An empirical psycholinguistic investigation of input processing and input enhancement in L1 English/L2 German: the acquisition of V2 and accusative case marking.

Submitted by:

Zoe Agiasophiti,

School of Modern Languages,

Newcastle University

Supervised by:

Professor Florence Myles,

Dr Carol Fehringer

Acknowledgements

I would like to thank my supervisors, Prof. Florence Myles and Dr Carol Fehringer for their valuable support and supervision throughout these years.

Many thanks to Prof. Bill VanPatten for our lengthy discussion and feedback about the design of the present study as well as for the example materials provided from earlier versions from the Culman et al. (2009) paper. I would also like to thank Prof. Michael Sharwood Smith and Dr Alessandro Benati for critically reviewing the materials used in the present study. Special thanks to Prof. Michael Sharwood Smith and Prof. John Truscott for allowing me to cite earlier versions of their manuscript 'Modular growth and use of language: a processing perspective' (in prep), as well as many presentations and draft articles regarding MOGUL. The manuscript has been a great inspiration and a valuable tool for the present study. I would also like to thank Prof. Michael Sharwood Smith for taking the time to explain to me how exactly MOGUL operates in greater detail.

Many thanks to Dr Simon Kometa for making statistics so easy to understand.

Special thanks to all the Schools, Teachers and Students that agreed to participate in this study.

Many thanks to Dr Clare Wright, Rola Naeb, Dr Nonie Chiang and Caroline Cordier for their advice and precious support. I am blessed to have such good friends.

Many thanks to my internal examiner, Dr Martha Young-Scholten, and my external examiner, Prof. Michael Sharwood Smith.

Last but not least I would like to thank my family: Dina, Kiki and George for their love and faith in me.

Without all these people, this thesis could not have been completed.

Thank you!

Zoe Agiasophiti

Abstract

Input Enhancement (IE) and Processing Instruction (PI) manipulate input in different ways in order to draw learners' attention to the target linguistic form. For IE the objective is to make input salient to make it more likely to become processed (Sharwood Smith, 1991; 1993). For PI the aim is to force learners to process the target form in order to decode the meaning of the sentence (VanPatten, 1996; 2004). Studies in PI and IE have shown positive effects for instruction, with explicit and more obtrusive types of PI and IE instruction being more effective than less explicit and less obtrusive (Norris and Ortega, 2000; Doughty, 2003; Lee and Huang, 2008). Despite promising results, the validity of PI and IE studies has been questioned because of small sample sizes and short time lapses before the administration of the delayed post-test (Norris and Ortega, 2000; Doughty, 2003; Lee and Huang, 2008). In addition, the theoretical underpinnings of PI (and to some extent IE studies [see Sharwood Smith and Trenkic, 2001] have been criticised for being vague and adopting outdated psycholinguistic theories (Carroll, 2004; DeKeyser et al. 2002; Collentine, 2004).

The present study examines the L2 acquisition of German V2 and case marking and investigates if and to what extent PI, IE, the combination of the two compared to no targeted instruction are effective in the acquisition of the target form in the short and long term. A hundred and thirty one secondary school English learners of German were randomly assigned to four groups, namely: +IE –PI, -IE+PI, +IE +PI, -IE –PI and received a two day instruction. An online pre-test, immediate post-test and delayed post-test including error correction, comprehension, production and interpretation tasks were administered. The +IE+PI group performed significantly better than the other groups in both immediate and delayed post-tests, according to the following hierarchy: +IE+PI>PI>IE>C. The results are discussed in the light of the theories traditionally thought to underpin PI and IE, and Modular Online Growth and Use of Language (MOGUL) is used to provide a more sophisticated and coherent interpretation of the results obtained (Sharwood Smith and Truscott, 2004, 2005; Truscott and Sharwood Smith, 2004; Sharwood Smith and Truscott, in prep.).

The present study's findings provide support that combining PI and IE is more effective as a teaching intervention than the sole application of the two and/or no instruction. PI can successfully alter learners' strategies when processing German OVS sentences by

forcing them to pay attention to word order and case marking. IE is successful in drawing learners' attention to the target linguistic form, although gains are short lived. However, it remains to be seen whether the benefits of the combined method, which are maintained to some extent in the delayed post-test, are still present in the longer term.

Table of Contents:

Table of Figures:

List of Tables:

List of abbreviations:

ACC/Acc: Accusative

AfS: Affective structure

AIT: Autonomous Induction Theory

ANOVA: Analysis of a Variance

APT: Acquisition by Processing Theory

BAAL: British Association for Applied Linguistics

C: Control (group)

CALL: Computer Assisted Language Learning

CEFR: Common European Framework of Reference for Languages

CM: Competition Model

CRB: Criminal Records Bureau

CS: Conceptual structure

CV: Communicative Value

DPT: Delayed-post test

EI: Explicit Instruction

Fem: Feminine

FNP: First Noun Principle

FonF: Focus on form

FonfS: Focus on forms

IE: Input Enhancement

IP: Input Processing

IPT: Immediate post-test

L1: First language

L2: Second Language

L3: Third Language

LAD: Language Acquisition Device

LFG: Lexical Functional Grammar

LTM: Long Term Memory

Masc: Masculine

MOG: Meaning-Oriented Group

MOGUL: Modular Online Growth and Use of Language

NH: Noticing Hypothesis

NOM/Nom: Nominative

OVS: Object Verb Subject

P&P: Principles and Parameters

P: Principle

PI: Processing Instruction

POpS: Perceptual Output Structures

PS: Phonological structure

PT: Processability Theory

RMA: Repeated Measures ANOVA

ROG: Rule-Oriented Group

SD: standard deviation

SI: Structured Input (activities)

SLA: Second Language Acquisition

SS: (Morpho-) Syntactical structure

SS: Sharwood Smith

SST: Sharwood Smith and Truscott

SVO: Subject Verb Object

TAP: Think-Aloud Protocols

TI: Traditional Instruction

TSS: Truscott and Sharwood Smith

TTIE: Typographical/Textual Input Enhancement

UG: Universal Grammar

VS: Visual structure

WM: Working Memory

1 Introduction

The present study aims to investigate the acquisition of case marking and word order by English learners of German. This study seeks to explore how input presented in different ways through focused type of instruction can be more effective in aiding language development. Therefore the main aim and objective is to provide a better insight into how the initial stage of language processing can have an impact on the final stage of language acquisition.

Defining Second Language Acquisition (SLA) requires considering the contribution of various fields, since SLA is a complex process which is not fully understoodto date. SLA is a process that involves learning and therefore individuals' mental states and changes, while it attempts to identify what constitutes L2 knowledge, i.e. 'property theory', and how L2 knowledge develops, i.e. 'transition theory' (Gregg, 2003; Carroll, 2001; Doughty and Long, 2003; Mitchell and Myles, 2004; Young-Scholten and Piske, 2009 in: Piske and Young-Scholten, 2009; Sharwood Smith and Truscott, in prep). As such, SLA is viewed as a cognitive process that takes place in a social context and therefore the contribution of various disciplines is necessary to understand and define this complex process (Robinson, 2001; Doughty and Long, 2003; Gregg, 2003; Mitchell and Myles, 2004; VanPatten and Williams, 2007).

According to Gregg (2003:839), SLA theory, amongst other things, needs to explain the initial state that is the knowledge state of the L2 learner prior to L2 exposure, and the final state after which input ceases to have any effect. This explanation requires a 'property theory' (see Gregg, 1993, 1996a, 2001, 2003; Cummings, 1983 for L1), such as a linguistic theory that can explain how linguistic knowledge is represented in mind, and a transition theory that could explain the processes that take place in the learner's cognitive system so that a representation can become part of it (see also Doughty and Long, 2003; Mitchell and Myles, 2004; Young-Scholten and Piske, 2009 in: Piske and Young-Scholten, 2009). In SLA there is no unified approach, and transition theories vary according to the property theory they are linked to. In this vein, the role of input (and intake) in acquisition, and what constitutes acquisition, differs greatly in the field (Rast, 2008).

From a Universal Grammar (UG) perspective, innateness plays a central role in acquisition, as humans inherit a mental language faculty like a ‘blueprint’, which constrains the shape that human languages can take (Chomsky, 1981, 1986, 2000; Mitchell and Myles, 2004; White, 1996, 2003). Based on UG propositions, the core of language is seen as a separate entity from other aspects of human cognition, although there is close interaction between the two. In this framework, the L2 acquisition process is seen as very similar to L1 acquisition until the ‘critical period’, i.e. children’s early development. On the contrary, it is believed that adult second language learners have to resort to other learning mechanisms (Bley-Vroman, 1989; Hawkins, 2001; Johnson and Newport, 1989; Mitchell and Myles, 2004; White, 2003; among many). Hence, teaching L2 learners morphosyntactic aspects of language can promote native-like fluency and or second language competence[1] (see Archibald 2000; Chomsky, 1980, 1981, 1995; Hawkins 2001a, 2001b; Herschensohn 2000; Mitchell and Myles, 2004; Schwartz 1993, 1998; White 1989, 1996, 2000, 2003). Within the field of SLA there are different perspectives about language development, adopting cognitive-processing or functional-cognitive perspectives (for detailed discussion please refer to section 2.4).From a cognitive-processing perspective, real time processing of language and how learners are accessing and processing linguistic information are central issues (see Carroll, 2004; Pienemann, 1998; Sharwood Smith and Truscott, 2004 and in prep).

From a functional-cognitive perspective, linguistic knowledge and learning mechanisms are located within the general cognitive system, while (second) language is acquired through usage and driven by input (Anderson 1983; Bates and MacWhinney 1989; DeKeyser, 1995, 1997, 2007; N.C. Ellis, 2003; Tomasello, 1998, 2000, 2003). From an interactionist approach language is constrained by cognition; however learning is mediated through communication, where social, affective and cognitive variables play a role (Long, 1981, 1983a, 1983b, 1996; Gass, 1998 inter alia).

Another aspect of defining SLA is the distinction between instructed (guided, formal, tutored), when a second language is learned/acquired under pedagogical guidance, versus uninstructed (aka naturalistic, spontaneous, unguided, informal, untutored) SLA,

[1] It should be noted that within the field of UG there is a controversy whether second language learners have the ability to reach native-like fluency and second language competence after the ‘critical period’. However, it is beyond the scope of this research to step into this debate.

when a second language is learned/acquired through spontaneous communication in authentic social situations (Ellis, 1985; 1994; Housen and Pierrard, 2005; Klein, 1986; Larsen-Freeman and Long, 1991; McLaughlin, 1987). The present study attempts to investigate the role of instruction and more specifically the role that type of instruction may have in SLA. It is therefore important to first define what instruction is and how instruction can affect SLA. Housen and Pierrard (2005:2) define instruction as 'any systemic attempt to enable or facilitate language learning by manipulating the mechanisms of learning and/or the conditions under which these occur'.

In the same vein, the present study is set to investigate three instructional methods, namely Input Enhancement (IE), Processing Instruction (PI) and a combination of the two in order to measure if and to what extent they are effective in acquiring German word order and case marking. The three proposed types of instruction assume that drawing attention to specific forms of the input by manipulating the input, either through saliency (IE), or by forcing learners to process the form in order to obtain meaning (PI) and/or the combination of the two (IE and PI), will have an impact on language development. If this is the case, are the effects also noticeable in the long-term? Can we speak of acquisition of the target form? Can type of instruction trigger the learning processes and mechanisms in order to promote restructuring of lexical competence? In short, one of the present study's motivations is to investigate how, and to what extent, types of instruction can influence SLA (see also Ellis, 1994; Housen and Pierrard, 2005; Klein, 1986; Sharwood Smith, 1991, 1993; VanPatten, 1996, 2004, 2007, 2009).

The proposed types of instruction are based on psychological concepts such as attention and consciousness. They therefore raise another important issue that requires further investigation, namely what their role and impact of psycholinguistic concepts in second language processing and acquisition is (Schmidt, 1990, 1993, 1995, 2001; Sharwood Smith, 1991, 1993; Tomlin and Villa, 1994; VanPatten, 1996, 2000, 2004, 2007, 2009). In the literature, these psycholinguistic concepts are considered to be generating specific types of language knowledge, namely they distinguish implicit vs. explicit knowledge. Implicit knowledge is characterized by being largely intuitive and abstract knowledge, which is acquired subconsciously and incidentally on the basis of unplanned communicative language use. On the other hand, explicit knowledge is considered to be knowledge about language and is largely regarded as a conscious type of knowledge that is learned intentionally. The latter can be further divided into ‘analysed vs. metalinguistic knowledge’. Metalinguistic knowledge is considered verbalized knowledge about the structure of language and its theoretical constructs. In the acquisition of metalinguistic knowledge, learning is deliberate and conscious and it involves higher levels of conscious awareness (see DeKeyser, 1998; Doughty and Long, 2003; Ellis, 1995, 2001; Paradis, 2009; Sharwood Smith 1991, 1993; 2008, 2009; Sharwood Smith and Truscott, 2004, 2005, and in prep; Truscott, 1998). According to Bialystok (1994),‘analytic knowledge’ derives from implicit knowledge manifesting itself in problem solving tasks, such as cloze tasks and grammaticality judgment tasks, where learners are required to pay focal attention in choosing the correct linguistic form.

Krashen (1981, 1985) was the first to distinguish acquisition from learning in SLA. For Krashen (1981, 1985) acquisition is a subconscious process that leads to ‘acquired’ knowledge, whereas learning requires conscious effort on behalf of the learner, resulting in ‘learned’ knowledge (explicit knowledge). The latter is considered to play a very limited role in the development of second language proficiency. Second language acquisition is described as an unstoppable and consciously inaccessible process that begins as soon as L2 learners attempt to decode messages in the second language. In order for the acquisition process to be successful, input should be ‘comprehensible’ and has to match the learner’s stage of development. If these two conditions are met, for example if a learner is at stage ‘i’ and there is comprehensible input with slightly advanced information, i.e. ‘i+1’ then acquisition can occur. For Krashen (1981, 1985), instruction is not significant for acquisition, as it leads to learned conscious knowledge.

