# 19 Attention and Memory during SLA

PETER ROBINSON

## 1 Introduction and Overview

Attention to and subsequent memory for attended language input are both essential for SLA, and are intricately related. Attention is the process that encodes language input, keeps it active in working and short-term memory, and retrieves it from long-term memory. Attention and memory structures can be viewed hierarchically. The focus of attention is a subset of short-term memory, and short-term memory is that part of long-term memory in a currently heightened state of activation. Long-term memory is where instances of encoded input are stored and assume (or confirm, in some innatist theories of SLA) the representational shape that recognition processes match to new instances of input in working memory during parsing and comprehension. These representations also form the basis of speech production "plans," which guide retrieval processes during grammatical and phonological encoding, and articulation of a message. Attention, then, can be viewed as a process for which memory provides structure and constraint.

Research into attention and memory during SLA has begun to accumulate in the last decade or so, addressing such issues as the following: what levels of attention and awareness are necessary for encoding L2 input in short-term working memory? What is the nature of the encoding, rehearsal, and retrieval processes that operate on attended input? How do L2 task demands affect the allocation of memory and attention? And is memory simply functionally differentiated, or also neurophysiologically differentiated, reflecting the operation of distinct learning and memory systems? Many of these issues also dominate recent debate in cognitive psychology concerned with distinctions between implicit, incidental, explicit, and intentional learning – issues that are discussed elsewhere in this volume, but not in any detail here (see Ellis, DeKeyser, and Hulstijn, this volume). Research into attention and memory during SLA is relevant to a transition theory (see Gregg, 2001, this volume) of the cognitive mechanisms that move L2 knowledge from point A to point B, and so has largely been concerned with specifying the universal cognitive architecture of attention and memory during learning. A transition theory predicates a property theory (how knowledge at points A and B is represented) and specifies mechanisms which can be activated by attentional processes and memory structures (spreading activation, parameter resetting, cue strengthening, etc.) that give knowledge at point B representational shape (see Ellis, O'Grady, and White, this volume, for substantive discussion of options in SLA property theories).

Arising, in part, out of interest in the architecture of attention and memory during SLA is resurgent interest in the implications of individual differences in attentional and memory resources. Issues this research addresses include the following: do individual differences in L2 working memory capacity affect skill development? Can the influence of age differences on SLA be explained by developmentally regulated changes in attentional and memory resources? Do individual differences in resource availability affect explicit but not implicit learning? And how are differences in attentional and memory resources related to language learning aptitude?

Attention and memory can be studied and measured at various levels, including ecological/adaptive (Reed, 1996), cognitive/information-processing (Sanders, 1998), and neurophysiological/biochemical (Carter, 1998; Posner and Petersen, 1990). This chapter presents a cognitive-level characterization of attention and memory that describes the information-processing operations and stages mediating stimulus input and response selection. This raises issues of both learning – the attentional and memory processes responsible for the acquisition of new and the restructuring of existing representations – and performance – the skilled deployment of existing knowledge to achieve task goals. Comprehensive accounts of human cognition view a theory of learning as embedded within, and commensurate with, a theory of action (Allport, 1987; Clark, 1997; Hazelhurst and Hutchins, 1998; Korteling, 1994; Shallice, 1978; Thelen and Smith, 1994), which describes how attentional and memory resources are drawn on in task and context analysis, and during adaptive responses to both. Consequently, I will describe the role of attention and memory in selection and maintenance of new information in memory (see also Schmidt, 1995, 2001; Tomlin and Villa, 1994, for reviews of this area), as well as in control of action, and sustained attention to the goals of action – areas where there has been less theoretical discussion of the role of cognitive factors in SLA research.

In what follows I focus on the interrelated areas of attention and memory separately, describing current theoretical issues and models of each, then summarizing research that has examined the influence of these cognitive factors on SLA, using a variety of methodologies.

# 2 Attention

## *2.1 Overview*

Three general stages of information processing at which attention operates are captured in figure 19.1. The stages correspond broadly to three current themes in attentional research and theory (Sanders, 1998; Sanders and Neumann, 1996; Sergeant, 1996); (i) auditory and visual information intake and processing; (ii) central control and decision-making functions, such as allocation of attention to competing task demands, and automatization; and (iii) response execution and monitoring via sustained attention. These three themes and stages also correspond to three uses of the concept of attention; to describe selection of information (we pay attention to things as a way of selecting them for further processing); to describe the capacity of attentional resources (sometimes we are able to pay a lot of attention to a task, while at other times we are not); and to describe the effort involved in sustaining attention to task goals (we can maintain the level of attention we pay to a task, or attention and performance can decline over time). These are distinct but related uses of the concept of attention; each one related to separate functions, which, however, often operate in conjunction with each other.

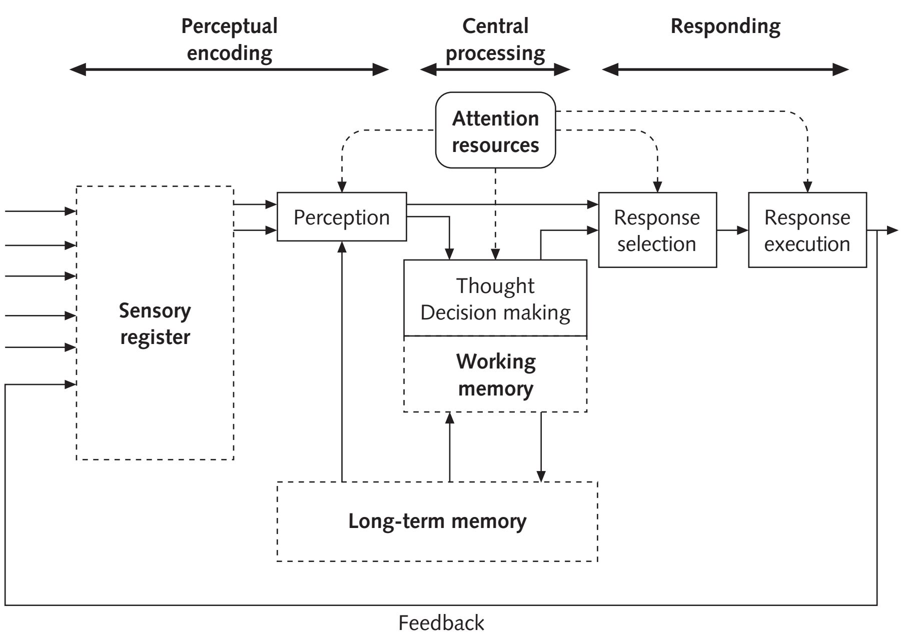


**Figure 19.1** A generic model of human information processing with three memory systems

*Source*: Wickens, Gordon, and Liu (1997, p. 147)

## 2.2 Attention as selection

Learning and performance both involve selection and subsequent encoding of information available in the environment. A traditional distinction in SLA theory is between linguistic "input" to the learner and "intake" or mental registration of the input (Corder, 1967). Recent SLA research and theory have examined the role of attention in mediating this process by studying, for example, the level of attention needed for selecting input for processing (S. Carroll, 1999; Chaudron, 1985; Gass, 1988, 1997; Leow, 1993; Tomlin and Villa, 1994); whether pedagogic intervention can facilitate switches of attention from meaning to aspects of the (syntactic, morphological phonological, semantic, and pragmatic) form of input which otherwise may lack saliency for learners and so remain unattended to during communication (Doughty, 2001; Doughty and Williams, 1998; Long, 1996); and what, if any, level of awareness must accompany or follow the selection process if intake is to be permanently registered in memory (Philp, 1998; Schmidt, 1990, 1995, 2001; Sharwood-Smith, 1981, 1991). Three important theoretical issues are: (i) when and how does selection of information happen; (ii) why is information selectively attended for further processing; and (iii) what mechanisms guide the selection process?

### 2.2.1 *When and how does selection happen?*

As figure 19.1 shows, during the first stage of information processing preattentively processed sensory information is detected and held temporarily in the sensory register, where it is selected for perceptual encoding by attentional mechanisms. Auditory and visual processing dominated early research into the role of selective attention in perception, and two issues largely divided early theories; whether the attention allocated to information selection from the sensory register is limited or unlimited in capacity, and whether information is selected early or late during processing. These issues concern the "why" and the "when" of selection. Broadbent (1958, 1971) assumed that attentional capacity is limited and that therefore auditory and visual information must be channeled and specific stimuli sequentially selected early, via a filtering operation, for further processing. These assumptions appeared necessary to explain findings such as the following: answering two different questions that overlap temporally interferes with performance, but prior knowledge that one question will be irrelevant enables it to be screened out, or inhibited, thereby facilitating performance on the relevant question, which receives subsequent full semantic analysis. Selection, that is, was viewed as a functional consequence of limited attentional capacity (Neumann, 1996, p. 395) and was thought to be made early on the basis of a partial analysis of specific features of the input. Once widely accepted, these assumptions were challenged by evidence from both letter discrimination tasks (Sperling, 1960) and dichotic listening tasks, in which different messages are presented simultaneously in each ear (Treisman, 1971), which showed that multiple sources of information can be processed in parallel (messages presented in either ear, all letters presented briefly in a visual display), and that selection of any one may be late, and based on full semantic, not partial feature analysis. Late selection theories, consequently, argue selection takes place in working memory after stimuli have been fully analyzed (Allport, 1987).

### *2.2.2 Why does selection happen?*

If many inputs can be processed in parallel, and the attention available to the sensory register and the central processor is potentially unlimited, then Broadbent's main reason for proposing the "when" (early) and "how" (a filter) of selection disappears. Why then is information selected if it is not a functional response to capacity limitations? It is as well to separate answers to this question that are principally concerned with the broader issue of general information processing, efficiency, and performance from those that are principally concerned with information specific to language learning, representational change, and competence. The performance argument made variously by Allport (1987, 1989), Korteling (1994), Neumann (1987, 1996), and van der Heijden (1992) is that selection serves as a means of action control rather than as a response to capacity limitations. Actions are responses to task demands, and allocation of attention to input with the goal of meeting these demands is the result of control processes, operationalized in short-term/working memory. Selection of input relevant to the dominant action also serves the important function of inhibiting and suppressing perception of the many other stimuli which are detected and held in the sensory register, and which may be called for by alternative, contradictory speech and action plans (Faust and Gernsbacher, 1996; Neely, 1977; Shallice, 1972, 1978; Tipper, 1992). Thus, the requirement for coherent speech and action, and continued adherence to a plan, not scarcity of resources, forces selective perception and thought.

### *2.2.3 What guides selection of L2 input?*

Carroll has recently argued that in SLA theory "the idea of attention as a selection function cannot be maintained" (S. Carroll, 1999, p. 343). Clearly it cannot be maintained as an *autonomous function*. Input is detected (via peripheral attention) and stored in the sensory register, then selected (via focal attention) from the stimulus array. But selection is at the same time a response to *control processes* such as attention allocation policy, scheduling and switching between concurrent task demands, and strategy monitoring. Selection of linguistic input is therefore just one aspect of action control, guided by the supervisory attentional system, and executive control mechanisms. There are a number of accounts of these control mechanisms in cognitive psychology, which can be broadly grouped into three categories: those involved in task analysis, in selection and control of the cognitive and metacognitive strategies for performing the task, and in monitoring the effectiveness of strategies (see Baddeley, 1986; Butterfield and Albertson, 1995; Case, 1992; Eslinger, 1996; Sternberg, 1985).

SLA theories diverge, as S. Carroll (1999) points out, on the role of control processes in guiding selection of input for language learning, and the mechanisms and parsing procedures available to them (see Harrington, 2001). In some views, selection is guided by innate representations of abstract phonetic and grammatical knowledge, which enable auditory cues in the input to be detected, analyzed, categorized, and parsed. Compatible with these views are models of L1 and L2 speech perception and word recognition that propose a categorical process of phoneme identification, drawing on knowledge of universal phonological "features" (Liberman and Mattingly, 1989) upon which L1 phonetic categories are based, and which may interfere with L2 speech perception (Flege and Munro, 1994). Such non-semantic representations may be modularly encapsulated (Fodor, 1983; Schwartz, 1999), distinct from (but interfaced with) the conceptual system (Jackendoff, 1997; White, this volume) or not (O'Grady, this volume), and activated early, automatically, and involuntarily. But automatic activation still requires attention (Boronat and Logan, 1997; Holender, 1986; Hsiao and Reber, 1998; Logan, 1990; Mulligan, 1997).

Alternatively, constructivist accounts of SLA argue no modular, encapsulated knowledge is available to guide language development, and recognition and selection of input. Compatible with these views are models of speech perception that propose a non-categorical, continuous process of pattern recognition, which is non-specialized (Massaro, 1987). Some argue that knowledge of language emerges out of an automatic distributional analysis of co-occurring features of the input (Broeder and Plunket, 1994; N. Ellis and Schmidt, 1997; Elman, 1990; Gasser, 1990; Gasser and Smith, 1998), contributing to chunk strength and knowledge of sequencing constraints (N. Ellis, 2001, this volume), represented as a pattern of associations over neurons, and that this occurs late during full semantic processing. In MacWhinney's Competition Model (1987, 2001) this distributional analysis is guided by selective attention to cues in the input (e.g., word order, case marking) which enable form–function relations to be mapped during L2 message comprehension. While they disagree on issues of whether speech perception is a specialized/categorical or general/continuous process of pattern recognition, and whether representations of language properties are modular and encapsulated or not, and innate rather than emergent, all agree that selection of detected auditory input happens (whether early or late), and that attention is required for it to happen, but that it need not (but very often does) implicate awareness (N. Ellis, 2001; Hsiao and Reber, 1998; Schmidt, 1990, 1995, 2001; Tomlin and Villa, 1994). Issues of the relation between detection, selection, and awareness during L2 learning are taken up again below in reporting findings from SLA research.

## 2.3 SLA research into attention as selection

### 2.3.1 Input, intake, and awareness

The role of attention, and awareness, in selecting input as intake for L2 learning has been a controversial issue in SLA theory for some time. Krashen (1985,

1994) has argued that adult learners have access back to the "unconscious" processes and innate mechanisms that guide L1 "acquisition," and that conscious "learning" is minimally influential on the ability to learn and use an L2 in communication. However, Schmidt (1990) argues that the critical notion of "unconscious" is inadequately described in Krashen's work, and can be used to describe three different things: learning without intention (unconscious learning is possible in this sense, since we can learn without intending to); learning without explicit metalinguistic knowledge (unconscious learning is possible in this sense, since nobody has metalinguistic knowledge of all the rules of their L2); and learning without awareness. It is in this last sense that learning must be conscious, Schmidt argues, since we must pay attention to input and also have the momentary subjective experience of "noticing" it, if we are to subsequently learn. Schmidt argues that a higher level of awareness than noticing, rule understanding, is not necessary for learning, but can be facilitative. Schmidt's "noticing" hypothesis has been the focus of recent debate. Two broad theoretical objections have been raised to it. It has been claimed that attention without awareness can lead to learning (Tomlin and Villa, 1994), and also that the noticing hypothesis is pre-theoretic, since it does not specify what properties of input are available for noticing and learning (S. Carroll, 1999). A third objection is methodological (Truscott, 1998): it has been argued the noticing hypothesis is unfalsifiable given the difficulties of precisely measuring awareness.

