

# The Effects of Task Type and Instructions on Second Language Acquisition



# The Effects of Task Type and Instructions on Second Language Acquisition

By

Hayo Reinders

**CAMBRIDGE  
SCHOLARS**

---

P U B L I S H I N G

The Effects of Task Type and Instructions on Second Language Acquisition,  
by Hayo Reinders

This book first published 2010

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data  
A catalogue record for this book is available from the British Library

Copyright © 2010 by Hayo Reinders

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-4438-1778-3, ISBN (13): 978-1-4438-1778-3

To my Wife, the light of my life.



# TABLE OF CONTENTS

Chapter One.....	1
Introduction	
Chapter Two .....	5
Cognitive Approaches to Second Language Acquisition	
Chapter Three .....	45
Input, Output, Interaction, and Intake in Second Language Acquisition	
Chapter Four .....	65
The Role of Tasks in Second Language Acquisition	
Chapter Five .....	81
Pilot Study	
Chapter Six .....	93
Methodology	
Chapter Seven.....	131
What are the Effects of the Instructional Treatments on Intake?	
Chapter Eight.....	145
What are the Relative Effects of Implicit and Explicit Treatments on Acquisition?	
Chapter Nine.....	165
What are the Relative Effects of the Individual, Collaborative, and Dictation Treatments on Acquisition?	
Chapter Ten .....	191
Discussion of the Results	
Chapter Eleven .....	205
Task and Instructional Types in Perspective	
References .....	215





# CHAPTER ONE

## INTRODUCTION

### 1.1 Motivation for the study

This research stems in large part from my own interest in a phenomenon that I have observed over many years of my own and others' language study. Why is it that under very similar circumstances (same language class, same teacher, same amount of tuition, even similar motivation/goals) some learners succeed and others do not? I have experienced this in secondary school French classes in Holland, in Arabic classes at universities in Cairo and Damascus, in Hebrew classes in Jerusalem, and in many other settings. As for myself, I knew it could not (only) have to do with the 'hard wiring' of the language learning system; I have always been poor at memorising vocabulary and even poorer at hearing differences in pronunciation, yet I have often been relatively successful at mastering languages. My own instinct told me it had to do something with what I used to call 'focus'. Now I would probably call this attention, although I still like the term focus as it is something that can be present (or absent) in a general sense rather than being related to a particular task or situation. A friend once referred to this as 'having your radar on' (which sounds a lot like Tomlin & Villa's "alertness"; 1994). Students who keep their eyes and ears open *at all times* for the new language, who constantly try to monitor others' and their own speech, and who actively hypothesise about the language as they go along, are the ones who do best. My studies in applied linguistics led me to the field of autonomy and self-directed language learning which further confirmed me in my thinking. However, in these areas it is not always clear what is meant when it is said that successful learners are "more proactive" or more "independent". I decided that to investigate more deeply what affects learning at this level, I had to turn my attention to, well, attention.

Much research has been done investigating the role of attention in second language learning (cf. Robinson, 1996; Schmidt, 1990, 1994, 2001). Although there continue to be some fierce debates, it appears that there is a

consensus that more attention leads to increased learning. One way to investigate attention is by looking at intake; the intermediate stage between input and acquisition. If learners take in information, it has to have been attended to. But if information has been attended to, is it learned? What factors affect this? How can learners' attention be increased, and can it be directed towards specific features in the language? It is these and other questions that prompted me to design the present study. I use tasks as tools to investigate the effects of different variables affecting 1) intake, 2) acquisition, and 3) I look at the relationship between intake and acquisition.

Recent years have seen a growing interest in the role of tasks in aiding second language learning and instruction (cf. R. Ellis 2003; van den Branden 2006; Willis & Willis 2007; van den Branden, van Gorp & Verhelst 2009). Tasks are now commonly researched for their relative effects on learning and their pedagogical contribution to classroom and out-of-class learning. Although advances have been made in isolating task characteristics beneficial to learning (e.g. various types of planning; Foster & Skehan, 1998, the effects of task repetition; Lynch & Maclean, 2001, the interaction between task and grammatical structure; Tarone, 1985), it is not yet clear what task characteristics aid in directing learners' attention to specific features of the input. A related question is whether successfully directing learners' attention to formal aspects of the language results in the development of explicit knowledge of those aspects. Another debate has focused on the roles of input, output and interaction in learning. Some have argued that providing learners with an opportunity to interact in the target language (Long, 1996; Swain, 1995), has a greater effect on acquisition than exposure to input only. Others have argued instead for the importance of learners processing the input before producing it (VanPatten, 1990). Other task characteristics such as the amount of time available to learners (Chaudron, 1985), whether the task is completed individually or collaboratively (Oxford, 1997), whether its completion requires attention to meaning (Hulstijn & Hulstijn, 1984), all have been shown to affect learning. However, the possibilities to be investigated are far from exhausted.

There is now also an increasing interest in the cognitive processes leading to learning and how tasks and activities in general affect those processes. Researchers such as Leow (1995, 1998) and Rosa & O'Neill (1999) have investigated the effects of tasks on both intake and learning. However, comparatively little empirical evidence exists for the specific workings of the relationship between the two. There is an obvious pedagogical interest

in determining the effects of certain task and instructional types on affecting immediate intake and (task) performance on the one hand, and eventual learning on the other (cf. Schneider, Healy, & Bourne, 1998).

