# 7 Vocabulary Acquisition: The Implicit Ins and Outs of Explicit Cognitive Mediation

Nick Ellis

*University of Wales Bangor*

## INTRODUCTION

There are now many demonstrations of the dissociation between implicit and explicit processes of learning. This division roughly parallels that between practical or tacit intelligence and academic or explicit intelligence. The research priority that necessarily follows from such distinctions is to determine which of human cognitive capabilities are acquired implicitly and which learned explicitly. This question is both of theoretical and practical pedagogic importance since teaching interventions are of less relevance to implicitly learned skills but are essential to explicitly learned ones. Nowhere has the role of consciousness been more a matter of debate than in the realm of human language skills, both in native (L1) and second (L2) languages (Ellis, 1993). There are two major causes of this continuing contention. The first is the 'slipperiness' of the term 'consciousness' both in its constitutive definitions and in their operationalisations (McLaughlin, 1990; Schmidt, 1993). The second is that 'language learning' is equally poorly defined, mainly because of its numerous facets. Although the particular issue of whether vocabulary acquisition reflects implicit or explicit learning processes appears more tractable, even this point is still far from resolved.

This article therefore reviews research related to L2 and L1 vocabulary acquisition in the areas of (i) incidental vocabulary learning, (ii) the associations between vocabulary and academic intelligence, (iii) implicit memory, (iv) global amnesia. It argues that these diverse areas of research all reveal a dissociation whereby the recognition and production aspects of vocabulary learning rely on implicit learning, but meaning and mediational aspects of vocabulary heavily involve explicit, conscious learning processes. The operation of the input and output lexical modules is next described, followed by an analysis of the explicit cognitive systems for mediation with semantics and conceptual systems. Finally the pedagogic implications of these findings are briefly summarised. Wherever possible I give examples from L1 and L2. When only evidence for L1 is presented it is because, to the best of my researches, parallel work remains to be done in this aspect of L2 vocabulary acquisition.

Applied linguistic perspectives on the status of vocabulary span the whole gamut, from the position where acquisition is entirely implicit to the extreme contrary view that we should be explicitly teaching large amounts of vocabulary to language learners. Krashen (1982; 1985; 1989; Krashen & Terrell, 1983) exemplifies the implicit position. His Input Hypothesis assumes that we acquire language by understanding messages. More precisely, comprehensible input is the essential environmental ingredient and a richly specified language acquisition device (LAD) assimilates vocabulary from the evidence provided in natural language. According to the Input Hypothesis, "when the LAD is involved, language is subconsciously acquired - while you are acquiring, you don't know you are acquiring; your conscious focus is on the message, not form. Thus, the acquisition process is identical to what had been termed 'incidental learning.' Also acquired knowledge is represented subconsciously in the brain - it is what Chomsky has termed 'tacit knowledge'." (Krashen, 1989, p. 440). At the other extreme there is a history of attempts to collate lists of a core vocabulary which L2 and L1 teachers are recommended to use to decide which words and meanings should be taught first (e.g. West's 1953 *General Service List* of 2,000 words where it was claimed that knowing these items gave access to about 80% of words in any written text. Thus students would be motivated to learn them since the words acquired could be seen by learners to have a demonstrably quick return). In parallel there have been developments of methods for explicit vocabulary learning and instruction. There is a now wide diversity of available teaching models (see Carter, 1987; Carter & McCarthy, 1988; McKeown & Curtis, 1987; Nation, 1987 for reviews) but an illustrative rich method of instruction is that of Beck and McKeown (1983). Here there is intensive instruction consisting of various involving and increasingly challenging activities designed to teach a set of 8 to 10 semantically related words over a period of a week. The activities include matching words with definitions, associating new words with a variety of contexts, creating contexts for words, comparing and contrasting words to discover relationships, generating affective responses to words, and using the words outside of class. Sometimes this can require as much as 30 minutes of instruction for each word taught.

What is the language practitioner to make of all of this? How can both of these positions be tenable? Does vocabulary come naturally or must it be taught and learned? Or else, what are the aspects of vocabulary that are amenable to instruction and study?

In psychology the role of consciousness in learning is discussed under such diverse rubrics as ‘implicit’ or ‘explicit’ learning; ‘incidental’ or ‘intentional’ learning; ‘conscious’, ‘unconscious’, or ‘subconscious’ processes; ‘implicit’ or ‘explicit’ memory; ‘metacognition’; ‘automatic’ and ‘controlled’ processing; ‘subliminal learning’; ‘attended vs. unattended learning’; memory of ‘how’, also called ‘procedural’, ‘skill’, and ‘habit’ memory, *vs.* memory of ‘that’, also called ‘declarative’ memory; ‘practical’ vs. ‘academic’ intelligence, etc. (Squire, 1987)

For example, Wagner & Sternberg (1985, 1986) differentiate between practical (or tacit) intelligence and academic (or explicit) intelligence. They claim that the type of knowledge used for success in most real-world settings is of the practical, tacit kind. Academic intelligence is formal, and is evaluated primarily by tasks usually found on mental ability tests and in academic settings; tacit knowledge is practical, informal, and usually acquired indirectly or implicitly. In addition, they argued that because much tacit knowledge is loosely organised and relatively inaccessible, the implicit type of knowledge may not lend itself to being directly taught.

Although the traditional stuff of education and psychology has been the measurement of explicit learning, only recently has Reber (1967, 1976) developed what has now become a standard paradigm for assessing implicit learning. Subjects are shown strings of letters generated from a finite state grammar. They are told simply to pay attention to these strings. After seeing perhaps 40 grammatical examples, the subjects are told for the first time that the strings are formed according to a complex system of rules and that they will be shown some further strings, some of which conform to the grammar and some of which violate it. Their job is to judge the grammaticality of the strings. Subjects are significantly accurate on this task, yet often they cannot explain the basis of their judgements (Reber & Allen, 1978). Reber thus argues that complex abstract knowledge of the structured stimulus domain is acquired implicitly, held tacitly, and used unconsciously to make accurate decisions about the well-formedness of novel items.

Reber, Walkenfield & Hernstadt (1991) have recently shown the dissociation between implicit and explicit learning in an analysis of individual differences. Subjects were given an implicit learning task following the standard artificial grammar-learning format (Reber, 1969; 1976; Reber & Allen, 1978). They also were given an explicit learning task - a series-completion task as found in many tests of problem solving ability - and a standard IQ test (a shortened WAIS). Substantial individual differences were found between subjects on the explicit task; relatively small individual differences were found on the implicit task. Moreover, in accord with Wagner & Sternberg's claims for a dissociation between implicit practical and explicit academic learning, performance on the explicit task correlated strongly with IQ, but performance on the implicit task did not.

Further evidence for the dissociations between explicit and implicit learning comes from studies of global amnesic patients (e.g. Korsakoff's syndrome) who, as a result of medial temporal lobe, midline diencaphalic, or basal forebrain lesions, show normal implicit learning yet anterograde amnesia for explicit, episodic memories (Strauss, Weingartner & Thompson, 1985; see Schacter, 1987 for review). The perennial anecdotal evidence for this comes from Claparède (1911) who reported that he once shook hands with a female Korsakoff patient while concealing a pin in his hand. This caused the patient some pain, and when he returned a few minutes later and offered his hand again, she refused to shake it. Her avoidance continued, even though she could give no explanation of why she was avoiding him. In this case implicit learning (behaviour which is changed as a result of a previous encounter) is preserved in the absence of any conscious, explicit recollection of the event.

Implicit learning is coming to learn the underlying structure of a complex stimulus environment by a process which takes place naturally, simply and without conscious operations. Simple attention to the stimulus suffices for implicit learning mechanisms to induce statistical or systematic regularities in the input environment (e.g. Berry, this volume; Berry & Broadbent, 1984; 1988; Ellis, 1993; Reber, 1967; 1969; 1976; Winter & Reber, this volume). Explicit learning is a more conscious operation where the individual makes and tests hypotheses in a search for structure. Concept attainment can thus take place in a number of different ways: (i) explicitly via given rules (assimilation of a rule following instruction), (ii) explicitly through selective learning (the learner searching for information and building then testing hypotheses), or (iii) implicitly (a nonconscious and automatic abstraction of the structural nature of the material derived from experience of specific instances).

There are thus a number of psychological methodologies which can be applied to answer the question of which aspects of vocabulary are learned implicitly and which explicitly.

## WHAT IS IT TO LEARN A WORD?

What is it to learn a new word, in either L1 or L2? Minimally we must recognise it as a word and enter it into our mental lexicon. But there are several lexicons specialised for different channels of Input/Output (I/O). To understand speech the auditory input lexicon must categorise a novel sound pattern (which will be variable across speakers, dialects, emphases and so on); to read the word the visual input lexicon must learn to recognise a new orthographic pattern (or, in an alphabetic language, learn to exploit grapheme-phoneme correspondences in order to access the phonology and hence match the word in the auditory input lexicon); to say the word the speech output lexicon must tune a motor programme for its pronunciation; to write it the spelling output lexicon must have a specification for its orthographic sequence (or, again in an alphabetic language, learn to exploit phoneme-grapheme correspondences in order to access the orthography and hence produce the spelling of the word as if it were regular). We must learn its syntactic properties: its part of speech and its syntactic subcategorisations. We must learn its place in lexical structure: its relations with other words. We must learn its semantic properties, its referential properties, and its roles in determining entailments. We must learn the conceptual underpinnings that determine its place in our entire conceptual system. Finally we must learn the mapping of these I/O specifications to the semantic and conceptual meanings: the relation between word form and word meaning is generally arbitrary (relics of onomatopoeic or pictographic origin aside).

## INCIDENTAL VOCABULARY ACQUISITION

We have not been taught the vast majority of the L1 vocabulary that we know, nor indeed have we looked up these words in dictionaries - most vocabulary is learned from context (Krashen, 1989; Miller & Gildea, 1991; Sternberg, 1987).

There are, perhaps, 600,000 words in the English language. Smith (1982) uses the estimates of Seashore and Eckerson (1940) that the average college undergraduate knows about 156,000 of these words (58,000 ‘basic’, 96,000 ‘derived’, and 2,000 ‘rare’) to emphasise that this knowledge could not have come from 156,000 trips to the dictionary, 156,000 flash cards, or 156,000 fill-in-the-blank exercises. Miller (1977) has estimated that children between the ages of six and eight pick up an average of fourteen basic words a day. Nagy and Herman (1987) recalculated older studies of vocabulary size correcting a number of errors that affected the estimates. These recalibrations “give good reason to believe that the average high school senior’s vocabulary is in the neighbourhood of 40,000 words” (p. 21). Additionally, Nagy and Herman estimate that the average student learn about 3,000 words a year during the school years. Whatever the exact number, it is clear that direct teaching of vocabulary cannot be the source of these gains and that the natural language learner must acquire considerable amounts of vocabulary without instruction. A rule of thumb estimate of the amount that a typical fifth grader learns simply from context is about 1,000 L1 words a year (Nagy, Anderson and Herman, 1987).

I will first consider infants' acquisition of L1 vocabulary before reviewing evidence for preschoolers' incidental vocabulary learning from context in certain television programs and the same process in older children and adults while they read in their L1 or L2.

By the age of six the child has acquired the meaning of some 14,000 words. Assuming such learning begins at around 18 months, the child learns an average of nine words a day, one per waking hour. After the age of two, conscious drill on words - for example by parents pointing at and naming referents - is not typical; therefore the child learns most of these 14,000 words by hearing people use them in their normal contexts. Carey (1978) was concerned with the child's acquisition of the meaning of these lexical items, and she took pains to distinguish between lexical, conceptual and semantic aspects of vocabulary. The conceptual domain is the mental representation by which one describes or understands the world or one's own actions. The lexical domain is the structured set of words that encodes aspects of the conceptual domain. Consider the terms of colours. There are many things that we know about colours that are not captured in the structure of the colour lexicon. We cannot, for example, differentiate between different shades of red without reference to particular perceptual objects, as in the following quotation from Conan Doyle: "every man who had a shade of red in his hair had trampled to the city to answer the advertisement. Fleet Street was choked with red-headed folk and Pope's Court looked like a coster's orange barrow... Every shade of colour they were - *straw, lemon, orange, brick, Irish-setter, liver, clay*; but, as Spaulding said, there were not many who had the real vivid *flame-coloured* tint." (Conan Doyle, 1957, p. 28). Similarly the Dani of Indonesian New Guinea, whose language has only two colour words, *mili* and *mola*, still process focal colours better than non-focal ones. Conversely, it is possible to have certain aspects of lexical and semantic knowledge without the corresponding perceptual experiences or conceptual knowledge (as in the cases of blind children who can learn a great deal about the lexical domain of colour without corresponding perceptual experience, or young children who can list the names of colour hues in response to such questions as "Do you know any colours?" before they can map these terms correctly onto the conceptual/perceptual domain).

Carey experimentally introduced a 'new' word, *chromium* (actually a colour name), for olive-coloured objects, in natural situations in a nursery school atmosphere. The 3- and 4-year olds originally applied either the label *green* or *brown* to the olive colour when they were asked to name the colour of objects at the beginning of the study. They changed their concepts over the next few months, restructuring their lexicon to find the right place for *chromium* - that it was associated with green but not a synonym of it, and restructuring their conceptual domain by learning that olive was not included in the category of green or brown but was a colour that had its own name. Carey developed a theory in which the children engaged in '*fast mapping*', which allows them to hold onto fragile new entries in their lexicon while the meanings are being established. Fast mapping did not reflect children's quick, full mastery of a word's meaning, but instead a partial mapping of the conceptual underpinnings of a word with the lexical form. Carey hypothesised that children quickly recognise a word as a word, know something about its grammatical role, and have a partial sense of the set of things or events to which the word refers. Thus development of the lexical aspects of vocabulary can occur in the process, and before full mastery of, the mappings between lexical items and their conceptual meanings. In infants it seems to occur in situations where exposure to a new word is limited to a few instances, the targeted words are object or attribute labels, and the learning occurs as an adult interacts with the child, manipulating joint attention to the targeted object while at the same time naming it, usually in a manner that contrasts the new word with a known word. Lexical development and conceptual development can occur separately, before any mapping of one domain onto the other occurs, and these distinctions between the lexical system, the conceptual system, and the processes of mapping between them are as crucial to the present as to any consideration of vocabulary acquisition.

Rice and Woodsmall (1988) had 3- and 5-year-old children individually view animated programs that introduced 20 unfamiliar words in a story context. After two viewings, the 5-year-olds comprehended on average five new words, the 3-year-olds 1.5 new words. Rice, Huston, Truglio, and Wright (1990) show that these effects additionally occur outside the laboratory in a longitudinal investigation of preschoolers' viewing of "*Sesame Street*". Cross-age multiple regressions demonstrated a positive effect of "*Sesame Street*" viewing on vocabulary development from 3 to 5 years old. There was, however, no effect of viewing other kinds of children's programs (such as cartoons), suggesting that "*Sesame Street"* is specially well suited to introducing word meanings to young viewers by virtue of its dialogue which closely resembles a mother talking to her child, with simple sentences, much talk about the here and now, repeated emphasis on key terms, and an avoidance of abstract terminology. Rice (1990) argues that these conditions are sufficient for the process of quick incidental learning (QUIL) of words which is internally driven with children able to infer new word meanings without explicit assistance from adults.

When reading, "whatever cognitive processes are engaged over word or word-group units (phonological coding, semantic activation, parsing, induction of new vocabulary items) are being exercise hundreds of times a day. It is surely to be expected that this amount of cognitive muscle-flexing will have some specific effects." (Stanovich & Cunningham, 1992). Saragi, Nation & Meister (1978) asked adults to read Anthony Burgess' novel, "*A Clockwork Orange"*. This contains a number of words from a Russian-based slang called *nadsat*. which few readers will have met before reading the book. There are 241 *nadsat* words in the novel and they are repeated on average 15 times. The subjects were simply asked to read the book (which crucially did not contain a dictionary of these words as an appendix), and were told that afterwards they would be given a test of comprehension and literacy criticism. However, within a few days of their finishing the book they were given a multiple choice test covering 90 *nadsat* words. Results showed that a considerable vocabulary acquisition had taken place. Scores ranged from 50-96% correct, with an average of 76%. Krashen (1989) concluded from this study that subjects has picked up at least 45 words simply by reading a novel and additionally cited three similar demonstrations of adults' L2 incidental vocabulary acquisition during reading novels such as *Animal Farm* and a *Clockwork Orange* (Ferris; Kiyochi; Pitts, White & Krashen; all in Krashen, 1989, p. 446).

Finally there is the correlational evidence that people who read more know more vocabulary (Anderson, Wilson, & Fielding, 1988). This relationship between print exposure and vocabulary appears to be causal in that it holds even when intelligence (figural analogies and Ravens matrices) and even reading comprehension ability - an excellent measure of general verbal ability - is controlled (Stanovich & Cunningham, 1992).

There is thus little doubt that reading affords vocabulary assimilation. It is an ideal medium for such growth. Moderate-to-low-frequency words - precisely those words that differentiate between individuals of high and low vocabulary size - appear much more often in common reading matter than they do in common speech. And new words are there on the page for potential scrutiny. There is opportunity for the reader to make regressive eye-movements and to study the context, to form hypotheses at leisure and cross validate them, to have time to problem solve about meanings. The word is frozen in time on the page, whereas in speech it passes ephemerally.

But what might be the details of the processes of vocabulary acquisition whilst infants fast-map or when older children and adults read? Without further experimentation it is impossible to resolve whether such vocabulary acquisition from context reflects *implicit learning*, *incidental learning*, or even *explicit learning* without *explicit instruction*.

One possibility is that it is a simple result of use: a new vocabulary item whose meaning is apparent from its context is simply understood and learned implicitly or incidentally, without conscious effort to commit it to memory. An *implicit vocabulary learning hypothesis* would hold that the meaning of a new word is acquired totally unconsciously as a result of abstraction from repeated exposures in a range of activated contexts. This is essentially the behaviourist position on language learning as proposed in Skinner (1957). An *incidental vocabulary learning hypothesis* is that we acquire the new vocabulary item without intent so to do - perhaps we *notice* that it is a novel word (Schmidt, 1990) but we attend to understanding the passage as a whole, and memory for the new word comes as a natural result of this process, a conscious efforts to learn being unnecessary.

But people are strategic, active processors of information. Contra Krashen (1989), it does not follow that vocabulary has been subconsciously acquired from the fact that we have not been taught the vast majority of the words that we know. That we have not been taught vocabulary does not entail that we have not taught ourselves. An *explicit vocabulary learning hypothesis* would hold that there is some benefit to vocabulary acquisition from the learner noticing novel vocabulary, selectively attending to it, and using a variety of strategies to try to infer its meaning from the context. Furthermore there may also be advantage from applying metacognitive strategies to remember new vocabulary, to consolidate a new understanding by repetition and associational learning strategies such as semantic or imagery (keyword) mediation techniques, these being well established as effective mnemonic techniques (Desrochers & Begg, 1987; Ellis & Beaton, 1993 a,b; Levin & Pressley, 1985). Sternberg (1985, p. 307) argues this case: "simply reading a lot does not guarantee a high vocabulary. What seems to be critical is not sheer amount of experience but rather what one has been able to learn from and do with that experience. According to this view, then, individual differences in knowledge acquisition has priority over individual differences in actual knowledge." Jensen (1980, pp. 146-147) argues this position even more strongly: "Children of high intelligence acquire vocabulary at a faster rate than children of low intelligence, and as adults they have a much larger vocabulary, not primarily because they have spent more time in study or have been more exposed to words, but because they are capable of educing more meaning from single encounters with words. ... The vocabulary test does not discriminate simply between those persons who have and those who have not been exposed to words in context... The crucial variable in vocabulary size is not exposure per se, but conceptual need and inference of meaning from context, which are forms of eduction. Hence vocabulary is a good index of (academic) intelligence."

Studies of vocabulary acquisition from reading demonstrate that neither dictionary look-up nor *explicit instruction* is *necessary* for vocabulary acquisition, The remainder of this paper analyses the degree of involvement of conscious learning processes in different aspects of vocabulary acquisition.

