauner, G. A., & Carlson, G. N. (1993). More on combinatory lexical information: Thematic structure in parsing and interpretation. In G. T. M. Altmann & R. C. Shillcock (Eds.), *Cognitive models of speech processing: The Second Sperlonga Meeting.* Hove: Erlbaum.

Tanenhaus, M. K., Carlson, G., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes, 4*(3/4), SI 211–234.

Tanenhaus, M. K., Garnsey, S. M., & Boland, J. (1990). Combinatory lexical information and language comprehension. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 383–408). Cambridge, MA: MIT Press/Bradford Books.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic context. *Journal of Verbal Learning and Verbal Behaviour, 18,* 427–441.

Tinker, M. A. (1946). The study of eye movements in reading. *Psychological Bulletin, 43,* 93–120.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language, 35*(4), 566–585.
Trueswell, J. C., & Tanenhaus, M. K. (1991). Tense, temporal context, and syntactic ambiguity resolution. *Language and Cognitive Processes, 6,* 303–338.
Trueswell, J. C., & Tanenhaus, M. K. (1994). Towards a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–179). Hillsdale, NJ: Erlbaum.
Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language, 33*(3), 285–318.
Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*(3), 528–553.
Tyler, L. K., & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behaviour, 16,* 683–692.
Underwood, N. R., & McConkie, (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly, 20,* 153–162.
Valian, V. (1979). The wherefores and therefores of the competence–performance distinction: In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 1–26). Hillsdale, NJ: Erlbaum.
van Orden, G. C. (1987). A rows is a rose: Spelling, sound, and reading. *Memory & Cognition, 15,* 181–198.
van Turennout, M., Hagoort, P., & Brown, C. (1997). Electrophysical evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 787–806.
van Turennout, M., Hagoort, P., & Brown, C. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science, 280,* 572–574.
Vonk, W., Hustinx, L., & Simons, W. (1992). The use of referential expressions in structuring discourse. *Language and Cognitive Processes, 11,* 301–335.
Vroomen, J., & de Gelder, B. (1997). Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 23*(3), 710—720.
Wanner, E. (1980). The ATN and the sausage machine: Which one is baloney? *Cognition, 8,* 209–225.
Wanner, E. (1987). The parser's architecture. In F. Kessel (Ed.), *The development of language and language researchers: Essays in honour of Roger Brown*. Hillsdale, NJ: Erlbaum.
Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & Miller, G. A. (Eds.), *Linguistic theory and psychological reality* (pp. 119–137 (partial)). Cambridge, MA: MIT Press.
Warrington, E. K., & McCarthy, R. A. (1983). Category-specific access dysphasia. *Brain, 106,* 859–878.
Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107,* 829–853.
Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*(5), 672–683.
Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development, 7,* 49–63.
Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34,* 311–334.
Wundt, W. (1900). *Die Sprache.* Leipzig: Englemann.
Yamazaki, M., Ellis, A. W., Morrison, C. W., & Lambon-Ralph, M. A. (1997). Two age of acquisition effects in the reading of Japanese Kanji. *British Journal of Psychology, 88*(3), 407–421.
Zurif, E., & Swinney, D. (1994). The neuropsychology of language. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 1055–1074). San Diego, CA: Academic Press.
Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32,* 25–64.