Input and intake are also interconnected with acquisition. From a behaviourist perspective input is considered to be the driving force for acquisition (Skinner, 1957). This view of course contradicts the generativist perspective, as Universal Grammar ‘UG’ is considered to be driving acquisition. However, since there is no unified approach to what constitutes acquisition, there cannot be a unified approach to what constitutes input and intake, as definitions vary according to the theoretical perspectives that researchers adopt (Rast, 2008). Regardless of the adopted perspective, there is a consensus among researchers about the importance of input in SLA. However, there are still many unanswered issues regarding the factors that may influence input and acquisition, such as what kind of input and how much input is necessary. Researchers working in the field of input in SLA have expressed a shared concern: the need for providing a well-defined theoretical framework of input in SLA (Carroll, 2001; Young-Scholten and Piske, 2009; Sharwood Smith and Truscott, 2004, 2005 and in prep; VanPatten, 2000). Truscott and Sharwood Smith (2009) point out that it is essential to provide a theoretical definition of input in order to be able to answer more thorny questions such as how we can convert input into intake, how we can know what part of the signal is processed or not and how we can best describe this processing. In other words, defining input and intake and the process of converting input into intake will provide the basis for defining what constitutes acquisition.

Defining second language acquisition is evidently a complex issue requiring a cross-disciplinary approach in investigating the processes involved for a structure to be acquired. The present study assumes that instruction does play a role in language development. It therefore seeks to explore whether and, to what extent, various types of teaching interventions can be effective in the acquisition of German case marking and word order by English learners. In the field of Input Enhancement (IE) it is assumed that manipulating the input externally may induce salience of the target form internally and possibly lead to further processing of the target form by the language learning mechanisms (Sharwood Smith, 1991, 1993). On the other hand, in the field of Processing Instruction (PI) it is assumed that forcing learners to process the target form can generate further processing of the target form and ultimately lead to the establishment of correct form-meaning connections. Both types of teaching intervention support the idea that psycholinguistic concepts such as attention, noticing and consciousness play a significant role in the various stages of the processing of the target form. A detailed presentation of the theoretical underpinnings of the psycholinguistic concepts, the instructional interventions and the theoretical frameworks and models is therefore presented in Chapter 2. The design and methodology used to operationalise the proposed research project are presented in detail in Chapter 3. Results from the present research study are provided in Chapter 4, while findings are discussed in Chapter 5. Finally, conclusions from the present study are presented in Chapter 6.

2 Input processing models and psycholinguistic concepts

The focus of the present study is to investigate the effectiveness of IP, PI, IE and FonF in the acquisition of German word order and case marking. It is therefore essential to provide an overview of the theoretical underpinnings of these approaches. In this way a better understanding about the psycholinguistic processes involved can be achieved, as well as a better understanding of how input manipulation is considered to be an effective teaching intervention through IP, PI, FonF and IE.

This chapter begins with a review of the psycholinguistic concepts underpinning IE, FonF and PI and therefore Schmidt's (1990, 1993, 1995, 2001) 'Noticing Hypothesis' and Tomlin and Villa's (1994) multistage definition of attention, namely alertness, orientation and detection are presented (section 2.1, 2.1. and 2.2.2). The review on these psycholinguistic concepts is important in order to understand the similarities, differences as well as overlaps in the theoretical underpinnings between typographical/textual IE, FonF (sections 2.2, 2.2.1 and 2.2.2), IP and PI (sections 2.3, 2.3.1. and 2.3.2). A review of IE and PI studies as well as conclusions and overall contribution to the purposes of the present study is provided in (sections 2.2.3, 2.2.4, 2.3. and 2.3.4). In the fourth part of this chapter language processing frameworks relevant to the purposes of the present study (section 2.4) are reviewed. Concluding this chapter the motivation, research questions and hypotheses of the present study are presented (section 2.5).

2.1 Psycholinguistic concepts underpinning IP, PI, IE and FonF

Introduction

In order to better understand the processes and mechanisms involved in IE and PI, it is vital to present the psycholinguistic concepts that underpin both instructional approaches. In the following section the various proposals on what constitutes attention, noticing, consciousness and awareness as well as their role within the proposed approaches will be described.

IE and IP through PI are two different but interrelated approaches that draw, to some extent, on common psycholinguistic theories. Both approaches are based on the notion that attention has a facilitative role in the acquisition of a target form. For IE, input should be manipulated – enhanced – in order to draw learner's attention to the target form (Sharwood Smith, 1991, 1993). For IP, input is manipulated through PI and more specifically through structured input activities in order to force learners to process the target form and make correct form-meaning connections (VanPatten, 1996, 2004, 2007, 2009). Attention therefore plays a crucial role in the way both IE and IP operate.

Attention is not an easy construct to define. Much of our understanding about the role of attention in SLA comes from cognitive psychology, neuroscience and psycholinguistics. There are six characteristics that define attention, namely that attention is limited; selective; partially subject to voluntary control; attention controls access to consciousness; attention is essential for the control of action; and attention is essential for learning. Among researchers there is no consensus about all six characteristics attributed to attention. A thorny issue in the literature about attention is whether attention and awareness are two overlapping but not isomorphic concepts that promote second language leaning through noticing understood as conscious processing (see Schmidt, 1990, 1993, 1995, 2001; see also FonF studies for instance Alanen, 1995; Jourdenais et al. 1995; Leow, 1997). Some proposals argue against this notion and dissociate attention from awareness. In this case attention is seen as a process involving three stages, namely alertness, orientation and detection. According to this proposal, it is during the process of detection that we can become conscious of processing without dismissing the fact that even within detection, processing occurs without conscious awareness (Tomlin and Villa, 1994; see also Sharwood Smith, 1991, 1993; VanPatten, 1996; Truscott 1998). One proposition is based on the belief that noticing, i.e. attended processing, is a necessary and sufficient condition for learning and acquisition to take place whereas according to the second proposition, learning and acquisition can occur with or without conscious awareness.

In the field of IE and PI, attention has a prominent role; however studies have been criticised for not linking theory to practice (see criticism from Sharwood Smith and

Trenkic, 2002 on IE/FonF studies; DeKeyser et al. 2002; Harrington, 2004, for PI studies among others; Truscott, 1998). This and other issues will be addressed below.

This section is divided in three subsections. In the first section, attention will be defined based on the literature from cognitive psychology and psycholinguistics. In the second section attention will be reviewed based on its practical applications in SLA. In this section the focus will be the definition of psycholinguistic concepts such as noticing, alertness, orientation and detection, as proposed by Schmidt (1990, 1993, 1995, 2001) and Tomlin and Villa (1994). The second section will also describe how these constructs have been operationalized in SLA, through IE and IP. The third subsection will review a novel proposal placing consciousness in a theoretical framework of language processing, thus addressing criticisms on IE and PI studies regarding the lack of a theoretical framework that can link psycholinguistic concepts used in cognitive research SLA theory to practical application. This is MOGUL, which attempts to bridge this gap by providing a theoretical framework giving attention and consciousness a central role in language learning.

2.1.1 Defining attention and its role in IE, FonF, IP and PI

J.R. Anderson (2004) defines attention as the cognitive process of selectively concentrating on one aspect of the environment while ignoring other things. Attention has also been referred to as the allocation of processing resources to input or to stimuli (ibid).

The dominant theory on attention is the idea that attention is a limited resource (Anderson, 2004; Broadbent, 1958; Kahneman, 1973; Posner and Snyder, 1975; Shiffrin and Schneider, 1977; Wickens, 1984; Tomlin and Villa, 1994; Schmidt, 2001; Miyake and Shah, 1999; McLaughlin et al. 1983; Towell and Hawkins, 1994; among many). Because of capacity limitations, attention cannot be allocated to more than one task, if it cannot be performed automatically. There are generally thought to be two general human information processing systems, namely ‘controlled’ versus ‘automatic’ (Anderson, 2004; DeKeyser, 1994, 1995, 1997; McLaughlin, 1990). In the case of ‘controlled’ processes (where information is yet to be learned, i.e. learning how to drive a car) they are effortful, of limited capacity, interfere with other processes and are attention-demanding. On the other hand ‘automatic’ processes (over-learned information, i.e. driving a car skilfully) have been established as capacity-free processes, requiring little or no attention and not interfering with other processes (Broadbent, 1958; Kahneman, 1973; McLaughlin et al. 1983; Miyake and Shah, 1999; Posner and Snyder, 1975; Schmidt, 2001; Shiffrin and Schneider, 1977; Tomlin and Villa, 1994; Towell and Hawkins, 1994; Wickens, 1984; among many). For IE and IP the idea that attention has a limited capacity is fundamental (Sharwood Smith, 1991, 1993; VanPatten, 1996, 2004, 2007, 2009).

A second characteristic of attention is that it is selective. According to selective attention theories (Broadbent, 1958; Deutsch and Deutsch, 1963; LaBerge, 1995; Schmidt, 2001; Tomlin and Villa, 1994), individuals have a tendency to process information from only one part of the environment with the exclusion of other parts, based on their arousal levels. A vexed issue in the literature remains whether the shifts in attention that accompany changes in the arousal level are automatic or deliberate. Researchers supporting limited capacity models propose that we all have a limited amount of mental capacity to allocate to various tasks, at any given time. Other researchers propose mechanisms that can orient attention, emphasizing mainly the concept of salience and therefore supporting deliberate changes in the arousal level of selective attention. Based on these proposals, attention can be oriented to the most salient of available locations (LaBerge, 1995; Corbetta and Shulman, 2002; Shipp, 2004). Schmidt (2001) recognises that there are also passive and involuntary forms of attention (see also LaBerge, 1995). For instance one can hear a loud noise whether one wants to or not.

IE and IP are mainly concerned with the idea that selective attention can be oriented. In the case of IE, manipulating the input is necessary in order to increase salience; while in the case of IP, input is manipulated in order to force learners to process the form when encoding the meaning.

A fourth characteristic of attention and another rather vexed issue is whether attention controls access to consciousness. Koch (2004, 2007) defines consciousness as a epiphenomenal experience. Based on the fact that when we pay attention to an object we become conscious of its various attributes and that the object fades from consciousness when we shift attention away, many researchers see selective attention and perceptual consciousness as two interconnected if not identical processes (Chun and Wolfe, 2000; Jackendoff, 1996; Merikle and Joordens, 1997; O'Regan and Noe, 2001; Posner, 1994; Prinz, 2004; Velmans, 1996). Others argue that attention and consciousness are two distinct phenomena with distinct functions and neural mechanisms (Baars, 1997, 2005; Block, 2005; Bachmann, 2006; Dehaene et al. 2006; Iwasaki, 1993; Kentridge et al. 2004; Koch, 2004; Lamme, 2003; Naccache et al. 2002; Woodman and Luck, 2003; Wundt, 1974). In SLA, many researchers view attention and awareness as two sides of the same coin (Neumann, 1996; Schmidt, 1990, 1993, 1995, 2001), whereas others argue that attention may be with or without consciousness (Carroll, 2006; Tomlin and Villa, 1994; Sharwood Smith, 1991, 1993, 2008, 2009; Sharwood Smith and Truscott, 2004, 2005, in prep; Truscott, 1998; VanPatten, 1996). A third proposal on attention and consciousness comes from Marcel (1983), who proposes 'focal attention' as the mechanism that divides attention, in two stages. In early processing stages attention produces non-conscious representations of all stimuli, while in a second, higher level of processing results in consciousness[2].

According to some cognitive psychologists and SLA researchers, a fifth characteristic attributed to attention is that it is essential for learning, although it is not necessarily conscious and therefore learning can be incidental (e.g. Baars, 1988, 1996; Carlson and Dulany, 1985; Carr and Curran, 1994; Gass, 1998; Kihlstorm, 1984; Logan, 1988; Posner, 1992; Schiffrin and Schneider, 1977; Sharwood Smith, 1991, 1993, 2008, 2009; Schmidt, 1995, 2001; VanPatten, 1996, 2004, 2007, 2009). The idea that attention has a facilitative role in the learning process has been a key notion for both IE and IP. However, in the field of SLA it is mainly interconnected with the former characteristic of attention, namely that attention holds (or not) access to consciousness. The debate in SLA is therefore whether conscious awareness of information, in other words 'noticing' or 'attended processing', is the only and sufficient condition for L2 learning to take place (Schmidt, 1990, 1993, 1995, 2001; see also section 2.2.2); or whether learning occurs with or without conscious registration of information, with detection of

[2] SLA researchers have been mainly informed from studies in the field of cognitive psychology. For further information about attention and awareness please refer inter alia to Posner, 1992; Posner and Petersen, 1990; Posner and Schneider, 1975; Schiffrin and Schneider, 1977; among many.

information being a prerequisite, as previously discussed (Sharwood Smith, 1991, 1993, 2008, 2009; Sharwood Smith and Truscott, 2004, 2005, and in prep; Tomlin and Villa, 1994; Truscott, 1998; VanPatten, 1993, 1996, 2004, 2007, 2009).

The debate on whether attention holds access to consciousness or not has had major implications in the field of IE research. Two strands of research have been created within the field of IE. More precisely, followers of the idea that attention holds access to consciousness have developed a Focus on Form (FonF)[3] approach (following Schmidt, 1990, 1993, 1995; see also Alanen, 1995; Doughty and Williams, 1998; Jourdernais, et al. 1995; for details please refer to section 2.2.2), while others have followed the original underpinnings of IE supporting that attention can occur with or without awareness (following Tomlin an Villa, 1994). Similar to IE, IP also argues that attention occurs with or without consciousness (VanPatten, 1996).