## VOCABULARY AND INTELLIGENCE

Because implicit learning is a more basic form of learning than that involving conscious processes it should show a different distribution in the population. Phylogenetically older systems display less variation than new ones. Systems and structures with considerable antiquity have shed the variability that characterises evolutionary newer systems. Once an adaptive functional system evolves, and the system is broadly operative in diverse environments, there is no adaptive value in change (Reber et al., 1991). Consider the innate skill of walking - just about everyone on the globe learns to do it at roughly the same time and moves through the same motor milestones on the way. Its ubiquity is supportive evidence for it being an implicitly learned skill which the brain is pre-programmed to acquire and refine. Chomsky used the same argument to support the idea that language is an independent faculty separate from non-linguistic cognitive abilities: “As far as we know, the development of human mental capacity is largely determined by our inner biological nature. Now in the case of a natural capacity like language, it just happens, the way you learn to walk. In other words language is not really something you learn. Acquisition of language is something that happens to you; it’s not something that you do. Language learning is something like undergoing puberty. You don’t learn to do it; you don’t do it because you see other people doing it; you are designed to do it at a certain time.” (Chomsky, 1988, pp. 173-174).

Notwithstanding wide variation in intelligence, just about everyone learns to talk their L1. Lenneberg (1967, pp. 156-157) suggests that children with non-verbal IQs as low as 30 (i.e. more than 99.99% of people) can still complete the single word stage of language. Lenneberg, Nichols and Rosenberger (1964) studied 61 Down’s syndrome children over three years and found that on language imitation tests these children performed like younger normal children. 75% had reached at least the first stage of language development; they had a small vocabulary and could execute simple spoken commands. It was the passing of particular motor milestones, and not the particular IQ level, that best predicted language development.

If we are simply interested in vocabulary aspects of language acquisition, there is plenty of evidence that the input/output of vocabulary is relatively independent of intelligence. Thus, for example, children severely mentally retarded as a result of hydrocephalus, may talk excessively with impressive vocabularies, even though their speech lacks content (Taylor, 1959; Hadenius, Hagberg, Hyttnäs-Bensch & Sjögren, 1962; Ingram & Naughton, 1962; Cromer, 1991 for review). Hadenius et al. coined the term ‘cocktail-party syndrome’ for the condition since there was a “peculiar contrast between a good ability to learn words and to talk, and not knowing what they are talking about" (p. 118). Yamada (1990) presents a very detailed case study of such a teenager, Laura, who despite very profound retardation in non-language skills and cognition (in many areas performing at a 3-year old level; IQ in the 40s) had intact language *qua* formal system of rules and representations. She demonstrated phonological, morphological and syntactic skills at a very high level, producing sophisticated grammatical forms and structures of a child at least 5 years of age or older , e.g. passivization, complementation and subordination. Yet in spite of their complex structural qualities, Laura's utterances were often semantically anomalous, with particular difficulty with respect time, manner, dimensionality, and number - she would often use words or phrases that fell within the correct semantic category or domain (a location, a date, a number, and so on) but her lack of conceptual knowledge led to an inappropriate response (as in her answer 'five' after she had just been given two pennies and been asked 'how many pennies do you now have?' ). Such cases clearly reinforce Carey's (1978) claims that some aspects of the lexicon in children can develop independently of conceptual knowledge.

Even the ability to develop normal reading and writing skills (as transcoding procedures) is found in the context of severe mental deficiency but syntactically well-developed language (Cossu & Marshall, 1990). This dissociation is also present at the other end of life. Clinically dementing adults' ability to explain what words mean declines in the face of retained ability to access word pronunciation (Nelson & McKenna, 1975). Thus scores on the National Adult Reading Test (NART) which has people read aloud words like *aisle*, *heir*, *banal*, and *aver* which are 'irregular' with respect of the common rules of pronunciation are used as an index of their pre-morbid IQ (Nelson, 1982).

In contrast with mere I/O, when the criterion of vocabulary knowledge additionally involves the understanding of words, i.e. the mapping between lexical, semantic and conceptual domains, then there are, of course, strong correlations between academic intelligence and vocabulary size. Learning disabled children produce particularly low scores on the Vocabulary subtest of the WISC-R (The Wechsler Intelligence Scale for Children) where children have to demonstrate an understanding of words by defining them (Kaufman, 1979). At ages between 6 and sixteen, WISC-R vocabulary scaled scores correlate an impressive 0.69 with full scale Stanford-Binet IQ (Wechsler, 1976). Mill-Hill Vocabulary Test scale scores correlate with Raven's Matrices ('non-verbal') IQ 0.60 in people under 30 years old (Table SPM I, Raven, Court & Raven, 1983). WAIS-R vocabulary scaled scores reliably correlate at greater than 0.80 levels with full scale WAIS-R IQ (Wechsler, 1981).

Other psychometric enterprises of L1 language development reveal the same dissociation. The Illinois Test of Psycholinguistic Abilities (McCarthy & Kirk, 1961) was designed to assess different aspects of L1 development in children between 2.5 and 10 years old. It contains a number of subscales measuring such components as Auditory Reception (understanding very simple sentences), Auditory Association ("A dog has hair. A fish has ___?"), Visual Association ('Which goes with which?' from pictures of a sock, a hammer, a shoe, a nail), Verbal and Manual Expression, Grammatical Closure ('Here is a dog, here are two ___?'), Auditory Closure (filling in missing parts of auditorily presented words, e.g. 'tele / one' for *telephone*, 'ta / le / oon' for *tablespoon*), Visual Closure (identifying the missing components of objects), Auditory Sequential Memory (digit span), Sound Blending (discrete sounds of a word are spoken every half second and the child has to blend them together to say what the word is), etc. Factor analyses (Pareskevopoulos & Kirk, 1969) of the intercorrelations between these variables demonstrate that they are described by two orthogonal divisions: that between the visual and auditory channels, and that between 'the *representational* level' (reception, expression, and association tests) and the '*automatic* level' (closure tasks, sound blending, sequential memory). Their representational level tests tap semantic and conceptual knowledge. Their automatic level tests tap lexical I/O functions. Not only are these two aspects not of like kind, but also they show different patterns of association with IQ. In Table 10-8 of Pareskevopoulos and Kirk (1969) it is clear that conceptual tasks such as Auditory Association and Auditory Reception correlate more highly with IQ than do 'automatic', more hygienic I/O tasks such as Auditory Closure and Sound Blending between the ages 2.5 years to 10 years.

The discussion so far has concerned associations between IQ and different aspects of L1 vocabulary acquisition. Because of a lack of relevant studies it is difficult to find similar direct evidence for these effects in L2. Genesse (1976) failed to find consistent association between IQ level and performance on *any* measure of L2 interpersonal communication skills - either listening, pronunciation, vocabulary, grammar or communicativeness - in either students in French-as-a-second-language or French immersion programmes. There is , however, some indirect evidence for the dissociation between I/O and mediational aspects of L2 vocabulary. Typical correlations between intelligence and language aptitude as measured with the Modern Language Aptitude Test (MLAT: Carroll & Sapon, 1955) are typically greater than 0.4 (e.g. Gardner & Lambert, 1972). In a group of 793 Canadian adult public servants, the correlation between MLAT total score and Thurstone and Thurstone Primary Mental Abilities IQ was found to be 0.67 (Wesche, Edwards and Wells, 1982). However, as Carroll, (1985, p. 97) states: “second language aptitude tests measure certain abilities, such as phonetic coding and rote memory, that are not tested by intelligence tests, and thus make a unique contribution to the prediction of second-language learning success over and above what can be contributed by intelligence tests alone. Validities of second-language aptitude tests in predicting second-language learning success are generally significantly higher than those of intelligence tests.” Thus the I/O of L2 vocabulary is separable from the semantic and conceptual aspects.

Other indirect evidence which reveals a temporal dissociation comes from the literature on critical periods and maturational constraints on L2 development. Long (1990, p. 274) states the conclusions of his recent review thus: “starting after age 6 appears to make it impossible for many learners (and after age 12 for the remainder) to achieve native-like competence in phonology; starting later than the early teens, more precisely after age 15, seems to create the same problems in morphology and syntax. Preliminary results suggest that similar generalisations will eventually be found to hold for lexis and collocation, and for certain discourse and pragmatic abilities.” Even if, as Long believes, there are maturational constraints on *all* aspects of L2 acquisition, they appear much earlier for I/O aspects like phonology than they do for the more cognitive aspects of vocabulary acquisition.

Taken together, these studies clearly reveal different, separable components of vocabulary acquisition both in L1 and L2. I/O processing neither correlates highly with cognitive mediational components nor with intelligence, yet these latter two abilities are inextricably interrelated.

## PRIMING STUDIES OF MONOLINGUAL IMPLICIT AND EXPLICIT MEMORY SYSTEMS

The main technique for studying implicit memory is *repetition priming*, i.e. facilitation of the processing of a stimulus as a function of a recent encounter with the same stimulus. Repetition priming has been observed on a variety of tasks that do not make explicit reference to a prior study episode - the subject at no point has to consciously recall the prior event. The tests most commonly used in priming research are *lexical decision*, *word identification*, and *word stem completion.* On the *lexical decision* task subjects have to rapidly decide whether a particular letter string is a word (e.g. *watch*) or not (e.g. *wetch*); priming is reflected by a decreased latency in the making of a lexical decision on the second presentation of a letter string relative to the first. The *word identification* test measures the minimum exposure necessary for correct recognition of the word. Priming is indexed here by a lower minimum exposure for repeated words. On *word completion* tasks, subjects are given a word stem (e.g. *tab__* for *table*) and are instructed to complete it with the first appropriate word that comes to mind. Here, priming is reflected by an enhanced tendency to complete test items with words exposed on a prior study list.

Priming has been used to investigate word recognition and lexical organisation in skilled adult L1 users. Four examples will illustrate these studies. (1) While the word *phrase* is primed by itself, it is not by the string *frays*, thus (i) the visual input lexicon is not phonologically coded and (ii) the input lexicon is separable from the output lexicon (Neisser, 1954). (2) Subsequent recognition of *bored* was primed by prior experience of *boring* but not by *born* - the degree of visual and acoustic similarity between the practice words and the test word was equivalent; the difference is that *boring* is morphologically related to *bored* but *born* is not. This led to the conclusion that visual input lexicon units (logogens) function at the level of the morpheme, not the word as such (Murrell & Morton, 1974). More recent work (see Monsell, 1985 for review) suggests that there are small priming effects of sub-lexical units (abstract graphemes, phonological constituents and their patterning), but that the effect associated with fully activated lexical units is orders of magnitude more persistent than, and may thus be distinguished from, those associated with sub-lexical or supra-lexical levels (e.g. Meyer and Schvanaveldt, 1971 demonstrated that lexical decision of a target word is facilitated by a prior semantic prime (e.g. bread-*butter*), but this effect is rapidly attenuated with the intervention of either a single item or a four second interval between prime and probe, whereas lexical repetition priming can last over hours or days). (3) While there are strong effects of repetition priming of real words, it is generally found that nonwords show either no or minimal priming effects. Again this suggests that these lexical priming effects are specific to *lexical* units (Kirsner & Smith, 1974). (4) A prior reading of a word primes its subsequent visual recognition, a prior hearing of a word primes its subsequent auditory recognition, but cross-modal priming is typically absent or minimal. This led to the dissociation of the input lexicon into two independent modules, the auditory input and visual input logogen systems (Morton, 1979; Murrell & Morton, 1974). It is now generally agreed that there are separate, independent, systems for auditory and visual recognition of words and for spoken and written word production (Ellis, 1984).

There is now a vast research industry concerning priming. One major ongoing debate concerns whether word identification is entirely a ‘‘bottom-up’, data driven process, with semantics only affecting post-lexical processing, or whether there are ‘top-down’ effects right down to the level of the lexicon where semantic context affects word identification itself. There is insufficient space here to properly review this issue and there remain strong advocates of both camps. Suffice it to say that the data from studies of priming and of amnesia (to be reviewed below) strongly suggest that lexical identification is modular and unaffected by meaning (for sake of balance the reader is referred to Balota, Ferraro & Connor, 1991, for contrary evidence).

The modular view is that lexical I/O operations work automatically and implicitly and that their functioning is cognitively impenetrable - stimulus analysis mechanisms are isolated from higher-level hypotheses and real-world knowledge (Fodor, 1983). Lexical access in fluent readers is a fast, obligatory, low capacity process largely uninfluenced by knowledge structures outside the lexicon or by higher level expectations - it has nothing to do with aspects of cognition like reasoning or problem solving (Stanovich, 1992). But of course, when the bottom-up processes that result in word recognition are deficient, either because of noisy stimuli or novice processors, then word identification becomes non-automatic, conscious explicit processes are marshalled, and other knowledge sources like contextual information are used.

Perhaps the clearest evidence for the separation of implicit I/O lexical systems from semantic aspects comes from the interaction of depth of processing and the types of operation that we ask subjects to do with words. I will review later the many demonstrations that led Craik and Lockhart to propose a *Depth of Processing* theory of learning and memory whereby the more subjects analyse material semantically and the more they elaborate upon its meaning, the better they will recall it in long-term tests of explicit memory (recognition and recall tasks). Jacoby and Dallas (1981) showed subjects a list of familiar words and had them perform either a study task that required elaborative processing (e.g. answering questions about the meaning of the target word) or a shallow study task that did not require elaborative processing (e.g. deciding whether or not the target word contained a particular letter). Explicit memory for the words was subsequently tested by *yes/no* recognition and implicit memory was assessed by savings in word identification tests. Recognition performance was higher following elaborative study than non-elaborative study. However, implicit memory was unaffected by the study manipulation: priming effects on word identification performance were about the same following the elaborative and non-elaborative tasks. Graf, Mandler and Haden (1982) report a similar pattern of results by using free recall as an index of explicit memory and stem completion as a measure of implicit memory. Taken together, experiments of this type (see Schacter, 1987 for review) clearly demonstrate that word identification operates according to implicit memory principles - it is affected by mere exposure and the frequency thereof. But explicit memory for words is clearly affected by the depth of processing and the degree to which subjects analyse their meaning.

Schacter reviews two other types of dissociation between implicit and explicit memory for words. The Morton (1979) experiment described above demonstrated that implicit priming effects were attenuated with changes of study modality. However changing modality of presentation from study (auditory) to test (visual) does not affect *yes/no* explicit recognition memory (Jacoby & Dallas, 1981). Finally on word identification tests, lexical decision and word identification (but not stem completion) priming effects persist with little change across delays of days or weeks, whereas recognition memory declines across these same delays (Jacoby & Dallas, 1981).

## PRIMING STUDIES OF BILINGUAL IMPLICIT AND EXPLICIT MEMORY SYSTEMS

The same dissociation can be found in the contrast between studies of implicit and explicit memory in bilinguals (Durgunoglu and Roediger, 1987; Heredia & McLaughlin, 1992). Concern over the organisation of the bilingual brain has a long history: Is there a single memory store for both languages (the *interdependence* or *compound* model) or a separate store for each (the *independence* or *co-ordinate* model)? The *interdependence* model assumes that items or concepts are stored in the bilingual's memory in the form of language-free concepts with a single conceptual or semantic representation subserving the two lexical entries. Evidence for the interdependence model typically comes from tests of explicit memory tasks, e.g. free recall experiments where exposure to the same concept in different languages is additive. The *independence* model contends that bilinguals' memory is organised with one memory for each language, with information in one language not readily available to the other system. Evidence for the independence view typically comes from implicit memory tasks such as word identification or stem completion tasks (e.g. Kirsner, this volume; Kirsner, Smith, Lockhart, King & Jain, 1984). Kolers in 1966 suggested a resolution to this debate, a compromise whereby bilinguals have neither separate nor shared memories: some information is restricted to the language of encoding while some is accessible to both linguistic systems. He has been proved correct in the light of subsequent evidence.

Durgunoglu and Roediger (1987) investigated implicit and explicit memory in bilingual subjects. Subjects saw words twice either (a) in English, (b) in Spanish, or (c) once in both languages. They saw other words (d) once in Spanish and had to generate their English equivalents in writing, or (e) twice in Spanish and also had to generate images of their referents. These latter two conditions require more elaboration and a greater depth of processing. Explicit memory for the words was tested with free recall, implicit memory with word-fragment completion in English. In free recall (i) the language studied was unimportant, and (ii) elaborations such as generating a translation or forming an image of the referent facilitated recall. In word-fragment completion (i) if the study language matched the test language fragment completion rates were significantly higher than the rate for non-studied control items; if the study language did not match the test language the fragment completion rates did not differ from the non-studied items, and (ii) elaboration during study did not improve word-fragment completion rates.

These results like those of Heredia & McLaughlin (1991) again emphasise the distinctions between implicit, data driven I/O modules which are language specific, and an explicit, conceptually driven cognitive system which supports the semantics and concepts which both languages describe. In bilinguals, implicit memory systems are independent, but there is one compound explicit memory for the conceptual representations for words in their two languages.

## EVIDENCE FROM GLOBAL AMNESIA

Global amnesia is a devastating disability. The loss of one's memory and the ability to learn is the loss of one's autobiography, one's identity, one's self (Warnock, 1987). There are two major components of the amnesic syndrome: (i) patients may have difficulty in recalling or recognising events and facts which have occurred since the onset of illness - *anterograde amnesia*; (ii) they may be impaired in recalling events and facts which occurred before the onset of their amnesic condition - *retrograde amnesia* (McCarthy & Warrington, 1990; Paradis, this volume; Squire, 1987). Their disability can inform us about the normal processes of vocabulary acquisition.

The first experimental investigations of a patient (H.M.) with severe and selective anterograde amnesia were conducted by Milner (Scoville & Milner, 1957; Milner, 1966). H.M. had intractable epileptic seizures which were finally treated by surgery which involved bilateral resection of the medial temporal lobes and ablation of the anterior two thirds of the hippocampal complex, the uncus, the amygdala, and the hippocampal gyrus. Unfortunately damage to these limbic structures (and to the associated midline diencaphalon which deteriorates in Korsakoff's syndrome as a result of chronic alcoholism) causes profound amnesia (Dudai, 1989; Squire, 1992). H.M. "could no longer recognise the hospital staff, apart from Dr. Scoville himself, who he had known for many years; he did not remember and could not relearn the way to the bathroom, and he seemed to retain nothing of the day to day happenings in the hospital... A year later, H.M. had not yet learned the new address, nor could he be trusted to find his way home alone, ... he is unable to learn where objects constantly in use are kept" (Milner, 1966, p. 113). "He remarked 'Every day is alone by itself, whatever enjoyment I've had, and whatever sorrow I've had.' Our own impression is that many events fade from him long before the day is over. He often volunteers stereotyped descriptions of his own state, by saying that it is 'like waking from a dream'. His experience seems to be that of a person who is just becoming aware of his surroundings without fully comprehending the situation because he does not remember what went before." (Milner, Corkin & Teuber, 1968). Yet despite these difficulties, he had a high average IQ of 112 and was alert and co-operative.

H.M. like other pure cases of anterograde amnesia had normal recall of events that occurred before his brain damage and his short-term memory was normal. His prior semantic knowledge, including that of vocabulary and concepts, was preserved as was evidenced by his continued high IQ and his lack of symptoms of language disability. But he had no memory for episodes that occurred after the operation. On the Wechsler Memory Scale his scores were very impaired: his immediate recall of stories fell far below the average level, he obtained zero scores in learning to associate arbitrarily paired words, "moreover, on all tests we found that when he had turned to a new task the nature of the preceding one could no longer be recalled, nor the test recognised if repeated." (Scoville & Milner, 1957, p.17). Other experimental demonstrations of his severe learning difficulties included: (i) his failure when presented with twelve faces and asked to select those that he had seen from a larger array after a delay of only 90 seconds filled with a distracter task, (ii) he was unable to learn a sequence of digits or light flashes that was longer than his short-term memory span despite repeated presentations, (iii) he failed on detailed recall of a complex drawing. Such findings are now taken as diagnostic of cases of anterograde amnesia.

Memory for a recent event can be expressed explicitly, as conscious recollection, or implicitly, as a facilitation of test performance without conscious recollection. Surprisingly, while amnesics such as H.M. show a severe deficit in explicit memory, they can learn implicitly, as evidenced by practice effects. They show normal classical conditioning; they can acquire motor skills such as mirror drawing as fast and as well as normal individuals; they show good perceptual learning (e.g. reading text in a mirror); and they show normal performance on tests of priming which are taken to indicate *normal implicit memory* abilities (Kirsner, this volume; Schacter, 1987; McCarthy & Warrington, 1990).