First, Tomlin and Villa (1994) argue "detection," not selection accompanied by noticing, is the attentional level at which SLA must operate, since detected information can be registered in memory, and dissociated from awareness. Experiments by Marcel (1983) appear to show this. In these experiments rapidly presented words which subjects cannot report awareness of, such as "doctor," prime and so speed the time taken for reading subsequent words, such as "nurse," to which they are semantically related, but do not prime others, such as "balloon," which consequently are read more slowly. As Schmidt (1995) has pointed out, however, these findings do not address the issue of "learning," or new memory for input, since subjects already know the priming and primed words. Such studies are evidence only of automatic, unaware, activation of existing knowledge that Schmidt does not deny could occur. In fact it must occur, as I make clear below.

While detection is clearly necessary for further processing of novel stimuli, Schmidt argues only that the subset of detected information that is selected via focal attention can be "noticed," and that this is the attentional level at which input becomes "intake" for learning. I have argued (Robinson, 1995b) that memory processes, such as maintenance and elaborative rehearsal, which allocation of focal attention activates, are coresponsible for noticing and the durability and extent of awareness that noticing is accompanied by. These relationships are illustrated in figure 19.4 below and are discussed more fully in the following section on memory in SLA. Importantly, however, as described previously, while focal attention and noticing are selective of input, they are also *inhibitory* of the much larger set of detected information, and suppress perception of it in the interests of maintaining continuity of action and preventing interference. Thus detection is necessary as a stage prior to intake, but cannot be coextensive with it. Yet on occasions, *involuntary* switches of focal attention do occur (Naatanan, 1992; Posner, 1980) when automatic activation of existing knowledge calls for them (e.g., when you pause in conversation because you notice a burning smell coming from the kitchen), or when an assumed regularity in the input (based on an internal model of, e.g., word order, pronunciation, or morphological affixation) is seemingly randomly violated (Prinz, 1986). Speech and other plans are important to maintain, but must be interruptible. These issues are important to understanding the role of attention during incidental learning and the rationale for "focus on form" described below, which aims to facilitate switches of attention from meaning to form during communication.

A second objection, made by S. Carroll (1999) and Truscott (1998), is that the noticing hypothesis is representationally empty, or pre-theoretic regarding properties of the input signal that "trigger" noticing. While a property theory is essential to a theory of SLA, these are not valid objections to the noticing hypothesis per se, which is not a comprehensive theory, and was not proposed as one. Schmidt describes what must be noticed as "elements of the surface structure of utterances in the input, instances of language, rather than any abstract rules or principles of which such instances may be exemplars" (2001, p. 5). Comments that "we do not notice and are not consciously aware of the properties or categories of our own mental representations of the signal . . . we do not notice and have no awareness of the internal organization of aspects of logical form or scope" (S. Carroll, 1999, pp. 354, 356) are thus irrelevant to the noticing hypothesis as stated. These two objections are linked, of course. If innate representational knowledge of the shape of possible grammars is accessible in adulthood, then positive evidence of the L2, detected outside of awareness, could prime and automatically activate it, as in Marcel's experiments described above, triggering learning mechanisms such as parameter resetting. Schmidt's noticing hypothesis stands as a simple challenge to these "Minimalist" accounts of the role of attention and awareness in SLA.

The third objection to the noticing hypothesis – the difficulty of measuring awareness precisely – cuts both ways: any counter-claim that learning is possible without the momentary subjective experience of awareness must also demonstrate its absence. Schmidt (1990) operationally defined "noticing" as the availability for verbal report. Admittedly, this raises complicated methodological and interpretive issues, since the contents of awareness are sensitive to, but not always coextensive with, what can be reported, given that awareness may be momentary and fleeting, that subjects differ in their propensity and ability to verbalize, and that some things that are noticed are easier to put into words than others (Faerch and Kasper, 1987; Jourdenais, 2001; Kasper, 1999; Schmidt, 1995, 2001; Shanks and St John, 1994). For this reason, recognition measures of awareness, such as those adopted in implicit memory studies (e.g., preference rating, word fragment completion tests; for discussion see Richardson-Klavehn and Bjork, 1988; Robinson, 1995b, 1996a) may be more sensitive measures than those requiring on- or off-line production and verbalization of the contents of awareness. Given this caveat, however, results of a number of recent studies using verbal reports as data appear to support Schmidt's hypothesis.

### *2.3.2 Operationalizing "noticing"*

Methodologies for studying the role of awareness and noticing in learning (in a variety of linguistic domains, across a variety of L2s) have included both *off-line* verbal report measures, such as diary entries, questionnaire responses, and immediate and delayed retrospection, and *on-line* measures such as protocols. Schmidt (Schmidt and Frota, 1986) found that diary entries describing aspects of L2 input (Portuguese) that he noticed in the input corresponded strongly with the subsequent appearance of these features in his production during interaction with a native speaker in planned, monthly conversations.

Robinson (1996a, 1997a) found that written questionnaire responses asking participants exposed to L2 input in an immediately prior experiment if they had searched for rules, and could say what the rules were, correlated positively and significantly with learning in an implicit (memorize examples) learning condition, and that ability to verbalize rules correlated positively and significantly with learning in a condition where participants were instructed to try and find rules during exposure to the input. In both conditions, positive correlations of language learning aptitude and awareness suggest that this is an ability variable that can trigger awareness at the levels of noticing, rule search, and verbalization.

Kim (1995) used immediate off-line retrospective verbal reports to examine the relationship between phonological awareness and L2 listening comprehension (measured as the ability to correctly match a picture to one of 30 aurally delivered texts). Finding slow speech rate resulted in greater comprehension than normal speech rate, Kim established a tentative implicational hierarchy of phonological awareness based on verbal reports of those clues in the speech stream learners attended to in arriving at answers to the comprehension questions: perception of key words > of phrases > of clauses > and of conjoined clauses. Coding learners based on this hierarchy, however, failed to distinguish level of awareness of learners exposed to slow vs. normal speech, though there was a trend to higher levels of phonological awareness for those exposed to slowed speech, who also demonstrated significantly greater comprehension.

Philp (1998) also used an immediate off-line simulated recall technique, in this case to assess whether learners had noticed the relevant properties of orally delivered recasts. Immediately following provision of a recast during dyadic NS–NNS interaction, the NS prompted recall via a signal (a knock on the table). Correct recall and repetition of the recast form was assumed to demonstrate noticing. Philp found that, in general, and particularly for higher-level learners, those who demonstrated greater noticing during the simulated recalls also demonstrated greater gain and development of question forms from pre- to immediate and delayed post-tests.

Other studies have used on-line measures of awareness, such as protocols (Alanen, 1995; Jourdenais et al., 1995; Leow, 1997, 2000; Rosa and O'Neill, 1999) to examine uptake and learning of information during treatments designed to draw learners' attention to forms while processing for meaning (these involved italicizing and underlining words in a text; completing a crossword puzzle; and completing a multiple choice textual jigsaw puzzle). Alanen (1995), Jourdenais et al. (1995), Leow (1997, 2000), and Rosa and O'Neill (1999) all reported that those subjects demonstrating greater noticing and awareness during the on-line protocols also demonstrated greater intake and gain, at least on some aspects of the targeted forms in each study (aspects of Finnish grammar in Alanen, 1995, and of Spanish grammar in Jourdenais et al., 1995; Leow, 1997, 2000, and Rosa and O'Neill, 1999) than those whose protocols demonstrated less noticing and awareness of the targeted forms.

While of theoretical interest, zero-point issues of whether learning is possible without attention or without "noticing" are of much less practical interest to SL pedagogy than the findings summarized above. Few would argue the zero-point issue with regard to attention. Gass (1997), however, claims that evidence of the generalizability of relative clause instruction on more marked (and complex) objects-of-preposition relative clauses to less marked subject and object relative clauses is evidence of non-attentional learning. Yet in both Gass (1982) and Eckman, Bell, and Nelson (1988), who found similar effects, there was pre-testing, and attended exposure to all forms of relative clause, before the instructional treatment, and there is additionally no guarantee that in their prior learning experience learners in these studies had not attended to the three forms of relative clause in question. In experimental studies, where such control is guaranteed, there are few advocates of the zero-point option for attention and learning. For example, most explanations of "implicit" learning of artificial grammars, or rules governing repeating sequences of letters or lights (Hsiao and Reber, 1998; Nissen and Bullemer, 1987; Stadler, 1992), clearly state that attention is required for processing the learned stimuli. As Hsiao and Reber observe, in implicit sequence learning experiments, increasing the structural constraints on and therefore the complexity of rules describing the repeating sequences also increases the probability of event sequences/letter strings occurring after other event sequences. The probability learning that is facilitated by and results from exposure to such sequences is merely *less* demanding of attention, not independent of it: "The fewer the constraints, the more attentional resources will be required to learn that sequence" (Hsiao and Reber, 1998, p. 475).

In summary, the necessity of noticing and awareness is more controversial than the necessity of attention for SLA (Schmidt, 1995, 2001) and is difficult to prove conclusively, given that no measurement instrument or technique can be assumed to be entirely coextensive with, and sensitive to, the contents of awareness and noticing. Nonetheless, cumulative findings from the studies reported above are predominantly in line with Schmidt's noticing hypothesis, and are certainly not contrary to it. Furthermore, many have argued that, even if it is not necessary, noticing certainly contributes to learning and retention, and that consequently consciousness raising (Rutherford, 1987), input enhancement (Sharwood-Smith, 1991), processing instruction (VanPatten, 1996), or focus on form (Long, 1991; Long and Robinson, 1998), which aim to induce it, are likely to be beneficial to learners.

### *2.3.3 Focus on form*

The noticing hypothesis offers a partial explanation of why a focus on meaning alone, with plentiful opportunities for exposure and processing of input, as in Canadian immersion classrooms, often results in levels of high comprehension ability and fluency, but poor accuracy in production (Harley, 1993; Harley and Swain, 1984). Learners did not selectively attend to and notice communicatively redundant, perceptually non-salient, or infrequent and rare forms in the input. In these and other cases, Long (1991) has argued *focus on form*, in the context of meaningful use of language, may be necessary to promote and guide selective attention to aspects of input which otherwise may go unnoticed, unprocessed and unlearned:

> Focus on form refers to how focal attentional resources are allocated . . . during an otherwise meaning-focussed classroom lesson, focus on form often consists of an occasional shift of attention to linguistic code features – by the teacher and/or one or more students – triggered by perceived problems in communication. (Long and Robinson, 1998, p. 23)

Undoubtedly, while processing oral L2 input for meaning, as in naturalistic or immersion environments and during L2 reading, learners do unintentionally attend to, notice, and learn many vocabulary or grammatical and pragmatic features of the L2 (incidental learning) (Gass, 1999; Huckin and Coady, 1999; Hulstijn, this volume; Rott, 1999; Schmidt, 1990, 1995). However, in those areas where unguided incidental learning is slow and inefficient (Long, 1996), or just not possible for learnability reasons (L. White, 1991), guided focus on form is widely accepted to be a necessary pedagogic intervention. More controversial is the nature of the pedagogic technique that intervention should adopt in order to be optimally effective, while being minimally intrusive on the communicative activity (Doughty and Williams, 1998). For example, is it more effective to proactively instruct learners in targeted features prior to communicative activities, via a brief rule explanation or metalinguistic summary (instructed learning)? Or is it better to adopt less communicatively intrusive techniques for focusing attention on form, by giving learners instructions to process for meaning (e.g., to read a news article in preparation for a debate) while drawing their attention, through underlining or highlighting, to targeted forms in the text (enhanced learning)? Alternatively, reactive techniques for focus on form, such as oral recasts of problematic learner utterances, involve no a priori decision about which forms to target.

Recent experimental laboratory research has investigated these issues by comparing differences in learning under incidental, instructed, and enhanced conditions across a variety of linguistic domains (see Hulstijn, 1997, for review). This research has often also been concerned to match the difficulty or complexity of the targeted instructional form to the best learning condition. While conceptualizations and/or operationalizations of rule complexity differ across studies (see Doughty, 1998; Hulstijn and DeGraaff, 1994; Robinson, 1996b, for discussion), a general summary of the laboratory research findings is that proactive rule instruction can lead to short-term rate advantages over incidental and enhanced learning in simple grammatical domains (DeGraaff, 1997a, 1997b; DeKeyser, 1995; N. Ellis, 1993; Robinson, 1996a, 1997a), but the positive effects of rule instruction are much less obvious for complex grammatical domains.

There is also evidence from experimental laboratory research (Robinson, 1997b; Williams, 1999) and classroom studies (Alanen, 1995; S. Carroll and Swain, 1993; Doughty, 1991; Doughty and Varela, 1998; Fotos, 1993; Iwashita, 1999; Jourdenais et al., 1995; Leeman, Arteagoitea, Fridman, and Doughty, 1995; Leow, 1997, 2000; Muranoi, 1996, 2000; Spada and Lightbown, 1993; J. White, 1998; L. White, Spada, Lightbown, and Ranta, 1991) that enhanced learning conditions, adopting (i) techniques for off-line, proactive, textual input enhancement of targeted forms and (ii) reactive, on-line, aural/interactive, or gestural enhancement of problematic aspects of production during communicative tasks (which are both assumed to induce selective attention and noticing) can positively affect learning, relative to unstructured and unenhanced exposure alone. However, relying, as they may, to a much greater extent on individual differences in cognitive ability variables such as aptitude (Robinson, 1997a, 2001b) or working memory capacity (Mackey et al., 2002; Philp, 1999; Robinson, 2001b, 2002; Robinson and Yamaguchi, 1999; Robinson, Strong, Whittle, and Nobe, 2001; Williams, 1999), group effects for input and output enhancement have been less robust than those for explicit rule instruction. Nevertheless, given the short-term nature of most of the experimental laboratory studies of the effects of rule instruction, it may be that the positive effects of input and output enhancement obtained in classroom studies – which are typically studied over much longer, and more ecologically valid, periods of exposure – while showing less immediate short-term gain, are more durable and permanent (see Doughty and Williams, 1998; N. Ellis and LaPorte, 1997; Long and Robinson, 1998; Norris and Ortega, 2000; and Spada, 1997, for extended reviews and interpretations of these findings).

## 2.4 *Attention as capacity*

### 2.4.1 *Overview*

Tasks differ in the demands they make on our attention. Elsewhere in this volume this issue is dealt with in terms of important distinctions between controlled and automatic L2 processing (DeKeyser, 1997, 2001; Segalowitz, this volume), the former being traditionally viewed as more attention demanding than the latter; and between explicit and implicit L2 learning (DeKeyser, this volume; N. Ellis, 1994), the former also being traditionally viewed as more attention demanding than the latter. Some argue that speed-up of control processes and withdrawal of attention (McLaughlin, 1990; Shiffrin and Schneider, 1977) and unconscious abstract rule induction (Reber, 1993) in a separate implicit memory store (Paradis, 1994; Schacter, 1996) can explain these differentials in attentional demands. Others have argued (Logan, 1988, 1990; Shanks and St John, 1994) that automatic decision making (Robinson, 1995b; Robinson and Ha, 1993) and implicit L2 learning (Robinson, 1995b, 1996a, 1996b, 1997b) are memory-based processes involving storage and retrieval of attended instances in memory.

Differentials in the attentional demands of L2 tasks, and of dual versus single-task performance have also been proposed as one cause of within-learner interlanguage variation. Complex tasks are more attention demanding than simpler tasks, and performing two tasks simultaneously is more attention demanding than performing one alone (Gopher, 1992; Heuer, 1996), and varying these attentional demands may systematically affect the accuracy, fluency, and complexity of learner speech (Crookes, 1988, 1989; Hulstijn, 1989; Robinson, 1995a, 2000, 2001a, 2002, forthcoming; Robinson and Niwa, 2001; Robinson, Ting, and Urwin, 1995; Skehan, 1998; Skehan and Foster, 2001; Tarone, 1985). The specific issue addressed here, then, is the relationship of task demands to attention used in the sense of *capacity*, since capacity limits are often invoked to explain the greater "mental load" and therefore difficulty of controlled processing, explicit learning, and L2 processing during complex and dual-task performance. Three positions on the structure and significance of capacity limits can be identified.