Many agree that metalinguistic knowledge (namely knowledge about what has been taught, i.e. grammar rules) promoted through awareness, can be an aiding factor in mastering the target form. However, metalinguistic knowledge is not synonymous with the acquisition process (Paradis, 2004, 2009; Sharwood Smith, 1991, 1993; Truscott, 1998; Sharwood Smith and Truscott, 2004 and in prep). Although the issue of conscious awareness has been central in the field of IE causing a dichotomy leading to the formulation of the FonF approach, in the field of IP a weaker view of ‘noticing’ as proposed by Schmidt (1990, 1993, 1995, 2001) has been adopted. This weaker view suggests that attention is the necessary condition to force learners to process the form therefore aiding acquisition. However, IP endorses the idea of detection and orientation of attention as proposed by Tomlin and Villa (1994). The distinction between Schmidt’s (1990, 1993, 1995, 2001) ‘Noticing Hypothesis’ and Tomlin and Villa’s (1994) proposal on attention is crucial for our understanding of the aims and objectives of IE and IP, as well as the dichotomy between IE and FonF. The two proposals are presented in the next section.

[3] The term FonF has been proposed by Long (1991) referring to an approach, where learner’s attention is drawn incidentally to linguistic items of the input, whose focus is on meaning or communication (for details see section 2.2.2).

2.1.2 Noticing Hypothesis vs. Alertness, Orientation and Detection of Attention

IE and IP have been largely influenced by Schmidt's (1990, 1993, 1995, 2001) 'Noticing Hypothesis' (NH) and Tomlin and Villa's (1994) notion of attention. Despite the fact that both IE and IP endorse Tomlin and Villa's (1994) notion of attention, researchers within the field of IE (FonF strand, see section 2.2 and sub-sections 2.2.1 and 2.2.2 for further details and explanations) adopt Schmidt's (1990, 1993, 1995, 2001) NH. In the following, the differences between the two propositions are reviewed.

Schmidt argues that attention and awareness are two overlapping but not isomorphic concepts. 'Noticing' is the by product of conscious awareness without which L2 learning/acquisition cannot occur, since 'noticing is the necessary and sufficient condition for converting input into intake' (Schmidt, 1993:209). In addition, Schmidt (1990, 1993, 1995, 2001) argues that attention controls access to awareness and is responsible for 'noticing'. Attention has limited capacity, is selective and voluntarily controlled[4], and is necessary to ensure that information is processed and stored in memory. For Schmidt (2001) attention has a dual role; it facilitates information (intake) for further processing, while it also selects the relevant information to be processed by excluding non-relevant information. Similarly, Schmidt holds that there are two levels of awareness: awareness at the level of noticing that leads to intake and a higher level of awareness, which he calls 'awareness at the level of understanding', i.e. awareness that leads to deeper learning and is characterised by the learner's ability to consciously analyse, compare and test hypotheses (Schmidt, 2001). In a nutshell, 'SLA is driven by what learners pay attention to and notice in target language input and what they understand the significance of noticed input to be' (Schmidt, 2001:3-4). According to the NH, consciousness is central to the learning process, since subliminal learning (learning without consciousness/awareness) is impossible. Noticing is the necessary and

[4] That said, this does not mean that there cannot be incidental learning for Schmidt. Voluntarily controlled attention can be best described to the 'cocktail phenomenon', where one can eavesdrop a conversation, which can be far away from another conversation that is taking place close by and is being ignored (Schmidt, 1995:9).

sufficient condition for converting input into intake; incidental learning[5], i.e. 'learning without the intent to learn' (Schmidt, 1994:16) is possible and effective when the demands of a task focus attention on what is to be learned; and paying attention is facilitative in acquiring redundant grammatical features (Schmidt, 1990, 1993, 1994, 1995).

Schmidt does not clearly define what qualifies as input, however he defines intake and acquisition. What constitutes incidental learning is problematic in the formulation of the NH, since learning cannot occur without noticing (see also Truscott, 1998; Sharwood Smith and Truscott, in prep). Truscott (1998) suggests that the NH should be limited to claims about the acquisition of metalinguistic knowledge, for which conscious noticing is important. Another issue raised is the lack of empirical evidence supporting the NH, since the hypothesis is based on anecdotal evidence from a diary study from Schmidt's 5-month stay in Brazil studying Portuguese formally for 5 weeks and interacting with native speakers (Schmidt and Frota, 1986). Data analysis showed a significant association between recorded noticings in the form of diary entries and Schmidt's use of linguistic forms (Schmidt and Frota, 1986). Comparing his performance on 21 verbal constructions it was found that he had been taught 14 of these. Analysis appeared to suggest that presence of forms, frequency in input and comprehensible input seem to have played a role to some extent. However, in later writings, Schmidt (2001) acknowledges the limitations of diary studies. He also acknowledges, based on evidence from priming and naturalistic studies, that acquisition is possible without formal explicit instruction. Thus he slightly revises his original strong view of noticing with a weaker view, allowing the possibility of implicit learning (see Schmidt 2001: 28, 31; Hulstjin, 2005).

Based on cognitive science research (inter alia Posner, 1990; Posner and Petersen, 1990; Posner and Rothbart, 1992) a model of attention has been proposed for second language learning by Tomlin and Villa (1994). According to them, attention is a limited capacity system and a process during which critical information is selected for further processing. It can be effortful or less effortful, i.e. automatic or controlled. Their more

[5] For Schmidt incidental learning is 'learning of one thing (e.g. grammar) when the learner's primary objective is to do something else (e.g. communication)' (Schmidt, 1994:16).

fine-grained definition of attention suggests that attention comprises three separate but interrelated components, namely alertness, orientation and detection, and is dissociated from consciousness and awareness (Tomlin and Villa, 1994; see also Posner and Petersen, 1990). Alertness is defined as the 'overall general readiness to deal with incoming stimuli or data' (Tomlin and Villa, 1994:190). Its levels can be manipulated and as a result processing or performance on tasks can be affected. Alertness is responsible for the speed of selection of information, and it can lead to rapid selection of information; however, this is sometimes at the cost of accuracy as the incoming information might not be adequately processed. For SLA, alertness can represent the readiness of the learner to deal with the input utterance, and it is also highly linked with motivation referring to the learners' interest in learning a language and/or the teacher's personality, rather than being directly associated with input.

Orientation holds a central role for SLA, as its key characteristic is to 'align' attention on a stimulus, in the same way as focused attention (Schmidt, 1990, 1993, 1995, 2001), suggesting that learners' attention can be oriented on specific aspects of the input. In this case, instructional techniques might be sufficient to cue the learner to allocate his/her attentional resources to the appropriate aspects of the input, thus making grammatical relationships easier to detect. For example, in IE, learners' attention might be directed through colours on the verb ending –s of the third person singular. Similarly, in PI, learners' attention is oriented to the target form through structured input (SI) activities that are considered to be forcing learners to pay attention and process the target form by eliminating redundant forms in the sentence. For example in order to decode when an action has taken place, a redundant form such as 'yesterday' is excluded from the target sentence *'Peter walked the dog'*. In this way learners are forced to process the *–ed* ending and establish that the action has taken place in the past. Although awareness is not necessary in this process, it might play a role in the process by enhancing either the learner's alert state or by specifically orienting the learner to the grammatical form (Tomlin and Villa, 1994). Like awareness, alertness and orientation may have an enhancing role in increasing the chances of detection; however neither is a necessary condition.

Detection is the most significant component of attention and is defined as ‘the cognitive registration of sensory stimuli’ that selects or engages a particular bit of information for further processing (Tomlin and Villa, 1994:192). Once input is available the learner must detect (1) the presence of some element of grammatical form; (2) a new or unusual character to the incoming information; (3) that there is a relationship between the two levels of grammatical form and mental representation. The learner must then send those observations off for further processing (Tomlin and Villa, 1994:196; see also Posner and Petersen, 1990).

Many researchers have adopted what has been defined as the strong view on noticing, i.e. that noticing is the only and sufficient condition for learning to take place (see Schmidt, 1990, 1993, 1995, 2001; Robinson, 2001, 2003; FonF studies such as Alanen, 1995; Doughty, 1991; Jourdenais et al. 1995 amongst others). Others have adopted a weaker view on noticing. They do not dismiss the importance of noticing, however, they dissociate attention from consciousness, meaning that there can be attention with or without conscious awareness, following Tomlin and Villa (1994; see also Carroll, 2001, 2004, 2006; Sharwood Smith, 1991, 1993 2008, 2009; Sharwood Smith and Truscott, 2004, 2005 and in prep; Truscott 1998; VanPatten, 1993, 1996, 2002, 2004, 2007 and PI studies). The essence of the debate therefore lies in the question of implicit versus explicit learning. Followers of the strong view argue that learning and therefore acquisition occurs through explicit learning. Followers of the weaker view, on the other hand, although they do not dismiss explicit learning, argue that learning and acquisition can be also implicit.

For the purpose of this study, I assume that learning and acquisition can be both implicit and explicit, and that detection can occur with or without conscious awareness. Detection without attention is synonymous to implicit learning, whereas detection with attention is equivalent to explicit learning. Explicit learning generates metalinguistic knowledge which is not synonymous with acquisition. However, I argue that it is a factor that promotes acquisition (Paradis, 2004, 2007; Sharwood Smith and Truscott, in prep; Truscott, 1998). IE and PI can be instructional methods that promote noticing in its weaker formulation, namely they can orient attention and therefore promote detection of the target form for further processing. However there is no guarantee that the internal cognitive mechanisms will perceive the incoming stimuli, nor is there a guarantee that the target form will be acquired (similar to Sharwood Smith, 1991, 1993, 2008; Sharwood Smith and Truscott, 2004, 2005, and in prep; Tomlin and Villa, 1994; VanPatten, 1996).

It is widely accepted within SLA that attention has a facilitating role in L2 development, while substantial research within the fields of IE and PI has attempted to provide further insight about the role of attention and how it can be operationalized in language learning. However, there are methodological difficulties with regard to evaluating whether learner's attention is at a conscious level or not during language learning. Despite the advances in technology, there are limitations on providing evidence regarding implicit processing and learning (see also Truscott, 1998). Another issue in the literature is the lack of a theoretical framework that could link the psycholinguistic concepts of attention to SLA (see also sections 2.2.4, 2.3.4 2.4, 5.2, 5.3, and 5.4 for further discussion).

2.2 Input Enhancement

The debate on the role of awareness in SLA has direct implications for the field of IE and more specifically the field of Textual/Typographical IE (TTIE), since past studies have been based on the theoretical underpinnings of Schmidt's idea of 'noticing'. It is therefore essential at this point to present and review focus on form (FonF) in order to disentangle the differences between TTIE and FonF, and gain an overview of the effects of TTIE. Therefore in this section a review of the theoretical underpinnings of IE and FonF[6] will be provided in order to underline the similarities and differences between the two approaches. In the second part of this section TTIE/FonF studies will be reviewed.

2.2.1 Theoretical underpinnings of Input Enhancement

IE was first proposed by Sharwood Smith (1991, 1993) and refers to a process by which input becomes salient to the learner either by external or internal factors. Hence, IE is a process that can be 'a result of deliberate input manipulation or it can be the natural

[6] FonF approach is limited in the present study only to the part that refers to Textual/Typographical Input Enhancement.

outcome of some internal learning strategy' (Sharwood Smith, 1991:1). Externally created salience happens when specific features of the L2 input are made more salient, i.e. enhanced; when the teacher manipulates the input, for example the ending –s of the third person singular in English verbs. Internally created saliency, refers to internal learning mechanisms that make specific features in the input salient, i.e. when the verb ending –s has become somehow (with or without external input enhancement) salient to the internal mechanisms and this is evident in the learner's output, e.g. correct use of the verb ending –s in the third person singular (see Sharwood Smith, 1991, 1993; and also Figure 2.1 below). By comparison the first process refers to a more explicit intervention, while the latter describes implicit processes taking place in language learning (see below for question mark).

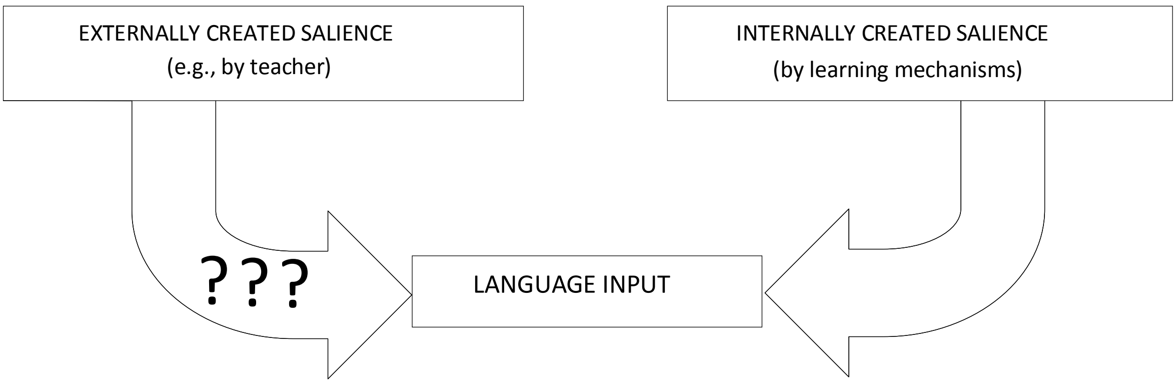


Figure 2.1: Input enhancement
(Source: Sharwood Smith, 1991:121)

The major theoretical underpinning of either perspective proposed by IE is that learners need to pay attention to specific items in the input before information can be further processed into the learner's language system. For IE, attention can have a facilitative role in language learning. The external manipulation of the target form through enhancement may induce saliency of the target form. IE does not guarantee that the input will become salient or that the learner's development will be affected. Indeed, learners may notice the signals but the input may not be salient to their learning mechanisms and therefore have no effect on development. Learners may also not notice the signals at all (Sharwood Smith, 1991). 'Noticing' in IE is similar to the propositions of Tomlin and Villa (1994), i.e. it is a process that does not require conscious registration, while it may involve alertness, orientation and detection of input. Sharwood Smith (1993:176) argues that when language is presented to the learner, a subset of signals may be either registered briefly in short-term memory and then lost without any further analysis, or registered by the learning device as violating the learner's current language system and therefore triggering some sort of restructuring of that system (similar to McLaughlin, 1990; Tomlin and Villa, 1994). These rationalisations explain the question mark in Figure 1, referring to the fact that we cannot know if the internal mechanisms will process the signal and how this is operationalized if they do. The fact that we cannot know how the internal mechanisms may perceive the externally created saliency may also support the use of metalinguistic explanation when applying IE. Sharwood Smith (1991:131) states that IE 'may work in ways unforeseen by the researcher' and that explicit metalinguistic explanation may be a necessary addition in providing a sophisticated rule explanation (Sharwood Smith, 1991). Another possible limitation proposed by Sharwood Smith (1991, 1993) can be the readiness of the learner for a growth in knowledge, according to which certain properties may be more noticeable than others at a given time (Sharwood Smith, 1991).