Priming is tested with word completion techniques or by assessing effects of repetition of stimuli. As described above, the word (or 'stem-') completion paradigm involves presenting lists of words and then testing retention either by standard recall or recognition techniques, or by presenting the first three letters of the target word and asking the subject to produce any word that begins with these three letters. The priming aspect is that normal subjects will be facilitated on this task and more likely to complete the stem with a word that they have recently studied. And amnesics show this effect too, even though they may deny that they had studied that word earlier (in the same way that Claparède's patient refused to shake hands a second time [i.e. had been affected by the stimulus], in the absence of any explicit memory of having done so and being jabbed before). This is a dissociation between implicit and explicit memory: amnesics are severely impaired on the recall and recognition tests which involve a conscious recollection of the prior episode, but they show normal practice effects as a result of prior exposure. Other demonstrations of this preservation of implicit memory in amnesics are to be found in Warrington & Weiskrantz (1982) who demonstrated that amnesics showed effects of repetition in naming, category identification (e.g., object/animal), and generation of opposites (e.g. black/white). Subjects were presented with target words and asked to respond as quickly as possible. On a second presentation of the same stimuli in these tasks, the reaction times of both the amnesics and the control subjects were faster, therefore showing facilitation.

So what does study of amnesia tell us about vocabulary acquisition and use? If amnesia is properly characterised as a deficiency in storage or in retrieval dependent on conscious voluntary procedures (explicit memory) while automatic procedures (implicit memory) are preserved (Schacter, 1987), then their successes at using and learning vocabulary inform us about the degree to which vocabulary learning is implicit. So what can they do in these respects?

## Lexical Access

### 1. Amnesics retain prior learned vocabulary and concepts and can access them normally.

They have no difficulty on vocabulary or naming tests, and their reaction times have been shown to be quite normal on word-retrieval tasks (Meudell, Mayes, & Neary, 1981).

### 2. They show normal facilitation in repetition priming experiments involving pre-existing memory representations such as common words or linguistic idioms

The results of Warrington & Weiskrantz (1982) described above show this. The readers is also referred to Schacter's (1987) review.

*Thus the recognition of pre-existing vocabulary and the access of known spoken or written word forms for expression relies upon automatic, implicit processes.*

## Learning New Lexical Units or Word Forms

### 3. Amnesics cannot explicitly recall new nonwords.

They cannot even explicitly recall *words* from an immediately subsequent verbal learning task.

### 4. Amnesics sometimes show implicit memory for novel words.

The evidence on this is mixed. In the priming experiments described above, amnesics showed normal priming for items with pre-existing unitary memory representations such as common words which they would have known before trauma. What about priming for nonwords that do not have any pre-existing representations as units in memory? If they show priming here it would demonstrate that recognition units at least for the word forms (i.e.

logogens) are automatically set up for novel vocabulary as implicit memories which simply result from experience.

Early studies failed to find priming for nonwords with amnesics, suggesting therefore that priming depends on activation of pre-existing information. For example, Diamond and Rozin (1984) claimed that amnesic patients do not show priming of nonwords in a stem completion task, and Cermak, Talbot, Chandler & Wolbarst (1985) produced similar results when implicit memory was tested with a perceptual identification task. However, more recent evidence (see Squire, 1992 for review) showed that priming can involve the acquisition of new information:

(i) Speeded perception of a novel visual shape is a laboratory analogue of learning a *new script* or a new ideogram (e.g. a new Kanji). Normal subjects and amnesic patients improve their ability to reproduce novel line patterns independently of their ability to recognise these patterns as having been presented previously (Musen & Squire, in press; Musen & Treisman, 1990). Gabrieli, Milberg, Keane and Corkin (1990) also show that H.M. exhibits this effect. Normal and amnesic subjects exhibit priming of unfamiliar visual objects, again independent of recognition memory performance (Schacter, Cooper, & Delaney, 1990).

(ii) One laboratory parallel of learning the visual form of new word in a *known script* is reading nonwords (of course, this task ignores any semantic aspects of vocabulary). Amnesic patients show normal practice effects in the acquisition of reading skill for regularly repeating nonwords (Musen & Squire, 1991).

*It therefore appears that the implicit learning capabilities of amnesic patients are sufficient to allow new pattern recognition networks to be established for the visual forms of new language, whether this involves a new script or new words in a known script. Both of these tasks also involve output modules - written copying of new shapes, spoken production of new words. Given that amnesics show a wide range of implicit learning of new motor routines, it seems likely that they show similar facilitation in their language output modules. However, it must be remembered that all of this implicit learning occurs in the absence of any explicit recall of the words. Unlike explicit memory, such learning is also gradual and slow.*

## Accessing Known Conceptual Associations

### 5. Amnesics can show normal learning of word pairs that are highly related.

Winocur and Weiskrantz (1976) and Shimamura and Squire (1984) showed that while amnesics' paired associate learning of unrelated word pairs was severely deficient, their performance on semantically highly related word pairs (*chair-bench*, *wealth-fortune*, etc.) was much better (although control subjects still significantly outperformed them). They can make use of activation of prior semantic associations in explicit learning but, as will be demonstrated below, they cannot explicitly learn new associations.

### 6. Amnesics show normal repetition priming effects of already known highly related paired associates.

Gardner, Boller, Moreines, and Butters (1973) showed Korsakoff's amnesics and controls a categorised word list. When subjects were later given category cues and asked to state the first category member that came to mind in a free association task, both amnesics and controls showed equivalent amounts of priming. Similar confirmation of priming of prior relations comes can be found in Schacter (1985) where amnesic patients showed normal levels of priming after studying a list of common idioms (e.g. *sour-grapes*) and then writing down the first word that came to mind on being given the stimulus word (*sour-?*).

*It seems therefore that implicit processes are sufficient to allow activation of pre-existing semantic associations*. Furthermore, at least to a degree, amnesics' normal implicit input module activation, their normal implicit access of prior semantic mediations, and their normal implicit output module activation can allow near normal performance on explicit recall of highly related word pairs (though their performance rapidly deteriorates as the semantic associations of the two words becomes less strong - see Schacter, 1985).

## Learning new Conceptual Associations

### 7. Amnesics are severely deficient at explicit recall of new pairs of associated words.

HM scored zero on this when the test required him to explicitly generate the second word of a previously studied pair when he was presented at test with the first. Generally, verbal paired associate learning of this type is very hard for amnesics even if they already know the words in question but are being asked to form a new association between them.

*To the extent to which vocabulary acquisition is learning of this type (e.g. that aardvark - isan - armadillo ....., 'what's an aardvark?'), then this deficit in amnesia tells us that this type of vocabulary acquisition is explicit. Recognition or recall of new semantic associations requires explicit memory.*

### 8. Amnesics do not seem able to implicitly acquire novel semantic associations.

The semantic priming experiments discussed in 6 above involved implicit activation of pre-existing memory associations between highly related word pairs. What about the priming of novel word associations, a laboratory analogue of implicit leaning of new meaning relations? Again, the evidence on this issue is mixed.

There were some early claims for normal priming of novel associations in amnesics. Moscovitch, Winocur and McLachlan (1986) assessed this with a task involving reading degraded pairs of unrelated words and observed normal priming of novel associations in amnesics. Schacter and Graf (1986) found that some amnesic patients - those with relatively mild memory disorders - showed normal implicit memory for a new association between unrelated words (e.g. study *bell-cradle*, test *bell-cra?*), whereas severely amnesic patients did not show such implicit memory for new associations. However, Squire's (1992) review concludes that later studies (e.g. Cermak, Bleich, & Blackford, 1988; Mayes & Gooding, 1989; Shimamura & Squire, 1989) demonstrate that amnesics do not exhibit this effect reliably.

Moscovitch et al. (1986) suggested that memory impaired patients could establish novel associations in a single trial on the basis of results from procedures where novel word pairs were presented one at a time, the subjects were asked to read as quickly as possible either (i) the same words that had already been presented, (ii) a new set of words pairs, or (iii) the old words presented in a recombined fashion in new pairs. The evidence that an association had been made between the word pairs was that the recombined word pairs were read more slowly than the repeated pairs. However, the effect was small and has proved difficult to replicate. In a recent improved study by Musen & Squire (1990) recombined word pairs were read just as quickly as old word pairs suggesting that the priming effects were at an input lexical level rather than a declarative associative one.

*The weight of the evidence is that amnesics' implicit learning is not sufficient to allow new associations between two semantically unrelated words* and Squire (1992) suggests that subjects may need to access a link between the two words that was formed declaratively (explicitly) at the time of study in order to do this.

### Combining all of these Aspects: Amnesics' Learning of New L1 Concepts and their Labels

Several studies have now provided somewhat conflicting findings concerning amnesics' learning of new labelled concepts. Gabrieli, Cohen and Corkin (1983, 1988) have shown that H.M. and a small group of other amnesic patients were unable to learn by means of rote repetition the meaning of ten English words that they did not know before.

However, amnesics can be taught new vocabulary by means of a clever technique of vanishing cues which capitalises on their preserved implicit learning abilities. Glisky, Schacter and Tulving (1986) taught amnesics a substantial amount of novel computer vocabulary by this method which is a variant of priming procedures. At the start the patient was presented with a definition (e.g. '*to store a program*') and the name of the command that enables that to happen ('*SAVE*'). On the next trial the definition was repeated, but only the first letter of the command (as in stem-completion tasks) was presented. If the patient could not answer, a second letter was presented, and so on until a correct response occurred. On the next trial the definition was presented again, alongside a fragment of the command containing one less letter than that needed for successful recall on the previous trial; thus, if a subject had been successful with *sav–*, he or she would see *sa-* on the next trial. Such learning was, of course, much slower compared to controls, and acquisitions were relatively inflexible. But nonetheless, the results were impressive in that all the patients were able to learn the appropriate commands for 15 different definitions without any cues being available. Similarly Dopkins, Kovner, and Goldmeyer (1990) have shown that Korsakoff amnesics could acquire a conceptual interpretation of a new colour name ('*bice*') but their conceptual information did not reach the same level of abstraction as that of controls and moreover it was not integrated with the rest of their colour knowledge.

According to Glisky et al. (1986) two factors could be responsible for the discordances between their study and that of Gabrieli et al. (1988): first, the use of the vanishing-cues method; second, the fact that the to-be-learned computer words in the Glisky et al. study were part of the patients' linguistic repertoire before the experiment (e.g. *save, run*) while the items used by Gabrielli et al. were uncommon words (e.g. *anchorite, manumit*). Consequently Gabrielli's patients had to learn the words and then associate them with their definitions.

Consider vocabulary learning, like paired-associate learning and other typical associative learning tasks using explicit instructions to memorise. These results suggest that amnesic patients with hippocampal damage may eventually be able to acquire new associations, by means of numerous repetitions, as in the development of a habit. But this is far from normal learning. Their rate of learning is grossly slow in comparison with normal subjects and the acquired knowledge is abnormal in other respects as well. For example, even after the knowledge is acquired, it is still relatively inflexible, i.e., accessible only when exactly the same cues are presented that were used during training. For amnesics it really is akin to *parrot fashion* learning - the patients have learned to produce a response, not to retrieve items from memory. Typically they can speak new vocabulary in the same way that a parrot can. By contrast, normal subjects learn quickly because they can apply a totally different strategy to the learning of new conceptual links. They can quickly memorise because they have an explicit, cognitive system of learning new associations which involves hippocampal and other limbic brain structures.

## Combining all of these Aspects: Amnesics' Learning of New L2 Concepts and Their Labels

Their has to my knowledge been only one published study attempting to teach an amnesic patient a foreign language. Hirst, Phelps, Johnson and Volpe (1988) describe the case of C.S., a 47 year old woman who became anterograde amnesic as a result of cardiopulmonary arrest at age 36. It is unfortunately difficult to generalise from this study because, before her accident, C.S was quite a linguist: she was a professor of Spanish and she also spoke English and had a passing knowledge of Italian. Both C.S. and her husband wanted to learn French because their daughter had recently moved to France and they wanted to visit her in the summer. Thus Hirst et al. encouraged her and her husband (who serves as the 'control' in subsequent comparisons) to take up tutoring sessions for Introductory French. There were seven sessions in the first two weeks, each lasting an hour and a half. Each of sessions 4 through 6 began with a written examination and was then followed by dictation exercises, free conversation, brief grammar lessons, and a great deal of structured question-and-answer. The question-and-answers were designed to teach grammatical rules and vocabulary through use, thereby de-emphasising rote memorisation. For example, the tutor might ask in French 'What is that?' as he pointed to a notebook, to which C.S. or her husband would respond (also in French) 'That is a notebook.' All new vocabulary was learned in this manner. Vocabulary questions were introduced five times on the day they were introduced and at least three times on the subsequent day. Repetition thereafter was sporadic and depended on the drift of the conversation. In the seven sessions the following aspects were taught: (i) present tense of the verbs '*to be*' and '*to have*' (*être* and *avoir*), (ii) indefinite articles, (iii) pluralisation, (iv) contractions ('*a les*' to '*aux*'), (v) vocabulary. At the beginning of the eighth session C.S. and her husband took an oral and written assessment designed to cover the material of the seven prior sessions. Both C.S. and her husband were exposed to 47 nouns in the course of the seven sessions. Their scores on the Vocabulary portions of the final assessment suggested that they retained about 90% of these. Moreover, such new knowledge was sufficiently flexible to permit C.S. to use newly acquired words in varied tasks not presented during the tutoring.

C.S.'s language learning seems much less laboured than that of other amnesics and her new words were open to more flexible use than is the norm. Why might these differences arise? Hirst et al. suggest a number of factors. Firstly C.S. might be a special case given her considerable linguistic expertise before her accident: they state (pp. 114-116): "Obviously, C.S. is building on her previous knowledge of languages as she learns French. Approximately 10% of the vocabulary terms could be classified as cognates of English, Spanish, or Italian.

C.S. also does not have to learn what it means to conjugate a verb or that nouns and pronouns have different genders, concepts that some native English speakers may need to acquire when learning French.". Secondly it might be a difference in procedures. C.S. clearly had trouble rote-learning unrelated pairs on both the Wechsler and the Randt (novel word pair learning like that used in Gabrielli et al., 1988), failing to retain even easy practice pairs on the Randt (e.g. 'city-town') after 5 min of the distraction, let alone 2 days later. However, instruction in French vocabulary was somewhat more implicit and naturalistic in character and involved more graded repetitions (see also Van der Linden, Meulemans, & Lorrain, 1992). Finally, Hirst et al. emphasise the difference in the relation between the new information and pre-existing knowledge. C.S. already had the concepts that the French vocabulary had to be attached to - she only had to integrate the new vocabulary into an existing contextual framework. Thus it is possible that the ease with which amnesics can acquire new information and the degree to which the resulting mnemonic representation is flexible depends upon the relation between the new information and pre-existing knowledge. Where there is little pre-existing knowledge to build on, acquisition may be difficult and the resulting representation may be inflexible or hyperspecific in use. When relevant pre-existing knowledge exists, acquisition may be less arduous and the resulting representation may be more flexible.

## Conclusions

What can we conclude about vocabulary acquisition from studies of amnesics who seem to retain implicit, automatic systems of learning in the absence of explicit, declarative learning?

1. They show normal implicit learning of the perceptual aspects of novel word forms. Thus input modules for recognising novel word forms are tuned by experience and therein new pattern recognition units (input logogens) develop simply, implicitly, and automatically as a result of frequency of exposure. (Of course, in all of these studies the learners have paid attention to the stimuli, there is no claim here that such learning can occur without the stimuli being attended to - see Schmidt, this volume).

2. They show normal implicit learning of new motor habits and the motor aspects of novel word forms that are necessary for language production. Thus output modules for producing novel word forms are similarly tuned by experience and new pattern recognition units therein (output logogens) also develop simply, implicitly, and automatically as a result of practice.
3. They are severely deficient at developing new conceptual information, at making new semantic links. In-between the implicit modules for receiving and producing language there is a conceptual system which operates according to cognitive principles, not those of habit. Vocabulary acquisition is as much concerned with meanings as it is word forms, and explicit learning is involved in acquiring and processing meanings.

What has been preserved in amnesia are the various, special purpose, relatively inflexible memory systems that permit one to behave differently as a result of experience, although usually only gradually over many trials. These deal with the abstraction of statistical regularities in the world (for perceiving new word forms) and our own behaviour (for producing new word forms). But amnesics' brain lesions do not produce loss of awareness without impairing some domain of information processing; they are unable to learn by means of the explicit cognitive system which is concerned with word meanings and which links input and output modules.

## THE IMPLICIT LANGUAGE MODULES AND THEIR ROLE IN VOCABULARY ACQUISITION

I/O modules are domain specific - e.g. visual word perception is not primed by prior listening. When automatised they are 'informationally encapsulated' (Fodor, 1983) - there are no top-down influences on the operation of input modules, and they are cognitively impenetrable - we have no conscious inklings into how they operate. They work automatically - e.g., as evidenced by the Stroop effect it is hard *not* to read a word (Stroop, 1935). Typically they are acquired by means of implicit learning: the underlying structure of a complex stimulus environment is acquired by a process which takes place naturally, simply and without conscious operations. Simple attention to the stimulus domain of words suffices for implicit learning mechanisms to induce statistical or systematic regularities in this input environment. And so our lexical systems become tuned to regularities in orthography (letter units and sequential letter probabilities), to regularities in phonology (phonemes and phonotactic sequences), to regular patterns of grapheme-phoneme and phoneme-correspondences, to high frequency words over low frequency ones, etc. And input modules recognise, and output modules produce, high frequency patterns faster as a result. The "golden rule of sensori-motor learning is much repetition" (Seibert, 1927, p. 309) - the more patterns are repeated, the more frequent they are, the better they are acquired. This is the sort of learning that connectionist models do very well, and frequency, recency, and regularity are the driving forces which tune such systems (Gasser, 1990; McLelland & Rumelhart, 1986; Plunket, this volume).

If we test vocabulary knowledge by tasks like lexical decision or word naming, which fairly cleanly tap I/O modules, then words which have a high frequency in the language, i.e. words which have had considerable life-span practice, are processed concomitantly faster. Kirsner & Speelman (in press) and Kirsner (this volume) propose a life-span practice model to explain these frequency effects whereby lexical performance in children and adults, both in L1 and L2, can be explained simply by reference to the power law of learning which Anderson (1982) uses to explain the relationships between practice and performance in the acquisition of cognitive skills generally, be they of sensory or motor nature. In so doing Kirsner is proposing that these lexical effects can be adequately explained in terms of general principles of implicit learning and skill acquisition without recourse to specifically 'lexical models'.

So it is practice that makes perfect in the input and output modules and these effects are clearly seen at a word level. We will next consider the effects of practice on statistical regularities specific to particular I/O modules.

## Reading and the Visual Input Lexicon

A reasonable consensus exists in the literature concerning the psychology of reading an alphabetic language that there are two ways of accessing the lexicon from written text, (i) a grapheme-to-phoneme route where the reader derives phonology from print by means of either grapheme-phoneme correspondence rules or analogies and uses this phonological representation for lexical access, and (ii) an orthographic route with access direct from the letter string (Coltheart, Patterson & Marshall, 1980; Ellis, 1984). Irregular words (e.g. *have*, *pint*) cannot be read correctly *with* grapheme-phoneme correspondences - i.e. exception words must be learned as exceptions. Novel words (or the experimental trick of pseudohomophones, e.g. 'kan yoo reid oar carnt u?') cannot be read *without* their application. Although these access routes may be differentially affective at different stages of reading acquisition (Frith, 1985; Ellis & Cataldo, 1990) and individuals may have particular biases (Treiman & Baron, 1983; Bryant & Impey, 1986), both routes to meaning are open to us. In skilled readers it seems that the orthographic route usually predominates since the time taken to make a lexical decision to a word of high frequency in the language is less than that to a low frequency word (Besner & Humphreys, 1991; Taft, 1991). But if the orthographic representation cannot be quickly located, i.e. if the word is of low frequency, then phonological recoding may occur. Such a mechanism may explain why the 'regularity' of the spelling-sound relationship influences measures of access time only for low-frequency words (Seidenberg, Waters, Barnes & Tanenhaus, 1984). Here we have two facets of the implicit learning of regularities by the input systems: the lexical effects of word frequency and the sub-lexical ones of spelling-sound regularity. The system also extracts orthographic regularity - the patterns of letter sequences in the language (e.g. *th* is a regular bigraph but *tz* is a rare combination), as is evidenced by the nonword legality effect (Taft, 1991) where lexical decision times to nonwords like *flink*, which have legal orthographic structure, are slower than to random letter strings like *lfkni.*

Tuning by practice is equally visible in studies which contrast people rather than words. Cunningham and Stanovich (1990) demonstrated that the amount of reading that children did, i.e. the variation in their exposure to print, accounted for significant variance in orthographic knowledge even when IQ, memory ability, and phonological abilities were partialled out.