### *2.4.2 Single capacity and multiple resources*

Kahneman (1973) proposed that a single finite volume of attention is available for allocation to competing task demands. Attention is allocated in working memory, and is selective of actions, not incoming messages. Capacity limits are not fixed and unchanging, but vary with the level of arousal. Task difficulty is defined in terms of capacity consumption, as reflected in physical indices of "effort" such as pupilliary dilation. More complex and less automatized tasks consume more attentional capacity, and require greater effort. Multiple-resource models (Wickens, 1984, 1989, 1992) go beyond simple single-capacity models by proposing distinctions between separate resource pools from which attention is allocated to different task dimensions, such as processing mechanisms required by the task (perceptual vs. response), codes of processing (spatial vs. verbal), and modality (auditory vs. visual; see figure 19.2). This modification to Kahneman's model is necessary since structural alterations in a secondary task, while keeping its difficulty constant, are known to affect performance. For example, when simultaneously performed tasks both require manual responses (steering and written recall of digits), there is more interference/worse

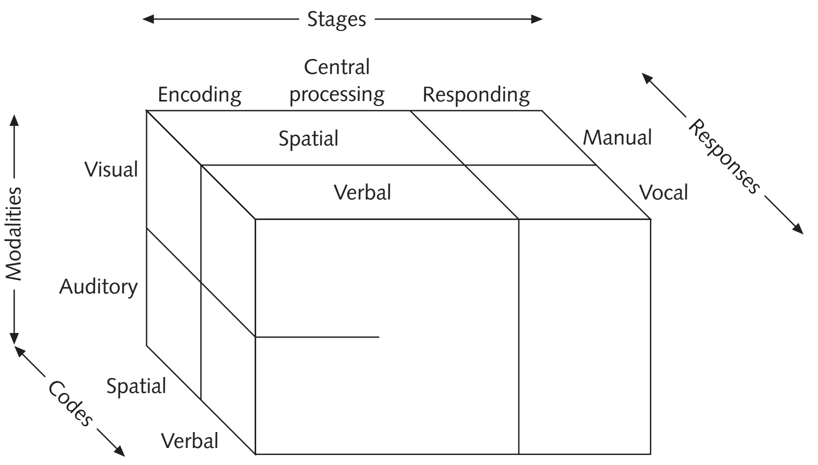


**Figure 19.2** The proposed structure of processing resources
*Source*: Wickens (1992, p. 375)

performance than when one requires a manual and the other a verbal response (steering and verbal recall of digits). Thus, resource competition was argued to exist within, but not between, separate attentional pools.

While multiple-resource theory has influenced some L2 research concerned with awareness and intake (Rosa and O'Neill, 1999), and the effects of task dimensions on input processing and production (see figure 19.3 below), and proven a productive framework for the study of human workload and workplace design (Wickens, 1992), it has a number of problems, one of which concerns its key theoretical assumption, that is, that attention is limited in capacity. First, performance limits are simply *ascribed* to capacity limits; it is not specified *how* or *why* capacity is limited – a key theoretical objection (Neumann, 1987). Second, interference between competing tasks is often more specific than is predicted on the basis of the resource pools identified (Navon, 1989). For example, performing arithmetic tasks simultaneously causes more interference than performing a spelling and arithmetic task simultaneously (Hirst and Kalmar, 1987), despite the fact that they are both classified, in Wicken's model, as drawing on the same resource pools. Third, inventing new classifications of resource pools to account for examples of successful and unsuccessful time-sharing is unsatisfactorily post hoc, and unconstrained. These issues of explaining interference beyond those predicted by classifications of resource pools have led to the development of an alternative account of the relation between attention and task demands.

### *2.4.3 Interference and/or capacity?*

Interference models argue that increasing the number of stimuli and response alternatives or the similarity between them will sometimes lead to confusion, reducing performance efficiency. This can be caused by competition for the same types of codes during information flow, or by "cross-talk" between similar codes. For example, while typing auditorily presented words, at the same time as shadowing (repeating) visually presented words (the same code), visually presented words are sometimes mistakenly typed (Shaffer, 1975). Interference is therefore caused by involuntary attention shifts, not by resource limitations, and is a breakdown in action control (Navon, 1989; Neumann, 1987). Within this approach:

> considerations of resource scarcity or the performer's ability to allocate sufficient processing efforts are irrelevant. The limits on task performance are not conceived in these terms. Attention control is constrained to a decision to engage, disengage and shift attention between tasks and the pursuit of intentions. In interference models the only limited resource is time and its derived scheduling constraints. (Gopher, 1992, pp. 279–80)

While there are clearly structural constraints on human information-processing ability, and limits on the information that can be stored in short-term working memory, these accounts nonetheless question the utility of the notion of "capacity limits" on attention in explanations of degraded task performance (Logan, 1992). Connectionist models of representation, processing, and attention (e.g., Phaf, van der Heijden, and Hudson, 1990; Schneider and Detweiler, 1988) complement these non-limited attention capacity accounts. Such models consist of "mutually activating and inhibiting units among and within various levels of processing. From this perspective, processing limits are due to interference, confusion and cross-talk among elements of a neural net and not to capacity constraints" (Sanders, 1998, p. 15).

### *2.4.4 Task demands, capacity, and interference*

Two points are worth making with regard to these three theoretical views of central processing, task demands, and attention. First, single-resource, limited-capacity models cannot explain many of the effects of structural alterations in task demands on task performance, whereas multiple-resource theory can. Second, multiple-resource theory may be able to accommodate interference models: interference models are lower-level (implementational) approaches to describing the causes of attention switching and task competition during control of information flow. This may mean that multiple-resource theory maintains the distinction between resource pools, but abandons the notion of capacity limitations (which interference models do not assume) within those pools.

Much SLA research within an information-processing framework assumes attentional capacity is limited, and, as a result, that accuracy, fluency, and complexity may compete for resource allocation during L2 task production (Skehan, 1998) or that "form" and "function" compete for scarce attentional resources during input processing (VanPatten, 1996). Tomlin and Villa (1994) have argued that these assumptions, and the single-resource, limited-capacity model of attention they are based on, are too "coarse-grained." Whether or not the notion of capacity limits during single- and multiple-task performance is retained in theories of attention, it seems clear that invoking limits on *undifferentiated* attention capacity as an explanation of various SLA processes (e.g., the inefficiency of input processing, transfer at a variety of levels, lapses in fluency and accuracy during task production) is unsatisfactory. As described above, current theories of attentional allocation to input are "rapidly moving away from the limited capacity processor" (Sanders, 1998, p. 356), and do not see this as a major constraint on why and when selection of input, or of action, takes place. Consequently, these trade-off effects (form vs. function, accuracy vs. fluency) may be better explained not in terms of a priori capacity limits on a single pool of attention but in terms of control functions during central processing (allocation policy, time constraints on scheduling attention allocation), and interference occurring during resource allocation to those specific *task demands* which central processing responds to. From the perspective of interference theory, explanations linking relative ease or difficulty of L2 comprehension, or different characteristics of L2 production, to task demands may be more legitimately framed in terms of confusion and cross-talk between codes (of L1, interlanguage, and L2 syntax, morphology, semantics, and phonology/orthography) within specific resource pools during task performance, rather than in terms of global capacity limitations.

Abandoning invocation of undifferentiated attention capacity limits to explain the effects of task demands on comprehension and production will require more precise specifications of constraints affecting attention allocation during language processing. Codes would have to be representationally specified, as would resource pools. The competition model (MacWhinney, 1987, 2001) offers one framework (there are others) for describing codes and their peaceful coexistence, or the competition and interference between them caused by task demands on comprehension and production – that is, the extent to which L1 and L2 differ in their cues to form–function relations. For example, pre-verbal positioning is a highly reliable and available cue to assigning agency in English, but less reliable in Spanish, and simply not available in verb-initial languages like Samoan (Samoan uses an ergative marker *e* to mark the subject of transitive verbs). Interference and misinterpretation (confusion) occur where the same cue is available but differs in reliability across languages (as with pre-verbal positioning in English and Spanish) (see MacWhinney, 2001). The search for a cue not available in the L2 (English) which is available in the L1 (Samoan), such as an ergative marker, can also lead to interference (cross-talk).

Investigating the structure of attentional resource pools drawn on in L2 processing is a recent area of SLA research, and some models have been proposed to guide research (Robinson, 2001a, 2001c). Task design features can *disperse* attention between pools (e.g., by requiring two task components to be performed concurrently), or *direct* attention to specific needed areas of the L2 within a pool (e.g., by requiring continual reference to events happening now, and so to use of the present tense), and it is possible that dispersal may lead to a higher probability of trade-offs and interference of the kind described above (Robinson, 2001c). These issues are further taken up in discussion of SLA findings below.

## 2.5 SLA research into attention as capacity

### 2.5.1 Resources, task demands, and language production

Recent SLA research has begun to examine the extent to which design features of L2 tasks make differential demands on attention, and here the notion of attention as capacity is most important to understanding the effects of these demands on perception and selection of input, as well as on production. Allocation of attentional capacity to task demands is a control process, and as task components and demands proliferate, so does the difficulty of managing allocation policy, with consequent lapses in perception and production. In studies of L1 and general intellectual development the relation between tasks, attentional capacity, and learning has most often been studied by contrasting performance on simple (less capacity-demanding) and complex (most capacity-demanding) versions of tasks at different ages, or stages of linguistic and cognitive development (see Case, 1985, 1992; Halford, 1993; Karmiloff-Smith, 1992; Nelson, 1996; Thelen and Smith, 1998). The same simple–complex task paradigm, along with studies of single vs. dual and multiple-task performance, has also been adopted in studies of the effects of task demands on attentional capacity in the acquisition of complex skills in adulthood (see Fleishman and Quaintance, 1984; Holding, 1989; Sanders, 1998; Wickens, 1992). Drawing to varying degrees on this research into L1 linguistic, cognitive, and skill development, as well as on previous classifications of the interactive demands of L2 tasks (Crookes, 1986; Duff, 1986; Long, 1989; Pica, Kanagy, and Falodun, 1993) and SLA research into the effects of attention to speech on L2 accuracy (Hulstijn, 1989; Tarone, 1985; Tarone and Parrish, 1988), SLA researchers have begun to theorize and operationalize the attentional demands of L2 tasks and to study their effects on production, comprehension, and learning (see Norris, Brown, Hudson, and Yoshioka, 1998; Robinson, 1996c, 2001a, 2001c; Skehan, 1996, 1998, for theoretical discussion and reviews of findings).

### 2.5.2 Attentional demands and task output

Figure 19.3 illustrates a number of dimensions of task demands that have been argued to affect attentional allocation and consequently the quality of L2 production and comprehension. Tasks where planning time and prior knowledge are available, and which involve only a single activity, are simpler and less attention demanding than dual tasks requiring simultaneous activities, and where no prior knowledge or planning time is available. Increasing complexity along these dimensions alone has the effect of depleting the attention available to perform the task, and dispersing it over many, non-specific linguistic aspects of production and comprehension. On the other hand, tasks which

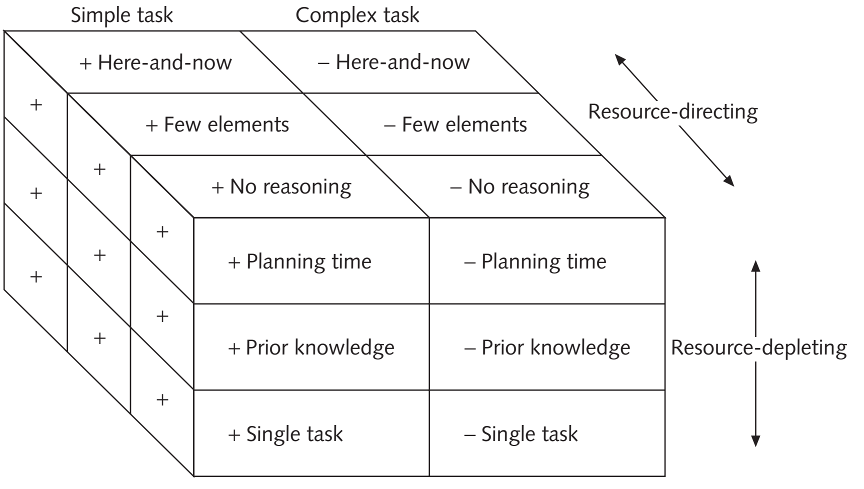


**Figure 19.3** Resource-directing and resource-depleting dimensions of task complexity

require reasoning and reference to many elements, and which are displaced in time and space, are more complex and attention demanding than their simpler counterparts, but these dimensions have the potential to direct learner attentional resources to needed aspects of language code, such as conjunctive coordinators to establish causality, past tense morphology and temporal expressions, and complex nominalizations to distinguish numerous similar elements. Increasing task complexity and attentional demands simultaneously along both these types of dimension have the effect of approximating the performance constraints of real-world task activity. For example, a simple task might involve giving directions to a partner after a period of planning using a simplified small map (few elements) of a known area (prior knowledge available) where the route from A to B was already marked on (making it a single task, as opposed to thinking up the route and describing it simultaneously). A complex, real-world version of this task would involve giving directions from an authentic map of a large area, which is unfamiliar to the information giver and receiver, without a route marked on, and with no planning time, as when a passenger gives directions to a driver from a road map as they travel quickly through an unfamiliar town.

A great deal of previous research has focused on the dimension of planning time (Crookes, 1989; R. Ellis, 1987; Foster and Skehan, 1996; Mehnert, 1998; Ortega, 1999; Skehan and Foster, 1997, 2001; Ting, 1996), finding, in general, that planned tasks result in greater fluency and complexity of production, with some studies also showing gains in accuracy on planned tasks. What has yet to be shown, but which might be expected, is the effect of planning time on the accuracy and complexity of production during tasks made complex along different resource-directing dimensions – such as tasks requiring reasoning vs.

tasks requiring reference to many similar elements. Planning time for the former could be expected to optimize time available for producing complex syntax to express logical causality (if-then), belief justification (X because Y), and conditionality (if it/I were, it/I would), for example: planning time for the latter to encourage planning of complex nominal predicates, use of relative clauses, and article and determiner use. These considerations point to the potential dangers of predicting global effects on accuracy and complexity of production on one resource-depleting dimension of complexity and attentional demands, such as the availability of planning time, regardless of its interaction with other dimensions that have the potential to differentially direct attentional resources to task-relevant aspects of language code.

Nonetheless, the tendency in this descriptive data-gathering period of research has been to examine the effects of differences in attentional demands along one dimension independently of others. Effects have been found for greater fluency but lower accuracy on narratives performed in the here-and-now (stories performed in the present tense, while looking at picture sequences illustrating the story) than in the more complex there-and-then (stories performed in the past tense while remembering the picture sequences; Rahimpour, 1997, 1999; Robinson, 1995a). Increasing the reasoning demands and number of elements that need to be referred to and described has also been shown to negatively affect fluency (Niwa, 1999; Robinson, 2001a, 2001c; Robinson and Niwa, 2001) while having positive effects on some aspects of lexical range and linguistic complexity (Brown, 1996; Brown, Anderson, Shillcock, and Yule, 1984; Niwa, 1999; Robinson, 2001a; Robinson and Niwa, forthcoming). Similarly, dual tasks have been shown to result in less fluent production than single tasks (Robinson and Lim, 1993; Robinson et al., 1995), as do tasks performed where no prior knowledge is available (Chang, 1999; Robinson, 2001a, 2001c). Lack of prior knowledge has also been shown to negatively affect comprehension during reading and listening tasks (Barry and Lazarte, 1998; Carrell, 1987; Clapham, 1996; Dunkel, 1991; Urwin, 1999), with prior knowledge of related content facilitating listening comprehension measured by inferencing questions, and prior knowledge of formal organizational schemas facilitating comprehension measured by recall questions (Urwin, 1999).