The purpose of the study is thus twofold:

- 1) To determine the relative effects of implicit and explicit instructions
- 2) To investigate the effects of different task characteristics

In addition I will investigate the effects of 1) and 2) on both intake and acquisition and the relationship between the two

## **1.2 Outline**

I start by reviewing the relevant literature. Chapter 2 specifically looks at cognitive theories of general and second language learning, as well as a number of related topics such as the roles of memory, attention, and awareness. Chapter 3 discusses input, output and interaction as sources of data available to learners, and intake as the mediating process between that data and learning. Chapter 4 discusses the rationale behind task-based learning, as well as the various types of tasks and their effects on learning. Chapter 5 presents the design, method and results of the pilot study. Chapter 6 provides an in-depth discussion of the method used for the main study. Chapters 7, 8 and 9 present the results of the main study. Chapter 7 looks at the effects of the treatments on intake, chapter 8 at the effects of the implicit and explicit instructions on acquisition, and chapter 9 at the effects of three different task types on acquisition. Chapter 10 summarises and interprets the results for chapters 7-9 and makes comparisons between the effects of the treatments on intake and acquisition. Chapter 11 draws a number of conclusions and discusses implications as well as limitations of the study.



# CHAPTER TWO

## COGNITIVE APPROACHES TO SECOND LANGUAGE ACQUISITION

### **2.1 Introduction**

This book deals with the effects of different task types on intake and acquisition, as well as the effects of the accompanying instructions. Its focus is predominantly cognitive in nature and the literature review reflects this. The first chapter deals with cognitive approaches to second language acquisition. The first part reviews general cognitive psychological theories of knowledge processing, storage and learning. The second part discusses cognitive approaches to second language acquisition. I will focus primarily on topics and concepts that will be of relevance in subsequent chapters such as different types of (second language) knowledge and learning and the roles of attention and awareness.

### **2.2 Cognitive theories of information processing, storage, and learning**

The field of second language acquisition has increasingly drawn on research in cognitive psychology for theory building and further research. This first section gives an overview of the main theories that have influenced second language acquisition research.

#### **2.2.1 Information processing**

Cognitive psychology is the study of how information is processed and investigates aspects of memory, awareness, and perception. An information-processing approach investigates how information is selected, processed, learned, and retained. McLaughlin and Heredia (1996, p.214) list six characteristics of the information processing approach:

- 1) Humans are viewed as autonomous and active.
- 2) The mind is a general-purpose, symbol processing system.
- 3) Complex behavior is composed of simpler processes. These processes can be modular.
- 4) Component processes can be isolated and studied independently of other processes.
- 5) Processes take time; therefore, predictions about reaction time can be made.
- 6) The mind is a limited-capacity processor.

One of the first limited-capacity theories (see point six above) was developed by Broadbent in the 1950's. He proposed (1958) that human processing is characterised by 1) limited attention (i.e. when dealing with information we can only attend to a limited number of stimuli), and 2) effortful attention (there is a finite pool of resources we draw on when attending to information). Because attentional capacity is limited, not all information is processed and attention is allocated selectively to filter out part of the input, on the basis of its physical characteristics. Dual-task performance studies, whereby participants are asked to listen to two auditory sources (for example two sound recordings, one in each ear) but ignore one, result in poor recall of information presented on the unattended channel. Broadbent took this as evidence that information that is not attended to is discarded. However, later studies (e.g. Allport, Antonis, & Reynolds, 1972) showed that this is not always the case and that it depends on the degree of similarity between the messages. Auditory and visual stimuli for example, can relatively easily be attended to simultaneously. In addition, Broadbent assumed that since participants had no awareness of the unattended information that it was not processed for meaning, but this was proven not always to be true (cf. Von Wright, Anderson, & Stenman, 1975)<sup>1</sup>.

In response to these criticisms, Treisman (1964) proposed an attenuation theory of information processing and suggested that unattended information is attenuated or reduced, with all information being processed first for its physical aspects, and then for meaning. Yet others (e.g. Deutsch & Deutsch, 1963) have proposed a late filter theory in which all information is completely analysed automatically, with attention only necessary to determine which aspects of the analysed information will

---

<sup>1</sup> It is, however, as Cowan (1995) points out, not always possible to be sure that the unattended channel, was in fact not attended to at all.

enter into the subject's response. There is now considerable evidence against the late filter theory (cf. Cowan, 1988).

Studies into dual-task performance resulting in impaired performance appeared to provide evidence for the limited capacity that humans have for dealing with information. Kahneman (1973), although largely in agreement with this, suggested however that this capacity is flexible to some extent. He stressed the importance of task demands (e.g. complexity, novelty) and their resulting arousal on determining the total pool of resources available for a certain task.