What are the implications for learning a L2? Quite simply, the more reading practice, the more these systems will become tuned to the L2. But the rate of acquisition will be affected by transfer from L1.

### Orthographic factors:(1): Orthographic regularity and different alphabets

A native speaker of a language using the Roman alphabet transfers more easily to another of the same script than one which uses different orthographic units or frames such as the Cyrillic alphabet or the logographs of Kanji (Carroll & Sapon, 1955). There have been similarly claims that transfer is easier if both scripts contain frames that move in the same way (e.g. in rows from left to right vs. the reverse, or vertically in columns - see Nation, 1987; Desrochers & Begg, 1987).

### Orthographic factors (2): Sequential letter probabilities

Orthographic regularity is the predictability of the letter sequences in a word. Different languages have different sequential letter probabilities, for example *ll* is common at the beginning of a Welsh word but never introduces an English word. Thus the learning of the orthography of FL words may be determined by the degree to which the sequential letter probabilities match those of the native language. The same holds at the individual word level: the degree to which a particular FL word accords with the orthographic patterns of the native language affects its ease of learning (Ellis & Beaton, 1993, in press).

### Familiarity of grapheme->phoneme mappings for reading

The L2 student has to learn how FL orthography maps onto FL pronunciation.

Scripts based on alphabetic writing systems reflect to a lesser (e.g. English) or greater (e.g. Korean, Serbo-Croatian, Welsh) degree the pronunciation of language units. There are rules of correspondence between graphemes and phonemes (e.g. see Venezky (1970) for the English ‘rules’). If the FL is regular in this respect then it is easier to learn to read it - an English learner of Maori can read sentences in this language aloud, without understanding them, after only a few minutes study since Maori uses the same letters as English and the relationship between spelling to sound is very regular (Nation, 1987). Yet these rules of correspondence can differ markedly between languages sharing the same script (*pace* the naïve English learner of Welsh who continues to pronounce *f* as /f/ rather than /v/. There are further difficulties of a different type if the script is logographic (e.g. Kanji) and contains few such cues for assembling phonology from script (thus denying use of the grapheme-

>phoneme route). It may be predicted for language and word levels that: (i) the less the overlap between the grapheme-phoneme correspondence rules of the native and the foreign language, the harder it will be for the FL learner to learn to read or write that language, (ii) the less the overlap between the grapheme-phoneme correspondence rules for the graphemes of the native and the foreign word, the harder it will be for the FL learner to learn to read or write that word.

### The development of automaticity

Not until these input operations have become automatic will the reader become sufficiently skilled to have spare attentional resource to properly deal with analysis of new word meanings. Skilled reading involves a number of cognitive processes, ranging from the identification of single letters and words to the higher order processes of inference and interpretation of text meaning. Bottom-up processes include perception of letters and words, activation of semantic representations, the encoding of these representations into the contextually appropriate meaning, and the encoding of basic sentence-level propositions. Top-down processes involve higher order knowledge structures (schemata) and metacognitive abilities. These processes occur simultaneously and interact at multiple levels of analysis. There are clear differences between skilled and unskilled readers in terms of their processing. Less skilled readers in both L1 and L2 tend to focus on graphic or phonological processes in reading at the expense of higher level semantic inferential processes (Laberge & Samuels, 1974; McLeod & McLaughlin, 1986). Skilled readers have automated these bottom-up processes and are thus able to devote their attention elsewhere (Perfetti, 1985). This controlled-automatic distinction (Anderson, 1983; McLaughlin, 1990) assumes that there are capacity limitations on the amount of information that can immediately be attended to in working memory. Relative efficiency in the lower-level processes of word recognition and the encoding of local propositional and syntactic structures allows the skilled reader to allocate more resources to higher-order processes.

A commonly used measure of working memory capacity is the Daneman & Carpenter (1980) reading span test. This is a concurrent memory test designed to tap both processing and storage functions in immediate processing. The subject is required to read several sentences in a set while simultaneously remembering the last word of each sentence. At the end of each set the subject attempts to recall all the final words in the sentences in the set. The number of sentences in a set - and thus the number of sentence-final words to be remembered - is steadily increased, thus placing an increasingly greater processing load on the subject. An individual's working memory capacity is indexed as the number of final words recalled.

Reading span measures have been shown to correlate extremely well with general reading proficiency such as the SAT and the Nelson-Denny Reading Test (r = 0.4 to 0.6) and with specific reading skills such as determining pronominal reference (r = 0.9) (Daneman & Carpenter, 1980). Constant replication of this effect across studies, coupled with parallel findings obtained for listening skill, has led some researchers to suggest a central role for working memory capacity in accounting for individual differences in native language comprehension skill (Carpenter & Just, 1989; Turner & Enge, 1989). The same has been shown with L2 - Harrington (1992) demonstrates correlations between 0.5 and 0.6 between Japanese students' English L2 reading span and their performance on the Test of English as a Foreign Language (TOEFL) Grammar and Reading Comprehension abilities. For the particular case of implicit vocabulary acquisition, Daneman & Green (1986) demonstrated that reading span correlated with learning the meaning of novel vocabulary words in a context with sufficient cues for inferring meaning (r = 0.69), a correlation that was highly significant even when vocabulary knowledge effects were statistically removed (r = 0.53). In line with this, Vanniarajan (1992) has shown that whilst native and non-native speakers are equally proficient at inferring the meanings of novel words from context given unlimited reading time, native speakers excelled non-natives under normal reading time conditions, presumably because the non-natives' lack of automaticity in lower-level L2 reading processes denied them working memory capacity for on-line higher-level inferencing.

### Spelling and the Written Output Lexicon

Exactly the parallel case can be put for spelling as for reading, both in terms of means of acquisition and cognitive mechanisms.

Just as for reading, so there are dual-route models of spelling. There must be a phonological route to spelling since we can spell nonwords, and an orthographic one because we can spell words that are irregular in their phoneme-grapheme correspondences (Ellis, 1984). However, phoneme-grapheme correspondence rules are not invariably the simple reverse of grapheme-phoneme correspondences. For example, in English the phoneme /O/ is only rarely (probability = 0.15) spelled *au* (as in *auction*), yet the graphemic option *au* is almost always (p=0.95) pronounced /O/ (Berndt, Reggia & Mitchum, 1987). Yet the learner must acquire these correspondences in order to spell an alphabetic FL using a phonological strategy.

Of course we have not been taught to spell every word that we can do so correctly. We can acquire regularities of spelling patterns and even pick up irregularly spelled items simply by dint of experience. Thus Krashen (1989) argues that our knowledge of spelling is acquired incidentally and that time spent teaching spelling is time wasted. He cites the case study of Goodman and Goodman's (1982) daughter Kay who, despite having received no formal instruction in reading, could spell correctly 58% of words from a third-grade reader and recognised 91% of the correct spellings on a multiple choice test. For the case of adults, Gilbert (1935) showed that students can improve their spelling simply by reading. He presented his subjects with a spelling pretest, a reading passage containing some of the words on the pretest, and a posttest. Subjects did much better on the posttest on those words that appeared in the reading passage. Studies like this do not typically show huge gains. For example, Ormrod's (1986) college students learned only to spell an average of 2.8 out of a possible eight pseudowords (e.g. *Aerodern* as the name of a character) in reading for meaning a passage that repeated each of them six times. Krashen, however, emphasises that this learning comes free - "after a brief exposure, subjects made progress in learning to spell unfamiliar words, even when not asked to do so" (1989, p. 447) - and that if we engage in enough reading these small gains will rapidly cumulate. Recent work by Stanovich and Cunningham (1992) confirms this speculation. People who read a lot (who have 'high print exposure' as measured on Author and Magazine Recognition Tests) are better spellers even when non-verbal intelligence is controlled in the analyses.

Brown, Watson and Loosemore (under submission) have modelled the acquisition of spelling using connectionist models where all knowledge is derived from experience with instances. The model shows the same classic U shaped learning curve that young children do - upon initial training it spells both regular and irregular words correctly; then it begins to abstract regularities from the instances, becomes a 'rule-governed speller' (although of course no rules have been explicitly provided) and overgeneralises to spell irregular words as if they were regular; only later does it deal correctly again with irregular words as exceptions.

Like the visual input module, the spelling output module is tuned implicitly, simply by dint of experience.

## Speech Input and Output Lexicons

It is clear that novice language learners are bound up in the orthographic and phonological aspects of vocabulary. While native speakers lexical entries are clustered semantically (as evidenced by free associations of the type 'snow' -> 'hill' -> 'valley' etc., learners often make associations driven by orthographic or phonological confusions, e.g. 'béton' -> 'stupide' (confusion with 'bête') or 'orchestre' (confusion with 'bâton') or 'téléphoner' (confusion with 'jeton') or 'Normandie' (confusion with 'breton') etc. (Meara, 1984). Henning (1974) claimed similarly that in a vocabulary recognition task, more advanced learners and native speakers made errors indicating semantic clustering of lexical items whereas less advanced learners showed evidence of a predominance of acoustic rather than semantic clustering. Why are phonological processes not yet automatic for the L2 learner who can already speak a first language? What new learning is necessary?

### 1. Familiarity of phonological features

There are three or four dozen different independent gestures of the articulatory apparatus that play distinctive roles in human speech (Wang, 1971). Different languages make use of different ranges of articulatory features. Thus difficulty arises when the FL learner is faced with features that are not exploited in the native language. For example, the contrast between /u/ and /y/ in French pronunciation differentiates between utterances of *au-dessous* 'below' and *au-dessus* 'above'. This contrast is not exploited in English and thus the English learner of French must (a) learn to identify these unfamiliar features to perceive speech and (b) develop new motor patterns to accurately reproduce these in their own speech (Desrochers & Begg, 1987). This leads to predictions at both language and word levels:

(i) the less the overlap between the feature set of the native and the foreign language, the harder it will be for the FL learner to learn to speak that language. This is exemplified by the great difficulty that native speakers of non-tonal English have with the tonal differences which distinguish the meaning of Mandarin characters. The following five characters are all pronounced as “ma”, but with five distinctive tones. The tone marks over the vowel *a* visually capture the contour of each pitch pattern: ¯ for the First, High and Level Tone, ´ for the Second or Rising Tone, ˇ for the Third or Low Tone, ` for the Fourth or Falling Tone, and no marking for the Neutral Tone:

| | | | | |
|---|---|---|---|---|
| 媽 | 麻 | 馬 | 罵 | 麼 |
| mā | má | mǎ | mà | ma |
| (mother) | (hemp) | (horse) | (to scold) | (a particle) |

*(ii) the less the overlap between the feature set of the native language and the particular foreign word, the harder it will be for the FL learner to learn that word. Thus, for example, a Chinese student of English has much more difficulty with the words* rice*,* regular *and* eighth *(which exploit contrasts not found in Chinese) than with* pen*,* see *and* sun *(Nation, 1987).*

## 2. Combinations of features: Phonotactic regularity

The pronunciability of a word is not only determined by its phonemes and their articulatory features, but also by their position in a spoken word. Both absolute and relative position are important. An example of absolute position is /=/ (the *ng* sound) which is very common in English at the end of words, but it never occurs at the beginning. In many languages such as Hopi, Eskimo or Samoan, it is a common beginning for a word. “Our patterns set up a terrific resistance to articulation of these foreign words beginning with /=/” (Whorf, 1956). With regard relative position, just as each language has its own set of phonemes so also does it have its characteristic sequential phoneme probabilities - the sequences that constitute phonotactic regularity. Rodgers (1969) demonstrated that Russian words which were more difficult for an English speaker to pronounce were learned more slowly than those that were easier to pronounce, even if they did not have to be spoken, and Ellis and Beaton (1993, in press) show this to be so for German vocabulary even when word length, frequency, part-of-speech, and imageability are controlled. Gathercole, Willis, Emslie and Baddeley (1991) demonstrated that the 'wordlikeness' of nonwords (e.g. 'defermication' is high in English wordlikeness compared to 'loddenapish') predicted 11% of variance in children's nonword repetitions even when word length was controlled. They conclude that not only word length but also phonological structure are important determinants of ease of repetition of novel words. This is a 'linguistic hypothesis' whereby the familiarity of a novel word's phonological structure determines its repetition accuracy, with phonological frames constructed from similar vocabulary entries in the learner's lexicon being used to support the temporary phonological representation.

However, such pronunciability effects can be countered if the learner has had practice with the sounds, sound combinations and spelling used in these words (Faust & Anderson, 1967). Seibert (1927) showed that for productive learning of French vocabulary saying the words aloud led to faster learning with better retention than silent rote repetition of vocabulary lists. She emphasised that learning the novel pronunciation of FL words is as much a matter of motor skill as of auditory perceptual memory, that "it is impossible to memorise speech material without articulating it in some form or another", and that this must be practised "since the golden rule of sensori-motor learning is much repetition" (p. 309).

### 3. The role of repetition in phonological short-term memory

Such rote-repetition, that is, 'parrot fashion learning' has long been out of fashion in instructional psychology because it is not the most effective strategy to learn meaningful material (see following section on *Levels of Processing*). Yet despite this there may well be a role for repetition in learning to *produce* new FL vocabulary since this is essentially a motor skill.

Recent work in cognitive psychology suggests that individual differences in ability to repeat novel phonological patterns (phonological STM span) play a part in determining long-term vocabulary acquisition. Gathercole and Baddeley (1989) demonstrated in a longitudinal study that five year old children's native receptive vocabulary acquisition was predicted by their short-term phonological memory ability (assessed by nonword repetition) one year earlier. The children's ability to repeat nonsense predicted their vocabulary growth.

Gathercole and Baddeley (1989, 1990) suggest that representation of the novel sound sequence of a new word in phonological STM promotes its longer term consolidation both for later articulation and as an entity with which meaning can be associated. The easier a novel word is in this respect, either because of its short length or because it conforms to the learner's expectations of phonotactic sequences of language, the easier it is to learn. Phonotactic regularity might allow the novel word to better match the learner's settings of excitatory and inhibitory links between sequential phonological elements (Estes, 1972) for input processes such as phonological segmentation or for output as articulatory assembly (Snowling, Chiat & Hulme, 1991), either per se or as expectations of phonological sequences as influenced by regularities in the learner's lexicons (Gathercole et al., 1991).

There are now a number of studies using different methodologies which converge on this conclusion. The first is a training study (Gathercole & Baddeley, 1990) where children poor on nonword repetition were found to be slower than children who were good on non-word repetition at learning new vocabulary (phonologically unfamiliar names such as *Pimas* for toys). They were not slower to learn a new mapping for familiar vocabulary (familiar names like *Thomas* for the toys). Thus it appears that temporary phonological encoding and storage skills are involved in learning new words. As Gathercole and Baddeley (1990) point out, "Acquiring a new vocabulary item... must minimally involve achieving a stable long-term representation of a sequence of sounds which is linked with other representations specifying the particular instance or class of instances. The locus of the contribution of phonological memory skills seems most likely to be in the process of establishing a stable phonological representation as, in order to do this, a temporary representation has presumably to be achieved first. Immediate phonological memory seems an appropriate medium for this temporary representation and, presumably, constructing the stable long-term memory representation of the novel event will interact with the adequacy of this temporary representation. By this analysis, the better the short-term representation, the faster the long-term learning." pp. 451-452).

A second source of evidence for a relationship between phonological memory and vocabulary acquisition comes from the study by Baddeley, Papagno and Vallar (1988) of an adult neuropsychological patient, PV, who appeared to have a highly specific acquired deficit of immediate phonological memory. PV was completely unable to make associations between spoken word and nonword pairs, despite showing normal phonological processing of nonword material. She had no difficulty, however, in learning new associations between pairs of words. In other words, temporary phonological memory is particularly involved in the long-term learning of *unfamiliar* phonological material.

This relationship holds for new words whether they are of native or foreign sources. Thus Service (1992) demonstrated that the ability to represent unfamiliar phonological material in working memory (as indexed by Finnish children's ability to repeat aloud pseudowords that sounded like English) predicted FL (English) acquisition two and a half years later.

However, theories of FL vocabulary learning and the role of phonological memory systems typically fail to make the important distinction concerning direction of translation. Ellis and Beaton (1993, in press) clearly demonstrate that phonological factors are more implicated in production where the student has a greater cognitive burden in terms of motor and sensory learning.

These effects parallel those of Papagno, Valentine and Baddeley (1991) who demonstrated that the articulatory loop is used in FL vocabulary acquisition when the material to-be-learned is phonologically unfamiliar and when semantic associations via native language cognates are not readily created, but it can be circumvented if the material readily allows semantic association.

In conclusion, there is a considerable body of evidence that phonological factors are involved in (particularly productive) LT vocabulary acquisition - (i) individuals deficient in phonological STM have difficulty in acquiring the phonological representations of unfamiliar words, (ii) phonological STM span predicts vocabulary acquisition in both L1 and L2, (iii) interfering with phonological STM by means of articulatory suppression disrupts vocabulary learning when semantic associations between the native and foreign word are not readily available, (iv) nonword length and 'wordlikeness' predict repeatability, and (v) FL word regularity in terms of L1 phonotactics determines long-term learnability.

How then should we conceptualise the development of productive vocabulary from the very beginnings of entry into a new language to full proficiency? What are the causal relationships between phonological STM and the phonological aspects of LTM for vocabulary? Gathercole and Baddeley (1989, 1990) demonstrate quite clearly that phonological STM predicts LT vocabulary acquisition. Yet at the same time there are robust demonstrations of a 'long term memory' component of STM span which is independent of speech rate, i.e. STM span is greater for FL lexical items which have been encountered more often (Hulme, Maugham & Brown, 1991; Brown & Hulme, 1992). The direction of causation is neither STM->LTM, nor LTM->STM, but rather it is reciprocal. Both directions apply because new skills or knowledge invariably initially build upon whatever relevant abilities or knowledge are already present, then, as they are used, they legitimate and make more relevant (Istomina, 1975) those prior skills and knowledge and so in turn cause their further development. This is the *normal* developmental pattern. The case of reading development is a clear example. Thus Ellis (1990; Ellis & Cataldo, 1990) demonstrate reciprocal interactions between reading, phonological awareness, spelling, and STM whereby, e.g., initial levels of implicit phonological awareness determine the child's entry into reading, but reading itself causes development of phonological awareness. Similarly Stanovich has persuasively argued the case for reciprocal relationships and bootstrapping effects in reading more generally: "In short, many things that facilitate further growth in reading comprehension ability - general knowledge, vocabulary, syntactic knowledge - are developed by reading itself" (Stanovich, 1986, p364). He refers to these as 'Matthew effects' - 'unto those who have shall be given' - the more you know, the easier it is to learn more - such is growth and development.

Vocabulary acquisition is no exception to this rule. "A further possibility is that nonword repetition ability and vocabulary knowledge develop in a highly interactive manner. Intrinsic phonological memory skills may influence the learning of new words by constraining the retention of unfamiliar phonological sequences, but in addition, extent of vocabulary will affect the ease of generating appropriate phonological frames to support the phonological representations." (Gathercole et al., 1991, pp. 364-365).

The novice FL learner comes to the task with a capacity for repetition of L1 words. This capacity is determined by (i) constitutional factors, (ii) metacognitive factors (e.g. knowing that repetitive rehearsal is a useful strategy in STM tasks), (iii) cognitive factors (phonological segmentation, blending, articulatory assembly). Such cognitive language processing skills occur at an implicit level in input and output modules which are cognitively impenetrable (Fodor, 1983) *but whose functions are very much affected by experience* (hence, for example, frequency and regularity effects in reading (Morton, 1969; Baron & Strawson, 1976; Seidenberg, Waters, Barnes & Tannenhaus, 1984; Brown, 1987; Paap, McDonald, Schvaneveldt, & Noel, 1987), spelling (Barron, 1980; Barry & Seymour, 1988), and spoken word recognition (Morton, 1969; Marslen-Wilson, 1987)).