### *2.5.3 Attentional demands and task intake*

One pedagogic motivation for examining the attentional demands of tasks and the effects of these on production has been a concern to design pedagogic tasks for learners which optimize production practice in the three areas of fluency, accuracy, and complexity of output (Bygate, 1996, 1999; Skehan, 1996, 1998; Skehan and Foster, 1997, 2001). A second motivation stems from proposals for "analytic" approaches to pedagogy, such as Long's proposals for Task Based Language Teaching (see Long, 1998, forthcoming; Long and Norris, 2000), which reject linguistic units of analysis (either grammatical rules, lexical items, or notions and functions, etc.) and associated criteria for grading sequencing units of instruction, in favor of a syllabus made up of a series of pedagogic tasks. One proposal for operationalizing such syllabuses is to base task sequencing on empirical evidence of differences in the cognitive demands of tasks, so that pedagogic tasks progressively approximate the full information-processing complexity of real-world target task demands over a course of instruction (Long, 1998; Long and Crookes, 1992; Robinson, 1996c, 2001a, 2001c).

A third and fundamentally important motive for studying the attentional demands of tasks lies in the effect these have on learning. Schmidt (1990) argued that along with input factors, such as perceptual saliency and frequency of forms, task demands are also powerful determinants of what is noticed and selected via focal attention for further processing. Unfortunately, to date there has been almost *no* research into the effect of task complexity and dual-task performance on selection and intake of new, previously unknown, task-relevant linguistic information. Two theoretical positions have been put forward, however, which promise to stimulate future research. One position is that increasing the complexity of tasks and their multiple components reduces a pool of generally available attention capacity (Kahneman, 1973), thus negatively affecting detection, selection, and subsequent memory for new linguistic forms in the input. This is compatible with VanPatten (1996) and Skehan's assumption of a single-resource, limited-capacity model of attention (Skehan, 1996, 1998; Skehan and Foster, 2001), which predicts that as learners' attentional limits are reached, learners prioritize processing for meaning over processing form. The researchers argue that this leads learners to adopt a strategy of paying attention to content words at the expense of grammatical morphology during message comprehension, and to an increasingly lexicalized, ungrammatical mode of speech production.

Alternatively, if different components of task demands draw on attention allocated from within separate resource pools (Wickens, 1984, 1989), then increasing the cognitive demands of tasks could, in a number of cases, be argued to increase the attention learners pay to input and output, and to process it more deeply and elaborately, without necessarily being constrained by capacity limits or competition for attentional resources. In this view, increasing task complexity along compatible, separately resourced, dimensions may increase the likelihood of detecting and selecting seeded aspects of the input (Robinson, 1995b, 2001a, 2001c), made salient through such techniques as flooding, visual enhancement (in the case of written text), or recasting (in the case of oral interaction).

To illustrate how tasks can be made complex along compatible, separately resourced dimensions, take the example referred to earlier, that of a simple direction-giving task (requiring reference to few elements where prior knowledge is available) vs. a complex direction-giving task (requiring reference to many elements, with no prior knowledge). I would argue the resource-directing dimensions of complexity identified in figure 19.3, such as *reference to few vs. many elements*, draw on the resource pools Wickens identified for *verbal encoding and vocal responding* (see figure 19.2), whereas this is not necessarily so of the resource-depleting dimensions. *Lack of prior knowledge* of an area described by a map, that is, would affect only visual encoding of the many elements (roads, buildings, other landmarks, etc.) the map contains and would draw on the *visual spatial encoding* resource pool in figure 19.2. Multiple-resource theory predicts little interference between the attentional demands of tasks which increase in complexity – as in this case – along two separately resourced dimensions, which can be time-shared successfully (i.e., no prior knowledge of an area illustrated by a map/visual-spatial-encoding resource pool; reference to many elements/vocal-responding resource pool).

Increasing task complexity may also lead to greater retention of noticed input. For example, Schneider, Healy, and Bourne (1998) showed that increasing the "intra task interference," and hence processing demands, of vocabulary word-list learning tasks (presenting words randomly vs. grouped into simplifying conceptual categories) led learners to process the randomly ordered group more elaborately (see Craik and Lockhart, 1972, and discussion in section 3.3 below), resulting in more retention for these words than for those grouped into categories. Further, since increasingly complex interactive tasks result in greater amounts of negotiation (see Robinson, 2001a) they also increase learner opportunities for, and maybe therefore the likelihood of, making cognitive comparisons between input and output, leading to noticing "gaps" or holes in production (see Doughty, 2001, this volume; Muranoi, 2000; Swain, 1995).

In summary, Skehan assumes attentional capacity is generally available and limited, and that increases in task complexity drain attentional resources and are therefore likely to have the effect of degrading the fluency, accuracy, and complexity of output, as well as perception of input and intake (see Skehan, 1998, p. 174). In cases where complex tasks make demands that exceed the learners' available attentional resources, Skehan argues additional task structure is necessary to attract learner attention to relevant aspects of form, which would otherwise not be processed. A similar rationale underlies VanPatten's (1996) proposals for processing instruction. The alternative position I have described lays less emphasis on capacity limits, and makes the prediction that where dimensions of task complexity are separately resourced, and can be time-shared, then increases in task complexity along multiple dimensions will not degrade output, perception of input, and intake, and may lead to qualitative increases in all three relative to performing simpler tasks. These issues are speculative, unresolved, testable, of great practical relevance to SL pedagogy and syllabus design, and in much need of further SLA research.

## 2.6 *Attention as effort*

Sustained attention to an activity over time is a third, separable use of the term "attention" and is a central notion in studies of vigilance, energetic states, and the causes of decline in performance on a task. Attention in this sense is a "state" concept referring to energy or activity in the processing system, not to structural processes such as selecting, allocating resources, and rehearsing information in memory. To maintain performance on a task, the attentional energy devoted to it must remain at a constant state. Three energetic pools have been proposed (Sanders, 1986), which correspond to the three information-processing stages in figure 19.1; the *arousal* pool (concerned with encoding and affected by variables such as cue salience, intensity, and novelty); the *activation* pool (concerned with central processing and affected by such variables as task preparedness and alertness); and the *effort* pool (concerned with responding, monitoring output, and the feedback it elicits, and affected by such variables as task complexity, time spent on task, and type of feedback provided). Since issues of attention in encoding and central processing are discussed above, I briefly focus here on sustained attention at the third stage of information processing in figure 19.1.

Failure to sustain attention to a task and maintain the level of *effort* expended results in a decline in performance over time. Failure to sustain attentional effort is caused not only by prolonged time on task, but also by the complexity of the task as determined by the number and compatibility of task components and sources of input (Koelega, 1996; Wickens, 1992). This is manifest in a decline in vigilance – failure to detect a target signal (in studies of visual search) and failure to correctly identify and interpret auditory input (in studies of comprehension), as well as failures in grammatical encoding and production leading to mistakes and speech errors (in studies of speech production). In psycholinguistic theories of speech production applied to SLA, failure to sustain attention to a communicative task can be identified as one cause of declines in self-repair and monitoring of output (Crookes, 1988; De Bot, 1996; Kormos, 1999; Levelt, 1989). Swain's notion of "pushed" output (Izumi, Bigelow, Fujiwara, and Fearnow, 1999; Swain, 1985, 1995; Swain and Lapkin, 1995) also appears to implicate the sense of attention as effort; pushed L2 production is more effortful than the normal production level of a learner. Coordinating joint attention to language through the provision and uptake of feedback during L1 child–caregiver (Tomasello and Farrar, 1986) or L2 interaction (Doughty, 2001, this volume; Gass, 1997, this volume; Iwashita, 1999; Lyster, 1998; Lyster and Ranta, 1997; Mackey, Gass, and McDonough, 2000; Muranoi, 2000; Oliver, 1995; Philp, 1998; Pica, 1988, 1992) also requires sustained, effortful attention, and this may increase as the number of participants in the interaction increases (Tomasello, Manle, and Kruger, 1989).

## *2.7 SLA research into attention as effort*

The relationship of attention as effort to attention as capacity is currently controversial. Kahneman's (1973) model of attention, implicitly adopted by Skehan (1996, 1998) and VanPatten (1996), assumed that sustaining attention to tasks which were high in their capacity demands was more effortful than sustaining attention to tasks which were low in their capacity demands. Kahneman argued that greater effort in sustaining attention was indexed by physiological measures, such as increased heart rate and pupilliary dilation, and by greater declines in vigilance and less freedom from distraction over time, compared to less effortful and capacity-consuming tasks. The alternative view I have described above, that is, that there are plentiful attentional resources within separate pools, suggests effortful tasks are those requiring coordinated attention to, and executive time-sharing between, task components drawing concurrently on the same resource pools. Where task components draw on separately resourced pools, sustained attention to a task will be less effortful, and performance will show less decline in vigilance (more effective monitoring and uptake and incorporation of feedback) over time than when there is competition for resources within the same pools. In this view, then, the effort involved in sustaining attention to L2 output and input results from the interaction of time constraints and coordination of attentional resources, not from their scarcity (Logan, 1992; Navon, 1989; Neumann, 1996).

Arousal and the effort pool are also related. Increases in stress lead to greater arousal and also to temporary increases in effort to perform the task, though there are limits to this equation, as described by the Yerkes-Dodson law. Levels of stress, arousal, and performance increase to a point beyond which performance declines, and this point is reached earlier on complex tasks than simpler ones. Attention as effort is therefore related to affective influences on SLA, such as motivation (Dörnyei, 1998, 2002), and the distinction between facilitating and debilitating anxiety (Holthouse, 1995; Horwitz, Horwitz, and Cope, 1986; Jacob, 1996), issues that are dealt with in more detail elsewhere in this volume (see the chapter by Dörnyei and Skehan).

# 3 Memory

## *3.1 Overview*

As the review above illustrates, research into the necessity of attention and awareness in selection of intake for learning has dominated recent SLA research. There is growing interest too in the issue of capacity constraints on attentional allocation. The role of memory has been less controversial, and so, perhaps, less studied. Recently, however, three issues have attracted theoretical interest: the relationship of selective attention and awareness to memory during noticing; the role of memory in implicit and incidental L2 learning; and the effect of individual differences in short-term, working memory capacity on SLA. A fourth issue – the organization and accessibility of information in long-term memory – has been addressed in a number of studies of listening and reading comprehension processes, as well as in studies of bilingual processing (Bialystok, 1991), lexical acquisition (Crutcher, 1998; Hulstijn , 2001), and lexical access and retrieval during SL production (Doughty, 2001), though full review of these areas is beyond the scope of this chapter.

## *3.2 Attention and memory*

It is uncontroversial that memory processes are functionally differentiated, and that the modal view of memory proposed by Atkinson and Shiffrin (1968), distinguishing between perceptual/sensory memory, short-term/working memory, and long-term/episodic and semantic memory, captures some of these functional distinctions (Pashler and Carrier, 1997). Short-term, working memory is capacity limited, whereas long-term memory is not. Information in short-term memory decays rapidly; information in long-term memory does not. Evolutionary explanations (e.g., Reber, 1993) argue capacity limitations on short-term, working memory are necessary – and so have evolved – if fast decisions (based on limited information), which are often necessary to survival, are to be guaranteed. In short, these accounts argue that capacity limits on short-term memory are the result of a decision-making trade-off in evolutionary development between speed (more necessary) and accuracy (less necessary).

More controversial in memory research is whether these functional distinctions correspond to neurophysiologically separate systems, or whether passive short-term and active working memory are distinct stores (Baddeley, 1986; Cowan, 1993, 1995; Nairne, 1996; Schachter, 1996; Shiffrin, 1993; Squire, 1992). However, most memory researchers do hold the view that short-term, working memory is that part of long-term memory in a currently heightened state of activation, and further, that awareness and working memory are isomorphic, and correspond to the contents of short-term memory which are within the focus of attention (Cowan, 1988, 1993, 1995; Nairne, 1996).

## 3.3 *Memory, rehearsal, and awareness*

Consistently with the position described above, figure 19.4 illustrates the following set of relationships between memory and attention. Detected information can briefly enter short-term memory and automatically access previously encoded information in long-term memory outside of awareness (as illustrated by subliminal exposure, priming experiments, such as those of Marcel, 1983, referred to earlier). Automatic, unaware activation of long-term memory representations is the result of categorization mechanisms which compute the similarity distance of the detected input to prior instances encoded in memory (see Estes, 1992; Nosofsky, 1992; Nosofsky, Krushke, and McKinley, 1992; Smith and Sloman, 1994). This is evidence, however, only of unaware recognition, not of learning, since the categories which are activated pre-exist the input. For newly detected information to be encoded in long-term memory, which is "uncalled for" by similarity computing mechanisms, and which needs, therefore, to be learned, the information must enter focal attention and so short-term working memory, where rehearsal processes operate prior to encoding in long-term memory. Rehearsal processes can be of two kinds; *maintenance rehearsal*, requiring data-driven, instance-based processing, and *elaborative rehearsal*, requiring conceptually driven, schema-based processing (Craik and Lockhart, 1972; Hulstijn, 2001).