Results from studies of dual-task performance were taken by some as evidence that attentional resources are distributed centrally by a single multi-purpose processor executive (e.g. Baddeley, 1986; Norman & Shallice, 1986). Others focused more on humans' apparent ability to engage in different tasks, if and when they are sufficiently different and suggested that this required the availability of a number of specific processing resources (e.g. Allport, 1989). Wickens (1984, 1989), suggested that information is processed perceptually based on different modalities (e.g. visual, auditory), then processed either verbally or spatially, and finally responded to manually or vocally. These different modalities, codes, and responses represent different resource pools and interfere with each other to varying degrees. This explained, according to Wickens, why some tasks can be performed more successfully together than others.

In the 1970's a number of researchers started to point out that the effortful processing in early information processing models was not always evident in human behaviour (which was not to say that it was not a precondition for that behaviour developing). Shiffrin & Schneider (1977, 1984) and Schneider & Shiffrin (1977) were among the first to distinguish between controlled and automatic processing. Controlled processing is a) of limited capacity, b) requires attention, and c) can (to some extent) be allocated flexibly. Automatic processing on the other hand is a) not bound by capacity limits, b) does not require attention, but c) is difficult to alter. Shiffrin & Schneider pointed out that a large part of human processing is automatic, for example during execution of everyday tasks. In order to perform complex tasks (such as driving a car, or speaking a second language) quickly and with a reasonable degree of accuracy, the necessary knowledge needs to be readily available for use. Their "automaticity model" explains skill development as an increasing efficiency with which

tasks are executed as a result of faster application of the underlying rules through practice. However, their theory has been criticised for having only descriptive and little explanatory power. Their experiments provide evidence of automatic processing, but do not explain how this takes place. Their suggestion that underlying patterns, similarities or rules pertaining to the stimuli are applied faster is not substantiated. Cheng (1985), for example points out that it may be that the nature of the knowledge changes rather than (just) its application.

Anderson's (1982, 1983) ACT (Adaptive Control of Thought) theory and its later version, the ACT\* (pronounced "act star") (1993) build on Shiffrin & Schneider's work. Anderson makes a distinction between declarative and procedural knowledge. The former is consciously accessible and comes in two shapes; episodic knowledge, which is autobiographical, and semantic knowledge, which is abstract knowledge, stored without reference to the context in which it was acquired. Procedural knowledge, on the other hand, is usually highly automatised and its application is dependent on specific context requirements. ACT\* suggests that all knowledge starts out in declarative form before it can become proceduralised. Recently Anderson has attenuated his position somewhat and has suggested that not all knowledge has to start out as declarative knowledge. It has been pointed out (Eysenck & Keane, 2000) that the model best fits fairly well-defined knowledge domains, and is less applicable to domains involving a greater degree of flexibility and creativity where it is difficult to identify the original declarative knowledge.

Logan's (1988) "instance theory" explains the development of automaticity not as a change from one type of knowledge to another but by suggesting that each encounter with a stimulus is stored separately, and obligatorily so once attended to. The more encounters occur, the more information is stored about the stimulus and the faster retrieval of that information becomes. Retrieval is automatic when encountering similar stimuli; 'Automaticity is memory retrieval: performance is automatic when it is based on a single-step direct-access retrieval of past solutions from memory' (p. 493). Automaticity, in this view, is thus not the speeding up of attentional control processes, but develops independently of such processes. Logan describes this retrieval process as a race between direct memory access and the outcome of a computational algorithm. When enough encounters with certain stimuli have been stored, access to them becomes direct, and outperforms any computational algorithm. It has become automatic. Robinson & Ha (1993) investigated some of the claims



made by instance theory, specifically for second language learning. They found evidence that new items required the activation of an algorithm-based operation, and old items were retrieved through direct access. They also found that whole chunks (“instances”) had been encoded by participants, rather than only parts of them, providing confirmation of Logan’s theory.

Similarly, in connectionist theories of learning, repeated encounters with a stimulus determine the extent to which it is retained. However connectionist theories see each encounter as strengthening existing connections, rather than as being stored separately. Connectionist theories do not view knowledge as stored discreetly, or symbolically (e.g. in rules), but as relational representations, or as distributed across a network of interconnected units, also referred to as a neural network (McClelland, Rumelhart, & Hinton, 1986). There is no rigid distinction between processing, learning, and the representation of information, as all are modelled in terms of units. Bechtel & Abrahamsen (1991) suggest that one of the main commonalities in connectionist theories is the connectivity of the units in a neural network. Activation (e.g. as the result of perceptual data entering the system) spreads through this network, thus relatively strengthening some connections and weakening others. There is no need for control of this activation by an executive processor and in this sense too, connectionist accounts differ from most symbolist accounts that do assign a role to controlled processing of stimuli. A term often used in relation to connectionism is parallel distributed processing (PDP) which forms one branch of connectionism research emphasising the importance of distributed representation, as opposed to localised representation (in the former view, information is stored across different points, in the latter it is stored separately, and wholly). An exhaustive discussion of connectionism is out of place here but for more information the reader is referred to Plunkett & Elman (1997).