The degree to which such relevant skills and knowledge (pattern recognition systems for speech sounds, motor systems for speech production) are transferable and efficient for L2 word repetition is dependent on the degree to which the phonotactic patterns in the L2 approximate to those of the L1, hence the phonotactic regularity effects at both language and individual word levels. Here then we have LT knowledge affecting phonological STM, i.e. the 'linguistic hypothesis' of Gathercole et al. (1991).

'Good language learners' (Naiman, Fröhlich, Stern, & Todesco, 1978) know that repetition and practice of new vocabulary are useful strategies (O'Malley, Chamot, Stewner-Manzanares, Kupper, & Russo, 1985). They do this and in so doing they acquire LT L2 vocabulary. Here we have phonological STM determining LT vocabulary acquisition (Gathercole & Baddeley, 1989, 1990).

As their L2 vocabulary extends, as they practise hearing and producing L2 words, so they automatically and implicitly acquire knowledge of the statistical frequencies and sequential probabilities of the phonotactics of the L2. Their input and output modules for L2 processing begin to abstract knowledge of L2 regularities, thus to become more proficient at ST repetition of novel L2 words.

And so L2 learning lifts itself up by its bootstraps.

To this extent vocabulary acquisition involves simple implicit learning processes. The sensory learning components involve the formation of pattern recognition units for speech sounds up to the word level (logogens) in the input module for speech perception and these are tuned by frequency and regularity of exposure. The motor learning components involve the formation of speech production units up to the word level (output logogens) in the output module for speech production and these are similarly affected by regularity and frequency of practice. The frequency of activation of these units can be increased by use of phonological STM to recycle phonological language representations.

## THE EXPLICIT CONCEPTUAL SYSTEM AND ITS ROLE IN VOCABULARY ACQUISITION

Items of experience are classified differently by different languages. The class corresponding to one word and one thought in Language A may be represented by Language B as two or more classes corresponding to two or more words and thoughts (Whorf, 1956). Thus, for example, Desrochers and Begg (1987) refer to the French distinction between *balle* - a spherical object which can be caught with one hand, and *ballon* - that requiring both hands; the English translation *ball* is insufficient to represent and distinguish these meanings. Terms for colour, temperature, divisions of the day, kinship and parts of the body are all semantic fields that are divided up in different ways in different languages (Carter & McCarthy, 1988). Navajo has a fourth person singular and plural, which is used to address someone in the room or within earshot without naming him or her directly, and many African languages have inclusive and exclusive forms of the first person plural (we including you to whom I am speaking vs. we not including you to whom I am speaking). Hopi has one noun that covers every thing that flies, with the exception of birds - Hopi Indians call insect, plane and aviator all by the same word and feel no difficulty about it. These few examples demonstrate the phenomenon of linguistic relativity (Whorf, 1956). Learning a new FL word is going to be easy if there is a 1:1 mapping of meanings represented by the native and foreign words. It is going to be harder if the same conceptual fields are covered by different lexical fields in different languages (Carter & McCarthy, 1988). Ijaz (1986) demonstrated that even advanced adult ESL learners differed substantially from native speakers in the semantic boundaries that they ascribed to English spatial prepositions, with word usage being heavily influenced by native language transfer. She concludes (p. 443) "the second language learners essentially relied on a *semantic equivalence hypothesis*. This hypothesis facilitates the acquisition of lexical meanings in the L2 in that it reduces it to the relabelling of concepts already learned in the L1. It confounds and complicates vocabulary acquisition in the L2 by ignoring crosslingual differences in conceptual classification and differences in the semantic boundaries of seemingly corresponding words in the L1 and L2."

The implications for FL learners are clear: where the native language does not encourage the distinction between concepts then students necessarily will have an additional conceptual chore in learning the FL which relies on these very distinctions. The greater the mismatch, the greater the problem: two French balls present less difficulty than 22 (or however many it is - Whorf, 1956; Lakoff, 1987) forms of Eskimo's snow.

### Word Class and Word Meaning: Semantics and Conceptual Representations

The part of speech of a word affects its learning: nouns are the easiest to learn, adjectives next, while verbs and adverbs are the most difficult to learn in explicit FL vocabulary list-learning experiments (Rodgers, 1969). These word-class effects are also found in other psycholinguistic performance measures, for example, Broca's aphasics have more difficulty in producing function words and inflexions in their speech than substantives (agrammatism - Ellis & Young, 1988, Ch. 9); deep dyslexic patients also have greater difficulty reading function words including auxiliary verbs, adverbs and pronouns (Morton & Patterson, 1980; Patterson, 1981); meaningful nouns produce substantially more interference in Stroop tasks than do relatively meaningless function words (Ehri, 1977; Davelaar & Besner, 1988); children acquire nouns before they do other parts of speech (Gentner, 1982). These effects may directly reflect grammatical word-class or they may stem from imageability (in general, nouns are more imageable than verbs - Davelaar & Besner, 1988; Ellis & Beaton, 1993) or meaningfulness (imageable items are more meaningful - Paivio, Yuille & Madigan, 1968; Ellis, 1991).

There are diverse psychological theories of meaning, but many posit that the element representing a word in semantic memory is associated with a number of features or, more fully, predicates. This assumption has been used to analyse work in sentence verification (e.g. Anderson, 1976), category prototypes (e.g. Rosch & Mervis, 1975), concepts (e.g. Schank & Abelson, 1977), basic categories (e.g. Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), similarity (e.g. Tversky, 1977), metaphor (e.g. Ortony, 1979), episodic memory (e.g. Tulving, 1983), semantic priming (e.g. Meyer & Schvaneveldt, 1971), and deep dyslexia (Jones, 1985). All of these models are concerned to represent meanings, and propositional representations are well suited to this end - knowledge is represented as a set of discrete symbols which are linked by associational relations to form propositions; concepts of the world are thus represented by formal statements, with the meaning of a concept given by the pattern of relationships among which it participates (see Rumelhart & Norman, 1985 for review). Meaningful concepts have many relations; less meaningful ones have few. When Rubin (1980) factor-analysed 51 psycholinguistic variables measured for 125 words, Paivio, Yuile & Madigan (PY&M) (1968) imageability, meaningfulness (m, associative frequency and categorizability), and concreteness loaded on the *same* factor, suggesting that these measures have much in common. Another way of operationalising this definition of meaning is to measure the 'ease of predication' of the word, i.e. the ease with which what the word refers to "can be described by simple factual statements" (e.g. a *dog* is a type of animal, a *dog* barks when angry, a *dog* has four legs, a *dog* wags its tail when pleased, a *dog* often lives in a kennel, etc., *vs*. an *idea* ..., Jones, 1985). When Jones (1985) had subjects rate 125 nouns for ease of predication, there was a very high correlation (r = 0.88) between this measure and PY&M imageability. When he chose a measure of predication time (the mean number of seconds taken to produce two predicates for each word) there was a correlation of r = -0.72 with PY&M imageability (Jones, 1988). These are high correlations - it seems that imageability and predicability go hand in hand. Schwanenflugel, Harnishfeger & Stowe (1988) and Schwanenflugel (1991) argue that the greater meaningfulness of imageable words arises from their greater 'context availability', a concept very similar to that of predicability. In this view imageable concepts, as a result of their experientially based cores, more easily allow access of relevant world knowledge, or 'inner provided contexts' that add meaning relations to the word. The common feature of all of these theories is that things experienced and analysed visually are imageable things are meaningful things about which we have co-ordinate and subordinate semantic information.

Gentner (1982) emphasises the parallelism of vision which allows for ready associations: good concrete objects are cohesive collections of percepts since objecthood is created by spatial relations among perceptual elements. These perceptual elements that are packaged into noun referents are highly cohesive (i.e. have many internal relations to one another), while the perceptual elements that are packaged into verb referents are distributed more sparsely through the perceptual field and have fewer internal relations with one another. Thus noun concepts are both richer, and nouns are more easily mapped onto discrete perceptual experiences. Hence nouns are more meaningful and more easily acquired in either first or second languages. Furthermore the nouns that appear in the child's first words are all either concrete or proper nouns; they centre on concrete percepts not abstract vagaries. Similarly Ellis (submitted) proposes that imageability effects in verbal learning reflect the fact that visual imageability confers meaning, or, as Lakoff & Johnson (1981) and Barsalou (1991) suggest, symbols are grounded in our perceptual experience, i.e. imageable items are meaningful items are memorable items.

## The Effectiveness of Explicit, Deep Processing, Mediational Strategies in L1 Word Learning

Over a century ago William James, one of the founders of modern psychology stated: "Of two men with the same outward experiences and the same amount of mere native tenacity, the one who thinks over his experiences most and weaves them into systematic relations with each other will be the one with the best memory" (James, 1890). In a similar emphasis, a similarly influential progenitor, Bartlett emphasised that ''memory is an effort after meaning" (Bartlett, 1932). These views are as relevant to learning vocabulary as they are to any other domain.

A more recent statement of this theme has been encapsulated in the *Levels of Processing* framework of Craik and Lockhart (1972, Craik & Tulving, 1975). In this model learners are seen as strategic - they can choose how to process incoming material. Information can be encoded in multiple forms in memory: for example, it could be in terms of semantic, phonemic, or visual features; in terms of verbal associates, or as an image. But different forms of encoding persist for different amounts of time. Our processing of information moves from a sensory level of analysis, through pattern recognition to semantic enrichment. Craik and Lockhart suggest that "memory trace persistence is a function of depth of analysis, with deeper levels of analysis associated with more elaborate, longer lasting, and stronger traces" (Craik and Lockhart, 1972, p. 675). Only the results of the deeper analyses are stored, while the by-products of preliminary analyses are discarded because the *meaning* of the information is what is typically needed later. We do not remember messages verbatim, we remember them for their gist. Highly familiar or meaningful material (like sentences, pictures, or common expressions) are so easily remembered because, Craik and Lockhart argue, these kinds of inputs are highly compatible with previously existing cognitive structures; therefore they are processed to a 'deep' level faster and more completely than unfamiliar information. In summary, *Levels of Processing* holds that Type I processing (e.g. maintenance rehearsal of an oral representation by means of repetition) does not necessarily lead to long-term retention and that 'deeper' Type II processing, whereby semantic associations are accessed and elaborated, is necessary for the formation of long-term representations.

A demonstration by Hyde and Jenkins (1969) illustrates these notions. They presented subjects with lists of words and various orienting tasks to perform on each word. One group of subjects had to make a pleasantness rating, one had to check for presence of the letter E, while a third group counted the number of letters in the word. In addition, a fourth group was merely presented with the word lists and the instructions that they were to learn them and would later be tested for recall. A clear pattern of results emerged on the explicit recall test. Group four subjects who were intentionally learning remembered on average 16.1 words. In contrast the three groups who were just analysing the material and whose learning came incidentally remembered in turn, E checking, 9.4; Number of letters, 9.9; Pleasantness rating, 16.3. Thus the results for those subjects with standard intentional learning instructions and those charged with making pleasantness rating (a semantic task) were almost identical, even though the subjects making pleasantness ratings were unaware of the impending recall test. The subjects who only had to analyse the words to shallower depth (checking for E's and counting letters) performed more poorly. Here then, as is generally the case with meaningful material, the more subjects analyse for meaning, the better they recall.

But there is another 'deep' form of processing. As is described above, word meaning can be semantic in terms of other words or propositions or it can be conceptual in terms of relations to the memories of perceptual experiences of the referents of the words. Imageable words, those that have referents which are easily imagined in the mind's eye, are much more meaningful.

When people are asked to explicitly learn lists of words, the greater the imageability of a word, that is the degree to which it arouses a mental image, the more likely it is to be recalled. This is a robust effect in free recall experiments (Paivio, 1971). It is even more reliable in paired-associate learning (PAL), a laboratory analogue of vocabulary learning, where the subject has to learn a novel association of a stimulus word experimentally paired with a response word (e.g. *fish-chair*) (Paivio, 1971; Rubin, 1980). This effect has withstood many attempts to demonstrate that its association with recall is spurious and attributable to *tertium quid* psycholinguistic attributes such as meaningfulness (Paivio, Yuille and Smythe, 1966; Dukes and Bastian, 1966; Christian, Bickley, Tarka and Clayton, 1978; Rubin, 1983), concreteness (Christian et al., 1978), familiarity (Paivio, 1968; Frinke, 1968) or age-of-acquisition (Gilhooly and Gilhooly, 1979). And this effect is equally true for the particular case of FL vocabulary learning: concrete FL words are generally learned earlier and more easily than abstract words (Carter, 1987; Ellis & Beaton, 1993, in press).

Within the literature on organisation and memory, mediation is discussed predominantly in terms of semantic links. A relationship might be made between the stimulus and response words because it taps into pre-existing semantic links, thus, at the extreme, highly associated pairs (e.g. *doctor-nurse*) are easy to learn (Jenkins, Mink and Russell, 1958). Otherwise subjects may choose from a wide variety of strategies. The experiment of Bower and Winzenz (1970) which investigates the effect of various learning strategies on paired-associate learning illustrates the usefulness of the two deeper strategies of semantic and imagery mediation. Subjects were all required to learn to associate 15 arbitrary pairs of words (e.g. *horse-cello*). They were allocated to one of four conditions: (i) Repetition: they were asked to repeat or rehearse each pair of words silently to themselves during the study time, (ii) Sentence Reading: subjects saw each pair of words as capitalised subject and object nouns in a simple declarative sentence, and were told to read the sentence aloud and use it to associate the two critical words, (iii) Sentence Generation: subjects were shown each pair of words and asked to construct and say aloud a meaningful sentence or phrase using and relating the two words in a sensible way, (iv) Imagery: subjects were asked to visualise a mental picture or image in which the two referents (denoted by the words) were in some kind of vivid interaction. The mean recall results in each condition were as follows: Repetition 5.6, Sentence Reading 8.2, Sentence Generation 11.5, Imagery 13.1. According to Bower and Winzenz, in the Repetition condition, the two words are related to each other in terms of temporal contiguity, but few relations other than contiguity are used for storing the word pair (i.e. it is a relatively shallow level of processing). Embedding the two words in a declarative sentence results in a richer memory representation involving the semantic relations implicit in the sentence. The difference between this condition and the Sentence Generation condition is that when subjects have to search out and construct a linking relationship between the two words they are more likely to activate and thereby associate the full meanings of the words than when they merely have to read the sentences. The Imagery condition leads to the formation of both verbal and imagery long-term representations - the instructions force the subject to generate a visual scene, which is stored in an imagery memory representation, which depicts an 'actor-action-object' relational bonding between the two words, and this underlying semantic relation is stored as a verbal memory representation.

But additional possibilities for mediation are endless - even with research on the explicit learning of nonsense syllables, Baddeley (1976, p.273) writes: "Anyone who has worked with nonsense syllables will know that, despite the effort put into scaling their association value, familiarity, pronunciability, and so forth, the correlation between these measures and the learning of individual syllables is low... The probable reason for this lack of consistency is that subjects will use any strategy they can devise to give meaning to an item or pair...

Given a flexible and ingenious subject, it is hardly surprising that no measure based on a single coding dimension has proved to be an accurate or reliable predictor of learning. Indeed, the best way of predicting the difficulty of a nonsense syllable pair is still to ask subjects (Prytulak, 1971), presumably because subjects can base their judgements on the whole range of possible coding strategies, whereas most measures are based on a single coding dimension.". With meaningful material we give the fertile mind even more scope, and with imageable material yet more again. The next section demonstrates that these ideas hold just as well for mediational aspects of L2 vocabulary learning.

## The Effectiveness of Explicit, Deep Processing, Mediational Strategies in L2 Vocabulary Learning

In language programmes the world over, teachers recommend a variety of mnemonic techniques for L2 vocabulary learning (Merry, 1980; Nattinger, 1988) and learners frequently learn lists of vocabulary whether encouraged to do so or not (Carroll, 1963; Rodgers, 1969; Nation, 1987). They do it because it works: without too much effort the student can learn well over 30 words per hour by studying lists of vocabulary (Thorndike, 1908; Webb, 1962; Crothers & Suppes, 1967; Nation, 1982; Krashen, 1989, Appendix II). Richards (1976), Brown (1980), Nation (1982, 1987), Carter (1987), Carter and McCarthy (1988), Chall (1987), Graves (1987), and Beck, McKeown and Omanson (1987) all advocate the teaching of vocabulary explicitly, but by a range of means far beyond mere parrot fashion learning: The strategies advocated include collocation (combining words in high probability sequences in clichéd phrases), clines (e.g. scales such as cold/warm/hot expressed diagramatically), cloze procedures, inference from context, consultation of dictionaries, card indices of new vocabulary, guessing skills and so on.

Carter's thorough review of the applied psycholinguistics of vocabulary instruction concludes that for most learning purposes, vocabulary needs to be taught for comprehension and for production: (i) in early stages of learning a language, teaching techniques which foster imagery associations across L1 and L2 can be valuable; (ii) the more advanced the learner becomes and the more emphasis is placed on production, then the more teaching of words in a network of semantic associations should be activated. The teaching of words in semantic sets or grids can be beneficial here; (iii) The skills of guessing and of using contextual cues to make inferences is important, especially in reading in a foreign language; (iv) teaching fixed expressions (collocations, idioms, etc.) can be valuable at all levels and is especially important to allow learners access to more routinized aspects of production; (v) teaching words in discourse fosters the development of advanced skills of production but also encourages appreciation of the syntactic, semantic, and pragmatic functions of lexical items at all levels. Too great a focus on learning vocabulary as individual decontextualized items may lead to neglect of these skills (Carter, 1987, pp. 187-189). Conclusion (i) emphasises the role of imagery mediation in L2 vocabulary acquisition, conclusion (ii) semantic mediational strategies, and conclusion (iii) the autonomy of the active learner and the effectiveness of their metalinguistic knowledge in inferencing and deep processing and mediation. I will deal with evidence for each of these in turn.

### 1. Imagery Mediation using Keyword Methods

As early as 1862 the potential for learning French vocabulary by means of an associative link between the English and French word had been realised by the Rev. J.H. Bacon (see Desrochers & Begg, 1987). He described learning the French word *arbre*, meaning *tree*, by thinking of the arbour at the foot of his garden which was in the shade of an overspreading tree. The mental process described by Bacon can be represented as follows:

*arbre* (word)->*arbour* (word)->*arbour* (image)->*tree* (image)->*tree* (word).

More recently Atkinson and Raugh (1975) reported an experiment in which they compared learning of FL vocabulary by means of mnemonics with a control condition in which subjects used their own strategies. In the experimental condition subjects were presented with a Russian word and its English translation together with a word or phrase in English that sounded like the Russian word. For example, the Russian word for *battleship* is *linkór*. American student subjects were asked to use the word *Lincoln*, called the keyword, to help them remember this. Atkinson and Raugh found that subjects who had used the keyword method learned substantially more English translations of Russian words than the control group and that this advantage was maintained up to six weeks later.

Numerous subsequent studies have confirmed the effectiveness of the keyword method in FL and native language vocabulary learning (see Pressley, Levin & Delaney, 1982; Levin & Pressley, 1985; Desrochers & Begg, 1987; Cohen, 1987 for reviews). The common explanation for the success of these systems is that the keyword enables subjects to combine in a single associative image the referent of one native word with that of a second native word which sounds like the foreign word.

This is a two-stage view. The first stage of recalling the meaning of a foreign word involves the subject remembering the native keyword which sounds like the foreign word. The second stage involves accessing an interactive image containing the referent of the keyword and 'seeing' the object with which it is associated (this is the equivalent of the Imagery mediation condition of Bower & Winzenz, 1970, above). By naming this object the learner accesses the native translation. The keyword does not have to sound similar to the foreign word - an approximation can serve as a retrieval cue for the FL word.

In a comprehensive survey of almost fifty studies of the keyword technique Pressley et al. (1982) conclude: (i) the keyword technique is superior to other techniques such as rote repetition, placing vocabulary in the context of a meaningful sentence, and using pictures as synonyms; (ii) it can be used effectively not only with concrete nouns, but also with verbs, abstract nouns, and adjectives; and (iii) the keyword technique does not slow down recall of the foreign word.

There are, however, some limitations of the keyword technique (Ellis & Beaton, 1993, in press).