Schmidt (1990, 1995) has argued that noticing and focal attention are essentially isomorphic. This position is illustrated in figure 19.4. However, unvariegated focal attention alone cannot explain the differential learning consequences

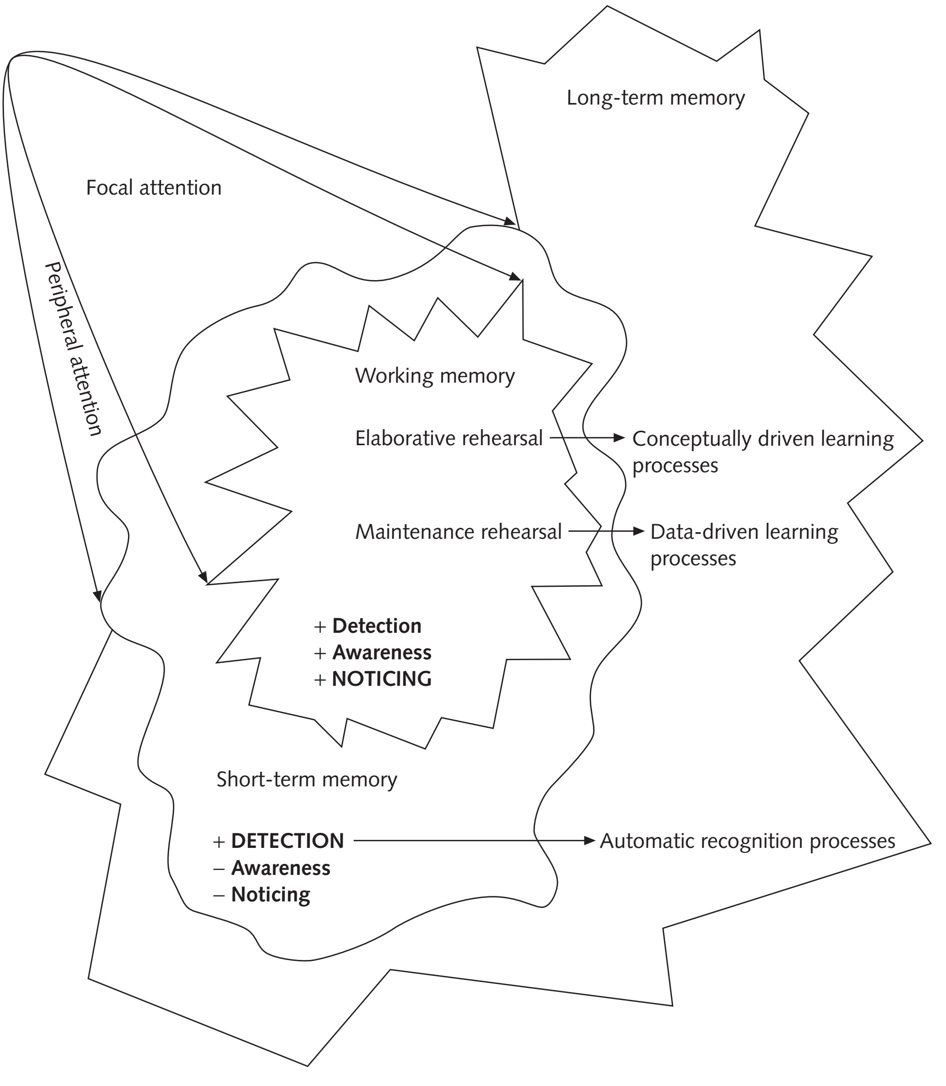


**Figure 19.4** "Noticing" as selective focal attention and rehearsal in working memory: "detection" as recognition outside of awareness in passive short-term memory

of noticing under different conditions of exposure, as revealed in laboratory studies of learning under implicit, incidental, and explicit conditions (see Hulstijn, 1997, for review; DeKeyser, this volume; box 19.1 below). In addition, one must invoke memory processes. I would argue that "noticing" involves that subset of detected information that receives focal attention, enters short-term working memory, and is rehearsed. Noticing and higher levels of awareness, that is, are the result of rehearsal mechanisms (maintenance or elaborative rehearsal) which send (however temporarily) information in short-term memory to long-term memory. It is these rehearsal processes that give rise to awareness, place limits on the extent of awareness, and constrain what can be verbalized during verbal reports. In this regard I have argued (Robinson, 1995b, 1996a, 1997b) that data-driven, instance-based processing and conceptually driven, schema-based processing correspond to those implicit and explicit learning processes that some, in contrast, (Krashen, 1985; Paradis, 1994; Reber, 1989; Schachter, 1987, 1996; Squire, 1992) argue result from neurophysiologically distinct implicit/explicit learning and memory systems.

## *3.4 SLA research into memory, rehearsal, and elaboration*

Williams (1999) addressed the issue of whether inductive SLA could be characterized as a data-driven learning process, requiring maintenance rehearsal of instances and "chunks" in unanalyzed form in working memory, as opposed to a conceptually driven learning process, requiring activation of schemas in long-term memory which are drawn on in elaboratively rehearsing and analyzing the input (see box 19.1 for further details). In a series of three (between-groups) computerized, experimental studies, Williams presented 40 sentences in a previously unknown language (Italian) in a display which illustrated the meaning of the sentences semi-graphically. Subjects both read and heard the sentences. The ability to recall each sentence verbatim following each presentation during training was used as a measure of memory, and was assumed to require predominantly maintenance rehearsal and data-driven

### Box 19.1 Williams (1999)

*Research question*: What is the relationship between verbatim memory for input and inductive learning of aspects of grammar?

The relationship between memory for language input (without awareness of, or intention to search for, grammatical rules) and subsequent induction of grammar has been a central issue in cognitive psychology, and in experimental SLA research throughout the 1990s. In a series of three laboratory experiments, Williams investigated the relationship between verbatim memory for input and inductive learning of aspects of grammar. The target grammar was Italian, a language none of the participants was familiar with. Williams explained the task as an exercise in memorizing sentences.

*Methodology*:
*The verbatim memory task – presentation phase*: Williams presented a semi-graphical display on a computer screen. Participants were asked to say aloud the sentence depicted by the display in English. After this they saw the correct English sentence on the screen. They then heard, and saw on the computer screen, an Italian sentence describing the graphical display. Following this participants heard segments

of the Italian sentences, accompanied by highlighted portions of the relevant aspects of the graphic. They were asked to repeat each segment aloud after they heard it. Finally, they heard and viewed (for 3 seconds) the whole sentence once more and were asked to repeat it.

Williams thus ensured that the meaning of each Italian sentence (a total of 40) had been understood, and had a taped record of the accuracy of recall of segments and the whole sentence.

*The verbatim memory task – recall phase*: After every two sentences presented and memorized following the above procedure, Williams presented the graphical representation of each sentence again with instructions to recall the Italian sentence aloud. Following this, to prompt further accuracy in recall, participants were allowed to view some letters of the Italian *content* words that appeared on the graphical display, and say the sentence aloud again if they wished to change their initial response. No aspects of Italian grammar, such as verb or article noun agreement inflections, were presented on the display during the recall phase.

*Testing learning*: Williams tested the learning that had occurred following the memorization task via a computerized translation test. This involved the presentation of a series of semi-graphical displays. Participants had to click the correct verb form, and noun and article forms, to construct the corresponding Italian sentence.

*Results*: Williams found considerable variance in accuracy of recall at the begining of training (individual differences in memory), for the first eight Italian sentences viewed and heard (block 1) but rather less variance at the end of training, on the last eight sentences viewed and heard (block 5). Williams also found the translation task revealed accurate learning of *some* of the aspects of grammar, such as verb inflections, but not of others, such as article–noun agreement.

Importantly, Williams found that accuracy of recall early during training (on block 1) correlated significantly and positively with accuracy during the translation post-test. In fact, accuracy of early recall correlated much *more* strongly than accuracy of later, block 5 recall. Williams concludes: "It would appear from this result that there is at least some sense in which knowledge of grammatical rules emerges out of memory for input and that individual differences in memory ability that are apparent even in the earliest stages of exposure have consequences for ultimate levels of learning. The results suggest that the learning occurring in this experiment can be characterized as data driven" (p. 22).

*Conclusion*: Three brief comments are worth making, considering the issues raised in this chapter. First, Williams's findings contradict the claims of Krashen (1985) and Reber (1993), reported in section 3.6, that incidental and implicit learning are insensitive to measures of individual differences. Second, the finding that verb inflections were learned more easily than article noun agreement rules is in line with the findings of DeKeyser (1995) and DeGraaff (1997a) that implicit and incidental learning processes interact with the complexity of the learning domain. Finally, Williams found (though not reported in my summary) that techniques for making salient aspects of the targeted structures also interacted with the complexity of the linguistic domain, and led to greater learning of some forms, supporting the claims reported in section 2.3.3 for the value of focus on form.

processing (since learners had no prior conceptual knowledge of the language to draw on in processing the input). Individual differences in the memory-recall task were then examined to see if they related to performance on a transfer translation task which was used as a measure of learning. The transfer translation task also presented a semi-graphical representation of a sentence, and subjects had to select words to make the matching Italian sentences, by clicking on an array of possible words presented on the computer screen.

Williams found strong significant correlations between accurate recall early in training, and performance on the transfer translation task. There were lower positive correlations with recall performance later in training (when learners, particularly those with greater grammatical sensitivity, might have been expected to switch to conceptually driven processing and hypothesis testing), suggesting that the inductive learning that had occurred did draw predominantly on data-driven processing and maintenance rehearsal. In a second experiment Williams introduced a technique to enhance, or make visually salient, aspects of form during presentation of the sentences. There were still significant positive correlations of memory, as measured by early training task recall ability, and learning, as evidenced by accurate performance on the transfer translation task. However, enhancement had the effect of dramatically increasing learning on some aspects of grammar (e.g., knowledge of morphemes for marking article–noun agreement) which had been imperfectly recalled during training. These results suggest that enhancement techniques, for selectively inducing learner attention to form during processing for meaning, are effective for *some* forms, in the short term, but that the learning processes they activate result from a more conceptually driven mode of processing and elaborative rehearsal, in contrast to the data-driven maintenance rehearsal reflected in accurate verbatim recall of sentences presented early in training. Inducing selective attention via enhancement, that is, induces noticing and elaborative rehearsal, resulting in a different pattern of learning outcomes than that which results from noticing and maintenance rehearsal (see box 19.1 for further discussion). However, in a third experiment Williams found that providing feedback on accuracy of recall attempts during training (a more explicit form-focusing technique than visual enhancement, which might be expected to facilitate greater conceptually driven processing, hypothesis testing, and more elaborative rehearsal) did not significantly alter the extent of learning, and led to worse learning on some forms than did the second experiment.

In short, Williams found evidence for a complex interplay of data-driven and conceptually driven processing during inductive second language learning, in which individual differences in working memory for written and aural input largely (and positively) predicted the extent of subsequent learning in all experiments. Compared to the unstructured, memorize-only training condition in the first experiment, inducing selective attention to form via visual enhancement in the second experiment facilitated greater learning of some forms, though explicit feedback about accuracy of recall during training (which might be expected to induce rule search) in the third experiment led to decreases in learning of some forms. These findings are consistent with those for more successful learning under enhanced conditions of exposure than in unenhanced, memorize-only conditions (Robinson, 1997b), and for the negative effects of explicit rule search where forms to be acquired are complex (Robinson, 1997a). Studies such as Williams (1999), motivated by attentional, learning, *and* memory theory, are therefore to be encouraged for the additional insight they provide into the cognitive processes activated by the focus-on-form techniques described earlier in this review.

## 3.5 Short-term and working memory

Passive measures of short-term storage, such as backward digit span, in which subjects hear and repeat, in reverse order, a list of numbers, are distinguished from active measures of short-term storage, such as reading span tests. In reading span tests (Daneman and Carpenter, 1983; Osaka and Osaka, 1992), subjects read sets of sentences aloud from written cue cards in which selected words are underlined; subjects are then instructed to recall the underlined words. These tests measure the extent to which information is actively maintained and periodically refreshed in short-term memory while other processing operations (reading, speaking) take place. While the relationship between them is controversial, both types of test are argued to reflect important memory processes. Baddeley (1986) has proposed a model of working memory to account for the effects of active measures of short-term memory, in which information is maintained via rehearsal in two slave-systems of working memory, the *phonological loop* and the *visuospatial sketchpad*, which are jointly coordinated by a *central executive*. However, figure 19.4 represents working memory as *within* short-term memory (Nairne, 1996), since priming and subliminal exposure experiments such as those of Marcel (1983) show automatic access to long-term memory, and may exhibit primacy and recency effects, but these cannot be attributed to working memory and the focus of attention (see Baddeley, 1986; Cowan, 1993, 1995; N. Ellis, 2001; Gathercole and Baddeley, 1993; Shiffrin, 1993, for discussion).

## 3.6 SLA research into the role of short-term and working memory

A number of studies have examined the relationship between individual differences in short-term, working memory and SLA. Cook (1977) found a closer relationship between performance on passive measures of short-term storage in the L1 and the L2 than on measures of long-term memory, and suggested short-term L1 memory capacity was more transferable to L2 learning and use than long-term L1 memory capacity. Harrington and Sawyer (1992) found active measures of L2 working memory, measured by reading span tests, predicted superior L2 reading skill more than did passive, digit span measures of short-term storage. Similarly, Geva and Ryan (1993) found a closer relationship between L2 proficiency and measures of L2 working memory than between L2 proficiency and passive measures of short-term storage. Miyake, Friedman, and Osaka (1998) also found a close positive relationship between working memory and L2 linguistic knowledge and L2 listening comprehension. Working within the framework of MacWhinney's (1987) competition model (see discussion and examples given in section 2.2 above), they found that Japanese learners of English with high working memory capacity, measured by a listening span version of Daneman and Carpenter's (1983) reading span test, demonstrated more accurate cue assignment strategies (correctly assigning agency to nouns in English sentences on the basis of word order) than did learners with low working memory capacity, who preferred the L1 (Japanese) based strategy of assigning agency on the basis of animacy. High working memory learners also demonstrated greater comprehension of complex sentences read at natural speed.

As might be expected, then, following Schmidt's "noticing" hypothesis, measures of working memory capacity, which affects the extent and efficiency of focal attention allocation, are closely and positively related to second language proficiency and skill development. Miyake and Friedman (1998) argue that for this reason, working memory measures should be included in tests of language learning aptitude. Surprisingly, this is not currently the case, since traditional measures of language learning aptitude, such as the Modern Language Aptitude Test (J. B. Carroll and Sapon, 1959) and tests based on it (Sasaki, 1996), use only rote, passive measures of short-term memory ability (see Robinson, 2001b; Dörnyei and Skehan, this volume). In addition to the *inferential* evidence provided by the correlational studies of working memory capacity and existing levels of L2 proficiency and L2 reading and listening skill reported above, recent experimental studies of second language learning lend more *direct* support to the claim that working memory is an important contributor to second language learning ability, under *some* conditions of exposure. This issue is taken up in the following section.

## 3.7 Implicit and explicit memory: individual differences and child and adult L2 learning

Direct tests of memory (in which subjects are instructed to attend to material presented in a study phase in order to complete a later recall test) are assumed to access explicit memory, whereas indirect tests (in which material is simply presented in a study phase with no instructions to remember the information for a later recall test) are assumed to access implicit memory (see Kelley and Lindsay, 1997; Merickle and Reingold, 1991; Robinson, 1993, 1995b). Separate systems accounts argue that attention and awareness regulate access to (recognition and retrieval from) explicit but not implicit memory, and that explicit and implicit learning encode new information differentially into each. However, this neat equation is problematic. One problem is the fact that implicit memory measures (Marcel's priming experiments are just one example) generally show strong task and modality effects, such that information presented during a study phase on one task, in one modality (written words), may often not be recognized or recalled when tested later in another task, in another modality (aurally presented words). However, when implicit memory study/test tasks are similar and modality is the same, recognition and recall are much higher (Buchner and Wippich, 1998). In contrast, implicit learning experiments (see Reber, 1989, 1993) have been argued to show robust *generalizability* of learned information across different study/test tasks, and across modalities (though many disagree; see DeKeyser, this volume; Shanks and St John, 1994).

These issues have interesting implications for Universal Grammar (UG) explanations of SLA (S. Carroll, 1999; Cook, 1994; Gregg, this volume; Krashen, 1985; Schwartz, 1999; Truscott, 1998; White, this volume). Some UG explanations argue innate representations of the shape of possible languages persist into adulthood and so pre-exist adult exposure to L2 input, and that "full access" (Schwartz, 1999) to these representations can be triggered by exposure to positive evidence of the L2 alone, during processing for meaning, with no conscious attention to form. In this case, adult access to UG *may* be interpreted as implicit memory for existing knowledge or unconscious "acquisition" (Krashen, 1985), and dissociable from the learning that draws on consciously accessed explicit memory (Paradis, 1994). If so, the available evidence suggests modal specificity, so that positive evidence presented in one modality (reading) will have limited transfer to another (listening). This, of course, would be consistent with *child* L1 or L2 language development (before the age of 6 and the onset of maturational constraints; see Long, 1990), which is aural/oral modality dependent. However, it poses problems for UG accounts of *adult* L2 development (Krashen, 1985, 1994; Schwartz, 1999) that claim positive evidence obtained via *reading*, with no conscious attention to or noticing of form, can trigger access (via implicit memory) to representations which subsequently promote grammatical development in other modalities, such as *speaking*. The interesting possibility raised by studies of implicit memory, then, is that if separate systems are proposed in a transition theory, "full access" to UG during adulthood may be modality specific.