### **2.2.2 Information storage**

The studies and theories discussed above all concern themselves with the way information is *processed*. Closely linked to this is the investigation of how information is *stored*. One of the earlier theories by Atkinson & Shiffrin (1968) proposed that there are three types of memory stores; a sensory, a short-term store, and a long-term store. The sensory store deals with perceptual data (sounds, visual, tactile, olfactory information). Attention to specific information in sensory memory is required for

information to enter short-term memory where it needs to be rehearsed in order to be retained in long-term memory. The sensory store deals with incoming perceptual information (in separate sub-stores, for example an iconic store for visual information, and an echoic store for auditory information). It has been found to be very limited in the amount of information it can contain. Short-term memory also has a very limited capacity (about seven items; Miller, 1956) and it too is prone to rapid decay or displacement, unlike long-term memory which has a seemingly unlimited capacity and can retain information indefinitely (although the information may become inaccessible over time).

There is evidence for the existence of separate memory stores but Atkinson & Shiffrin's model was later found to be overly simplified. For example, studies involving patients with brain damage showed that different types of new information were apparently stored in different short-term memory stores. Cowan (1995) points out that information also does not necessarily proceed from sensory memory to short-term, and from short-term memory to long-term memory in a linear fashion. Information can enter sensory memory, be matched with information already stored in long-term memory, the result of which may be sent to short-term memory which processes it (with the help of long-term memory) and sends the result to long-term memory, in a recursive process. Cowan (1993) referred to short-term memory as 'the interface between everything we know and everything we perceive or do' (p. 166).

Baddeley & Hitch (1974) proposed that the concept of short-term memory should be replaced with that of working memory. Working memory has three components:

- a central executive or supervisory attentional system that regulates information flow within working memory, allocates attention to particular input modalities or long-term memory systems, activates or inhibits whole sequences of activities guided by schemata or scripts and resolves potential conflicts between ongoing schema-controlled activities.
- a phonological loop comprising a phonological store and a process of articulatory rehearsal where inner speech can be used to refresh the decaying representations in the phonological store.
- a visuo-spatial sketchpad that generates images, mental maps, etcetera.

The central executive processor activates the phonological loop and the visuo-spatial sketchpad. Further research (Baddeley, Thomson, & Buchanan, 1975) showed that the former consists of 1) a passive phonological store directly concerned with speech perception, 2) an articulatory process linked to speech production that gives access to the phonological store. Presentation of auditory information directly activates the phonological store; visual information only indirectly if there is subvocal articulation. The visuo-spatial sketchpad was found to consist of a visual cache, which stores information about visual form and colour, and 2) an *inner scribe* which deals with spatial and movement information (Logie, 1995).

A distinction is frequently made between implicit and explicit memory. According to Schacter (1987) implicit memory is revealed 'when previous experiences facilitate performance on a task that does not require conscious or intentional recollection of those experiences' (p. 501) and conversely, explicit memory does involve conscious recollection. Evidence for the latter has been found in studies of priming (in which participants are exposed to input, such as words they need to complete a task later, without being told so beforehand). Butler & Berry (2001) have pointed out, however, that in order to demonstrate implicit memory is drawn on it is necessary to provide proof (amongst others) that performance reflects an unintentional retrieval strategy (this is called the "intentionality criterion"; cf. Schacter, 1989) and that performance is not accompanied by conscious recollection. Performance would also have to be affected equally on both perceptual tasks (tasks where the original stimulus and the retrieval stimulus are similar) and conceptual tasks (where the original and retrieval stimuli are only *conceptually* related). Butler & Berry (ibidem) reviewed a number of studies and concluded that 'considerably more studies using perceptual, rather than conceptual, implicit memory tasks have met the intentionality criterion' (p. 195) and that none had met both criteria. Following Richardson-Klavehn, Clarke, & Gardiner (1999), they suggested the use of the terms intentional versus incidental memories, rather than explicit/implicit memories. This would then make the distinction different from that proposed by Cohen & Squire (1980) between declarative (knowing that) and procedural (knowing how) memory. Tulving (1972), in addition, proposes a distinction between episodic (relating to specific memories of actual events) and semantic memory (relating to general knowledge about the world).

Some have argued that the different types of memory are related to different memory stores (e.g. one for implicit and one for explicit memory; cf. Schacter & Tulving, 1982; Paradis 1994) but others have claimed that the two need not be dissociated (for a review, see Robinson, 1995). Memories could be activated (as opposed to stored) differently. For example, a memory could be activated automatically without activation of its contextual (i.e. episodic) information (Graf & Mandler, 1984). Robinson (*ibidem*) holds the view that ‘...level of attentional awareness during retrieval is a function of task demands and automatic processes, which jointly determine access to a single long-term memory store’ (p. 317).

Tulving (1993) attempted to provide a synthesis of research until that time and suggested that there were five types of memory:

- 1) procedural memory (involved in skill learning, implicit)
- 2) perceptual representation (implicit)
- 3) short-term memory (explicit)
- 4) semantic memory (knowledge system, categorical memory, implicit)
- 5) episodic memory (autobiographical, personal memory, explicit)

Although there is increasing evidence for the existence of different types of memory, a great deal of research has so far only provided descriptions rather than actual explanations of the findings. Nonetheless, research into memory storage has greatly aided our understanding of the processes of learning.