Similarly, Pienemann (1981, 1984, 1987, 1998) argues that learners' interlanguages tend to develop along predictable paths, whereas skipping these predictable stages is not possible (see studies by Clahsen et al. 1983; Meisel et al. 1981; Pienemann et al. 1999, 2005). According to his 'Teaching Hypothesis' (Pienemann, 1987, 1995, 1998), teachers should therefore teach what the learners are ready to learn, targeting the next stage of development.

In the literature there are various types of IE such as Input Flood; Corrective Feedback; Recasts; Elicitation; Metalinguistic Clues; Clarification Requests; Repetition of learner's error; Textual/Typographical IE and PI (see Wong, 1995, 2003; Lee and Benati, 2007). The focus of this research project is Textual/Typographical IE (henceforth TTIE) and PI. TTIE involves written input that is enhanced by visually altering its appearance in the text (see also Rutherford and Sharwood Smith, 1987; Wong, 2003:49). The target item can be manipulated in various ways: it can bolded, italicized, capitalised, underlined, highlighted or enhanced with colours. A further distinction between positive and negative IE is available. Positive evidence refers to naturally occurring samples of grammatical language and provides information about what is possible in the L2. For instance, highlighting in colour the ending of the German masculine definite article in the accusative case will draw the learner's attention to it and signal at the same time to the learner that this is the important feature of the form which he/she has to focus on. On the other hand, negative evidence provides information about what is not possible in the L2 grammar. For instance, highlighting in colour, underlining or providing corrective feedback of incorrect forms in the input can be types of negative evidence. Despite the possibility that positive and/or negative evidence may trigger changes in the learner's grammar there is, however, no guarantee that this will occur (Sharwood Smith, 1991).

IE techniques may vary in degrees of explicitness and elaboration. Explicitness refers to the sophistication and detail of the signal. At the highly explicit end we find metalinguistically sophisticated rule explanation. At the less explicit end a highlighted target item with colour for written input is possible. Elaboration refers to the amount of time involved in employing the enhancement technique. For instance, IE can be explicit and less elaborate (e.g. one short explanation about word order and case marking in German), more elaborate but less explicit (e.g. highlighting the ending of the German masculine definite article in the accusative every time that it occurs in a lesson), or it can be elaborate and explicit (e.g. a long explanation about the case system in German and the role of word order each time a noun appears in a lesson) and less elaborate and less explicit (e.g. underlining an error made by a learner once). The key in each possible situation is the externally induced signal that will draw learners' attention to the target item in question. Table 2.1 provides an illustration of the possible types of elaboration and explicitness in IE techniques.

Table 2.1: Elaboration and explicitness of consciousness raising in IE
(Sharwood Smith, 1991:120)

	Less Explicit	More Explicit
Less Elaborate	signal once when errors occur	short explanation once when error occurs
More Elaborate	short signal each time error occurs	long explanation each time error occurs

As a teaching intervention, IE endorses the possibility of both explicit and implicit learning. This is evident from the scale of elaboration discussed previously. Attention has a facilitative role in language learning, although it is difficult to determine the extent of this role (Sharwood Smith, 1991, 1993; Tomlin and Villa, 1994; VanPatten, 1996). For IE, attention is not synonymous with conscious attending, thus also allowing learning and acquisition through implicit processes that the learner does not become consciously aware of (Tomlin and Villa, 1994). For the purposes of the present study I argue that IE, as presented in this section, follows the theoretical underpinnings of attention as proposed by Tomlin and Villa (1994).

2.2.2 Focus on Form (FonF)

In a seminal paper, Long (1991) proposed a tripartite pedagogical intervention that aimed to draw learners' attention to form, namely focus on form (FonF), focus on forms (FonFS) and focus on meaning (FonM). FonF and FonFS are particularly relevant to this study. FonF has been defined as what 'draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication' (Long, 1991:45-46). Long and Robinson (1998:23) add that '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 with comprehension or production'. FonFS on the other hand, 'always entails isolation or extraction of linguistic features from context or from communicative activity' (Doughty and Williams, 1998a:3; see also Norris and Ortega, 2000:437-439). Long's (1991) use of the term FonF excludes any pedagogic practices that would require proactive planning, as FonF occurs incidentally and while learners are engaged in meaningful communication (Long and Robinson, 1998). This means that all classroom activities need to be based on communicative tasks, while any treatment of grammar should arise from difficulties in communicating meaning (Sharwood Smith and Trenkic, 2001; Sheen, 2005; Wong, 2005). On the other hand, FonFS is described as a 'traditional forms-in isolation type of grammar teaching' (Sheen, 2005:282). Long (1991:45-46) provides further explanations in defining the differences between the two approaches:

"Whereas the content of lessons with a focus on forms is the forms themselves, a syllabus with a focus on form teaches something else-biology, mathematics, workshop practice, automobile repair, the geography of a country where the foreign language is spoken, the culture of its speakers, and so-on and overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication" (Long, 1991:45-46).

The major theoretical underpinning of this type of FonF, especially in textual/typographical FonF studies, has been that focusing learners' attention on form promotes 'noticing', which is claimed as a necessary condition for any learning (Schmidt 1990, 1995, 2001; see also Doughty and Williams, 1998). These types of studies adopt a broader definition of FonF, which departs greatly from the original idea proposed by Long (1991). For Doughty and Williams (1998:4): '…focus on form entails a focus on formal elements of language […]. The fundamental assumption of focus on form instruction is that meaning and use must already be evident to the learner at the time that attention is drawn to the linguistic apparatus needed to get the meaning across'. DeKeyser (1998:62) adds 'This focus on form does not need to imply the return to a structural syllabus but, in my opinion, can, for certain learners, imply the explicit teaching and systematic practicing of certain forms'. Lightbown (1998:194) argues for 'a role for "grammar instruction" that is separate from communicative activities, and yet is integral to the lesson as a whole'. Clearly, the expanded definitions provided for FonF are quite similar to the original propositions of IE (see Sharwood Smith, 1991, 1993; Sharwood Smith and Trenkic, 2001; Wong, 2005).

Consequently, it is evident that there are two different types of FonF described: one focuses learners' attention to form in the context of incidental learning (Long, 1991; and Long and Robinson, 1998) and a second, broader term that focuses learners' attention in a predetermined focused instruction (Doughty and Williams, 1998 and studies in this volume; Doughty, 2001; Spada 1997; Lightbown, 1998; see also studies in Schmidt, 1995 volume; Wong, 2005). In the first case, learning is based on interaction, while in the second case, learning is promoted through noticing of the target grammatical form.

The ‘FonF approach’ proposed by Doughty and Williams (1998) posits that the key to learning is detection or noticing of the target form, following Schmidt’s (1990, 1993, 2001) propositions on noticing and acquisition. According to Doughty and Williams (1998) there are FonF techniques with low obtrusiveness, in which learners can merely detect the target form in the input, as well as obtrusive FonF methods, where noticing is promoted, resembling Sharwood Smith’s (1991) explicitness of IE techniques. Based on these assumptions, Doughty and Williams (1998:258) propose the following taxonomy of FonF tasks and techniques (Figure 2.2). According to this classification, IE is considered a less obtrusive teaching intervention than Input Processing, IP (to be discussed below). Interestingly, although Doughty and Williams (1998) argue for a FonF approach, in their schema (Figure 2.2) this is presented as IE. More important is the fact that they argue about obtrusive vs. unobtrusive techniques of IE, which can be put in parallel with the concepts of explicit vs. implicit techniques. However, based on the original propositions of the Noticing Hypothesis (Schmidt, 1990, 1993, 1995) they argue that such a distinction cannot be possible, as learning and as a result, acquisition can only occur with explicit teaching interventions.

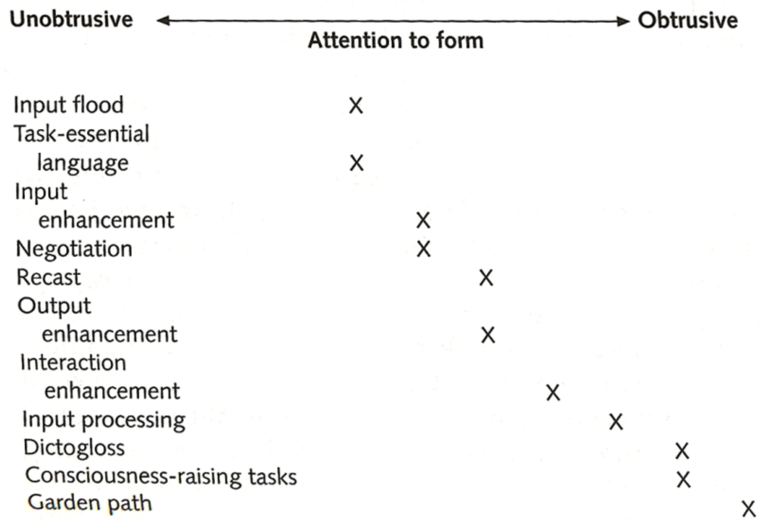


Figure 2.2: Degrees of obtrusiveness of Focus on Form techniques
(Source: Doughty and Williams, 1998:258)

Concluding, I argue that, contrary to IE’s theoretical underpinnings, the proponents of FonF adopt the strong view of ‘noticing being the only and sufficient condition for learning/acquisition to occur’ as proposed by Schmidt (1990, 1993, 1995, 2001). Long’s (1991) original use of the term FonF excluded pedagogical practices that required proactive approaches (see also Long, 2007). On the contrary, Sharwood Smith's (1991, 1993) IE definition is not restrictive in its pedagogical approaches and is not limited to interaction. Clearly, the TTIE studies that are reviewed next are not based on interaction but are rather designed and actualised in a classroom setting with predetermined grammatical forms in focus and pre-planned instruction. Thus, they depart greatly both in design and methodology from the definition that Long (1991) proposes and they resemble more closely the constructs of IE (see also Sharwood Smith and Trenkic, 2001; Wong, 2005). For these reasons, in the present study, the term adopted for these studies is TTIE, however reference to the terms FonF is going to be maintained and the two definitions will be used interchangeably for parity reasons.

2.2.3 IE studies

This section reviews IE studies that have investigated the role of TTIE. Many of these studies have been presented as FonF studies; however I have already extensively discussed the reasons why they are considered as IE studies in the present study. This section will provide further arguments supporting the TTIE definition, as well as an overview of the practical applications to date of TTIE for SLA purposes.

The first study that has investigated possible effects of TTIE was carried out by Doughty (1991). She investigated the effects of TTIE on explicit rules on the acquisition of relative clauses. It was assumed that instruction that targets marked relative clauses would generalize to unmarked contexts of relativization. The target form was the object of a preposition type of relative clauses e.g. *'This is the book that I was looking for'*. Twenty English learners of diverse L1 backgrounds without prior knowledge of relative clauses and an average length of stay in the US 3.7 months were randomly assigned to three groups: a meaning-oriented (MOG), a rule-oriented (ROG) and a control group (COG). A ten-hour exposure to the target form was comprehension-based. All three groups received reading passages, 5-6 sentences of a 'story' containing relative clauses formed on the object of a preposition type. Materials were administered in a Computer-Assisted-Language-Learning (CALL) environment.

The differences in the three groups were as follows: The meaning-oriented group received paraphrases and clarifications of the text in addition to the passages and the target form was also visually enhanced through underlining, colour and capital letters. The rule-oriented group received along with the text explicit grammar instruction through 'animated grammar', i.e. text sentences were first decomposed in two simple sentences. Identification and labelling of the relative and main clauses has followed. In turn, information explaining where the relative clause came from was presented. In the next step the relative clause was replaced from another main clause followed by information about the duplicated noun with respect to the verb phrase, where necessary. In the last step, the original relative and main clauses were identified and then the two clauses were separated into simple sentences. Once this process was over, a reverse process of a step by step re-composition of the text sentence was presented. The whole process lasted for four minutes for each sentence presented. At each step, both a rule and an animated process were presented simultaneously. The control group received only an unhighlighted version, no animation of the text and time to read the text. It is therefore evident that the treatment design for the meaning-oriented group included two variables, i.e. paraphrases and clarifications with textual enhancement of the target form through underlining, colour and capitalisation, thus making impossible to trace which of the two variables would be the most effective in the acquisition of the target form. On the other hand the rule-oriented group received an enhanced and highly focused version of explicit instruction, where each sentence was re-and de-composed in a process lasting for four minutes. Unfortunately, the treatment design shows therefore that groups were not equally distributed, thus potentially compromising the validity and reliability of the study.

Assessment tasks included: a free recall task on passage content; a grammaticality judgement task; a sentence combination task; a guided-sentence completion; and an oral test following a pre and immediate post-test design. Results of the study showed that all three groups (including the control group) improved in written and oral production of relative clauses. Significant differences were reported for the two groups when compared to the control, however no significant differences were revealed between the two instructional groups, thus, suggesting that both instructional techniques are equally effective with respect to improvement on relative clauses. The meaning-oriented group significantly outperformed rule-oriented group and control groups in comprehension, while there were no significant differences between rule-oriented group and the control groups. With regard to markedness, results suggested a positive effect for all three instructional treatments, as well as an apparent trend for improvement on relativization.