Some non-UG explanations of SLA assume that access to innate representations and cognitive mechanisms available in childhood is attenuated or not possible for adults, and that child and adult L2 learning are fundamentally different (see Bley-Vroman, 1988; Long, 1990; Skehan, 1998). Compatible with this assumption, effects for adult implicit learning are likely attributable not to separate systems, but to a preponderance of data-driven processing, maintenance rehearsal, and instance learning of noticed information (accounting for the difficulty of verbalizing the contents of awareness during implicit learning), whereas adult explicit learning likely results from a preponderance of conceptually driven processing, and elaborative rehearsal of noticed information (see Doughty, this volume, for another view). The preponderance of one or the other is largely a consequence of the way the study tasks force the material to be processed. Evidence for implicit memory and learning reflects study/test overlaps in data-driven processing, and evidence for explicit memory and learning reflects study/test overlaps in conceptually driven processing – different systems are not involved (cf. Healy et al., 1992; Jacoby and Dallas, 1981; Roediger, Weldon, and Challis, 1989; Shanks and St John, 1994, for similar proposals). To this extent, implicit and explicit learning in adulthood are fundamentally similar, requiring focal attention and rehearsal of input in memory, and are both sensitive to individual differences in relevant cognitive capacities. Differences in consciously attended task demands, together with individual differences in relevant cognitive variables, such as working memory capacity or speed (N. Ellis, 1996, 2001; N. Ellis, Lee, and Reber, forthcoming; Gathercole and Thorn, 1998; Harrington and Sawyer, 1992; Miyake and Friedman, 1998; Robinson, 2002), or language learning aptitude (Robinson, 1997a, 2001b; Sasaki, 1996; Sawyer and Ranta, 2001; Skehan, 1998), cause differences in learning outcomes. However, those who adopt separate-systems explanations of implicit/explicit learning argue implicit learning will be unaffected by individual differences, and be much more homogeneous across populations of learners than explicit learning, where individual differences in cognitive abilities *are* predicted to play a significant role in determining the extent of learning (Krashen, 1985; Reber, 1993; Reber, Walkenfield, and Hernstadt, 1991; Zobl, 1992, 1995). These and related issues have begun to be addressed in recent SLA research.

## *3.8 SLA research into the role of memory and individual differences*

Reber (1989; Reber et al., 1991) and Krashen (1985; cf. similar proposals by Zobl, 1992, 1995) have argued that individual differences in cognitive abilities will affect consciously regulated explicit learning, but not unconscious implicit or incidental learning. In support of this claim, Reber et al. showed nonsignificant correlations of intelligence (using the Wechsler adult intelligence scale) with implicit learning of an artificial grammar, but significant positive correlations of intelligence and explicit learning during a forced choice series-solution task. Similarly, Robinson (1996a, 1997a) found that learning in an incidental, process-for-meaning condition alone was unrelated to measures of aptitude (the MLAT measures of rote memory for paired associates and grammatical sensitivity), in contrast to learning under explicit instructed and rule-search conditions. This finding appears to support Krashen's claims for the aptitude independence of supposedly unconscious acquisition processes activated by incidental learning conditions, and is in line with Reber et al's. findings that unaware, implicit learning is insensitive to measures of individual differences. Robinson (2002), however, argued that this may have been because the measure of paired-associates rote memory used in the earlier study did not reflect the active nature of the processing demands of the incidental learning task – processing sentences for meaning while incidentally noticing grammatical information.

In contrast to the earlier finding, in an extended replication of Reber et al. (1991), Robinson showed that incidental learning of a previously unknown L2 (Samoan) by Japanese learners (exposed during training to 450 sentence strings, which were processed in order to answer comprehension questions over a period of three hours; training took place on two separate days, one to three days apart) did correlate positively and significantly with measures of L1 working memory capacity (Osaka and Osaka's, 1992, reading span test). This was so for one-week and six-month delayed (not immediate) post-tests using grammaticality judgment measures of responses to novel Samoan sentences not encountered during training, as well as post-test measures of production of sentences (a word ordering test of sentence construction). However, in contrast to the findings of Reber et al. (1991), in a repeated-measures design using the same implicit and explicit training conditions operationalized in Reber et al., there were significant *negative* correlations of intelligence and implicit artificial grammar learning, as well as significant positive correlations of intelligence and explicit learning of the series-solution task. Learning during the implicit and explicit tasks, however, unlike incidental L2 learning, did not correlate positively and significantly with working memory. Once again, as in Robinson (1996a, 1997a), language learning aptitude, as measured by Sasaki's (1996) Language Aptitude Battery for the Japanese (based on J. B. Carroll and Sapon's, 1959, MLAT), did not correlate significantly with learning in the incidental L2 learning condition on the immediate post-test, and even correlated significantly and negatively with learning of some of the complex rules present in the training sentences. However, aptitude did positively predict six-month delayed post-test performance, but only as measured by the ability to produce sentences on the word-ordering task.

This study suggests, then, that incidental L2 learning, contrary to arguments put forward by Krashen (1985) and also by Zobl (1992, 1995), *is* sensitive to measures of individual differences in cognitive abilities, but that the measures used must correspond with, and be sensitive to, the processing demands of the particular training condition under which exposure takes place, and, importantly, that individual differences will be most likely to show delayed (not immediate) effects on incidental learning. To this extent, current L2 aptitude measures, as Miyake and Friedman (1998) suggest, and as this study shows, may appear to lack treatment validity, if perfomance on immediate post-tests is the measure of learning. The results of delayed post-tests, however, show aptitude to be a predictor of learning, but only when the measure is a productive one. Further, while the results of the implicit and explicit learning experiments partially replicate the findings of Reber et al. (showing intelligence to be positively related to explicit learning), the study demonstrates (what has often been asserted though not directly shown) that claims based on evidence of implicit learning of artificial grammars cannot be validly generalized to incidental L2 learning. Incidental L2 learning shows a different pattern of correlations with individual difference measures than learning in the other two conditions.

# 4 Conclusion

If this chapter has been relatively long, and inconclusive, in part it is because discussion of the role of cognitive variables in transition theories of SLA is often short and conclusive – learning is "triggered" (somehow) by input; implicit learning (or "acquisition") happens automatically, outside of awareness, and is insensitive to individual differences in cognitive capacities. What, then, could there be to discuss in a chapter on the role of attention and memory in SLA, except their irrelevance?

I have argued, however, that current views of attentional resources, and the contribution of individual differences in memory processes and capacities, are underrepresented, little explored, and sometimes misconceptualized in SLA research that has referred to them. Much recent attentional theory questions the oft-invoked SLA notion of "capacity" constraints. Why are there capacity constraints on attention? If these are general and undifferentiated across task demands, why then have multiple-resource theories been able to predict successful and unsuccessful competition for, and time-sharing between, attentional resources as a function of different task demands? Clearly memory "structures" are capacity limited, and impose constraints on attentional processes, but what learning processes do these structures and constraints give rise to: implicit/explicit learning, or different kinds of attentionally regulated rehearsal during processing, which appear to correspond to different learning systems? These and other questions raised (but not answered) in this review will hopefully prompt further research into cognitive processes and the role of cognitive variables during SLA, adding to, and refining, the findings which have begun to accumulate in the field, and resonating with research findings that have accumulated in other related fields of psychological inquiry and learning theory.

## ACKNOWLEDGMENT

Preparation of this chapter was made possible by support provided by the Aoyama Gakuin Soken Research Institute.

## REFERENCES

Alanen, R. 1995: Input enhancement and rule presentation in second language acquisition. In R. Schmidt (ed.), *Attention and Awareness in Foreign Language Learning and Teaching*. Second Language Teaching and Curriculum Center Technical Report No. 9. Honolulu: University of Hawai'i Press, 259–302.

Allport, D. A. 1987: Selection for action: some behavioral and neurophysiological consequences of attention and action. In H. Heuer and A. Sanders (eds), *Perspectives on Perception and Action*. Hillsdale, NJ: Lawrence Erlbaum Associates, 395–419.

Allport, D. A. 1989: Visual attention. In M. Posner (ed.), *Foundations of Cognitive Science*. Cambridge, MA: MIT Press.

Atkinson, R. and Shiffrin, R. 1968: Human memory: a proposed system and its control processes. In K. Spence (ed.), *Psychology of Learning and Motivation: Advances in Research and Theory. Vol. 2*. New York: Academic Press, 89–195.

Baddeley, A. 1986: *Working Memory*. Oxford: Oxford University Press.

Barry, S. and Lazarte, A. A. 1998: Evidence for mental models: how do prior knowledge, syntactic complexity, and reading topic affect inference generation in a recall task for nonnative readers of Spanish? *Modern Language Journal*, 82, 176–93.

Bialystok, E. (ed.) 1991: *Language Processing in Bilingual Children*. Cambridge: Cambridge University Press.

Bley-Vroman, R. 1988: The fundamental character of foreign language learning. In W. Rutherford and W. Sharwood-Smith (eds), *Grammar and Second Language Teaching*. New York: Newbury House, 19–30.

Boronat, C. and Logan, G. 1997: The role of attention in automatization: does attention operate at encoding or retrieval or both? *Memory and Cognition*, 25, 36–46.

Broadbent, D. 1958: *Perception and Communication*. London: Pergamon Press.

Broadbent, D. 1971: *Decision and Stress*. New York: Academic Press.

Broeder, P. and Plunkett, K. 1994: Connectionism and second language acquisition. In N. Ellis (ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press, 421–54.

Brown, G. 1995: *Speakers, Listeners and Communication*. Cambridge: Cambridge University Press.

Brown, G., Anderson, A., Shillcock, R., and Yule, G. 1984: *Teaching Talk: Strategies for Production and Assessment*. Cambridge: Cambridge University Press.

Buchner, A. and Wippich, W. 1998: Differences and commonalities between implicit learning and implicit memory. In M. Stadler and P. Frensch (eds), *Handbook of Implicit Learning*. Thousand Oaks, CA: Sage, 3–46.

Butterfield, E. and Albertson, L. 1995: On making cognitive theory more general and developmentally pertinent. In F. Weinert and W. Schneider (eds), *Research on Memory Development*. Hillsdale, NJ: Lawrence Erlbaum Associates, 73–99.

Bygate, M. 1996: Effects of task repetition: appraising the developing language of learners. In J. Willis and M. Willis (eds), *Challenge and Change in Language Teaching*. Oxford: Heinemann, 136–46.

Bygate, M. 1999: Task as a context for the framing, reframing and unframing of language. *System*, 27, 33–48.

Carrell, P. 1987: Content and formal schemata in ESL pedagogy. *TESOL Quarterly*, 21, 461–81.

Carroll, J. B. and Sapon, S. 1959: *The Modern Language Aptitude Test*. San Antonio, TX: Psychological Corporation.

Carroll, S. 1999: Putting "input" in its proper place. *Second Language Research*, 15, 337–88.

Carroll, S. and Swain, M. 1993: Explicit and implicit negative feedback. *Studies in Second Language Acquisition*, 15, 357–86.

Carter, R. 1998: *Mapping the Mind*. London: Weidenfeld and Nicolson.
Case, R. 1985: *Intellectual Development: Birth to Adulthood*. Norwood, NJ: Ablex.
Case, R. 1992: *The Mind's Staircase: Exploring the Conceptual Underpinnings of Children's Thought and Knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Chang, Y. F. 1999: Discourse topics and interlanguage variation. In P. Robinson (ed.), *Representation and Process: Proceedings of the 3rd Pacific Second Language Research Forum. Vol. 1*. Tokyo: PacSLRF, 235–41.
Clapham, C. 1996: *The Development of IELTS: A Study of the Effect of Background Knowledge on Reading Comprehension*. Cambridge: Cambridge University Press.
Clark, A. 1997: *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
Cook, V. 1977: Cognitive processes in second language learning. *International Review of Applied Linguistics*, 15, 1–20.
Cook, V. 1994: The metaphor of access to Universal Grammar in L2 acquisition. In N. Ellis (ed.), *Implicit and Explicit Learning of Languages*. New York: Academic Press, 477–502.
Corder, S. P. 1967: The significance of learners' errors. *IRAL*, 5, 161–70.
Cowan, N. 1988: Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, 104, 163–91.
Cowan, N. 1993: Activation, attention and short-term memory. *Memory and Cognition*, 21, 162–7.
Cowan, N. 1995: *Attention and Memory: An Integrated Framework*. Oxford: Oxford University Press.
Craik, F. and Lockhart, R. 1972: Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–84.
Crookes, G. 1986: *Task Classification: A Cross-Disciplinary Review*. Technical Report No. 4, Center for Second Language Research, Social Science Research Institute, University of Hawai'i at Manoa.
Crookes, G. 1988: *Planning, Monitoring and Second Language Development*. Technical Report No. 6, Center for Second Language Research, Social Science Research Institute, University of Hawai'i at Manoa.
Crookes, G. 1989: Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11, 183–99.
Crutcher, R. 1998: The role of prior knowledge in mediating foreign language vocabulary acquisition and retention: a process analytic approach. In A. Healy and L. Bourne Jr (eds), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention*. Mahwah, NJ: Lawrence Erlbaum Associates, 91–112.
Daneman, M. and Carpenter, P. 1983: Individual differences in integrating information within and between sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9, 561–83.
De Bot, K. 1996: The psycholinguistics of the output hypothesis. *Language Learning*, 46, 529–55.
DeGraaff, R. 1997a: The eXperanto experiment: effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19, 249–97.
DeGraaff, R. 1997b: *Differential Effects of Explicit Instruction on Second Language Acquisition*. HIL Dissertations No. 35. The Hague: Holland Institute of Generative Linguistics.
DeKeyser, R. 1995: Learning second language grammar rules: an experiment with a miniature linguistic

system. *Studies in Second Language Acquisition*, 17, 379–410.

DeKeyser, R. 1997: Beyond explicit rule learning: automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19, 195–224.

DeKeyser, R. 2001: Automaticity and automatization. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 125–51.

Dörnyei, Z. 1998: Motivation in foreign and second language learning. *Language Teaching*, 31, 117–35.

Dörnyei, Z. 2002: The motivational basis of language learning tasks. In P. Robinson (ed.), *Individual Differences and Instructed Language Learning*. Amsterdam: John Benjamins, 137–58.

Doughty, C. 1991: Second language instruction does make a difference: evidence from an empirical study of relativization. *Studies in Second Language Acquisition*, 13, 431–69.

Doughty, C. 1998: Acquiring competence in a second language: form and function. In H. Byrnes (ed.), *Learning Foreign and Second Languages*. New York: MLA, 128–56.

Doughty, C. 2001: Cognitive underpinnings of focus on form. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 206–57.

Doughty, C. and Varela, E. 1998: Communicative focus on form. In C. Doughty and J. Williams (eds), *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press, 114–38.

Doughty, C. and Williams, J. 1998: Pedagogical choices in focus on form. In C. Doughty and J. Williams (eds), *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press, 197–262.

Duff, P. 1986: Another look at interlanguage talk: taking task to task. In R. Day (ed.), *Talking to Learn: Conversation in Second Language Development*. Rowley, MA: Newbury House, 147–81.

Dunkel, P. 1991: Listening in the native and second/foreign language: towards an integration of research and practice. *TESOL Quarterly*, 25, 431–57.