### **2.2.3 Theories of learning**

Memory research is of course closely linked with the study of learning. Craik & Lockhart were two memory researchers who proposed a theory called ‘depth of processing’ to explain the effects of different types of processing on retention of information. The original theory was developed in 1972 and later expounded in a number of articles and books (e.g. Craik & Tulving, 1975; Cermak & Craik, 1979). Depth of processing theory claims that the way information is processed will to a large extent determine learning, with more elaborate forms of processing leading to more learning. What one remembers can be ‘[...] regarded as the byproduct of perceptual processing; [...] the resulting memory trace may be more or less elaborate depending on the number and qualitative nature of the analyses carried out on the stimulus (Craik & Tulving, 1995, p. 270),

and: ‘stimuli that are attended to, fully analyzed, and enriched by associations or images yield a deeper encoding of the event, and a long-lasting trace’ (ibidem p. 270). After publication of the original theory Craik and associates (e.g. Craik & Tulving, 1975), have suggested that “spread” and “elaboration” may be better descriptors than “depth” as the latter may be taken to imply quantity (e.g. time) rather than quality. They summarise the effects of processing as follows:

The memory trace could usefully be regarded as the byproduct of perceptual processing; just as perception may be thought to be composed of a series of analyses, proceeding from early sensory processing to later semantic-associative operations, so the resultant memory trace may be more or less elaborate depending on the number and qualitative nature of the analyses carried out on the stimulus. It was further suggested that the durability of the memory trace is a function of depth of processing. That is, stimuli which do not receive full attention, and are analyzed only to a shallow sensory level, give rise to very transient memory traces. On the other hand, stimuli that are attended to, fully analyzed, and enriched by associations or images yield a deeper encoding of the event, and a long-lasting trace. (p. 270)

A number of problems with the depth of processing model have been identified. Baddeley (1990) notes that it is difficult to determine to what depth information has been processed. Even where it is possible to do so, it is not always clear how to interpret this as what may be shallow processing in one situation may be deep and meaningful processing in another, depending on the type of retrieval the learner expects to have to perform. In addition, research has showed improved learning without a change in processing depth (for example through task repetition). Baddeley (ibidem) points out that for some types of learning the forming of associations is more important for an increase in performance. This may well be true for a great deal of language learning too. Another problem is that the tests used to measure learning were biased towards deep processing. Morris, Bransford, Franks, Morris & Stein (1979) showed that whether information is stored depends on its relevance to the subsequent test and labelled this the “transfer-appropriate processing theory”. The type of memory test used in most of the earlier levels-of-processing studies involved deliberate and conscious recall of information (i.e. explicit memory), but results were not always corroborated on tests of implicit memory.

The process of information retention can take place with or without conscious awareness of either the information (implicit or explicit

memory), or the process of retention (implicit or explicit learning). Reber is one of the best-known researchers to have investigated especially implicit learning. He conducted a large number of experiments (e.g. 1967, 1969, 1976, 1989) whereby participants were typically exposed to strings of letters conforming to a Markovian finite-state grammar. When subsequently asked to classify novel strings as grammatical or ungrammatical, participants performed above chance but could not verbalise the rules of the underlying grammar. A famous experiment was conducted by Berry & Broadbent (1984) in which participants had to manage complex processes to maintain certain production levels of a fictional sugar factory. Participants learned to perform this task quite well but could not report how they were able to do so. Sequence learning experiments expose participants to a sequence of (traditionally) lightflashes with recurrent patterns. Participants are able to predict patterns, without being aware of those patterns (cf. Curran & Keele, 1993).

Reber claimed that his own and others' research provided evidence for implicit learning. He defined implicit learning as 'a primitive process of apprehending structure by attending to frequency cues' as opposed to a 'more explicit process whereby various mnemonics, heuristics, and strategies are engaged to induce a representational system' (1976, p. 93). He also claimed that the knowledge resulting from such learning is abstract.

However, researchers have pointed out that in many cases it is difficult to be certain that participants did not develop any explicit knowledge in the course of the experiment. Not being able to verbalise knowledge may not be the same as not having any discrete knowledge of it. Shanks & St. John (1994) argue that for learning to be taken as implicit, two criteria would have to be met:

- 1) the information criterion (i.e. the information participants are asked to provide on an awareness test must be the information that has caused improved performance.).
- 2) The sensitivity criterion (i.e. the measure of awareness needs to be sensitive to *all* possible types of awareness of information.).

One well-known attempt to show that a great deal of learning labelled as implicit is in fact explicit learning, was made by Perruchet & Pacteau (1990). They carefully investigated some of the learning in Reber's experiments and showed that increased performance could be due to a

development of explicit knowledge of bigrams or trigrams (simple relationships between two or three items of information), rather than the abstraction of rules. They pointed out that a lot of research investigating implicit learning may not have found evidence for the development of explicit knowledge because the tests used to measure explicit learning favoured certain types of explicit knowledge only. Pothos & Bailey (2000), for example, found that decisions in Reberian experiments were often made not on the basis of grammaticality but on the basis of similarity between items and the strength (e.g. reliability) of certain chunks. It is also not clear that gained knowledge transfers to new exemplars; 'the Granada of unconscious rule learning' (DeKeyser, 2003). Dekeyser (*ibidem*) cites some studies that appear to have found such learning to take place (e.g. Brooks & Vokey, 1991). Redington & Chater (1996) argue, however, that such transfer phenomena can be explained if it is assumed that participants learn fragments (bigrams and /or trigrams) during the training phase of the experiment, and only abstract across the fragments at test time.