Firstly it is much less effective with abstract vocabulary and keywords of low imageability Ellis & Beaton, 1993). Thus, e.g., Raugh, Schupbach & Atkinson (1977) found that the keyword method worked better with nouns and verbs than with adjectives in the acquisition of a large Russian vocabulary and suggested that this was a result of the adjectives being more abstract (less imageable). Atkinson and Raugh (1975) reported that the probability of remembering the image-based link between keyword and native word in one set of subjects correlated 0.49 with the relative recall of the native words (given the foreign word) by other students learning FL vocabulary under keyword instructions. The method rests on an imagery mediation between the L1 word and the keyword, if either of these are abstract and difficult to image then the method fails.

Secondly, it is much less effective in productive vocabulary learning than in learning to comprehend the L2 (Ellis & Beaton, 1993). This is because, as Pressley, Levin, Hall, Miller, and Berry (1980) admit, there is no mechanism in the keyword method to allow retrieval of the whole L2 word from the keyword. Thus it is not surprising that the use of the method does not increase whole foreign word recall. The imagery association in the keyword technique allows retrieval of the keyword which is merely an approximation to the L2 form. There is thus no reason to expect keyword techniques to help spelling or pronunciation once an approximation to the correct response has been retrieved since the technique does not have any in-built 'mnemonic tricks' to help spelling or pronunciation. Therefore, to be more effective in productive vocabulary learning than other retrieval techniques, the form of the keyword may have to have a close connection with the form of the foreign word. If care is taken with this then the keyword technique is good for productive vocabulary learning. Hence the first of Raugh and Atkinson's (1975, p.2) criteria for a good keyword is that it 'sounds as much as possible like a part (not necessarily all) of the foreign word' and their demonstration of a correlation of 0.53 between the probability of a keyword being remembered given a Russian word and the probability of the English translation being remembered by different subjects using the same keyword as a mnemonic (Atkinson & Raugh, 1975). While it may be relatively easy to find English keywords which sound like some foreign words (e.g. for the German words: *Blech, Böttcher, Decke, Flitter*), others are considerably more problematical (e.g. the German nouns: *Abhilfe, Bleiarbeiter, Durchschlag, and Geschluchze*) (Desrochers & Begg, 1987; Nation, 1987; Hall, 1988). Word recall is likely to be best if the keyword or part of it overlaps with the initial part or cluster of the foreign word to be recalled (Horowitz, Chilian & Dunnigan, 1969; Loess & Brown, 1969; Desrochers & Begg, 1987). However, even in an experiment by Hall, Wilson and Patterson (1981, Experiment 2), where the keywords were identical to the first syllables of the foreign words, there was a marginally significant *inferiority* of keyword group performance in productive learning of the spelling of Spanish vocabulary.

Pressley and Levin (1981) reasoned that the keyword method might have a facilitative effect on productive retrieval from the moment the unfamiliar responses were integrated or available in memory. The learners were thus pre-familiarised with the unfamiliar target responses. When cued by the definitions, later recall was higher in the keyword than in the free-strategy control condition. Although Paivio and Desrochers (1979), using a variation of the key word method in which interactive imagery was employed but no explicit acoustic relation existed between the word to be learned and the nature of the keyword image, found superior recall for subjects using imagery as compared with those using rote repetition in productive recall of French words, the same authors in 1981 (p. 784) conclude "The keyword method does not have an immediate facilitating effect on productive recall, unless the new response items are already available in memory". Similarly Desrochers and Begg (1987, p. 67) argue, '"access with the keyword fragment is insufficient for the reproduction of the whole unfamiliar item, unless it has been learned previously" - but there are few FL learning situations where the learner knows the words of the FL, their pronunciation and their spelling, but not yet their translations (Cohen, 1987).

In conclusion, the keyword technique can be used to achieve one necessary component of vocabulary learning - the mediation of L1 and L2 forms - but there is some reason for concern over the sufficiency of keyword strategies for productive vocabulary learning. Thus Ellis & Beaton (1993) demonstrate that for effective productive vocabulary learning the keyword technique must be complemented with repetitive practice at producing the L2 word forms. Imagery mediation does not contribute to the lexical productive aspects of L2, but it does forge L1-L2 linkages.

## 2. Semantic Mediation

### *(a) Using keywords*

Sometimes FL words just remind us of the native word, a factor which usually stems from the languages' common origins or from language borrowing. Cognates are pairs of words that show sound-meaning correspondences indicating their historical relationship (Banta, 1981). Thus the German *Hund* (dog) may be more easily retained than the French *chien* because of its etymological and sound similarity with the English *hound* (Nation, 1982). Such reminding, whether based on orthography, phonology, etymology or ‘borrowing’ (e.g. ‘le hot-dog’) typically facilitates the learning of that FL word (Anderson & Jordan, 1928). There can, of course, be interference when such reminding is inappropriate (e.g. the Englishman mentally groping for a French hug might be happily surprised to get more than he bargained for if he lunged at *embrasser*).

If the reminding is not naturally there, one can create it using keywords and semantic rather than imagery mediation. By simply remembering the keyword and the native word in a mediating sentence it is possible to derive the translation (the equivalent of the Sentence Generation condition of Bower & Winzenz, 1970). Associations of either verbal or imaginal or both types may thus underlie the effectiveness of the keyword technique. Thus, in one study, Presley, Levin and Miller (1981) found that fifth grade school children learning ten abstract and ten concrete words in Spanish were just as successful learning concrete nouns when they used the imagery as when they used the verbal mnemonic approach. On abstract nouns, however, the group using verbal mnemonics was more successful than the one using the imagery technique. In contrast, Atkinson (1975) working with adults and Kasper (1983) working with middle school children did find some advantage of interactive imagery over use of a linking sentence in promoting recall of Spanish - English word pairs. These results, together with those of Ellis & Beaton (1993), suggest that there is an advantage of imagery mediation with concrete vocabulary and keywords, but semantic linkage is more effective when either of this pair are abstract.

### *(b) Deep Processing and Elaboration*

Beck, McKeown, and Omanson (1987) advocate that learners focus on the meaning of the new word and that they should act upon this meaning in a way that is considered integrative in relation to already existing semantic systems. This is more than merely associating a word with its definition, rather, “students should be required to manipulate words in varied and rich ways, for example by describing how they relate to other words and to their own familiar experiences” (p.149). Beck et al. are thus urging students to be, in Craik and Lockhart’s terms, deep and elaborative, Type II processors. Crow and Quigley (1985) evaluated the effectiveness for ESL students of several such semantic processing strategies (such as the 'semantic field' approach where subjects manipulated synonyms along with the target words in meaningful sentences) and found them to be superior to 'traditional methods' over long time periods.

It can be advantageous to combine all of these aspects of (a) use of keyword reminders and (b) elaborative processing. Thus Brown and Perry (1991) contrasted three methods of instruction for Arabic students' learning of English vocabulary. The keyword condition involved, presenting the new word, its definition, and a keyword, and learners were given practice in making interactive images; the semantic condition presented the new word, its definition, two examples of the word's use in sentences, and a question which they were required to answer using the new word; the keyword-semantic condition involved all of these aspects. A delayed testing over a week later demonstrated that the combined keyword-semantic strategy increased retention above the other conditions.

### 3. Metalinguistic Strategies for Inferencing and Remembering

Rubin (1981) and O'Malley, Chamot, Stewner-Manzanares, Kupper and Rosso (1985) identified a number of strategies which learners adopt to aid their vocabulary acquisition. These include *Clarification/ Verification*: asking for examples of how to use a word, asking for the correct form to use, etc.; *Monitoring*: correcting pronunciation or vocabulary with respect to appropriateness for the setting; *Memorization*: attempting to acquire words through associations (Imagery, Auditory practice, Keyword, Elaboration; Note-taking, etc.) designed to assist storage and retrieval; *Guessing/ Inductive Inferencing*: using hunches derived from clues from the surrounding language context or from an item's repeated use in different contexts to guess meaning; *Practice*: experimenting with new sounds, drilling self on words in different forms, making use of new words in speaking, etc. Of course, the fact that learners adopt these strategies does not necessarily mean that the strategies are effective. What is needed for this conclusion is at least correlational evidence whereby it is the better language learners who report using these strategies and finding them helpful, or, better still, causal evidence that people trained to use these techniques improve as a result (see Chamot & O'Malley, this volume).

Early correlational evidence of this type is available from Rubin (1981) and Naiman, Fröhlich, Stern and Todesco (1978). More recently O'Malley, Chamot and Küpper (1989) recorded the think-aloud protocols of high school students as they were listening to academic presentations in English. These demonstrated that effective listeners monitored their own comprehension, associated new information to prior knowledge, and made inferences about unknown words significantly more often than ineffective listeners.

More persuasive training study evidence can be found in Cohen and Aphek (1980) and O'Malley, Chamot, Stewner-Manzanares, Rosso, and Kupper (1985). Cohen and Aphek (1980) showed that students who were instructed to look for inter- and intra-lingual mnemonic associations generally retained new words with greater efficacy. O'Malley et al. (1985b) randomly allocated ESL high school students into either a control group, a group receiving both metacognitive and cognitive strategy instruction, and a group receiving only cognitive strategy instruction. Instruction lasted one hour a day for two weeks. At the end of this training the metacognitively trained group outperformed the other on a transactional speaking task and on some aspects of daily listening comprehension tests.

Let us finally consider the particular strategy of decontextualisation which allows students to infer word meanings from context, thus to increase their vocabulary (Sternberg, 1987). Twadell (1973) suggests that 'knowing a word' is a continuous rather than discrete process: we may 'know' a very large number of words with various degrees of vagueness; at any given time there are words we know well, words we do not know, and words that are in-between; vagueness is reduced bit by bit as we read more and encounter unfamiliar words more and more.

Sternberg (1987) presents a thorough analysis of learning vocabulary from context. He identified three basic subprocesses: *selective encoding* (separating relevant from irrelevant information for the purposes of formulating a definition), *selective combination* (combining relevant cues into a workable definition), and *selective comparison* (relating new information to old information already stored in memory). He categorised the types of available cue and the following moderating variables: (i) the number of occurrences of the unknown word. (ii) the variability of contexts in which multiple occurrences of the unknown word appear, (iii) the importance of the unknown word to understanding the context in which it is embedded, (iv) the helpfulness of the surrounding context in understanding the meaning of the unknown word (e.g. an equivalence cue such as ‘an *ing* is a low-lying pasture’ is most effective, a spatial cue such as ‘the cows grazed the *ing* in the shadows of the surrounding mountains’ is more effective than a temporal cue such as ‘at dawn the cows grazed the *ing’* for a spatial concept, etc.), (v) the density of unknown words (too high a proportion of unknown words will thwart attempts to infer meaning), (vi) the usefulness of previously known information in cue utilisation. Subjects trained in use of these moderating variables or given practice in the processes of inferencing from context showed marked gains over control subjects in vocabulary acquisition from texts in a pretest-posttest design similar to the *Clockwork Orange* studies mentioned above.

There are now a number of demonstrations which suggest that children derive benefit from metacognitive training programs which promote the metalinguistic skills underlying the inferencing and guessing of new vocabulary from its context (Graves, 1987; McKeown & Curtis, 1987; Yuill & Oakhill, 1988; Fawcett & Nicolson, 1991). In particular, Buikema and Graves (1993) compared a control class of junior high school L1 students of English with a training group who were instructed over five days thus: (i) they were told that they could learn words on their own and that they were going to be shown how to use context cues to determine the meanings of unknown words they ran across as they were reading; (ii) they were given experience in guessing words from riddles (e.g. “I am a colour that symbolises wealth. I am often seen on the robes of kings and queens. I am also seen on the petals of flowers which have African in their name. My name is included in the title of Prince’s most famous movie. What word am I?” Answer: “*purple*”); (iii) they were told to box in an unknown word in a passage, to write it below, to list the words and phrases that were cues to the possible meaning of the word, to think of cognates, to analyse the part of speech, and thence to guess. In posttests the training group significantly outperformed the control group in inferring the meanings on difficult or novel words in a variety of passages. Not only does such training promote inferencing from context, but also this active derivation of meaning makes the vocabulary more memorable. Thus, for the particular case of L2, Hulstijn (1992) provides some experimental support for a *Levels of Processing* or 'mental effort' hypothesis of vocabulary acquisition whereby inferred word meanings were retained better than those given to the reader through the use of marginal glosses.

### 4. Summary Conclusions

*Taking these results together it is clear that it truly matters what learners do in order to acquire the meaning of a new word. Sophisticated metacognitive knowledge allows them to choose suitable cognitive learning strategies appropriate for particular tasks. And the strategies that are relevant to vocabulary acquisition include inferring word meanings from context, semantic or imagery mediation between the FL word (or a keyword approximation) and the L1 translation, and deep processing for elaboration of the new word with existing knowledge.*

## Cognitive Aspects of the Involvement of Explicit, Episodic Memories in the Representation of Features of Vocabulary Meaning

There are other reasons, besides those of depth of processing and the problem-solving of meaning, underlying why the acquisition of the meanings of words requires explicit processing.

Young children, like L2 acquirers, do not completely understand a word on its first occurrence. They may well, as Chomsky (1988) urges, possess rich conceptual representations to which new words will refer, but initially children only represent some of the semantic features that characterise the full meaning of words. Some children may over-extend *horse* to other mammal-shaped objects, while others under-extend and restrict it to a particular toy. Some children over-extend a relationship term like *off* to cover any situation in which two objects move apart or are moved away from each other, while others may restrict it to the removal of clothes (Bowerman, 1976; Clark & Clark, 1977). According to missing-feature theory (e.g. Clark, 1973), the process of lexical development consists of adding features until the full entry is achieved. Consider, for example, spatial adjectives. *Big* and *little* are the first spatial adjectives learned in English L1. *Big*, *tall*, *wide*, *high*, etc., might initially be marked only [Adj] [comparative] [±pole] [spatial extent]; with the other features such as [vertical], and [primary] only gradually being added. On this view, *tall*, *wide*, *high*, etc., are all first synonyms of *big*; later *tall* and *high* might be synonyms, [vertical] having been added to each entry. Similarly, at first *short*, *narrow*, low, and so on, are all synonyms of *little*; later *thin* and *narrow* might be synonyms (Clark, 1972).

But how might these features be added? Carey (1978) proposes a theory of *Missing Features and Haphazard Examples*. The immature lexical entries for spatial adjectives might contain information about some particular objects to which each adjective applies. The entry for *tall* might include that it applies to buildings and people, of *short* that it applies to hair, people, and distances, etc. Thus sample lexical entries might be:

tall: [Adj] [comparative] [+pole] [___building, ground up; ____person, head to toe]

short: [Adj] [comparative] [-pole] [___person head to toe; ___hair, root to end; ___distance, direction of motion]

Carey suggests that the child learns object by object what spatial adjective applies to what kinds of variation. At the very beginnings of the child's experience only one or two of such aspects of particular objects to which the adjective applies will be represented. Those included will reflect the child's haphazard encounters with the word. The process of lexical growth then has two components: (i) the discovery of semantic features: the child must discover what aspects of the conceptual system are relevant to the structure of the semantic domain, i.e. which features of these contexts serve as lexical organisers, (ii) the mapping of semantic features onto words: the child must work out how the semantic features apply to all of the words of the domain. "If the semantic features relevant to the lexical domain are not already known to the child, then remembering the specific non-linguistic context in which the word was first acquired provides a basis for future uses of the word." (Carey, 1978, p. 292).

In this and related views, the understanding of words relies on episodic memories for the perceptual contexts of word use - retrievable, explicit memories of instances of the word's prior use associated with the particular perceptual memories of the referents that were then present. These factors probably constitute the major advantage of L2 Naturalistic methods of language acquisition - it is easier to learn words in a world of things which provides ready perceptual memories than it is in a world of other words.

## Neuroanatomical Aspects of Explicit, Episodic, Cross-Modal Memories in Vocabulary Representation

Finally, what of the neuroanatomy of all of this? Is it possible to bridge linguistic, cognitive and neuropsychological evidence and theory at an anatomical level of analysis? The following tentative speculations are at least consistent with current knowledge of brain function.

Mishkin & Appenzeller (1990) and Squire (1992) review research on the role of circuits involving the limbic system and structures linked to it (the hippocampus, amygdala, diencephalon [thalamus and mamillary bodies], prefrontal cortex, and basal forebrain) in the formation of long term memories in monkeys. These are the same structures which are damaged in cases of human amnesia. Animals with normal occipital and infero-temporal lobes but damaged hippocampus and amygdala can perceive visual patterns normally, but had impaired long-term episodic memories for visual stimuli. Furthermore, damage to these structures resulted in a global anterograde amnesia - these animals were equally impaired on touch recognition. Damage to the limbic system leaves old memories largely intact but prevents the normal development of explicit memory for new information. Thus it appears that while the occipital and inferotemporal lobes might subserve perception and be the locus of pre-existing LT visual memories, the subcortical memory circuits must engage in a feedback whereby after a processed sensory stimulus activates the hippocampus (and possibly the amygdala), the memory circuits play back on the sensory area, strengthening and so perhaps storing the neural representation of the sensory event that has just taken place. The amygdala is a kind of crossroads in the brain with extensive connections with all the sensory systems in the cortex and also deeper into the brain to the hypothalamus which is thought to be the source of emotional responses. Monkeys with amygdala damage cannot form LT cross-modal memories - they cannot for example learn to relate the touch of an object with its sight. Because they are also slow in learning to relate an object to reward, Mishkin &

Appenzeller suggest that the amygdala allows association between stimuli and their emotional associations.

The conceptual meaning of a word is a conspiracy of perceptual memory traces. Think on the word 'Grandmother'. In so doing you awaken (whether consciously or not) your memories of all of the times you saw her, you heard her voice, you felt the touch of her hand and smelt her characteristic perfume. You remember the happy times and the sad. Conceptual meanings draw on rich cross-modal associations. Damage to the limbic structures which subserve these processes may thus deny the formation of these conceptual associations. This damage also prevents the formation of explicit LT memories between the perception of a word and the perceptual memories of its co-occurring referent.

## CONCLUSIONS

In summary, the evidence presented for unconscious processes in the acquisition of I/O (receptive/productive) aspects of vocabulary included:

- Child first language vocabulary development is essentially ubiquitous (following the frequency distribution of *implicit learning* systems).
- Child first language vocabulary development is relatively insensitive of 'academic intelligence' (it does *not* correlate with *explicit learning* abilities).
- Read vocabulary is so well preserved in dementia that it is taken as an index of premorbid IQ (in loss it thus behaves like *implicit memory* abilities).
- Repetition priming effects in lexical decision and word identification tasks demonstrate implicit vocabulary I/O activation (*implicit memory*).
- Priming studies show that I/O lexical modules for bilingual individuals are independent (like modularised *implicit memory* systems).
- Amnesics who are deficient at explicit memory abilities show normal lexical priming effects for both old and new lexical items (*implicit memory*).
- Effects of word regularity and the proportions of 'friends and enemies' demonstrate implicit acquisition of grapheme<->phoneme correspondences and spelling patterns for processing written vocabulary (*implicit learning*).

- Spoken word production is like other motor skills in that it is affected by frequency and statistical regularities in the subcomponent phonotactic sequences (*implicit learning*).
- Analyses of effects of exercise, practice, frequency of use, and life-span practice show that vocabulary acquisition, like implicitly-acquired skills, conforms to the power law of learning (*implicit learning and implicit memory*).
- Connectionist (Parallel Distributed Processing) modelling is a medium for investigating implicit learning in humans. Such models of conceptual, vocabulary, morphology, and reading and spelling acquisition can all reproduce to a remarkable degree the characteristics of people learning language - behaviours previously assumed to be characteristic of rule-governed systems even though the connectionist nets do not contain explicit rules (*implicit learning*).

The evidence presented for conscious processes in the acquisition of meaning aspects of vocabulary included:

- When people are assessed for their *understanding* of vocabulary there are high correlations between academic IQ and, respectively, reading ability and adult breadth of vocabulary (correlations with *explicit learning and memory* abilities)
- Free recall studies show that conceptual systems for vocabulary in bilinguals are interdependent (like *explicit memory* influenced by a wide range of conscious, cognitive factors).
- Human amnesics (who demonstrate normal implicit memory in the absence of explicit, episodic memories) have severe difficulty in acquiring vocabulary-concept pairings (*explicit memory*).
- Explicit memory for words is affected by Depth of Processing (*explicit memory*).