Doughty (1991:457-458) attributes obtained findings to instruction and exposure to marked data. She argues that they had an impact on the participants' relativization ability in other contexts. This remark is quite significant for the purposes of the present study, since the target structure is German OVS, which is also considered a marked structure. Equally important is the significance that Doughty places on the perceptual saliency achieved by IE (i.e. for MOG implicitly through TTIE, while for ROG through the animation), as well as the two levels of redundancy (i.e. for MOG textual enhancement and explanation, metalinguistic explanation and grammar animation for ROG) and the frequency of presentation of marked relative-clause forms were reasons for improvement in performance.

Doughty's study, although self-defined as a FonF study, is clearly a well pre-defined study in a set laboratory with clearly-defined target groups and in a highly controlled computerised environment. Thus, learners' attention to the target form does not occur incidentally; instruction does play a role. The overall design of the study therefore contradicts the original propositions of Long (1991) in order to establish the study as a FonF study. Furthermore, the study clearly contrasts implicit over explicit learning and no instruction, a design that is more closely related to the propositions of IE (Sharwood Smith, 1991, 1993). The fact that implicit learning is considered as a possibility to facilitate learning and acquisition is an additional contradiction to the main argument of the 'Noticing Hypothesis' (NH), as proposed by Schmidt (1990, 1993). More important is the fact that in both experimental groups, MOG and ROG, there is great effort to alert and orient learners' attention to detect the target linguistic form's properties, a fact that is more evident in the treatment of the ROG group, where each sentence is presented for four minutes using re- and de-composition of sentences and a step by step explanation of sentence structure and argument roles. The task design is clearly based on propositions of attention, as suggested by Tomlin and Villa (1994). However, the study fails to provide evidence supporting positive effects of TTIE, as the particular variable is not isolated but used in combination with lexical explanation. Long term effects for these types of instruction cannot be reported, as a delayed post-test was not conducted;

moreover the sample size of the study is too small to provide generalizable, valid and reliable conclusions for any possible effects of the proposed interventions.

Shook (1994) examined the effects of TTIE on Spanish relative pronouns and the Spanish present perfect. A sample size of 125, sixty first year and sixty five second year university level learners of Spanish were assigned into three groups: a group receiving TTIE passages with the target forms being enhanced through capitalisation, increase of character size and use of bold format; a group receiving passages with the same TTIE as the first group plus FonF rule instruction[7]; and a control group receiving passages without TTIE or FonF instruction. Treatment included two reading passages, one for each of the target forms (relative pronouns and present perfect). The study followed a pre and post-test design. Assessment tasks involved a multiple choice recognition and a cloze form production test. Results of the study showed no significant differences between the two experimental groups, although both outperformed the control group. According to Shook (1994), these results suggested an effect of TTIE rather than FonF, leading him to conclude that explicitly asking participants to pay attention to the target form did not significantly affect the results. On the other hand, the argument could be contradicted, since there was no difference between TTIE and TTIE plus FonF. The type of linguistic form tested appeared to have an effect on results, as findings showed that despite the fact that both the Spanish relative (*que* and *quien*) pronouns were typographically enhanced, gains after instruction were significant only for the present perfect. Shook (1994) therefore argues that present perfect verb forms are more important to understand the content of the passage than relative pronouns.

Overall, Shook's study is important as it explores the effectiveness of TTIE over TTIE plus FonF instruction, while its findings contradict the meta-analyses of explicit versus implicit TTIE studies (see Doughty, 1998, 2003; Lee and Huang, 2008; Norris and Ortega, 2000). TTIE, a more implicit and less obtrusive type of instruction than TTIE plus FonF, is more effective than TTIE plus FonF. More importantly, Shook's (1994) findings contradict the basic claim of the NH, in terms of learning/acquisition being a

[7] Shook (1994) oriented learners' attention of this group to the target form by informing participants to pay attention to the bold items, as they would be requested to formulate a rule for them. In addition, he oriented participants' attention to the meaning of the passage, as they were informed that a written recall of the passage would follow.

by-product of conscious processing (Schmidt, 1990, 1993, 1995). The study includes a sufficient sample size that could provide valid, reliable and generalizable conclusions for instructional effects; however lack of a delayed post-test does not allow claims for long-term effects of instruction.

Alanen (1995) investigated the effects of two types of IE, namely, TTIE (use of italics) and explicit rule explanation on the acquisition of locative suffixes. Thirty-six adult English learners of semi-artificial Finnish were assigned in four groups: a textual enhancement group; an explicit rule instruction group; a combined-textual and explicit rule instruction group; and a control group. In order to reinforce the focus of learners' attention on content, all participants were informed in advance that they would be tested on their comprehension of the given passages. During the two-day study phase, learners were asked to think aloud and verbal reports were recorded in order to measure noticing and rule awareness. Assessment tasks after the study phase included a surprise sentence completion task that tested learners' knowledge of the target features and a grammaticality judgement task, in which learners were required to explain why the sentence was ungrammatical and to provide and/or restate the rule that in their opinion governed the use of linguistic items.

Results from the sentence completion production task revealed that both the rule-oriented and the combined groups outperformed the TTIE only and the control groups; however no significant differences were reported between the rule-oriented and combined groups (as in Doughty, 1991). A positive effect for rules was revealed, since the rule-oriented group differed significantly from the TTIE only group. Comparisons between the TTIE and the control groups on correct rule statement suggested no significant differences. Thus, one can conclude that the variable TTIE in the production and rule statement tasks did not prove to have an effect on learners' performance. Alanen argues that closer examination of the production task showed that typographical manipulation of the target feature had a facilitative effect on learners' recall of locative suffixes, based on the fact that the enhanced group did learn one of the suffixes contrary to the control group. However, a stronger effect has been revealed for explicit rule instruction. Alanen attributes findings to the chosen textual manipulation (use of italics), concluding that it may not have been as salient as other types of TTIE. The analysis of the think aloud protocols suggests that the overall outcome in performance is highly linked to the learners' focused attention, as those who have shown improvement and acquisition of the target form appeared also to have more comments in the think-aloud, rule formulation and grammaticality judgement tasks. Given the obtained results it is therefore valid to conclude that Alanen's study findings provide support to Schmidt's (1990, 1993, 1995, 2001) claims of noticing at the level of understanding.

Alanen's study design can be defined as an IE rather than FonF study, as it takes place in a controlled environment, where the target form does not occur through spontaneous interaction. The fact that the target form is part of a semi-artificial language produced in a language laboratory further supports the claim. Alanen's use of think aloud protocols has been widely applied in studies defining themselves as FonF, as a valid testing measurement of learner' s 'noticing' of the target form (see studies Alanen, 1995; Jourdernais et al. 1995; Leow, 2001, 2003; Leow and Morgan-Short, 2004 among others). Think aloud protocols may indeed provide a 'window into mental processes', however they can also carry 'positive and/or negative reactivity', i.e. they may cause learners to perform more accurately and systematically (positive reactivity); or in the case of 'negative reactivity' they may negatively affect learners' performance (Godfroid, Housen and Boers, 2010:174). Thus, verbalisation of cognitive processes may influence the very cognitive processes one is aiming to describe (ibid). Godfroid, Housen and Boers (2010) point out that non-metalinguistic verbalisations require verbalisation of learner's thoughts, whereas metalinguistic verbalisations require additional information from the learners such as justifications and explanations for what they are thinking (see also Ericsson and Simon, 1993; Bowles and Leow, 2005). Studies and analyses conducted on reactivity in L2 reading tasks have found that concurrent verbalisation is not reactive for accuracy; however it increases the time required to complete the task (Bowles, 2008; Bowles and Leow, 2005; Leow and Morgan-Short, 2004; Sachs and Polio, 2007; Sanz et al. 2009). One could argue that think aloud protocols should not only be employed but also analysed with care, while the inclusion of response time reaction can provide more accurate information about cognitive processes involved (Godfroid, Housen and Boers, 2010; Sachs and Polio, 2007; Sanz et al. 2009). In this vein, the deployment of think aloud protocols in Alanen's (1995) study fully justifies the adoption of Schmidt's (1990, 1993, 1995, 2001) NH, regarding language acquisition. Findings of her study do not support Schmidt's claims but rather show that external input manipulation may not be effective in internally enhancing the target form, as Sharwood Smith (1991, 1993) and Tomlin and Villa (1994) argue. However, the study fails to provide data for long-term effects, since it applies an immediate post-test only, while the sample size of thirty six participants divided into four groups is not sufficient in order for conclusions to be generalizable.

Jourdenais et al. (1995:183) also examined the effects of TTIE in making the target form more noticeable to learners. Fourteen English learners of Spanish at university level participated in a study which investigated their ability to detect Spanish preterit and imperfect verbs in a written text. Participants were already familiar with the target structure. In order to avoid priming effects, instead of a pre-test, prior knowledge of the target form was measured through a midterm examination. Participants were randomly assigned to two groups, a TTIE, and a control group. The enhancement applied in the study was shadowing for the preterit verbs and bold format for the imperfect verbs, while both forms were additionally underlined and enlarged in font size. The treatment included a written narration from the story of 'Little Red Riding Hood' (enhanced vs. unenhanced versions). Assessment tasks included a think aloud protocol and a written pictured-based narration task. Analysis of the think aloud protocols reported significantly more episodes with preterit and imperfect verbs for the enhanced group participants than for the control group. Moreover, the enhanced group outperformed the control group in the written production task. Jourdenais et al. (1995:206) attribute these results to priming caused by increased registration of the input stimulus through the typographical manipulation. They conclude that the IE applied is effective in promoting salience and in turn noticing, in terms of conscious registration of the target form (as proposed by Schmidt, 1990, 1993, 1995, 2001), which in turn impacts on learners' production and output.

Jourdenais et al.'s study is another example of a TTIE that has been labelled as a FonF study. However, based on previously stated arguments the study clearly follows an IE approach in its treatment design: it is controlled and the target form does not surface through interaction. Moreover, the deployment of think aloud protocols is more closely linked to the exploration of conscious cognitive processing, as metalinguistic knowledge, use and explanation of the target form is required when thinking aloud. The results from the think aloud protocols verify the 'positive reactivity' effects of think aloud protocols (Bowles and Leow, 2005; Ericsson and Simon, 1993; Godfroid, Housen and Boers, 2010). Results however, should be interpreted with caution, as the target form investigated is not a novel form for participants; thus not only priming effects but as well individual differences may have played a significant role in the obtained findings, factors that are also acknowledged by Jourdenais et al. 1995 (see also Stratman and Hamp-Lyons, 1994 for think aloud protocols and priming effects). Moreover, non-inclusion of delayed post-test data and lack of sufficient sample do not allow generalizable conclusions.

Overstreet (1998) carried out a conceptual replication of Jourdenais et al. (1995) while adding the variable 'content familiarity'. Fifty adult English learners of Spanish were randomly assigned to four different groups: a group that received textual enhancement and was familiar with the content; a group that received textual enhancement and was not familiar with the content; a group that received an unenhanced version of the text but was familiar with the content; and a group that received an unenhanced version of the text and was not familiar with the content. The target form, the Spanish preterit, was typographically manipulated through underlining, shadowing and the use of different and larger fonts; while the imperfect tense was manipulated through underlining, bolding and using a larger and different font than the rest of the text. Two versions, an enhanced and an unenhanced version, of two passages, 'Red Riding Hood' and 'A letter to God' were developed. A pre-, post-test design was applied, while the assessment tasks measured comprehension via a true/false comprehension quiz and intake via a narration and a circle-the-verb task. Results from the study showed a negative effect for TTIE in the comprehension task. This was attributed to the multiple enhancements, leading to the hypothesis that multiple textual enhancements distracted learners (Overstreet, 1998). In addition, no significant differences were revealed in terms of comprehension and intake of the target form favouring either textual enhancement or content familiarity.

Overstreet's findings are in line with VanPatten's (1990) earlier findings supporting the idea that learners have difficulties focusing on forms and meaning simultaneously, especially when the form in focus is not crucial to the meaning. Processing of two forms at the same time as in Overstreet's study may have been too cognitively demanding. Moreover, the fact that texts contained many enhanced forms seems to have a negative effect on processing and acquisition of the target form. The pre- post-test design followed in this study does not provide data regarding long-term effects (if any) of the proposed intervention, while the sample size is not sufficient to provide generalizable conclusions.

Leeman et al. (1995) compared two types of content-based instruction. A FonF instruction that used textual enhancement to orient learners' attention to the target forms while keeping attention to meaning and a purely communicative type of instruction that focused solely on meaning without drawing learners' attention to the target form. Thus, the aim of the study was to investigate whether IE techniques could improve L2 accuracy of L2 forms, while keeping meaning in focus. Again the target structure was Spanish preterit and imperfect verbs. Twenty-two English advanced learners of Spanish at university level divided into two intact classes participated in the study. They were randomly assigned to one of the two groups, namely FonF versus communicative type of instruction. The treatment materials were the same for both groups with the only difference being the enhanced forms (use of underlined, highlighted and colour-coded enhancement) directions and feedback for the FonF group. Treatment materials included reading comprehension tasks, debates and content-based error corrections on classmates' performance. Testing tasks included a battery of tasks, namely: reading passages; in-class debates that required home preparation; essays; a picture-based judgement task; and a cloze task. A pre-test, a two day treatment and a post-test (one week after treatment) design was followed. Results from the study revealed that the FonF group significantly outperformed the purely communicative group in all tasks. The communicative group improved only on the essay task, while only the FonF group showed significant use of the target forms in obligatory contexts. Leeman et al. (1995) concluded that the FonF type of instruction containing both TTIE and corrective feedback is more effective than communicative only types of instruction and they attributed the beneficial effects of FonF instruction to the enhancement of the target form.