However, there appears to be some research that satisfies the above points as well as the information and sensitivity criteria postulated by Shanks & St. John (e.g. Howard & Howard, 1992). Additional evidence comes from research into the activation of different brain areas (cf. Gazzaniga, Ivry, & Mangun, 1998) under implicit and explicit learning conditions. Recent research (Willingham & Goedert-Eschmann, 1999) suggests that the two types may develop in parallel rather than one after the other, lending support to earlier suggestions by Graf & Schacter (1985) that the two systems (implicit and explicit) may not be dissociable because they operate concurrently.

Another common distinction is that made between inductive and deductive learning. Deductive learning proceeds from the general (e.g. a rule) to the specific (e.g. examples of the rule in use). Inductive learning proceeds from the specific to the general, resulting in the development of patterns and generalisations (Decoo, 1996).

Holland, Holyoak, Nisbett, & Thagard (1987) emphasised especially the importance of induction in their framework of cognitive processing which they described as 'all inferential processes that expand knowledge in the face of uncertainty' (*ibidem*, p. 1). An important element of their framework is its flexibility and its ability to incorporate experiences from failed processings. People avoid drawing up large numbers of hypotheses because 'they see constraints that can be derived from the general nature

of an information-processing system that pursues goals in a complex environment and receives feedback about its success in attaining its goals' (p. 2). In order to do this learners draw on events which inform their expectations about future events.

Holland and associates see a role for rules as the building blocks for mental operations but argue that for efficiency, rules need to be organised in relation to each other. This implicit organisation results from patterns of conditions and actions dependent on context. The resulting rule cluster can operate quickly in response to future events that are similar enough to trigger its activation. The authors point out that this explains how sometimes complex sets of interrelationships are learned more quickly than a single relationship between two events.

Processing in their framework is in the form of condition-action rules (IF/THEN statements) which operate when the condition part is matched by salient features of the environment. Specific rules gain dominance on the basis of their strength, their past performance, their specificity, and the amount of support they receive from other rules (the extent to which situational features affirm the relevance of the selected rule). Their framework differs from connectionist models in that it allows for both the strengthening of existing rules and the generation of new rules. It may thus result in immediate and radical behaviour changes.

### **2.2.4 Recent cognitive research**

Recent research in the areas of information processing, storage, and learning, looks particularly promising and relevant to the field of second language acquisition. Neurobiological research is already increasing our understanding of processes of (second) language intake (see the following chapter) and new technologies such as PET (Positron Emission Tomography) and CAT (Computerised Axial Tomography) scans allow us to identify brain activation during language use and learning, aiding in our understanding of how language is processed. One recent area of investigation is that of the effects of different task types and instructions on dual-task performance. One finding from early experiments was that dual-tasks (in which participants are engaged in two tasks simultaneously or in sequence) result in slowed and/or impaired performance. The response to the second of a pair of stimuli is slower than that to the first stimulus. The time difference is called the *psychological refractory period*. Initially this phenomenon was explained with reference to the central



bottleneck theory which holds that certain mental operations cannot be performed in parallel but instead must be performed one at a time (Welford 1952). More recently, however, alternative explanations have been put forward. Meyer & Kieras (1997a, 1997b) proposed a model called Executive-Process/Interactive Control (EPIC) which suggests that it is possible to perform multiple tasks at the same time without loss in performance. In this model, task execution is to some extent under the control of executive processes (e.g. task selection, monitoring) that will determine how and in what order tasks are executed (see also Baddely, 1986; Logan 1985; Norman & Shallice, 1986; Shiffrin & Schneider, 1977, 1984). If there is no incentive to carry out tasks in parallel, they will be executed sequentially.

Subsequent research has found that tasks can influence this process to the extent that in some cases dual-task slowing has been eliminated (Schumacher et al. 1997, reported in Levy & Pashler, 2001). Levy & Pashler (2001) termed this the instruction dependent hypothesis (IDH) and conducted a replication study of Schumacher et al.'s research to further investigate the findings. This involved participants responding with a keypress to the location of a figure appearing in one of three locations on a computer monitor and producing a vocal response to the pitch of a computer-generated tone. Participants were asked to give equal weight to each task and were given a financial incentive for speedy and accurate performance. These conditions were thought to favour parallel processing. As a result, in Schumacher et al.'s study the psychological refractory period was in some cases reduced to zero. In Levy and Pashler's study some of the variables were altered; in one experiment participants were asked to produce a keypress in response to the tone and a vocal response to the location of the figure. In another experiment, participants had to produce two responses to separate attributes of one stimulus. All experiments included items that required only one response (as a control). On the direct replication of Schumacher et al.'s study, no significant difference was found between items involving one or two stimuli/responses, which seems to support the instruction dependent hypothesis. However, the authors point out that in the case of the task requiring both a keypress and a vocal response, the keypress was made so quickly that the required processing could have been completed before the vocal response was made, but that it was likely that at least some overlap took place. In other experiments there were significant differences, conforming expectations based on a central bottleneck theory. In conclusion, there appears to be some preliminary evidence that dual-tasks

slowing may be eliminated, but not in all cases. The authors suggest further research be conducted varying, for example, stimulus onset of the different stimuli used.