- Explicit, deep processing, mediational strategies (semantic and imagery elaboration) are highly effective in vocabulary learning (*explicit learning and memory*).
- Training in such metacognitive strategies as word-analysis or inferring meanings from contexts (*explicit learning and memory*) facilitates learners' vocabulary acquisition when reading.

The L2 or L1 language learner must acquire the I/O of new vocabulary: the pronunciation elements and their compounds in the tongue as well as the graphemes and their patterns of orthographic combination in the script. We have shown that there are specialised modules, the input and output lexicons, which acquire the word forms and regularities of the surface form of language by implicit learning processes. Like other sensory or motor skill systems, these modules do so automatically and they are tuned by practice - by frequency, recency, and regularity. To the extent that vocabulary acquisition is learning these surface forms of language then vocabulary acquisition is an implicitly acquired skill. Even amnesics who have very impaired explicit memory can acquire new vocabulary in this sense. In saying this I am not denying that the tunings of these systems cannot be guided by practice governed by explicit knowledge. In the same way that verbal declarative knowledge can coach the learner driver ('ease off the accelerator, down with the clutch, etc.'), so it can the learner speller ('*i* before *e* except after *c*...'). In the early stages of any skill we use conscious declarative knowledge on the way to automatization (Anderson, 1982, 1993). But essentially we learn to drive by driving itself, just as we learn to spell on the job of spelling or speak by speaking. In the main, these aspects of vocabulary acquisition reflect implicit learning.

However, the function of words is meaning and reference. And the mapping of I/O to semantic and conceptual representations is a cognitive mediation dependent upon explicit learning processes. It is heavily affected by depth of processing and elaborative integration with semantic and conceptual knowledge. Amnesic subjects have great difficulty in these aspects of vocabulary acquisition. But metacognitively sophisticated language learners excel because they have cognitive strategies for inferring the meanings of words, for enmeshing them in the meaning networks of other words and concepts and imagery representations, and mapping the surface forms to these rich meaning representations. To the extent that vocabulary acquisition is about meaning, it is an explicit learning process.

For any learning environment to be effective it must cater to all of these aspects. The I/O systems are tuned by practice, so the programme must encourage this as much as possible. There is little doubt that naturalistic settings provide maximum opportunity for exposure and motivation. Reading provides an ideal environment for the implicit acquisition of orthography, and also, in individuals tutored in metacognitive and cognitive skills for inferring meanings from contexts, explicit acquisition of meanings. But many are the times when we have discovered a word's meaning, either from text or from a dictionary, only for it to fade from our memory. Explicit, deep, elaborative processing concerning semantic and conceptual/imaginal representations prevents this. Learners can usefully be taught explicit skills in inferencing from context and in memorising the meanings of vocabulary.

## ACKNOWLEDGEMENTS

I thank the following people for useful comments on an earlier draft of this paper: Gordon Brown, Andrew Mayes, Barry McLaughlin, Paul Meara, Elisabeth Service, Mark Williams. A willingness to comment does not imply endorsement.

## REFERENCES

Anderson, J.P., & Jordan, A.M. (1928). Learning and retention of Latin words and phrases. *Journal of Educational Psychology*, **19**, 485-496.

Anderson, J.R. (1976). *Language, Memory, and Thought*. Hillsdale, N.J.: Erlbaum.

Anderson, J.R. (1982). Acquisition of a cognitive skill. *Psychological Review*, **89**, 369-406.

Anderson, J.R. (1983). *The Architecture of Cognition*. Cambridge Mass.: Harvard University Press.

Anderson, R.C., Wilson, P., & Fielding, L. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, **23**, 285-303.

Atkinson, R.C., & Raugh, M.R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, **104**, 126-133.

Baddeley, A.D. (1976). *The Psychology of Memory*. New York: Harper and Row.

Baddeley, A.D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, **27**, 586-595.

Balota, D.A., Ferraro, F.R., & Connor, L.T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. Schwanenflugel (Ed.) *The Psychology of Word Meanings* (pp. 187-222). Hillsdale, N.J.: Erlbaum.

Banta, F.G. (1981). Teaching German vocabulary: The use of English cognates and common loan words. *Modern Language Journal*, **65**, 129-136.

Baron, J., & Strawson, C. (1976). Use of orthographic and word specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 386-393.

Barron, R.W. (1980). Visual and phonological strategies in reading and spelling. In U. Frith (Ed.), *Cognitive Processes in Spelling* (pp. 195-214). London: Academic Press.

Barry, C., & Seymour, P.H.K. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *Quarterly Journal of Experimental Psychology*, **40(A)**, 5-40.

Barsalou, L.W. (1991). Structure and flexibility in concepts. Keynote address at the International Conference on Memory: Lancaster University, July 1991.

Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press.

Beck, I.L., & McKeown, M.G. (1983). Learning words well: A program to enhance vocabulary and comprehension. *The Reading Teacher*, **36**, 622-625.

Beck, I.L., McKeown, M.G., & Omanson, R.C. (1987). The effects and use of diverse vocabulary instruction techniques. In M.G. McKeown & M.E. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 147-163). Hillsdale, N.J.: Lawrence Erlbaum.

Berndt, R.S., Reggia, J.A., & Mitchum, C.C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behaviour Research Methods, Instruments & Computers*, **19**, 1-9.

Berry, D.C. (this volume). Implicit and explicit learning of complex tasks. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press.

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology,* **36A**, 209-231.

Berry, D. C., & Broadbent, D. E. (1988). Interactive Tasks and their Implicit-Explicit Distinction. *British Journal of Psychology,* **79***,* 251-272.

Besner, D., & Humphreys, G.W. (1991). *Basic Processes in Reading: Visual Word Recognition*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bower, G.H., & Winzenz, D. (1970). Comparison of associative learning strategies. *Psychonomic Science*, **20**, 119-120.

Bowerman, M. (1976) Semantic factors in the acquisition of rules for word use and sentence construction. In D.M. Morehead & A.E. Morehead (Eds.), *Normal and Deficient Child Language* (pp. 99-179). Baltimore: University Park Press.

Brown, D. (1980). Eight Cs and a G of Vocabulary Teaching. *Guidelines for Vocabulary Teaching*. *RELC Journal Supplement*, **3**. Singapore: RELC.

Brown, G.D.A. (1987). Resolving inconsistency: A computational model of word naming. *Journal of Memory and Language*, **23**, 1-23.

Brown, G.D.A., & Hulme, C. (1992). Cognitive psychology and second-language processing: the role of short-term memory. In R.J. Harris (Ed.), *Cognitive Processing in Bilinguals* (pp. 105-122). North Holland: Elsevier.

Brown, G.D.A., Watson, F.W., & Loosemore, R.P.W. (under submission). Normal and dyslexic spelling: A connectionist approach.

Brown, T.S., & Perry, F.L. Jr. (1991). A Comparison of three learning strategies for ESL vocabulary acquisition. *TESOL Quarterly*, **25**, 17-32.

Bryant, P.E., & Impey, L. (1986). The similarities between normal readers and developmental and acquires dyslexia. *Cognition*, **24**, 121-137.

Buikema, J.L., & Graves, M.F. (1993). Teaching students to use context cues to infer word meanings. *Journal of Reading*,

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan and G.A. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 264-293). Cambridge, Mass.: MIT Press.

Carpenter, P.A., & Just, M.A. (1989). The role of working memory in language comprehension. In I. D. Klahr & K. Kotovsky (Eds.), *Complex Information Processing: The Impact of Herbert A. Simon*. Hillsdale, N.J.: Erlbaum.

Carroll, J.B. (1963). Research on teaching foreign languages. In N.L. Gage (Ed.), *Handbook of Research on Teaching* (pp. 1060-1100). New York: Rand McNally.

Carroll, J.B. (1985), Second language abilities. In R.J. Sternberg (Ed.), *Human Abilities: An Information Processing Perspective* (pp. 83-101). New York: W.H. Freeman.

Carroll, J.B., & Sapon, S.M. (1955). *Modern Language Aptitude Test*. New York: The Psychological Corporation / Harcourt Brace Jovanovich Inc.

Carter, R. (1987). *Vocabulary: Applied Linguistic Perspectives*. London: Allen & Unwin.

Carter, R., & McCarthy, M. (1988). *Vocabulary and Language Teaching*. London: Longman.

Cermak, L.S., Bleich, R.P., & Blackford, S.P. (1988). Deficits in the implicit retention of new associations by alcoholic Korsakoff patients. *Brain and Cognition*, **7**, 312-323.

Cermak, L.S., Talbot, N., Chandler, K., & Wolbarst, L.R. (1985). The perceptual priming phenomenon in amnesia. *Neuropsychologia*, **23**, 615-622.

Chall, J.S. (1987). Two vocabularies for reading: recognition and meaning. In M.G. McKeown & M.E. Curtis (Eds.), *The Nature of Vocabulary Acquisition*. (pp. 1-18). Hillsdale, N.J.: Lawrence Erlbaum.

Chamot, A.U., & O'Malley, J.M. (this volume). Language learner and learning strategies. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press.

Chomsky, N. (1988). *Language and Problems of Knowledge*. The Managua Lectures. Cambridge, Mass.: MIT Press.

Christian, J., Bickley, W., Tarka, M., & Clayton, K. (1978). Measures of free recall of 900 English nouns: Correlations with imagery, concreteness, meaningfulness, and frequency. *Memory and Cognition*, **6** (4), 379-390.

Claparède, E. (1911). Recognition et moiité. *Archives de Psychologie Genève*, **11**, 79-90.

Clark, E.V. (1972). On the child's acquisition of antonyms in two semantic fields. *Journal of Verbal Learning and Verbal Behaviour*, **11**, 750-758.

Clark, H.H. (1973) Space, time, and semantics in the child. In T.E. Moore (Ed.), *Cognitive Development and the Acquisition of Language*. New York: Academic Press.

Clark, H.H., & Clark, E.V. (1977). *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt Brace Jovanovich.

Cohen, A.D. (1987). The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition*, **9**, 43-62.

Cohen, A.D., & Aphek, E. (1980). Retention of second language vocabulary over time: investigating the role of mnemonic association. *System*, **8**, 221-236.

Coltheart, M, Patterson, K., & Marshall, J.C. (Eds.) (1980). *Deep Dyslexia*. London: Routledge and Kegan Paul.

Conan Doyle, A. (1957). *The Red Headed League*. In *The Adventures of Sherlock Holmes*, London: John Murray. First published, 1892.

Cossu, G. & Marshall, J. (1990). Are cognitive skills a prerequisite for learning to read and write? *Cognitive Neuropsychology*, **7**, 21-40.

Craik, F. I. M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour*, **11**, 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General,* **104**, 268-294.

Cromer, R.F. (1991). *Language and Thought in Normal and Handicapped Children*. Oxford: Basil Blackwell.

*Crothers, E., & Suppes, P. (1967).* Experiments in Second-Language Learning. *New York: Academic Press.*

Crow, J.T., & Quigley, J.R. (1985). A semantic field approach to passive vocabulary acquisition for reading comprehension. *TESOL Quarterly*, **19**, 497-513.

Cunningham, A.E., & Stanovich, K.E. (1990). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge and spelling. *Journal of Educational Psychology*, **83**, 264-274.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Memory and Verbal Behaviour*, **19**, 450-466.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, **25**, 1-18.

Davelaar, E., & Besner, D. (1988). Word identification: imageability, semantics, and the content-functor distinction. *Quarterly Journal of Experimental Psychology*, **40A** (4), 789-799.

Desrochers, A., & Begg, I. (1987). A theoretical account of encoding and retrieval processes in the use of imagery-based mnemonic techniques: The special case of the keyword method. In M.A. McDaniel & M. Pressley (Eds.), *Imagery and Related Mnemonic Processes* (pp. 56-77). New York: Springer-Verlag.

Diamond, R., & Rozin, P. (1984). Activation of existing memories in the amnesic syndrome. *Journal of Abnormal Psychology*, **93**, 98-105.

Dopkins, S., Kovner, R., & Goldmeier, E. (1990). Acquisition of a new colour name in amnesics and normal controls. *Cortex*, **26**, 189-200.

Dudai, Y. (1989). *The Neurobiology of Memory*. New York: Oxford University Press.

Dukes, W.F., & Bastian, J. (1966). Recall of abstract and concrete words equated for meaningfulness. *Journal of Verbal Learning and Verbal Behaviour*, **5**, 455-458.

Durgunoglu, A.Y., & Roediger, H.L. (1987). Test differences in accessing bilingual memory. *Journal of Memory and Language*, **26**, 377-391.

Ehri, L.C. (1977). Do adjectives and functors interfere as much as nouns in naming pictures? *Child Development*, **48**, 697-701.

Ellis, A.W. (1984). *Reading, Writing, and Dyslexia: A Cognitive Analysis*. London: Lawrence Erlbaum Associates.

Ellis, A.W., & Young, A.W. (1988). *Human Cognitive Neuropsychology*. Hove: Lawrence Erlbaum.

Ellis, N.C. (1990). Reading, Phonological Processing and STM: Interactive Tributaries of Development. *Journal of Research in Reading*, **13** (2), 107-122.

Ellis, N.C. (1991). In verbal memory the eyes see vividly, but ears only faintly hear, fingers barely feel and the nose doesn’t know: Meaning and the links between the verbal system and modalities of perception and imagery. In R.H. Logie & M. Denis (Eds.), *Imagery and Cognition* (pp. 313-330) . Amsterdam: North Holland.

Ellis, N.C. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology,* **5**. 289-318.

Ellis, N.C., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology*. **46A**, 533-558.

Ellis, N.C., & Beaton, A. (in press). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*.

Ellis, N.C., & Cataldo, S. (1990). The Role of Spelling in Learning to Read *Language and Education*, **4**, 1-28.

Estes, W.K. (1972). An associative basis for coding and organisation in memory. In A.W. Melton & E. Martin (Eds.), *Coding Processes in Human Memory* (pp. 161-190). Washington, D.C.: Winston.

Faust, G.W., & Anderson, R.C. (1967). Effects of incidental material in a programmed Russian vocabulary lesson. *Journal of Educational Psychology*, **58**, 3-10.

Fawcett, A.J., & Nicolson, R.I. (1991). Vocabulary training for children with dyslexia. *Journal of Learning Disabilities*, **24**, 379-383.

Fodor, J.A. (1983). *The Modularity of Mind*. Cambridge, Mass.: MIT Press.

Frinke, G. (1968). Word characteristics, associative-relatedness and the free-recall of nouns. *Journal of Verbal Learning and Verbal Behaviour*, **7**, 366-372.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, M. Coltheart and J. Marshall (Eds.), *Surface Dyslexia* (pp. 301-330). London: Erlbaum.

Gabrieli, J.D.E., Cohen, N.J., & Corkin, S. (1983). Acquisition of semantic and lexical knowledge in amnesia. *Society for Neuroscience Abstracts*, **9**, 28.

Gabrieli, J.D.E., Cohen, N.J., & Corkin, S. (1988). The impaired learning of semantic knowledge following bilateral medial temporal lobe resection. *Brain and Cognition*, **7**, 157-177.

Gabrieli, J.D.E., Milberg, W., Keane, M.M., & Corkin, S. (1990). Intact priming of patterns despite impaired memory. *Neuropsychologia*, **28**, 417-427.

Gardner, H., Boller, F., Moreines, J., & Butters, N. (1973). Retrieving information from Korsakoff's patients: Effects of categorical cues and reference to the task. *Cortex*, **9**, 165-175.

Gardner, R.C., & Lambert, W.E. (1972). *Attitudes and Motivation in Second Language Learning*. Rowley, Mass.: Newbury House.

Gasser, M. (1990). Connectionism and universal of second language acquisition. *Studies in Second Language Acquisition*, **12**, 179-199.

Gathercole, S.E., & Baddeley, A.D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, **28**, 200-213.

Gathercole, S.E., & Baddeley, A.D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, **81**, 439-454.

Gathercole, S.E., Willis, C., Emslie, H., & Baddeley, A.D. (1991). The influence of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, **12**, 349-367.

Genesse, F. (1976). The role of intelligence in second language learning. *Language Learning*, **26**, 267-280.

Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S.A. Kuczaj II (Ed.), *Language Development, Volume 2: Language, Thought, and Culture* (pp. 301-334). Hillsdale, N.J.: Erlbaum.

Gilbert, L. (1935). Study of the effect of reading on spelling. *Journal of Educational Research*, **28**, 570-576.

Gilhooly, K.J., & Gilhooly, M.L. (1979). Age-of-acquisition effects in lexical and episodic memory tasks. *Memory and Cognition*, **7** (3), 214-223.

Glisky, E.L., & Schacter, D.L. (1987). Acquisition of domain specific knowledge in organic amnesia: Training for computer-related work. *Neuropsychologia*, **25**, 893-906.

Glisky, E.L., Schacter, D.L., & Tulving, E. (1986). Computer learning by memory-impaired patients: Acquisition and retention of complex knowledge. *Neuropsychologia*, **24**, 313-328.

Goodman, K., & Goodman, Y. (1982). Spelling ability of a self-taught reader. In F. Gollasch (Ed.), *Language and Literacy: The Selected Writings of Kenneth S. Goodman* (pp. 221-226). London: Routledge.

Haist, F., Musen, G., & Squire, L. (1991). Intact priming of words and nonwords in amnesia. *Psychobiology*, **19**, 275-285.

Graf, P., Mandler, G., & Haden, P. (1982). Simulating amnesic symptoms in normal subjects. *Science*, **218**, 1243-1244.

Graves, M.F. (1987). The roles of instruction in fostering vocabulary development. In M.G. McKeown & M.E. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 165-184). Hillsdale, N.J.: Lawrence Erlbaum.

Hadenius, A.M., Hagberg, B., Hyttnäs-Bensch, K., & Sjögren, I. (1962). The natural prognosis of infantile hydrocephalus. *Acta Paediatrica*, **51**, 117-118.

Hall, J.W. (1988). On the utility of the keyword mnemonic for vocabulary learning. *Journal of Educational Psychology*, **80**, 554-562.

Hall, J.W., Wilson, K.P., & Patterson, R.J. (1981). Mnemotechnics: some limitations of the mnemonic keywords method for the study of foreign language vocabulary. *Journal of Educational Psychology*, **73**, 345-357.

Harrington, M. (1992). Working memory capacity as a constraint on L2 development. In R. Harris (Ed.), *Cognitive Processes in Bilinguals* (pp. 123-135). North Holland: Elsevier.

Henning, G.H. (1974). Remembering foreign language vocabulary: acoustic and semantic parameters. *Language Learning*, **23**, 185-196.

Heredia, R., & McLaughlin, B. (1992). Bilingual memory revisited. In R. Harris (Ed.), *Cognitive Processing in Bilinguals* (pp. 91-104). North Holland: Elsevier.

Hirst, W., Phelps, E.A., Johnson, M.K., & Volpe, B.T. (1988). Amnesia and second language learning. *Brain and Cognition*, **8**, 105-116.

Horowitz, I.M., Chilian, P.C., & Dunnigan, K.P. (1969). Word fragments and their redintegrative powers. *Journal of Experimental Psychology*, **80**, 392-394.

Hulme, C., Maughan, S., & Brown, G.D.A. (1991). Memory for words and non-words: Evidence for a long-term memory contribution to short-term memory tasks. *Journal of Memory and Language*, **30**, 685-701.

Hulstijn, J. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and Applied Linguistics*. (pp. 113-125). London: Macmillan.

Ijaz, I.H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning*, **36**, 401-451.

Ingram, T.T.S., & Naughton, J.A. (1962). Paediatric and psychological aspects of cerebral palsy associated with hydrocephalus. *Developmental Medicine and Child Neurology*, **4**, 287-292.

Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.

James, W. (1890). *The Principles of Psychology*. Volumes 1 and 2. New York: Dover Publications.

Jenkins, J.J., Mink, W.D., & Russell, W.A. (1958). Associative clustering as a function of verbal association strength. *Psychological Reports*, **4**, 127-136.

Jensen, A. (1980). *Bias in Mental Testing*. New York: Free Press.

Jones, G.V. (1985). Deep dyslexia, imageability, and ease of predication. *Brain and Language*, **24**, 1-19.

Jones, G.V. (1988). Images, predicates, and retrieval cues. In M. Denis, J. Engelkamp, & J.T.E. Richardson (Eds.), *Cognitive and Psychological Approaches to Mental Imagery* (pp. 89-98). Dordrecht: Martinus Nijhoff.