This study fails to investigate TTIE as an isolated variable; however, findings reveal that enhancement that promotes attention to form and meaning is more beneficial than enhancement promoting solely attention to form. It is valid to assume that the view adopted on ‘noticing’ is more closely linked to the ideas proposed by Tomlin and Villa (1994), while treatment and testing design is more closely linked to proposals developed by Sharwood Smith (1991, 1993). A delayed post-test could have provided evidence of long-term effects of the proposed type of instruction, while a greater sample size could have provided valid, reliable and generalizable conclusions.

Leow (1997a) investigated the effects of text length and textual enhancement on learners’ comprehension and intake of Spanish formal imperative verb forms. Eighty-four adult English learners of Spanish were randomly assigned to one of the four conditions: a long enhanced text; a long unenhanced text; a short enhanced text; and a short non-enhanced text. The TTIE applied in the study was underlining and bolding the target forms in the texts. Treatment involved four versions of four enhanced and unenhanced, short and long passages for each of the four group conditions. However, only fifteen of the common target forms in the short and long passage version were typographically enhanced during treatment. Assessment tasks involved a short-answer comprehension task that required participants to respond in their L1 and a multiple choice recognition task that measured intake. Results of the study showed a significant effect of text length on comprehension, favouring the short version. However, no significant differences were reported for TTIE on either comprehension or intake. Perhaps longer exposure to the target form might have resulted in different findings, as learners’ exposure was very brief, since participants read the enhanced and/or unenhanced versions of the text only once and the typographical cues were limited.

White (1998) combined input flood with textual enhancement. She examined the acquisition of English possessive determiners over a ten-hour-treatment period of two weeks. Eighty-six Francophone primary school level children were distributed in three groups: a group with exposure to a reading package with the target forms typographically enhanced and an additional supplementary reading program over a course of five months; a group with exposure to a reading package with the target forms typographically enhanced without the supplementary reading; and a control group with exposure to the reading package without enhancement or supplementary reading, yet non target form *–ed* endings of the verbs were enhanced. The target form was enhanced through italics, underlining, bolding and enlarging of font size. A pre-test, an immediate and a delayed post-test five weeks post-treatment were administered, including a passage-correction, a multiple-choice and a picture-description task. However only results from the picture-description task are reported in White. Results indicated an overall quantitative difference between groups in the use of possessive determiners from pre- to immediate post-test condition, as all groups used both more correct and incorrect target form features in the immediate post-test condition. However, only the enhanced treatment groups were significantly higher. White (1998) concluded that TTIE promotes increased frequency of use of target forms. However, the great variability in correct and incorrect use of the target form suggests that the learners have not yet acquired the form.

White's study provides evidence regarding the impact of IE in SLA in young learners, as most of the studies reviewed look at adult SLA. The pre-defined study design leads to the valid conclusion that this too is an IE study rather than a FonF, as extensively argued in this section. Treatment duration compared to previous studies reviewed is far longer while delayed post-test data are included in the particular study. Unfortunately, substantial findings from the study, such as results from the passage-correction and multiple choice tasks are not reported, which could have provided greater insight about comprehension and accuracy performance as well as an account for performance variability.

Wong (2003) examined the effects of TTIE and simplified input on learners' comprehension of gender agreement in French past participles in adjectival relative clauses. Eighty-one adult English learners of French were assigned to one of four groups: a textual enhancement group with simplified passages; a textual enhancement group with unsimplified passages; an unenhanced and simplified group; and an unenhanced and unsimplified group. The enhanced version manipulated the entire noun phrase, as the target structure was not only enlarged, but also bolded, italicised and underlined; while the definite article of the head noun was also enlarged, bolded and italicised but not underlined. In addition, a simplified version of the text was created focusing only on lexis and sentence structure without including the target structure.

Assessment tasks were a free-writing recall task measuring comprehension and an error-correction task measuring intake. Data measuring comprehension were analysed for total number of idea recalls and for total number of target form recall. Results showed a significant effect for TTIE of target form recall as well as significant improvement in the error correction task. However, since textual enhancement as a variable is not isolated in this study, i.e. it was applied in combination with simplified vs. unsimplified versions of the text, improvement in error correction performance was attributed to exposure to input and not enhancement or simplification, given the fact that no significant differences between the two conditions were revealed.

Wong clearly adopts the term IE, as proposed originally by Sharwood Smith (1991, 1993); however she does not support Tomlin and Villa's (1994) dissociation of attention and awareness, but on the contrary adopts Schmidt's (1990, 1993, 1995, 2001) position that awareness is necessary for intake derivation (similar also to Robinson, 1995; see also Simard and Wong, 2001 for a detailed discussion). The off-line measures applied in her study i.e. the recall tasks, have been widely criticised in terms of not being a valid and reliable way of measuring noticing of the target form (see Bowles, 2008; Godfroid, Housen and Boers, 2010; Leow, 2006; Sanz et al. 2009). With this in mind it is valid to assume that Wong (2003) does not investigate the role of noticing, in terms of conscious processing and registration in language learning, as proposed by Schmidt (1990, 1993, 1995, 2001). This is also evident from the fact that recall tasks were applied to measure comprehension. Thus it can measure effects of TTIE. Wong's IE study is therefore relevant to the purposes of the present study as it investigates whether there is recall and by extension an effect of textual manipulation on SLA.

2.2.4 Conclusion on IE studies and contributions of the present study towards IE

Many TTIE studies have been self-defined as FonF studies. In this chapter it is argued that all studies that take place in a controlled setting (language laboratories and/or classrooms)[8], where the researcher has a pre-determined set of treatment materials and testing tasks and where the target form does not emerge through interaction, as originally proposed by Long (1991), are viewed as IE studies based on the propositions

[8] I am not claiming that studies that take place in a classroom setting cannot be FonF studies. However, if the target form is not the outcome of social interaction then the present study adopts the IE construct.

of Sharwood Smith (1991, 1993; see also Polio, 2005; Sharwood Smith and Trenkic, 2001; Wong, 2005). With this in mind, note that many of the studies reviewed above investigate morphosyntactical acquisition. The target form is manipulated, in many cases provided through explicit instruction, thus activities are not based on purely communicative tasks, nor does grammar treatment and/or 'noticing' arise from difficulties in communicating meaning (Sharwood Smith and Trenkic, 2001; Sheen, 2005; Wong, 2005). This can be attributed to the fact that numerous studies investigating the possible effects of instruction have embarked the critical issue of helping learners direct their attention to form and notice a linguistic mismatch between their interlanguage and the target language (e.g., Ellis 1998, 1999; Sharwood Smith 1981, 1991; Swain 1995; Swain and Lapkin 1995; VanPatten and Cadierno 1993a, 1993b) based on the assumption that acquisition of grammatical features requires noticing a language form (similar to Schmidt 1990, 1992, 1993, 1995, 2001; and/or Tomlin and Villa, 1994). Muranoi (1996:3) attributes the contraction of FonF research in morphosyntax to 'the need to determine optimal ways to incorporate form-focused instruction into meaning-oriented communicative language teaching'. A further possible explanation of the direction of FonF studies can be attributed to the sharp contrast between FonF and FonfS, which is conceptualised as the contrast between communicative language teaching versus traditional grammar instruction (see Long, 1988, 1991a, 1991b; Doughty and Williams, 1998; VanPatten, 1996; among many).

The majority of the studies claim to be investigating the role of noticing in SLA as proposed by Schmidt (1990, 1993 1995, 2001). Operationalizing and measuring awareness in SLA has been a vexed issue because of the methodological difficulties in investigating such complex cognitive processes while processing L2 input. In the studies reviewed above, noticing has been measured based on offline/retrospective procedures, i.e. data collected after exposure to the L2 data (such as Doughty, 1991; Leeman et al. 1995; Overstreet, 1998; Shook, 1994; White, 1998; Wong, 2003); and/or online/introspective procedures, i.e. data collected while learners are interacting with the L2 data (such as Alanen, 1995; Jourdenais et al. 1995; Leow, 1997; Leow, 2001). Offline measures have been widely criticised for the low internal validity of the data, as they may not reflect truly what learners became aware of while exposed to the input (see Leow, 1997, 2001; Robinson, 1995a; Shanks and St. John, 1994). Online procedures have been operationalized through the use of think-aloud protocols (TAP), which are also limited in terms of the information they are able to provide, i.e. ‘positive and negative reactivity’. Furthermore, analysing TAP is considered a very delicate and sensitive procedure with questionable outcomes, regarding validity and reliability (see Bowles, 2008; Bowles and Leow, 2005; Godfroid, Housen and Boers, 2010; Ericsson and Simon, 1993; Leow and Morgan-Short, 2004; Sachs and Polio, 2007; Sanz et al. 2009)[9]. The present study is limited to the investigation of the possible effects of noticing in SLA post-exposure. However, this study will discuss the possible impact that the three attention-drawing types of instruction (PI, IE and the combination) might have on input processing and acquisition.

Most of the studies have reported a positive effect of IE in language learning, while few (Leow, 1997, 2001; Overstreet, 1998) have reported no effects of IE. However, conclusions could become robust, generalisable and reliable if delayed post-tests were also included in the test design of the majority of the studies[10]. Furthermore, in spite of positive and/or negative effects of IE reported in studies, there is no theoretical explanation describing the cognitive processes involved for the type of enhancement applied and the implications of the effect of enhancement on these processes. Moreover, there are no explanations regarding why and how one aspect of TTIE is more effective than another type, how much, what type and for which target form IE exposure is more effective. These are some of the questions that remain unanswered (see also Sharwood Smith and Trenkic, 2001).

Williams and Evans (1998:139) argue that ‘it has not been clear exactly what it means to draw a learner’s attention to form or how this is to be accomplished’. From a pedagogical perspective TTIE has failed to provide a consensus both at a theoretical and practical level regarding which features of the linguistic forms should be enhanced, i.e. the ending of the verb, or the whole verb, or the verb in combination with the subject, when using TTIE for teaching, for example, the third person singular –s; and how to enhance them better in order for the target form to become salient for the learner. For

[9] Perhaps a better insight about cognitive processes involved, such as noticing, could be provided with the use of eye-tracking, which can enhance the validity and reliability of studies (see also Godfroid, Housen and Boers, 2010; Sanz et al, 2009). However, this is beyond the research design of the present study.

[10] The reader is reminded that apart from White (1998), all remaining and reviewed studies in the previous section have adopted a pre-, immediate post-test design.

example, should a specific colour be used for the ending of the verb and the same or a different colour for the subject? (see also Sharwood Smith and Trenkic, 2001). With this in mind, although IE has been presented as a promising teaching intervention in foreign language teaching and learning it has failed to inform stakeholders.

The present study attempts to explore possible effectiveness of TTIE in SLA by isolating the variable colour in TTIE. Findings of the present study will be discussed based on the theoretical platform provided in this chapter. In this way, the present study will attempt to address criticism that IE studies to date have received, as well as explain the impact that TTIE teaching intervention might have on this instance, the acquisition of German word order and case marking.

2.3 Input Processing

Introduction

Input Processing (IP), proposed by VanPatten (1990, 1996, 2002, 2004, 2007, 2009), provides the theoretical framework of Processing Instruction (PI) also applied in the present study. IP is based on the fact that learners must simultaneously understand an utterance while their internal processors must map a meaning to a form when encountering input. IP is therefore concerned with the initial processes involved in input processing for SLA purposes. Assuming that there is an integral part of language processing and potentially acquisition in making correct form-meaning connections which is not explained, IP claims to be addressing three key points: firstly, under what conditions do learners make initial form-meaning connections? Secondly, why do learners make at given times some and not other correct form-meaning connections? Thirdly, what are the internal strategies that learners use for comprehension and how can these affect language acquisition (VanPatten 2004, 2007:116, 2009)?

In this section, the theoretical underpinnings of Input Processing followed by the theoretical underpinnings of Processing Instruction will be reviewed. The second part of this sub-section will review the practical application of IP as it is expressed through PI, while the third part will provide an overview of the PI studies conducted to date which are relevant to the purposes of the present study.

2.3.1 Theoretical underpinnings of Input Processing

Given the fact that IP is concerned with the initial stage of input processing, comprehension is the key element, since acquisition is believed to be similar to Krashen (1985) the by-product of comprehension (VanPatten, 2009:59). However, for VanPatten comprehension alone cannot guarantee acquisition, if the learner cannot make appropriate form-meaning connections during the act of comprehension (similar to White, 1987; Carroll, 2001; see VanPatten, 2007, 2009), thus departing from Krashen. In this process, attention holds a prominent role, since forcing learners to pay attention promotes the decoding of information, i.e. comprehension, which in turn leads to acquisition. The way attention is viewed in IP is similar to the propositions of Tomlin and Villa (1994)[11]: it is seen as effortful and of limited capacity therefore the selection of input is necessary. Comprehension is therefore an effortful process that consumes a great deal of attentional capacity. Because of these limitations, learners first allocate attentional capacity to detect content words in the input in order to decode the meaning. Thus, grammatical forms that convey little semantic information will remain undetected, and as a consequence unprocessed, as learners will rely on lexical items rather than grammatical forms in order to decode meaning (similarly to the propositions of the Competition Model, Bates and MacWhinney, 1989 see section 2.1.2). For example, in a sentence like *'Yesterday I walked the dog'*, the lexical item *'yesterday'* has a high Communicative Value (CV) because it contains a great amount of semantic information. If learners decode the semantic meaning of yesterday' they do not need to process the grammatical ending *–ed* of the verb in order to establish when the action takes place in this example. Thus, the ending of the verb will remain undetected and therefore not processed.