Eckman, F., Bell, L., and Nelson, D. 1988: On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics*, 9, 1–20.

Ellis, N. 1993: Rules and instances in foreign language learning: interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, 5, 289–318.

Ellis, N. 1994: Implicit and explicit learning: an overview. In N. Ellis (ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press, 1–16.

Ellis, N. 1996: Sequencing and SLA: phonological memory, chunking and points of order. *Studies in Second Language Acquisition*, 18, 91–126.

Ellis, N. 2001: Memory for language. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 33–68.

Ellis, N. and LaPorte, N. 1997: Contexts of acquisition: effects of formal instruction and naturalistic exposure on second language acquisition. In A. De Groot and J. Kroll (eds), *Tutorials in Bilingualism: Psycholinguistic Perspectives*. Hillsdale, NJ: Lawrence Erlbaum Associates, 53–83.

Ellis, N. and Schmidt, R. 1997: Morphology and longer distance dependencies: laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, 19, 145–71.

Ellis, N., Lee, M., and Reber, A. forthcoming: Phonological working memory in artificial language acquisition. MS.

Ellis, R. 1987: Interlanguage variability in narrative discourse: style shifting and use of the past tense. *Studies in Second Language Acquisition*, 9, 1–20.

Elman, J. 1990: Finding structure in time. *Cognitive Science*, 14, 179–212.

Eslinger, P. 1996: Conceptualizing, describing and measuring components of executive function. In G. Reid Lyon and N. A. Krasnegor (eds), *Attention and Executive Function*. Baltimore, MD: Paul H. Brookes, 395–421.

Estes, W. 1992: Mental psychophysics of categorization and decision making. In H. Geissler, S. Link, and J. Townsend (eds), *Cognition, Information Processing and Psychophysics: Basic Issues*. Hillsdale, NJ: Lawrence Erlbaum Associates, 123–41.

Faerch, C. and Kasper, G. 1987: From product to process – introspective methods in second language research. In C. Faerch and G. Kasper (eds), *Introspection in Second Language Research*. Clevedon: Multilingual Matters, 5–24.

Faust, M. and Gernsbacher, M. 1996: Cerebral mechanisms for suppression of inappropriate information during sentence comprehension. *Brain and Language*, 53, 234–59.

Flege, J. and Munro, M. 1994: The word as a unit in second language speech perception and production. *Studies in Second Language Acquisition*, 16, 381–411.

Fleishman, E. and Quaintance, M. 1984: *Taxonomies of Human Performance: The Description of Human Tasks*. New York: Academic Press.

Fodor, J. 1983: *The Modularity of Mind*. Cambridge, MA: MIT Press.

Foster, P. and Skehan, P. 1996: The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299–324.

Fotos, S. 1993: Consciousness raising and noticing through focus on form: grammar task performance versus formal instruction. *Applied Linguistics*, 14, 385–407.

Gass, S. 1982: From theory to practice. In M. Hines and W. Rutherford (eds), *On TESOL '81*. Washington, DC: TESOL, 129–39.

Gass, S. 1988: Integrating research areas: a framework for second language studies. *Applied Linguistics*, 9, 198–217.

Gass, S. 1997: *Input, Interaction, and the Second Language Learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. 1999: Incidental learning: a commentary. *Studies in Second Language Acquisition*, 21, 319–34.

Gasser, M. 1990: Connectionism and universals of second language acquisition. *Studies in Second Language Acquisition*, 12, 179–99.

Gasser, M. and Smith, L. 1998: Learning nouns and adjectives: a connectionist account. In K. Plunkett (ed.), *Language Acquisition and Connectionism*. Hove: Psychology Press, 269–306.

Gathercole, S. and Baddeley, A. 1993: *Working Memory and Language*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gathercole, S. and Thorn, A. 1998: Phonological short-term memory and foreign language learning. In A. Healy and L. Bourne Jr (eds), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention*. Mahwah, NJ: Lawrence Erlbaum Associates, 141–60.

Geva, E. and Ryan, E. 1993: Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning*, 43, 5–42.

Gopher, D. 1992: Analysis and measurement of mental workload. In G. d'Ydewalle, P. Eelen, and P. Bertelson (eds), *International Perspectives on Psychological Science. Vol. 2. State of the Art*. Hillsdale, NJ: Lawrence Erlbaum Associates, 265–91.

Gregg, K. 2001: Learnability and second language acquisition theory. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 152–82.

Halford, G. 1993: *Children's Understanding: The Development of Mental Models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Harley, B. 1993: Instructional strategies and SLA in early French immersion. *Studies in Second Language Acquisition*, 15, 245–60.

Harley, B. and Swain, M. 1984: The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper, and A. Howatt (eds), *Interlanguage: Studies in Honor of S. Pit Corder*. Edinburgh: Edinburgh University Press, 219–311.

Harrington, M. 2001: Sentence processing. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 91–124.

Harrington, M. and Sawyer, M. 1992: L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25–38.

Hazelhurst, B. and Hutchins, E. 1998: The emergence of propositions from the coordination of talk and action in a shared world. In K. Plunkett (ed.), *Language Acquisition and Connectionism*. Hove: Psychology Press, 373–424.

Healy, A., Fendrich, D., Crutcher, R., Wittman, W., Gesi, A., Ericsson, K., and Bourne, L. Jr 1992: The long term retention of skills. In A. Healy, S. Kosslyn, and R. Shiffrin (eds), *From Learning Processes to Cognitive Processes: Essays in Honor of William Estes. Vol. 2*. Thousand Oaks, CA: Sage, 1–29.

Heuer, H. 1996: Dual-task performance. In O. Neumann and A. Sanders (eds), *Handbook of Perception and Action. Vol. 3: Attention*. New York: Academic Press, 113–48.

Hirst, W. and Kalmar, D. 1987: Characterizing attentional resources. *Journal of Experimental Psychology: General*, 116, 68–81.

Holding, D. (ed.) 1989: *Human Skills*. New York: John Wiley.

Holender, D. 1986: Semantic activation without conscious identification in dichotic listening tasks, parafoveal vision and visual masking: a survey and appraisal. *Behavioral and Brain Sciences*, 9, 1–66.

Holthouse, J. 1995: Anxiety and second language learning task type. M.A. dissertation. University of Queensland.

Horwitz, E. K., Horwitz, M. B., and Cope, J. 1986: Foreign language classroom anxiety. *Modern Language Journal*, 70, 125–32.

Huckin, T. and Coady, J. 1999: Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 21, 181–94.

Hulstijn, J. 1989: A cognitive view on interlanguage variability. In M. Eisenstein (ed.), *The Dynamic Interlanguage*. New York: Plenum, 17–32.

Hulstijn, J. 1997: Second language acquisition research in the laboratory: possibilities and limitations. *Studies in Second Language Acquisition*, 19, 131–44.

Hulstijn, J. 2001: Intentional and incidental second language learning: a reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 258–86.

Hulstijn, J. and DeGraaff, R. 1994: Under what conditions does explicit knowledge facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review*, 11, 97–112.

Iwashita, N. 1999: The role of task-based conversation in the acquisition of Japanese grammar and vocabulary. Ph.D. dissertation. University of Melbourne.

Izumi, S., Bigelow, M., Fujiwara, M., and Fearnow, S. 1999: Testing the output hypothesis: effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21, 421–52.

Jackendoff, R. 1997: *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Jacob, A. 1996: Anxiety and motivation during second language task performance in Singaporean schools. M.A. dissertation. National University of Singapore/RELC.

Jacoby, L. and Dallas, M. 1981: On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 3, 306–40.

Jourdenais, R. 2001: Cognition, instruction, and protocol analysis. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 354–75.

Kahneman, D. 1973: *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Karmilloff-Smith, A. 1992: *Beyond Modularity*. Cambridge, MA: MIT Press.

Kasper, G. 1999: Self report data in pragmatics research. In N. O. Jungheim and P. Robinson (eds), *Pragmatics and Pedagogy: Proceedings of the 3rd Pacific Second Language Research Forum. Vol. 2*. Tokyo: PacSLRF, 1–15.

Kelley, C. and Lindsay, D. S. 1997: Conscious and unconscious forms of memory. In E. L. Bjork and R. Bjork (eds), *Memory*. San Diego: Academic Press, 31–63.

Kim, H. 1995: Intake from the speech stream: speech elements that learners attend to. In R. Schmidt (ed.), *Attention and Awareness in Foreign Language Learning and Teaching*. Second Language Teaching and Curriculum Center Technical Report No. 9. Honolulu: University of Hawai'i Press, 65–84.

Koelega, H. 1996: Sustained attention. In O. Neumann and A. Sanders (eds), *Handbook of Perception and Action. Vol. 3: Attention*. San Diego: Academic Press, 277–322.

Kormos, J. 1999: Monitoring and self repair in L2. *Language Learning*, 49, 303–42.

Korteling, J. 1994: *Multiple Task Performance and Aging*. Groningen: Bariet, Ruinen.

Krashen, S. 1985: *The Input Hypothesis: Issues and Implications*. Oxford: Pergamon Press.

Krashen, S. 1994: The input hypothesis and its rivals. In N. Ellis (ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press, 45–78.

Leeman, J., Arteagoitia, I., Fridman, B., and Doughty, C. 1995: Integrating attention to form with meaning: focus on form in content-based Spanish instruction. In R. Schmidt (ed.), *Attention and Awareness in Foreign Language Learning and Teaching*. Second Language Teaching and Curriculum Center Technical Report No. 9. Honolulu: University of Hawai'i Press, 217–58.

Leow, R. 1993: To simplify or not to simplify: a look at intake. *Studies in Second Language Acquisition*, 15, 333–55.

Leow, R. 1997: Attention, awareness and foreign language behavior. *Language Learning*, 47, 467–505.

Leow, R. 2000: A study of the role of awareness in foreign language behavior: aware vs. unaware learners. *Studies in Second Language Acquisition*, 22, 557–84.

Levelt, W. J. M. 1989: *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Liberman, A. and Mattingly, I. 1989: A specialization for speech perception. *Science*, 243, 489–94.

Logan, G. D. 1988: Toward an instance theory of automatization. *Psychological Review*, 95, 492–527.

Logan, G. D. 1990: Repetition priming and automaticity: common underlying mechanisms? *Cognitive Psychology*, 22, 1–35.
Logan, G. D. 1992: Attention and memory. In L. Squire (ed.), *Encyclopedia of Learning and Memory*. New York: Macmillan, 63–7.
Long, M. H. 1989: Task, group, and task–group interactions. *University of Hawai'i Working Papers in ESL*, 8, 1–25.
Long, M. H. 1990: Maturational constraints on language development. *Studies in Second Language Acquisition*, 12, 251–85.
Long, M. H. 1996: The role of the linguistic environment in second language acquisition. In W. Ritchie and T. Bhatia (eds), *Handbook of Research on Second Language Acquisition*. New York: Academic Press, 413–68.
Long, M. H. 1998: Focus on form in task-based language teaching. *University of Hawai'i Working Papers in ESL*, 161, 49–61.
Long, M. H. forthcoming: *Task Based Language Teaching*. Oxford: Blackwell.
Long, M. H. and Crookes, G. 1992: Three approaches to task-based syllabus design. *TESOL Quarterly*, 26, 55–98.
Long, M. H. and Norris, J. 2000: Task based language teaching. In M. Byram (ed.), *Encyclopedia of Language Education*. London: Routledge, 597–603.
Long, M. H. and Robinson, P. 1998: Focus on form: theory, research, practice. In C. Doughty and J. Williams (eds), *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press, 15–41.
Lyster, R. 1998: Recasts, repetition and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51–81.
Lyster, R. and Ranta, L. 1997: Corrective feedback and learner uptake: negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37–66.
Mackey, A., Gass, S. M., and McDonough, K. 2000: Do learners recognize implicit negative feedback as feedback? *Studies in Second Language Acquisition*, 22, 471–97.
Mackey, A., Philp, J., Eri, T., Fuji, A., and Takaguchi, T. 2002: The outcomes of implicit feedback in conversational interaction: an exploration of the role of aptitude in phonological short-term memory. In P. Robinson (ed.), *Individual Differences and Instructed Language Learning*. Amsterdam: John Benjamins, 181–210.
MacWhinney, B. 1987: The competition model. In B. MacWhinney (ed.), *Mechanisms of Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates, 249–308.
MacWhinney, B. 2001: The competition model: the input, the context and the brain. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 69–90.
Marcel, A. 1983: Conscious and unconscious perception: experiments on visual masking and word recognition. *Cognitive Psychology*, 15, 197–237.
Massaro, D. 1987: *Speech Perception by Ear and by Eye: A Paradigm for Psychological Inquiry*. Hillsdale, NJ: Lawrence Erlbaum Associates.
McLaughlin, B. 1990: Restructuring. *Applied Linguistics*, 11, 113–28.
Mehnert, U. 1998: The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 83–108.
Merickle, P. and Reingold, E. 1991: Comparing direct (explicit) and indirect (implicit) measures to study unconscious memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 224–33.

Miyake, A. and Friedman, N. 1998: Individual differences in second language proficiency: working memory as language aptitude. In A. Healy and L. Bourne Jr (eds), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention*. Mahwah, NJ: Lawrence Erlbaum Associates, 339–64.

Miyake, A., Friedman, N., and Osaka, M. 1998: Cue acquisition and syntactic comprehension in second language learning. Findings reported in A. Miyake and N. Friedman 1998: Individual differences in second language proficiency: working memory as language aptitude. In A. Healy and L. Bourne Jr (eds), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention*. Mahwah, NJ: Lawrence Erlbaum Associates, 339–64.

Mulligan, N. 1997: Attention and implicit memory tests: the effects of varying attentional load on conceptual priming. *Memory & Cognition*, 25, 11–17.

Muranoi, H. 1996: Effects of interaction enhancement on restructuring of interlanguage grammar: a cognitive approach to foreign language instruction. Ph.D. dissertation. Georgetown University, Department of Linguistics.

Muranoi, H. 2000: Focus on form through interaction enhancement: integrating formal instruction into a communicative task in EFL classrooms. *Language Learning*, 50, 617–73.

Naatanen, R. 1992: *Attention and Brain Function*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nairne, J. 1996: Short-term/working memory. In E. L. Bjork and R. Bjork (eds), *Memory*. San Diego: Academic Press, 102–30.

Navon, D. 1989: The importance of being visible: on the role of attention in a mind viewed as an anarchic intelligence system. *European Journal of Cognitive Psychology*, 1, 191–238.

Neely, J. 1977: Semantic priming and retrieval from lexical memory: roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, 106, 226–54.

Nelson, K. 1996: *Language in Cognitive Development: The Emergence of the Mediated Mind*. New York: Cambridge University Press.

Neumann, O. 1987: Beyond capacity: a functional view of attention. In H. Heuer and A. Sanders (eds), *Perspectives on Perception and Action*. Berlin: Springer, 361–94.

Neumann, O. 1996: Theories of attention. In O. Neumann and A. Sanders (eds), *Handbook of Perception and Action. Vol. 3: Attention*. San Diego: Academic Press, 389–446.

Nissen, M. and Bullemer, P. 1987: Attentional requirements of learning: evidence from performance systems. *Cognitive Psychology*, 19, 1–32.

Niwa, Y. 1999: Reasoning demands of L2 tasks and L2 narrative production: effects of individual differences in working memory, intelligence, and aptitude. M.A. dissertation. Aoyama Gakuin University, Department of English.