Kasper, L. F. (1983). The effect of linking sequence and interactive picture mnemonics on the acquisition of Spanish nouns by middle school children. *Human Learning and Journal of Practical Research and Applications*, **2,** 141 - 156.

Kaufman, A.S. (1979). *Intelligence Testing with the WISC-R*. New York: John Wiley.

Kirsner, K. (this volume). Implicit processes in second language learning. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press.

Kirsner, K., & Smith, M.C. (1974). Modality effects in word identification. *Memory & Cognition*, **2**, 637-640.

Kirsner, K., Smith, M.C., Lockhart, R.S., King, M.L., & Jain, M. (1984). The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behaviour,* **23**, 519-539.

Kirsner, K., & Speelman, C.P. (in press). Is lexical processing just another ACT*? In A. Collins, M. Conway, S. Gathercole, & P. Morris (Eds.), *Theories of Memory*. London: Erlbaum.

Kolers, P.A. (1966). Interlingual facilitation of short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, **5**, 314-319.

Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.

Krashen, S.D. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.

Krashen, S.D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the Input Hypothesis. *The Modern Language Journal*, **73**, 440-464.

Krashen, S.D., & Terrell, T. (1983). *The Natural Approach: Language Acquisition in the Classroom*. Oxford: Pergamon.

Laberge, D., & Samuels, S.J. (1974). Towards a theory of automatic information processing in reading. *Cognitive Psychology*, **6**, 293-323.

Lakoff, G. (1987). *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By.* Chicago: University of Chicago Press.

Lenneberg, E.H. (1967). *Biological Foundations of Language*. New York: John Wiley & Sons.

Lenneberg, E.H., Nichols, I.A., & Rosenberger, E.F. (1964). Primitive stages of language development in mongolism. In D. McK. Rioch & E.A. Weinstein (Eds.), *Disorders of Communication 12: Research Publications*. Baltimore: Wiliams and Wilkins.

Levin, J.R., & Pressley, M. (1985). Mnemonic vocabulary acquisition: What's fact, what's fiction? In R.F. Dillon (Ed.), *Individual Differences in Cognition* (Vol. 2, pp. 145-172). Orlando, FL: Academic Press.

Loess, H., & Brown, A. (1969). Word fragments as aids to recall a whole word. *Journal of Experimental Psychology*, **80**, 384-386.

Long, M.H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, **12**, 251-285.

Marslen-Wilson, W.D. (1987). Functional parallelism in spoken word-recognition. In U.H. Frauenfelder & L.K. Tyler (Eds.), *Spoken Word Recognition* (pp. 71-102). Cambridge, Mass.: MIT Press.

Mayes, A.R., & Gooding, P. (1989). Enhancement of word completion priming in amnesics by cueing with previously novel associations. *Neuropsychologia*, **27**, 1057-1072.

McCarthy, J.J., & Kirk, S.A. (1961). *The Illinois Test of Psycholinguistic Abilities*. Urbana, Ill.: University of Illinois Press.

McCarthy, R.A., & Warrington, E.K. (1990). *Cognitive Neuropsychology: A Clinical Introduction*. San Diego: Academic Press.

McClelland, J.L., & Rumelhart, D.E., and the PDP Research Group. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*. Cambridge, Mass.: MIT Press.

McKeown, M.G., & Curtis, M.E. (Eds.) (1987). *The Nature of Vocabulary Acquisition*. Hillsdale, N.J.: Lawrence Erlbaum.

McLaughlin, B. (1990). "Conscious" versus "unconscious" learning. *TESOL Quarterly,* **24**, 617-634.

McLeod, B., & McLaughlin, B. (1986). Restructuring or automaticity? Reading in a Second Language. *Language Learning*, **36**, 109-124.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper and A. Howatt (Eds.), *Interlanguage*. (pp. 225-235) Edinburgh: University of Edinburgh Press.

Merry, R. (1980). The keyword method and children's vocabulary learning in the classroom. *British Journal of Educational Psychology*, **50**, 123-136.

Meudell, P.R., Mayes, A.R., & Neary, D. (1981). Amnesia is not caused by cognitive slowness. *Cortex*, **16**, 413-419.

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

Miller, G.A. (1977). *Spontaneous Apprentices: Children and Language*. New York: Seabury.

Miller, G. A., & Gildea, P.M. (1991). How children learn words. In. W. S.-Y. Wang (Ed.), *The Emergence of Language: Development and Evolution (Readings from Scientific American)* (pp. 150-158). New York: W.H. Freeman.

Milner, B. (1966). Amnesia following operations on the temporal lobes. In C.W.M. Whitty & O.L. Zangwill (Eds.), *Amnesia*. (pp. 109-133). London: Butterworth.

Milner, B., Corkin, S., & Teuber, H.L. (1968) Further analysis of the hippocampal amnesic syndrome: A 14-year follow up study of H.M. *Neuropsychologia*, **6**, 215-234.

Mishkin, M., & Appenzeller, T. (1990). The anatomy of memory. In R.R. Llinás (Ed.), The *Workings of the Brain: Development, Memory, and Perception. Readings from Scientific American*. New York: W.H. Freeman and Company.

Monsell, S. (1985). Repetition and the lexicon. In A. Ellis, (Ed.), *Progress in the Psychology of Language* (Volume 2 pp. 147-195). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, **76**, 165-178.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen models. In P.A. Kolers, M.E. Wrolstad, & H. Bouma (Eds.), *Processing of Visible Language* (Vol. 1, pp. 259-268) New York: Plenum.

Morton, J., & Patterson, K. (1980). 'Little words - No!'. In M. Coltheart, K. Patterson & J.C. Marshall (Eds.), *Deep Dyslexia* (pp. 270-285). London: Routledge & Kegan Paul.

Moscovitch, M., Winocur, G., & McLachlan, D. (1986). Memory as assessed by recognition and reading time in Alzheimer's disease and other neurological disorders. *Journal of Experimental Psychology: General*, **115**, 331-347.

Murrell, G.A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology*, **102**, 963-968.

Musen, G., & Squire, L.R. (1990). Implicit memory: No evidence for rapid acquisition of new associations in amnesic patients or normal subjects. *Society for Neuroscience Abstracts,* **16**, 287.

Musen, G., & Squire, L.R. (1991). Normal acquisition of novel verbal information in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **17**, 1095-1104.

Musen, G., & Squire, L.R. (in press). Intact priming of non-verbal material in amnesic patients. *Memory & Cognition*.

Musen, G., & Treisman, A. (1990). Implicit and explicit memory for visual patterns. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 127-137.

Nagy, W.E., & Herman, P.A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M.G. McKeown & M.E. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 19-35). Hillsdale, N.J.: Lawrence Erlbaum.

Nagy, W.E., Anderson, R.C., & Herman, P.A. (1987). Learning new words from context during normal reading. *American Education Research Journal*, **24**, 237-270.

Naiman, N., Fröhlich, M., Stern, H.H., & Todesco, A. (1978). *The Good Language Learner: A Report*. Toronto : Ontario Institute for Studies in Education.

Nation, I.S.P. (1982). Beginning to learn a foreign language vocabulary: A review of the research. *RELC Journal*, **13**, 15-36.

Nation, I.S.P. (1987). *Teaching and Learning Vocabulary*. Wellington: Victoria University of Wellington English Language Institute, Occasional Publication No. 7.

Nattinger, J. (1988). Some current trends in vocabulary teaching, In R. Carter & M. McCarthy, (Eds.), *Vocabulary and Language Teaching* (pp. 60-82). London: Longman.

Neisser, U. (1954). An experimental distinction between perceptual processes and verbal response. *Journal of Experimental Psychology*, **47**, 399-402.

Nelson, H.E. (1982). *National Adult Reading Test (NART) Test Manual*. Windsor: NFER Nelson.

Nelson, H.E., & McKenna, P. (1975). The use of current reading ability in the assessment of dementia. *British Journal of Social and Clinical Psychology*, **14**, 259-267.

O'Malley, J.M., Chamot, A.U., Stewner-Manzanares, G., Kupper, L., & Russo, R.P. (1985a). Learning strategies used by beginning and intermediate ESL students. *Language Learning*, **35**, 21-46.

O'Malley, J.M., Chamot, A.U., Stewner-Manzanares, G., Russo, R.P., & Kupper, L. (1985b). Learning strategy applications with students of English as a second language. *TESOL Quarterly*, **19**, 557-584.

Ormrod, J. (1986). Learning to spell while reading: A follow-up study. *Perceptual and Motor Skills,* **63**, 652-654.

Ortony, A. (Ed.) (1979). *Metaphor and Thought*. Cambridge: Cambridge University Press.

Paap, K.R., McDonald, J.E., Schvaneveldt, R.W., & Noel, R.W. (1987). Frequency and pronunciability in visually presented naming and lexical decision tasks. In M. Coltheart (Ed.), *Attention and Performance,* **XII** (pp. 221-244), London: Lawrence Erlbaum Associates.

Paivio, A. (1968). A factor-analytic study of word attributes and verbal learning. *Journal of Verbal Learning and Verbal Behaviour,* **7**, 41-49.

Paivio, A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart and Winston.

Paivio, A., Yuille, J.C., & Madigan, S.A. (1968). Concreteness, Imagery and Meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement,* **76** (1,2), 1-25.

Paivio, A., Yuille, J.C., & Smythe, P.C. (1966). Stimulus and response abstractness, imagery and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology*, **20**, 362-377.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological Short-Term Memory and Foreign-Language Vocabulary Learning. *Journal of Memory and Language*, **30**, 331-347.

Paradis, M. (this volume). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press.

Paraskevopoulos, J.N., & Kirk, S.A. (1969). *The Development of the Revised Illinois Test of Psycholinguistic Abilities*. Urbana: University of Illinois Press.

Patterson, K.E. (1981). Neuropsychological approaches to the study of reading. *British Journal of Psychology*, **72**, 151-174.

Perfetti, C. (1985). *Reading Ability*. Oxford: Oxford University Press.

Plunket, K. (this volume). Connectionist approaches to Second Language Learning. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press.

Pressley, M., & Levin, J.R. (1981). The keyword method and recall of vocabulary words from definitions. *Journal of Experimental Psychology: Human Learning and Memory*, **7**, 72-76.

Pressley, M., Levin, J.R., & Delaney, H.D. (1982). The mnemonic keyword method. *Review of Educational Research*, **52**, 61-91.

Pressley, M., Levin, J.R., Hall, J.W., Miller, G.E., & Berry, J.K. (1980). The keyword method and foreign word acquisition. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 163-173.

Pressley, M., Levin, J.R., & Miller, G.E. (1981). The keyword method and children's learning of foreign vocabulary with abstract meanings, *Canadian Journal of Psychology*, **35**, 283-287.

Prytulak, L.S. (1971). Natural language mediation. *Cognitive Psychology,* **2**, 1-56.

Raugh, M.R., & Atkinson, R.C. (1975). A mnemonic method for learning a second-language vocabulary. *Journal of Educational Psychology*, **67**, 1-16.

Raugh, M.R., Schupbach, R.D., & Atkinson, R.C. (1977). Teaching a large Russian vocabulary by the mnemonic keyword method. *Instructional Science*, **6**, 199-211.

Raven, J,C., Court, J.H., & Raven, J. (1983). *Manual for the Raven's Progressive Matrices and Vocabulary Scales. SPM I.* London: H.K. Lewis & Co. Ltd.

Reber, A. S. (1967). Implicit Learning of Artificial Grammars. *Journal of Verbal Learning and Verbal Behaviour,* **6***,* 855-863.

Reber, A. S. (1969). Transfer of Syntactic Structure in Synthetic Languages. *Journal of Experimental Psychology,* **81**, (1), 115-119.

Reber, A. S. (1976). Implicit Learning of Synthetic Languages: The Role of Instructional Set. *Journal of Experimental Psychology, Human Learning and Memory,* **2***, (1)*, 88-94.

Reber, A.S., & Allen, R. (1978). Analogy and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition,* **6***,* 189-221.

Reber, A.S., Wakenfield, F.F. & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **17**, 888-896.

Rice, M.L. (1990). Pre-schoolers' QIL: Quick incidental learning of words. In S. Conti-Ramsden & C.E. Snow (Eds.), *Children's Language, Vol. 7* (pp. 171-195). Hillsdale, N.J.: Lawrence Erlbaum.

Rice, M.L., Huston, A.C., Truglio, R., & Wright, J. (1990). Words from "Sesame Street": Learning vocabulary while viewing. *Developmental Psychology*, **26**, 421-428.

Rice, M.L., & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development*, **59**, 420-429.

Richards, J. (1976). The role of vocabulary teaching. *TESOL Quarterly*, **10**, 77-89.

Rodgers, T.S. (1969). On measuring vocabulary difficulty: An analysis of item variables in learning Russian-English vocabulary pairs. *International Review of Applied Linguistics*, **7**, 327-343.

Rosch, E., & Mervis, C.B. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, **104**, 192-233.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382-439.

Rubin, D.C. (1980). 51 properties of 125 words: a unit analysis of verbal behaviour. *Journal of Verbal Learning and Verbal Behaviour,* **19**, 736-755.

Rubin, D.C. (1983). Associative asymmetry, availability and retrieval. *Memory and Cognition*, **11** (1), 83-92.

Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics*, **11**, 117-131.

Rumelhart, D.E., & Norman, D.A. (1985). Representation of knowledge. In A.M. Aitkenhead & J.M. Slack (Eds.) *Issues in Cognitive Modelling* (pp. 15-62). Hillsdale, N.J.: Lawrence Erlbaum.

Saragi, T., Nation, P., & Meister, G. (1978). Vocabulary learning and reading. *System*, **6**, 70-78.

Schacter, D.L. (1985). Priming of old and new knowledge in amnesic patients and normal subjects. *Annals of the New York Academy of Sciences*, **444**, 41-53.

Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 501-518.

Schank, R.C., & Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding*. Hillsdale, N.J.: Lawrence Erlbaum.

Schacter, D.L., & Graf, P. (1986). Preserved learning in amnesic patients: Perspectives from research on direct priming. *Journal of Clinical and Experimental Neuropsychology*, **8**, 727-743.

Schacter, D.L., Cooper, L.A., & Delaney, S.M. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General*, **119**, 5-24.

Schank, R.C., & Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding*. Hillsdale, N.J.: Lawrence Erlbaum.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, **11** (2), 129-158.

Schmidt. R. (this volume). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press.

Schwanenflugel, P (1991). Why are abstract concepts hard to understand? In P. Schwanenflugel (Ed.) *The Psychology of Word Meanings*. Hillsdale, N.J.: Erlbaum.

Schwanenflugel, P., Harnishfeger, K., & Stowe, R. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, **27**, 499-520.

Scoville, W.B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, **20**, 11-21.

Seashore, R., & Eckerson, L. (1940). The measurement of individual differences in general English vocabularies. *Journal of Educational Psychology*, **31**, 14-31.

Seibert, L.C. (1927). An experiment in learning French vocabulary. *Journal of Educational Psychology*, **18**, 294-309.

Seidenberg, M.S., Waters, G.S., Barnes, M.A., & Tannenhaus, M.K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behaviour*, **23**, 383-404.

Service, E. (1992). Phonology, Working Memory, and Foreign-language Learning. *Quarterly Journal of Experimental Psychology*, **45A**, 21-50.

Shimamura, A.P., & Squire, L.R. (1984). Paired-associate learning and priming effects in amnesia: A neuropsychological study. *Journal of Experimental Psychology: General*, **113**, 556-570.

Shimamura, A.P., & Squire, L.R. (1989). Impaired priming of new associations in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 721-728.

Skinner, B. (1957). *Verbal Behaviour*. New York: Appleton-Century-Crofts.

Smith, F. (1982). *Understanding Reading*. Hillsdale, N.J.: Erlbaum.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, & Baddeley. *Applied Psycholinguistics*, **12**, 369-373.

Squire, L.R. (1987). *Memory and Brain*. New York: Oxford University Press.

Squire, L.R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, **99**, 195-231.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, **XXI**, 360-407.

Stanovich, K.E. (1992). The psychology of reading: evolutionary and revolutionary developments. *Annual Review of Applied Linguistics*, **12**, 3-30.

Stanovich, K.E., & Cunningham, A.E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory and Cognition,* **20**, 51-68.

Sternberg, R.J. (1985). *Beyond IQ: A Triarchic Theory of Human Intelligence*. Cambridge: Cambridge University Press.

Sternberg, R.J. (1987). Most vocabulary is learned from context. In M.G. McKeown & M.E. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 89-105). Hillsdale, N.J.: Lawrence Erlbaum.

Strauss, M.E., Weingartner, H., & Thompson, K. (1985). Remembering words and how often they occurred in memory impaired patients. *Memory and Cognition*, **13**, 1507-1510.

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

Taft, M. (1991). *Reading and the Mental Lexicon*. Hove and London: Lawrence Erlbaum Associates.

Taylor, E.M. (1959). *Psychological Appraisal of Children with Cerebral Defects*. Cambridge, Mass.: Harvard University Press.

Thorndike, E,L. (1908). Memory for paired-associates. *Psychological Review*, **15**, 122-138.

Treiman, R., & Baron, J. (1983). Individual differences in spelling: The Phoenician - Chinese distinction. *Topics in Learning and Learning Disabilities*, **3**, 33-40.

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford: Clarendon Press.

Turner, M.L., & Engle, R.W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, **28**, 127-145.

Tversky, A. (1977). Features of similarity. *Psychological Review*, **84**, 327-352.

Twadell, F. (1973). Vocabulary expansion in the TESOL classroom. *TESOL Quarterly*, **7**, 61-78.

Van der Linden, M., Meulemans, T., & Lorrain, D. (1992). Acquisition of new concepts by amnesic patients. Technical Report. Unité de Neuropsychologie Cognitive Université de Louvain, Belgium.

Vanniarajan, S. (1992). *Working Memory and Native-Non-native Differences in Inferencing the Meanings of Unknown Words in Context: An Experimental Study in Interactive Model Perspective*. Unpublished Applied Linguistics Ph.D. thesis, U.C.L.A..

Venezky, R.L. (1970). *The Structure of English Orthography*. The Hague: Mouton & Co.

Wagner, R.K., & Sternberg, R.J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology,* **49**, 436-458.

Wagner, R.K., & Sternberg, R.J. (1986). Tacit knowledge and intelligence in the everyday world. In R.J. Sternberg & R.K. Wagner (Eds.), *Practical Intelligence: Nature and origins of competence in the everyday world* (pp. 51-83). Cambridge: Cambridge University Press.

Wang, W. S.-Y. (1971). The basis of speech. In C.E. Reed (Ed.), *The Learning of Language*. NY: Appleton-Century-Crofts.

Warnock, M. (1987). *Memory*. London: Faber and Faber.

Warrington, E.K., & Weiskrantz, L. (1982). Amnesia: A disconnection syndrome. *Neuropsychologia*, **20**, 233-249.

Webb, W.B. (1962). The effects of prolonged learning on learning. *Journal of Verbal Learning and Verbal Behaviour*, **1**, 173-182.

Wesche, M.B., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence, *Applied Psycholinguistics*, **3**, 127-140.

Wechsler, D. (1976). *WISC-R UK Manual. Wechsler Intelligence Scale for Children - Revised*. Sidcup, Kent: The Psychological Corporation Ltd.

Wechsler, D. (1981). *WAIS-R Manual. Wechsler Adult Intelligence Scale - Revised*. Cleveland, Ohio: The Psychological Corporation Ltd.

West, M. (1953). *A General Service List of English Words*. London: Longman.

Whorf, B.J. (1956). *Language, Thought and Reality. Selected Writings of Benjamin Lee Whorf.* New York: John Wiley.

Winocur, G., & Weiskrantz, L. (1976). An investigation of paired-associate learning in amnesic patients. *Neuropsychologia*, **12**, 437-445.

Winter, W., & Reber, A.S. (this volume). Implicit learning and natural languages. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press.

Wood, F.B., Brown, I.S., & Felton, R.H. (1989). Long term follow up of a childhood amnesic syndrome. *Brain & Cognition*, **10**, 76-86.

Yamada, J.E. (1990). *Laura: A Case for the Modularity of Language*. Cambridge, Mass.: MIT Press.

Yuill, N., & Oakhill, J. (1988). Effects of inference training on poor reading comprehension. *Applied Cognitive Psychology,* **2,** 33-45.