The Communicative Value (CV) of a form refers to the meaningfulness of the form in contributing to the overall sentence meaning. CV has two features, namely [± inherent

[11]Conscious noticing as proposed by Schmidt (1990, 1993, 1995, 2001) is interpreted in IP in terms of a necessary condition in drawing learners' attention to the form in order to process grammatical forms to decode the meaning of a sentence (see VanPatten, 1996, 2004).

semantic value] and [± redundancy], suggesting that the communicative value of the form can be defined as high or low depending on whether meaning can be retrieved from the form itself. According to this proposition, a form's communicative value is greater when it is [+semantic value, -redundancy] than when it is [+semantic value, +redundancy], whereas forms with [-semantic value] regardless of redundancy do not contain any communicative value. However, there are exceptions to this rule. Some grammatical items are considered of high Communicative Value (CV). For example, in English verbal morphology the *–ing* ending is considered to be of high CV because it encodes progressive aspect, i.e. *–ing* = in progress and secondly because in discourse it rarely is redundant, i.e. no lexical information precedes and/or co-occurs that can provide cues to aspect (VanPatten, 1996:24). Moreover, there are cases where the learner's ability to obtain meaning from the input is effort-free (automatized), thus freeing attentional resources for the detection and processing of grammatical forms that have been previously skipped (VanPatten, 1996, 2004; similar to McLaughlin, 1990).

In the same way as Slobin's (1973) 'Universal Operating Principles'[12], VanPatten (1996, 2000, 2004, 2007, 2009) proposes a set of Input Processing (IP) Principles, that predict the strategies[13] that learners use when processing the input and propose alternative ways to process input in order for learners to make correct form-meaning connections (see also Peters, 1985). From a psycholinguistic perspective, IP Principles predict the cognitive steps that learners take when processing input, and from a pedagogical point of view, they help instructors produce materials for Processing Instruction (PI). These principles aim to alter learners' strategies in order for them to be able to make correct form-meaning connections. VanPatten (2004) presents the revised set of IP Principles and Sub-Principles (the full list of Principles is available in the Appendix A) from the original VanPatten (1996). For the purposes of the present study the review will be limited to Principle 2, the 'First Noun Principle' (FNP), and its sub-principles. The First Noun Principle states that 'learners tend to process the first noun or pronoun they encounter in a sentence as the subject/agent' regardless of their L1 grammar (see MacWhinney et al. 1985; Ervin-Tripp, 1974; Gass, 1989; LoCoco, 1987;

[12] According to Slobin (1973), when children are learning their first language they rely on certain basic and universal strategies while processing input, which he refers to as 'Universal Operating Principles' (see also Peters, 1985).

[13] Strategies in this context are synonymous with the cognitive steps that learners take when processing input (VanPatten, 1996, 2002, 2004, 2007, 2009).

Pienemann, 1987; while the sub-principles refer to lexical items (sub-principle P2a); event probabilities (sub-principle P2b) and/or context (sub-principle P2c) that may prohibit processing of the target form, as learners will rely on them in order to decode the meaning (see Figure 2.3 below):

P2. The First Noun Principle. Learners tend to process the first noun or pronoun they encounter in a sentence as the subject/agent.
P2a. The Lexical Semantics Principle. Learners may rely on lexical semantics, where possible, instead of word order to interpret sentences.
P2b. The Event Probabilities Principle. Learners may rely on event probabilities, instead of word order to interpret sentences.
P2c. The Contextual Constraint Principle. Learners may rely less on the First Noun Principle if preceding context constrains the possible interpretation of a clause or sentence.

Figure 2.3: The First Noun Principle and Sub-Principles
(VanPatten, 2004:18)

Based on the fact that languages vary in their word order, as some are strictly Subject-Verb-Object (SVO), others Object-Verb-Subject (OVS), while others allow SVO, OVS and Subject-Object-Verb (SOV) structures, VanPatten argues that in IP second language learners assign grammatical meaning (subject vs. object) and semantic roles (agent vs. non-agent) to the nouns during sentence processing in order to decode the message of the incoming information (VanPatten, 2004:15; similar to Bates and MacWhinney, 1989; Carroll, 1991, 2001, 2007 among many). The FNP predicts for example that an English learner of German will process the first noun as the subject/agent in both SVO and OVS German sentences, as English has a strict SVO structure, as the examples in the next page illustrate:

(1a) Der Mann küsst die Frau.
The man, subject/ agent kisses, Verb the woman, object/recipient.
(The man kisses the woman)

(1b) Den Mann küsst die Frau.
the man, object/recipient kisses, Verb the woman, subject/agent.
(The woman kisses the man)

According to the FNP predictions, in both sentences the English learner of German will interpret the first noun as the agent of each sentence. In the case of 1a, the learner will make a correct form-meaning connection, without necessarily processing the case marking but relying on the FNP. In the second example, 1b, the learner will make an incorrect form-meaning connection, as he/she will fail to process case marking of the nouns, relying again on the FNP. Consequently, if English learners are left to their own devices, there will be delays in the acquisition of German OVS structures and case marking (VanPatten, 2004). It is therefore valid to assume that IP makes the same assumptions for the role of 'cues' as the Competition Model (Bates and MacWhinney, 1989) and feature unification as Carroll (1991, 2001, 2006, 2007) and Pienemann (1984, 1998, 2007).

For VanPatten (2009), all learners, regardless of the word order of their L1, will be confronted with this parsing problem predicted by the FNP , since 'the tagging of the first noun as the subject is a universal processing procedure and not one derived from the L1'(VanPatten, 2009:51-52). Carroll (2004) argues against the universal application of the FNP, referring to topic prominent languages such as Chinese, where studies have shown that topic prominent learners are not always map the Agent semantic role onto the first noun phrase of the sentence (see Carroll, 2004:305; also Rutherford, 1982, 1983, 1987, 1988; Bates et al. 1982; MacWhinney et al. 1984; MacWhinney and Bates, 1989; Xiao, 2002). She argues that the FNP must be considered as a strategy typical for English speakers or other speakers of subject-prominent languages and not as a 'universal principle' of input processing (Carroll, 2004:304). On the other hand there have been studies, initiated from other theoretical frameworks, such as the 'Processability Theory' and the 'Competition Model', that lend support to VanPatten's claims regarding FNP being a universal principle (see from the Processability Theory cross-linguistic studies from DiBiase, 2005; Håkanson et al. 2002; Kawaguchi for Japanese, 2005; Pienemann and Håkanson, 1999; Pienemann et al. 2005; and from the Competition Model, studies from Bates and MacWhinney, 1989; MacWhinney, 2005, among many). Perhaps, VanPatten's claim could be reinforced if he provided further explanations regarding the role and implications of UG and L1 access and/or transfer in the L2 in his IP model. Nonetheless, VanPatten (2009) is in agreement with Carroll's (1991, 2001, 2007) view on acquisition being 'failure driven', as he argues that L2 learners will be able to overcome the FNP problem when 'the facts of the real world do not match the parsing' (VanPatten, 2009 cf. White, 1987).

The effects of Principle 2 (FNP) can be constrained by the sub-principles 2a (the Lexical Semantics Principle), 2b (the Event Probabilities Principle) and 2c (the Contextual Constraint Principle), as the following examples illustrate:

(2a) The zebra was kicked by the horse.

(2b) The fence was kicked by the horse.

In the first example (2a), the two animate nouns pose a processing difficulty for the learner in deciding which noun is the subject/object of the sentence. Principle 2 would predict in this case that the learners will assign the subject/agent role to the first noun of the sentence, in this case ‘the zebra’, given the FNP, and will fail to process case. In example (2b), on the other hand, lexical semantics, event probability and context allow the learner to make the correct form-meaning connections and assign correctly the role of the nouns. The appearance of one animate and one inanimate noun enables the learner to assign correctly the subject-object- agent-recipient roles by simply processing the words according to the lexical semantics principle-Principle 2a). Thus, the learner can conclude that the ‘fence’ cannot kick ‘the horse’ and therefore assume that the horse is the one kicking the fence. Through the same example the Event Probabilities Principle (2b) could also be illustrated, as the learner can simply, through the encoding of lexical semantics, fall under the influence of this sub-principle when assigning role to the nouns of the sentence, by his/her ‘world knowledge’. In this case, semantics and pragmatics can assist the learner in reaching the conclusion that the fence cannot kick the horse. In other words, if the sentence that the learner needs to process entails elements that can provide hints and/or can be helpful through world knowledge, then the learner will rely on this knowledge to establish the event that the sentence is describing rather than process the case marking and word order of the noun. It should be also noted that in this way evidence could be provided in order for the learner to change his grammar. The same is suggested by the sub-principle 2c, namely that if the context can provide information for the learner to decode the message then the learner will rely less on the FNP in interpreting the message of the sentence.

IP Principles do not operate in isolation, but quite often several principles redundant interact together, and one may take precedence over the other, as is illustrated in the presentation of Principle 2 and sub-principles (VanPatten, 1996, 2004, 2007, 2009). According to VanPatten (1996, 2004, 2007, 2009), the Principles do not attempt to account for all aspects of acquisition, neither do they constitute a total model. Instead, they form part of a model that attempts to account for the conversion of input to intake by psycholinguistic explanations.

VanPatten (2004) conceptualises SLA as the result of internal mechanisms that consist of a set of processes, similar to McLaughlin (1990). The first process involves input processing in which input is converted into intake (Phase 1). Intake for VanPatten (1996, 2002, 2004, 2007, 2009) is synonymous to filtered input, as originally proposed by Corder (1967). This means that learners during the first process of input processing filter the input they are exposed to, suggesting that not all the input is acquired and thus acquisition is not instantaneous (VanPatten, 1996). The second phase involves ‘accommodation’, in which the converted input, i.e. intake, is incorporated into the system; while ‘restructuring’ might be the result of the ‘accommodation’ process, depending on the data. This second process of accommodation and restructuring takes place in working memory. The third phase, access, involves linguistic data that have been integrated into the developing system and can be accessed through output and/or production, as the following figure illustrates (Figure 2.4):

$$\text{input} \xrightarrow{\text{I}} \text{intake} \xrightarrow{\text{II}} \text{developing system} \xrightarrow{\text{III}} \text{output}$$

I = input processing

II = accommodation, restructuring[14]

III = access, production procedures

Figure 2.4: A Sketch of Basic Processes in Acquisition
(Source: VanPatten, 2004:26)

[14]The terms accommodation and restructuring refer to how data makes it into the developing system and the impact this has on the grammar (VanPatten, 2004:7).

According to VanPatten (1996:31), ‘the job of input processing is to detect linguistic data and make initial form-meaning mapping, even if they are incomplete or partial’. The Figure (Figure 2.5) next page depicts how input processing works for grammatical form and role assignment according to the propositions of IP, having as an example the First Noun Principle (FNP)[15]. As information enters the processors, the processors will first search for content words and if resources are minimal then these lexical content words are delivered as intake into the developing system. In the event that resources are not depleted, then the processors undergo incremental stages of processing of the intake data and may make partial form-meaning mappings, which can in turn become form-meaning connections (subject to resource availability) before entering the developing system. This intake will contain grammatical features that were not processed initially and that contain only lexical items (Figure 2.5):

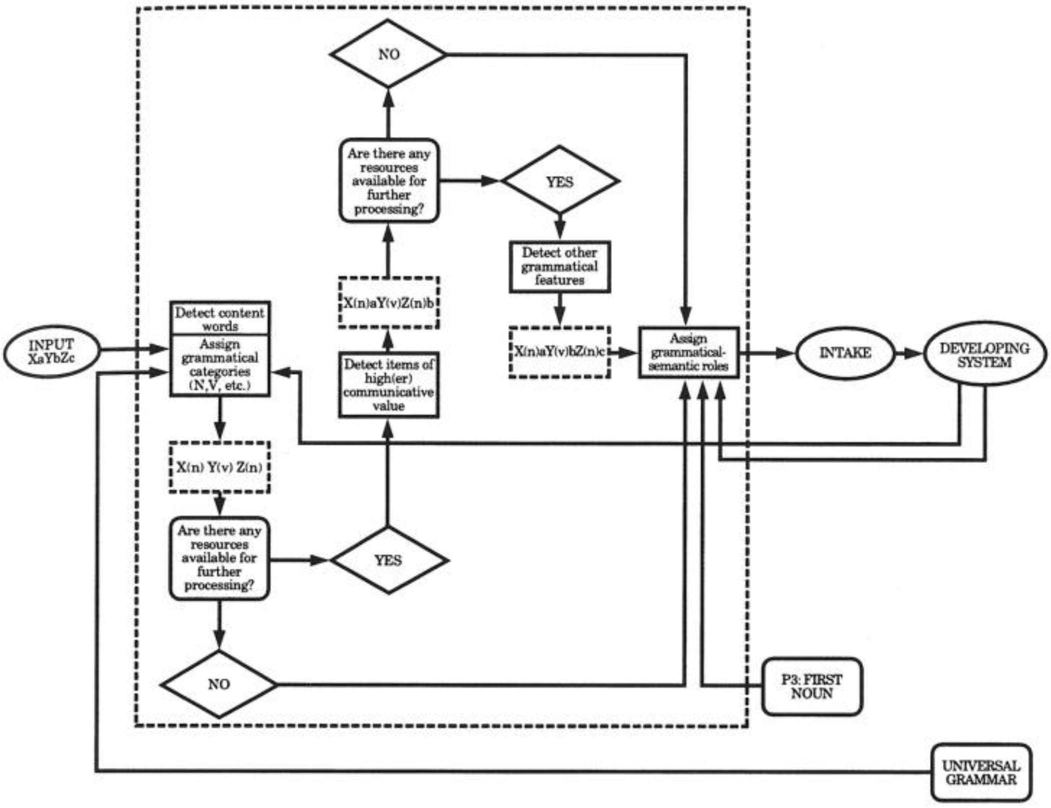


Figure 2.5: An integrated model of second language input processing for grammatical form and role assignment.
(Source: VanPatten, 1996:43)

[15] Please note that the formulae X(n) Y(v) Z(n) in the diagram symbolise the First Noun Principle. X(n) therefore symbolises the Object noun, Y(v) the verb and Z(n) the Subject noun.