Norris, J. and Ortega, L. 2000: Effectiveness of L2 instruction: a research synthesis and meta-analysis. *Language Learning*, 50, 417–528.

Norris, J., Brown, J. D., Hudson, T., and Yoshioka, J. 1998: *Developing Second Language Performance Tests*. University of Hawai'i Second Language Teaching and Curriculum Center Technical Report No. 19. Honolulu: University of Hawai'i Press.

Nosofsky, R. 1992: Exemplar-based approach to relating categorization, identification and recognition. In F. G. Ashby (ed.), *Multidimensional Models of*

*Perception and Categorization*. Hillsdale, NJ: Lawrence Erlbaum Associates, 363–95.

Nosofsky, R., Kruschke, J., and McKinley, S. 1992: Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning Memory and Cognition*, 15, 282–304.

Oliver, R. 1995: Negative feedback in child NS/NNS conversation. *Studies in Second Language Acquisition*, 17, 459–83.

Ortega, L. 1999: Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109–48.

Osaka, M. and Osaka, N. 1992: Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 30, 287–9.

Paradis, M. 1994: Neurolinguistic aspects of implicit and explicit memory: implications for bilingualism. In N. Ellis (ed.), *Implicit and Explicit Learning of Languages*. New York: Academic Press, 393–419.

Pashler, H. and Carrier, M. 1997: Structures, processes, and the flow of information. In E. L. Bjork and R. Bjork (eds), *Memory*. San Diego: Academic Press, 3–29.

Phaf, R., van der Heijden, A., and Hudson, P. 1990: SLAM: a connectionist model for attention in visual selection tasks. *Cognitive Psychology*, 22, 273–341.

Pica, T. 1988: Interlanguage adjustments as an outcome of NS–NNS negotiated interaction. *Language Learning*, 38, 45–73.

Pica, T. 1992: The textual outcomes of NS–NNS negotiation: what do they reveal about learning? In C. Kramsch and S. McConnel-Ginet (eds), *Text and Context*. Cambridge, MA: Heath, 198–237.

Pica, T., Kanagy, R., and Falodun, J. 1993: Choosing and using communication tasks for second language teaching and research. In G. Crookes and S. Gass (eds), *Tasks in Language Learning: Integrating Theory and Practice*. Clevedon: Multilingual Matters, 1–42.

Posner, M. 1980: Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.

Posner, M. and Petersen, S. 1990: The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25–42.

Prinz, W. 1986: Continuous selection. *Psychological Research*, 48, 231–8.

Rahimpour, M. 1997: Task condition, task complexity and variation in oral L2 discourse. Ph.D. dissertation. University of Queensland.

Rahimpour, M. 1999: Task complexity and variation in interlanguage. In N. Jungheim and P. Robinson (eds), *Pragmatics and Pedagogy: Proceedings of the 3rd Pacific Second Language Research Forum. Vol. 2*. Tokyo: PacSLRF, 115–34.

Reber, A. 1989: Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219–35.

Reber, A. 1993: *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. Oxford: Oxford University Press.

Reber, A., Walkenfield, F., and Hernstadt, R. 1991: Implicit and explicit learning: individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 888–96.

Reed, E. 1996: *Encountering the World: Toward an Ecological Psychology*. New York: Oxford University Press.

Richardson-Klavehn, A. and Bjork, R. 1988: Measures of memory. *Annual Review of Psychology*, 39, 475–543.

Robinson, P. 1993: Problems of knowledge and the implicit/explicit distinction in SLA theory. *University of*

*Hawai'i Working Papers in ESL*, 12, 99–146.
Robinson, P. 1995a: Task complexity and second language narrative discourse. *Language Learning*, 45, 99–140.
Robinson, P. 1995b: Attention, memory and the "noticing" hypothesis. *Language Learning*, 45, 283–331.
Robinson, P. 1996a: Learning simple and complex second language rules under implicit, incidental, rule search and instructed conditions. *Studies in Second Language Acquisition*, 18, 27–67.
Robinson, P. 1996b: *Consciousness, Rules and Instructed Second Language Acquisition*. New York: Lang.
Robinson, P. 1996c: Connecting tasks, cognition and syllabus design. In P. Robinson (ed.), *Task Complexity and Second Language Syllabus Design: Data-Based Studies and Speculations*. *University of Queensland Working Papers in Applied Linguistics*, 1, 1–16.
Robinson, P. 1997a: Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47, 45–99.
Robinson, P. 1997b: Generalizability and automaticity of second language learning under implicit, incidental, enhanced and instructed conditions. *Studies in Second Language Acquisition*, 19, 223–47.
Robinson, P. 2000: The cognition hypothesis of task-based language development and its implications. In *Proceedings of the 3rd Acquisition of Japanese as a Second Language Conference, Ochanomizu University, Tokyo, December*, 10–15.
Robinson, P. 2001a: Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics*, 22, 27–57.
Robinson, P. 2001b: Individual differences, cognitive abilities, aptitude complexes and learning conditions: an aptitude complex/ability differentiation framework for SLA research. In Larry Selinker and Usha Lakshmanan (eds), *Second Language Research*, 17, 368–92: Special Issue on Explanation and Scientific Method in SLA Research: Festschrift for Paul van Buren.
Robinson, P. 2001c: Task complexity, cognitive resources, and syllabus design: a triadic framework for investigating task influences on SLA. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 287–318.
Robinson, P. 2002: Effects of individual differences in working memory, intelligence and aptitude on incidental second language learning: a replication and extension of Reber, Walkenfield and Hernstadt 1991. In P. Robinson (ed.), *Individual Differences and Instructed Language Learning*. Amsterdam: John Benjamins, 211–66.
Robinson, P. forthcoming: *Task Complexity, Second Language Development and Syllabus Design*. Cambridge: Cambridge University Press.
Robinson, P. and Ha, M. 1993: Instance theory and second language rule learning under explicit conditions. *Studies in Second Language Acquisition*, 13, 413–38.
Robinson, P. and Lim, J. J. 1993: Cognitive load and the Route-Marked/Not-Marked Map Task. Unpublished data. University of Hawai'i at Manoa, Department of ESL.
Robinson, P. and Niwa, Y. 2001: Task complexity and reasoning demands of second language narratives: effects on production and perceptions of task difficulty. Ms.
Robinson, P. and Yamaguchi, Y. 1999: Communication, corrective feedback and aptitude: interactions of classroom, task and learner factors. Paper presented at the 13th AILA

World Congress of Applied Linguistics, Waseda University, Tokyo, August.

Robinson, P., Strong, G., Whittle, J., and Nobe, S. 2001: The development of academic discussion ability. In J. Flowerdew and M. Peacock (eds), *Research Perspectives on English for Academic Purposes*. Cambridge: Cambridge University Press, 347–59.

Robinson, P., Ting, S., and Urwin, J. 1995: Investigating second language task complexity. *RELC Journal*, 25, 62–79.

Roediger, H., Weldon, M., and Challis, B. 1989: Explaining dissociations between implicit and explicit measures of memory: a processing account. In H. Roediger and F. Craik (eds), *Varieties of Memory and Consciousness*. Hillsdale, NJ: Lawrence Erlbaum Associates, 3–41.

Rosa, E. and O'Neill, M. 1999: Explicitness, intake and the issue of awareness: another piece to the puzzle. *Studies in Second Language Acquisition*, 21, 511–56.

Rott, S. 1999: The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition*, 21, 589–620.

Rutherford, W. 1987: *Second Language Grammar: Learning and Teaching*. London: Longman.

Sanders, A. 1986: Energetical states underlying task performance. In G. R. Hockey, A. W. Gaillard, and M. G. Coles (eds), *Energetics and Human Information Processing*. Dordrecht: Martinus Nijhoff, 139–54.

Sanders, A. 1998: *Elements of Human Performance*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sanders, A. and Neumann, O. 1996: Introduction. In O. Neumann and A. Sanders (eds), *Handbook of Perception and Action. Vol. 3: Attention*. San Diego: Academic Press, 1–5.

Sasaki, M. 1996: *Second Language Proficiency, Foreign Language Aptitude and Intelligence*. New York: Lang.

Sawyer, M. and Ranta, L. 2001: Aptitude, individual differences and program design. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 319–53.

Schachter, D. 1987: Implicit memory: history and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13, 501–18.

Schachter, D. 1996: *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books.

Schmidt, R. 1990: The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–58.

Schmidt, R. 1995: Consciousness and foreign language learning: a tutorialon the role of attention and awareness in learning. In R. Schmidt (ed.), *Attention and Awareness in Foreign Language Learning and Teaching*. Second Language Teaching and Curriculum Center Technical Report No. 9. Honolulu: University of Hawai'i Press, 1–64.

Schmidt, R. 2001: Attention. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 3–32.

Schmidt, R. and Frota, S. 1986: Developing basic conversational ability in a second language: a case study of an adult learner of Portuguese. In R. Day (ed.), *Talking to Learn: Conversation in Second Language Acquisition*. Rowley, MA: Newbury House, 237–322.

Schneider, W. and Detweiler, M. 1988: The role of practice in dual-task performance: toward workload modeling in a connectionist-control architecture. *Human Factors*, 30, 539–66.

Schneider, W., Healy, A., and Bourne, L. Jr 1998: Contextual interference effects in foreign language vocabulary acquisition and retention. In A. Healy and L. Bourne Jr (eds), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention*. Mahwah, NJ: Lawrence Erlbaum Associates, 77–90.

Schwartz, B. 1999: Let's make up your mind: "special nativist" perspectives on language, modularity of mind, and nonnative language acquisition. *Studies in Second Language Acquisition*, 21, 635–56.

Sergeant, J. 1996: A theory of attention: an information processing perspective. In G. Reid Lyon and N. A. Krasnegor (eds), *Attention and Executive Function*. Baltimore, MD: Paul H. Brookes, 57–69.

Shaffer, L. 1975: Multiple attention in continuous verbal tasks. In P. Rabbitt and S. Dornic (eds), *Attention and Performance. Vol. 4*. New York: Academic Press, 157–67.

Shallice, T. 1972: Dual functions of consciousness. *Psychological Review*, 79, 383–93.

Shallice, T. 1978: The dominant action system: an information processing approach to consciousness. In K. Pape and J. Singer (eds), *The Stream of Consciousness*. New York: Plenum Press, 117–57.

Shanks, D. and St John, M. 1994: Characteristics of dissociable human systems. *Behavioral and Brain Sciences*, 17, 367–447.

Sharwood-Smith, M. 1981: Consciousness-raising and the second language learner. *Applied Linguistics*, 2, 159–68.

Sharwood-Smith, M. 1991: Speaking to many minds: on the relevance of different types of language information for the L2 learner. *Second Language Research*, 7, 118–32.

Shiffrin, R. 1993: Short-term memory: a brief commentary. *Memory & Cognition*, 21, 193–7.

Shiffrin, R. and Schneider, W. 1977: Controlled and automatic human information processing II: perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127–90.

Skehan, P. 1996: A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38–62.

Skehan, P. 1998: *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skehan, P. and Foster, P. 1997: Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1, 185–212.

Skehan, P. and Foster, P. 2001: Cognition and tasks. In P. Robinson (ed.), *Cognition and Second Language Instruction*. New York: Cambridge University Press, 183–205.

Smith, E. E. and Sloman, S. 1994: Similarity-versus rule-based categorization. *Memory & Cognition*, 22, 377–86.

Spada, N. 1997: Form-focussed instruction and second language acquisition: a review of classroom and laboratory research. *Language Teaching*, 29, 1–15.

Spada, N. and Lightbown, P. 1993: Instruction and the development of questions in the L2 classroom. *Studies in Second Language Acquisition*, 1 (15), 205–21.

Sperling, G. 1960: The information available in brief visual presentations. *Psychological Monographs*, 74 (11) (no. 498).

Squire, L. 1992: Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, 4, 232–43.

Stadler, M. 1992: Statistical structure and implicit serial learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 318–27.
Sternberg, R. 1985: *Beyond IQ: A Triarchic Theory of Human Intelligence*. Cambridge: Cambridge University Press.
Swain, M. 1985: Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (eds), *Input in Second Language Acquisition*. Rowley, MA: Newbury House, 235–53.
Swain, M. 1995: Three functions of output in second language learning. In G. Cook and B. Seidlhoffer (eds), *Principle and Practice in Applied Linguistics: Studies in Honor of H. G. Widdowson*. Oxford: Oxford University Press, 125–44.
Swain, M. and Lapkin, S. 1995: Problems in output and the cognitive processes they generate. *Applied Linguistics*, 16, 370–91.
Tarone, E. 1985: Variability in interlanguage use: a study of style-shifting in morphology and syntax. *Language Learning*, 35, 373–403.
Tarone, E. and Parrish, B. 1988: Task-based variation in interlanguage: the case of articles. *Language Learning*, 38, 21–44.
Thelen, E. and Smith, L. 1998: *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.
Ting, S. C.-C. 1996: Tasks and planning time in the acquisition of Chinese as a second language. In P. Robinson (ed.), *Task Complexity and Second Language Syllabus Design: Data-Based Studies and Speculations*. Brisbane: University of Queensland Working Papers in Applied Linguistics, 30–63.
Tipper, S. 1992: Selection for action: the role of inhibitory mechanisms. *Current Directions in Psychological Science*, 1, 105–9.
Tomasello, M. and Farrar, M. 1986: Joint attention and early language. *Child Development*, 57, 1454–63.
Tomasello, M., Manle, S., and Kruger, A. 1989: Linguistic environment of 1- to 2-year-old twins. *Developmental Psychology*, 22, 169–76.
Tomlin, R. and Villa, V. 1994: Attention in cognitive science and SLA. *Studies in Second Language Acquisition*, 16, 185–204.
Treisman, A. 1971: Shifting attention between the ears. *Quarterly Journal of Experimental Psychology*, 23, 157–67.
Truscott, J. 1998: Noticing in second language acquisition: a critical review. *Second Language Research*, 14, 103–35.
Urwin, J. 1999: Second language listening task complexity. Ph.D. dissertation. Monash University.
van der Heijden, T. 1992: *Selective Attention in Vision*. London: Routledge.
VanPatten, B. 1996: *Input Processing and Grammar Instruction: Theory and Research*. Norwood, NJ: Ablex.
White, J. 1998: Getting the learner's attention: a typographical input enhancement study. In C. Doughty and J. Williams (eds), *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press, 85–113.
White, L. 1991: Adverb placement in second language acquisition: some effects of positive and negative evidence in the classroom. *Second Language Research*, 7, 122–61.
White, L., Spada, N., Lightbown, P., and Ranta, L. 1991: Input enhancement and L2 question formation. *Applied Linguistics*, 12, 416–32.
Wickens, C. 1984: Processing resources in attention. In R. Parasuraman and D. Davies (eds), *Varieties of Attention*. New York: Academic Press, 63–102.

Wickens, C. 1989: Attention and skilled performance. In D. Holding (ed.), *Human Skills*. New York: John Wiley, 71–105.

Wickens, C. 1992: *Engineering Psychology and Human Performance*. Second edition. New York: HarperCollins.

Wickens, C., Gordon, S. E., and Liu, Y. 1997: *An Introduction to Human Factors Engineering*. New York: Longman.

Williams, J. 1999: Memory, attention and inductive learning. *Studies in Second Language Acquisition*, 21, 1–48.

Zobl, H. 1992: Sources of linguistic knowledge and uniformity of nonnative performance. *Studies in Second Language Acquisition*, 14, 387–403.

Zobl, H. 1995: Converging evidence for the acquisition/learning distinction. *Applied Linguistics*, 16, 35–57.