This type of research may not have direct implications for studies of second language acquisition but aids in a greater understanding of processes of task performance and the effects of instructions. If evidence for unimpaired dual-task is further established, then it may in future be possible to identify language learning task demands and instructions that will enable learners to engage in dual processing, such as attending to both formal and meaning-related aspects of the language simultaneously.

## **2.3 Cognitive approaches to second language acquisition**

Findings from cognitive psychology have increasingly influenced second language acquisition theory building and research. In this section, first a number of such theories are reviewed, followed by a discussion of more specific topics such as the role of memory, attention, and awareness in second language acquisition.

### **2.3.1 Cognitive theories of second language acquisition**

Over time a number of theories of second language acquisition have been proposed that are predominantly cognitive in nature (e.g. Faerch & Kasper, 1980; Gass, 1991; Krashen, 1982, 1985; McLaughlin, 1978; Sharwood Smith, 1986; Skehan 1998; Schmidt, 1990, 1994, 2001; VanPatten 1990, 1996). A number of pedagogical applications in second language instruction, such as certain task-based approaches (e.g. Long & Crookes, 1992; R. Ellis, 1993), also draw heavily on cognitive theories.

Krashen's theory of second language acquisition (1982, 1985) has been highly influential. His monitor model encompasses a number of hypotheses, most prominently that there is a distinction between acquisition and learning of a language. The former is a relatively effortless inductive process and is similar to the way children learn their first language. Learning is a more conscious and controlled effort, such as often found in classrooms. The products of these two types of learning are used differently; acquired knowledge is used for language production, learned knowledge is used to monitor language processes (e.g. to check for accuracy). Simple rules may be learned but the use of that knowledge for monitoring purposes depends on the situation and factors such as the

availability of time, and task demands (e.g. in formal language tests, learned knowledge may be called for). The two types of knowledge are separate; one cannot become the other. The second part of Krashen's theory, the natural order hypothesis, claims that language is acquired in a predictable order, regardless of instruction. Difficult rules are acquired later than easier rules. The monitor hypothesis claims, as mentioned above, that learned knowledge can be applied to 'edit' language production. In order for this to happen time, a focus on form, and knowledge of the rule are necessary. Monitor use, according to Krashen, is difficult to encourage. Krashen identifies three types of learners: monitor over-users, under-users, and optimal monitor users who make use of the monitor when it does not interfere with communication. The input hypothesis claims that acquisition is driven by input that is slightly beyond the learner's linguistic competence (called "i+1"). The input feeds into a Language Acquisition Device, an innate mental structure for language. Learners make use of non-language cues such as context, prior knowledge and expectations to understand the message. Krashen gives the example of caretaker speech, which is roughly tuned to the child's level of understanding. The advantage of rough tuning (as opposed to, for example, the targeting of a specific language feature), is that "i+1" is likely to occur automatically. In classrooms, rough tuning is likely to benefit more students. The final hypothesis is the affective filter hypothesis, which claims that (mainly) acquisition can be impeded by anxiety, a lack of motivation, or self-confidence. The role of the classroom is to provide a low-anxiety environment.

A large number of criticisms have been levelled at the monitor model over the years (cf. Gregg, 1984; McLaughlin, 1987; Skehan, 1998), and counter-evidence has been presented against some of its claims. It has also been suggested that the model is not testable and / or falsifiable, and that some of its claims depend on other claims (e.g. its argumentation is circular). Rather than dealing with each of these points in isolation, alternative theories are discussed below.

McLaughlin was one of the first second language researchers to propose a detailed cognitive theory of second language acquisition. In his seminal 1978 paper he suggested, based on current thinking in cognitive psychology at that time, that second language acquisition consists of a number of competing processes, all vying for attention. In this framework, learning involves the transfer of information to long-term memory and is regulated by controlled processes each of which requires attention

(although not necessarily conscious awareness), which in turn takes time. Automatic processing, on the other hand, results from routinisation and does not require attention, and thus frees up resources to be allocated elsewhere. Automaticity develops over time, but is necessarily preceded by controlled processing. A distinction was also made between focal and peripheral attention (McLaughlin, Rossman, & McLeod, 1983), reflecting opposite ends of a continuum of degrees of attention. When automatic versus controlled processing and focal and peripheral attention are combined into a matrix, there are four possible distinct types of learning; from focal controlled (performance based on formal rule learning), to focal automatic (performance in a test situation), to peripheral controlled (performance based on implicit learning or analogic learning), to peripheral automatic (performance in communication situations).

Gass's model (first presented in 1988 and later expanded in 1997) is an attempt to capture the entire acquisition process and includes several stages from apperceived input, to comprehended input, to intake, integration, and finally, output. Gass defines apperception as 'a process of understanding by which newly observed qualities of an object are related to past experiences' (1997, p. 4). Or: 'the recognition (by the learner) that there is something to be learned, that is, that there is a gap between what the learner already knows and what there is to know' (p. 4). Input that is apperceived is processed for meaning and results in *comprehended input*. Not all input that is comprehended is made available to the developing system, i.e. not all of it becomes intake. For example, intake can be processed for meaning only to meet communicative demands, and not be processed for syntax (cf. Faerch & Kasper, 1980). Integration occurs when hypotheses are tested against available intake, or existing knowledge is strengthened. Knowledge may be integrated into existing knowledge by adding to it (for example in the case of an exception to a rule), but also by replacing existing knowledge or otherwise restructuring the interlanguage. The role of output is important to Gass for testing hypotheses, the result of which can feed back into the intake component.

Bialystok (1982, 1985, 1994; Bialystok & Mitterer, 1987) has proposed, and over the years offered various versions of a framework aiming to account (amongst others) for the implicit/explicit distinction in (both first and second) language knowledge and learning. The framework has two components, analysis, 'the process by which linguistic and conceptual representations become more explicit, more structured, and more accessible to inspection' (1994, p. 561) and control, related to the degree of

intentionality in language use. These components draw on universal linguistic knowledge, on conceptual knowledge (the individual's knowledge of the world), and on language-specific details. These various types of knowledge and information interact with each other. In Bialystok's view, explicit knowledge 1) can be derived from implicit knowledge through analysis, 2) is organised around formal categories, and 3) can be uniquely accessed (p. 561). It also can be conscious and may be accessed automatically. Language knowledge becomes increasingly explicit as it is analysed throughout development, however explicit knowledge does not become implicit. Instead, access to it changes, facilitating a process of automatisisation. 'The nature of representation must not be confused with access to that representation' (p. 567). For Bialystok, learning is a process of increasing explicitness: 'the explicit knowledge dynamically evolves from the implicit knowledge through development, as the whole system moves towards a state of increasing explicitness' (p. 567).

This position is different from that taken by a number of cognitive theorists, such as Anderson (1982), who see a development from a controlled declarative stage, to an automatised procedural stage, as well as that of most second language acquisition researchers. Hulstijn writes: 'it should be emphasized that the acquisition of language skills is not merely a speeding up of the execution of essentially the same procedures originally formed from declarative knowledge. Rather, language acquisition essentially consists of the establishment of new procedures which *reorganize* a body of facts and rules previously acquired' (1990, p. 32; emphasis in original). Hulstijn (ibidem) criticised Bialystok's framework for 1) failing to explain development, but instead only applying to task demands, 2) the assumption that learning must start with unanalysed knowledge, 3) the assumption that learning starts with low control (implicit knowledge, in later work), 4) its complexity due to its dual nature, 5) for being too domain-unspecific, 6) for failing to take into account the quantitative difference between novices and experts, and 7) indeterminacy of the concept of automaticity. (See Bialystok, 1990, for a response).

The theories discussed above have in common that they take a limited-capacity view of learning and see it as effortful allocation of attention. However, capacity limited models and SLA theories based on them are not without criticism (cf. Truscott, 1998). Several SLA researchers (e.g. DeKeyser, Salaberry, Robinson & Harrington, 2002) have pointed to

recent developments and sometimes a resulting abandonment of limited-capacity models in the field of cognitive psychology (cf. Neumann, 1987; Sanders, 1998; see also the first section of this chapter). Contrasting theories have proposed that attentional capacity is unlimited but that demands that stimuli, especially conflicting ones, place on the system cause performance to be reduced. Particularly where the information that the system acts upon is similar or if the information is confusing in some way, efficiency and speed are affected<sup>2</sup>. DeKeyser et al. (ibidem) cite Gopher (1992, pp. 279-280): ‘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 ... [such] models the only limited resources are time and its derived scheduling constraints’ (p. 808).

An interesting quote from James in this respect is the following:

The number of *things* we may attend to is altogether indefinite, depending on the power of the individual intellect, on the form of the apprehension, and on what the things are. When apprehended conceptually as a connected system, their number may be very large. But however numerous the things, they can only be known in a single pulse of consciousness for which they form one complex ‘object’, so that properly speaking there is before the mind at no time a plurality of *ideas*, properly so called. (1890, p. 262)

One researcher working within a competing (connectionist) paradigm is MacWhinney (1987, 2001a, 2001b; MacWhinney & Bates, 1989). In his competition model input plays the role of providing “cues” for comprehension. The availability (e.g. frequency, saliency) and reliability (e.g. regularity) of these cues determine in what order they are acquired, and to what extent they aid in comprehension of the input. Learner differences such as brain development and “plasticity”, differences in working memory, learners’ first language and transfer between the L1 and the L2, and time on task all influence acquisition, as does (support derived from) the context. However, cues inevitably impinge on the system and do not necessitate conscious attention at time of encoding.

---

<sup>2</sup> The difference may be merely theoretical; it may be difficult to distinguish between reduced performance as a result of processes that have slowed down, or reduced performance as a result of sequential processing.