Interestingly, VanPatten (2004, 2007, 2009) argues that IP is not a model of acquisition, nor does it attempt to describe L2 parsing and processing. IP is concerned with the initial stage of input processing and more specifically with the process of how learners parse the input in order to establish correct form-meaning connections (without dismissing or diminishing the role of output). IP does not provide a detailed explanation of how data are incorporated into the developing system (accommodation) or how they affect the system (restructuring); nor does it provide explanation about the role of UG, access and/or transfer, the role of output, the role of interaction and/or other factors that might stimulate or hinder acquisition (VanPatten, 2004, 2007, 2009). Yet, for these reasons, IP has been heavily criticised as not being a "comprehensive theory of second language grammar development" and being vague and inadequate in explaining how the form-meaning connections become part of the learners' developing system. It has also been judged as a poor basis for interpreting the findings of the PI approach (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Harrington 2004:80-81; Salaberry, 1997). The dated literature references that IP theory cites for key psycholinguistic notions such as attention and processing have been also heavily criticised (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Harrington 2004; Salaberry, 1997). It is clear that the IP model adopts a cognitive-processing approach to language learning and as such it should provide updated literature supporting and/or interpreting its claims on language acquisition or, as Carroll (2004:297) argues, IP should explicitly commit to an existing functional theory of language processing architecture of the language faculty as well as a parsing model, in order to be able to provide clear and detailed interpretations for data obtained from PI studies. Indeed, these are fair criticisms that merit attention from the IP and PI framework. Given the fact that VanPatten claims that IP is not a model of acquisition but PI has been shown to have a positive effect on language development, a theory of language should be adopted in order to explain in greater detail the theory of learning provided by PI (similar to Gregg's, 2003 property and transition theory claims). In this way a better understanding can be provided on how exactly IP and PI have an impact in the language learning mechanisms. As a result PI could be operationalised in such a way in order to be as effective (or even more effective) as originally proposed and its merits could be better appreciated.

VanPatten (1996:52, 2004) argues that the ‘Competition Model’ (Bates and MacWhinney, 1989) can be seen as useful in understanding how Principle P3(b)[16] develops; however it cannot explain how the remaining Principles and sub-Principles operate, including the First Noun Principle (FNP) relevant to the present study, since it cannot provide information about the initial stages of input processing, especially the process of detection and the overall role of attention (VanPatten, 1996, 2004, 2007, 2009). VanPatten (2009) has lately attempted to explain IP with reference to the Autonomous Induction Theory (AIT); see Carroll, 2001). Although he does not dismiss AIT, in terms of parsing being failure driven, he argues that the AIT account departs greatly from the FNP predictions. AIT sees the FNP predictions as the outcome of parsing failure processes whereas, in the IP model, the FNP predictions derive from the learner’s assumptions for sentence initial position, often based on their L1. However, it is not clear whether and to what extent VanPatten (2009) adopts Carroll’s (2001) propositions, by arguing that for learners to overcome the predicted Principle, ‘the facts of the real world [should not] match the parsing’ (VanPatten, 2009:52). Therefore the need to provide sufficient explanations of what the basic architecture of the language faculty is for IP remains unfulfilled (Carroll, 2004). This has a direct impact on PI studies which, despite the promising data obtained, have received criticism on the basis that they fail to link theory with practice, i.e. explain and analyse results based on the theoretical underpinnings of IP (see also Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Harrington, 2004). As Collentine (2004: 172) stresses, fundamental questions remain unanswered, since neither IP nor PI can explain how learners make form-meaning connections during and/or after PI treatments.

Following criticism, in a more recent publication VanPatten (2009) attempts to explain IP in more detail by contrasting IP mainly with Carroll’s (2001) Autonomous Induction Theory perspectives on processing and acquisition. According to VanPatten (2009:51), IP can be compared to Carroll’s ‘Autonomous Induction Theory’ (2001), as it involves parsing of sentences and draws notions from cognitive psychology. He adds that the IP Processing Principles can be directly linked to Carroll’s (2001) view that acquisition is ‘failure driven’. IP Principles with their predictions assist learners in overcoming the

[16]‘Learners will adopt other processing strategies for grammatical role assignment only after their developing system has incorporated other cues (e.g. case marking, acoustic stress)’ (VanPatten, 1996:52). In later publications this Principle is referred to as P2c (see VanPatten, 2004).

parsing problems they encounter, since learners realise that incoming information does not match their beliefs (VanPatten, 2009). Comparing the IP framework with another processing framework may answer questions that remain open in the field. However, VanPatten (2009), though he does not dismiss Carroll's (2001) perspectives on parsing and acquisition, does not explicitly adopt them. Therefore it still remains open which particular model of parsing has been adopted in IP, as IP has not to date attempted to propose an alternative parsing model.

For the purposes of the present study, I argue that PI has been shown to be an effective teaching intervention. However, on a theoretical basis, IP has failed to provide sufficient explanation regarding the mechanisms and processes involved during PI and as a result in language development. There are fundamental issues that IP needs to address in detail, such as the need to describe the basic language architecture and where IP and PI fit. In the IP literature there are various references to UG, restructuring, attention, noticing, the Competition Model and working memory that need to be properly addressed with additional reference to the IP model, e.g. what their role is in the model, how they interact with the model, what is their impact on the IP model and its practical application, namely PI. Addressing these fundamental issues will provide the stepping stones in appreciating the beneficial effects that PI has on language development.

2.3.2 Processing Instruction (PI)

Processing Instruction (PI) has been defined as 'a type of grammar instruction whose purpose is to affect the ways in which learners attend to input data' (VanPatten, 1996:2). It is therefore a pedagogical teaching intervention whose theoretical framework is based on IP. The aim of PI is to force learners to attend to elements in the input that they might otherwise miss. This is supposed to be achieved by altering learners' IP strategies. PI argues that through this type of instruction learners are encouraged to make better form-meaning connections; thus richer grammatical intake can take place (VanPatten 1996, 2004, 2007). According to VanPatten (2009:54), PI 'induces failure in the parser and forces readjustment' similar to what for Carroll occurs naturally if and when acquisition takes place (1999, 2001, 2007).

PI consists of four main components, namely explicit grammar instruction, referential activities and affective activities (mainly referred to in the literature as ‘structured input activities’) and a set of guidelines (e.g. they assist instructors in developing the materials; a complete list and comments of PI guidelines are available in the Appendix A). The instructor must firstly identify the processing problem that needs to be altered by identifying the IP Principle that will be the focus of instruction (VanPatten, 2009). Failure in identifying the processing problem and associating it with the appropriate processing principle is a common shortcoming of studies that have attempted to investigate the effectiveness of PI (see critique on DeKeyser and Sokalski, 1996; Salaberry, 1997; mainly by Allen, 2000; VanPatten and Wong, 2004; Wong, 2004).

Explicit Instruction (EI) in PI is presented to learners very briefly and always prior to the practice activities (i.e. structured input activities). According to PI guidelines, only one thing is presented at a time, excluding the presentation of whole paradigms at once as in the case of traditional instruction. For example if the target form is case marking in German, learners will not be presented with a table showing the case system in German but will be presented with two cases in opposition, e.g. the nominative versus the accusative case. Unlike traditional types of grammar instruction, in PI there should be two contrasting structures so that learners can develop ‘cue strengths’ for the target structure, similar to the propositions of the ‘Competition Model’ (Bates and MacWhinney, 1989). The target feature is explained in terms of the form-meaning connections learners must make, i.e. nominative is used for subjects, while accusative for objects in a sentence. Then there are statements about the possible errors and the reason(s) why those occur, referring to the IP Principle under focus. For example, the flexibility of German word order (SVO and OVS possibilities) versus the strict English SVO order might be a possible reason for not assigning correct roles when interpreting the meaning of a German sentence. This is followed by one or two examples, such as: ‘Der Mann küsst die Frau’ (the man-subject kisses the woman-object) versus ‘Den Mann küsst die Frau’ (the man-object kisses the woman-subject). The aim is to ‘inform the learners of the potential problematic outcomes of processing strategies’ (VanPatten, 1996:62).

Explicit instruction (EI) is also noticeable in other parts of the PI package, e.g. at the end of an activity in the form of implicit feedback: 'Did you notice that the word order is not the same in the two examples?'Although feedback and its role is not clearly defined in PI, implicit feedback is also available during structured input activities. However, Sanz (2005) distinguishes between explicit and implicit feedback in PI. Explicit feedback can be available in the form of why response *a)* is correct and why response *b)* is incorrect (although such feedback is not the norm in PI). Implicit feedback is more often provided in PI studies, where the instructor provides feedback during the referential activities in the form of 'yes, answer *b* is correct, or no answer *a* is incorrect' (see VanPatten and Cadierno, 1993; see also Lee and VanPatten, 1995; VanPatten, 2004)[17].

Structured input (SI) activities occur after the presentation of explicit instruction (EI). Based on the guidelines, in developing SI activities, it is important to 'Keep the psycholinguistic processing mechanisms in mind', i.e. SI activities should be drawing learners' attention to the target form forcing them to process the grammatical form in order to decode the meaning of the sentence (Lee and VanPatten, 1995; VanPatten, 1996). During these activities EI is not available. The sequence in PI is rigid. This means that after EI, learners are presented with SI activities, which are distinguished in 'referential' and 'affective' activities. Referential activities are considered of higher significance than affective activities because they force learners to process the target structure. At the same time they indicate whether or not learners make correct form-meaning connections, since in these types of activities there is a right or wrong answer. On the other hand, affective activities provide learners' opinions and beliefs and there are no right or wrong responses (Lee and VanPatten, 1995; VanPatten, 1996, 2004, 2007, 2009). During SI activities the examples always refer to a third person singular and learners are not required to produce the target structure. The purpose of presenting examples only in the third person singular is not explained in detail and it is something that PI needs to address. Regarding the role of output, VanPatten (1996, 2002, 2004,

[17]Please note that in PI the terms explicit and implicit feedback do not correspond to the typical distinction available in SLA ,i.e. as proposed for example by DeKeyser (1995, 1997, 1998, 2003, 2005). As Sanz (2005) explicitly describes there is by comparison a distinction between explicit and implicit ways of feedback provision and as described in this paragraph. Perhaps a more accurate definition could be elaborate vs. less elaborate and or obtrusive vs. less obtrusive feedback (see also Sharwood Smith, 1991, 1993; Doughty and Williams, 1998; Doughty, 2003). Clearly, this is an issue that PI needs to address.

2007, 2009) does not dismiss its role nor Swain's (1985, 1995) 'Output Hypothesis' in terms of oral and/or written production of language being also under circumstances a process of second language learning (see also Swain, 2005:471). However, as VanPatten notes PI and IP are concerned with the initial stage of language learning and therefore the role of output in IP is constrained in terms of helping learners altering their processing behaviours. For example, when English learners of German produce German OVS sentences interaction should lead the learner to realise if he/she has misinterpret an OVS sentence. If interaction does not lead to such realisations then output cannot have an effect in terms of acquisition (see VanPatten, 2007).

The underlining property of the referential activities is that the explicit comparison of the two structures helps learners to notice the target structure and its linguistic properties. Based on this notion, the end result is that learners will avoid confusion and possible causes for errors will be avoided. It could be argued that during referential activities learners go through various stages during which questions and hypothesis testing regarding the target structure are asked and formulated, leading to the establishment of correct form-meaning connections and acquisition of the target form (see also Lee and VanPatten, 1995; Marsden, 2006; VanPatten, 2009).

Focusing on the FNP, I will briefly present possible formats that referential activities can take, with examples taken from the present study. Since the focus is German OVS sentences, learners should be presented during SI activities with both SVO and OVS examples in order for them to distinguish between the two possibilities and make (correct) form meaning connections (VanPatten and Fernandez, 2004; VanPatten, 1996, 2004, 2009; see also Bates and MacWhinney, 1989). Taking as an example 'Den Freund (Accusative-Noun) vermisst das Mädchen (Nominative-Noun)' (the girl misses the friend), learners may be presented with a set of pictures, one depicting a man missing a girl (SVO) and another depicting a girl missing a man (OVS), as well as the target sentence in German. Learners in this case are required to choose which picture corresponds best to the sentence (picture matching task-Example 3). An alternative possibility is to present the target sentence, in this case the German example provided above, and two possible interpretations in English (one interpreting the sentence as an SVO and the other interpreting the sentence as an OVS sentence) requiring learners to choose which English rendition corresponds best to the German sentence (Example 3), as presented below:

Example (3)

Circle which picture corresponds to the sentence:
Den Mann vermisst das Mädchen.

a) 

b) 

Example (3)

Choose which English sentence corresponds best to the German sentence Den Mann vermisst das Mädchen.

a) The man misses the girl.
b) The girl misses the man.

In this case, it is important to remember the sub-principles of the FNP, namely the 'Event Probabilities', 'Contextual Constraint' and the 'Lexical Semantics' sub-principles, when developing the materials for PI activities. This means that items should not contain information that is more likely to occur in real life, or contain words and contexts that would help learners to decode the meaning of the sentence without necessarily processing the target form, e.g. the man misses the table, or the table misses the man (Lee and Benati, 2007; Lee and VanPatten, 1995, 2003; Wong, 2005).

As the final step in PI, affective activities follow referential activities and as previously stated affective activities require learners' affective response about phenomena associated to their 'own world' knowledge and reality, indications of learners' opinions, beliefs, feelings and personal circumstances. Hence, learners' responses are often associated with their personal life, e.g. relatives and friends, tutors, well-known personalities, a joke or a cartoon, or affective activities can also consist in ranking (i.e. putting sentences in chronological, logical order etc. see Lee and Benati, 2007; Lee and VanPatten, 1995, 2003; VanPatten, 1996:64, 2009). The role of affective activities is defined in PI as a 'reminder' to the learner to pay attention to the target form (VanPatten, 1996:74). They also do not require learners to produce the target form and they usually take the form of agreement-disagreement, true or not true for me, check boxes in surveys etc. Therefore there are no right or wrong answers for the affective activities, since they refer to personal opinion. The following example shows a possible affective activity in which the target form German OVS is provided (Example 4):

Example (4)

Put the sentences into a logical order to make a story:

a) Den Mann vermisst die Frau. ________
(the lady misses the man[18])
b) Den Mann ruft die Frau an. ________
(the lady calls the man)
c) Der Mann trifft die Frau. ________
(the man meets the lady)
d) Den Mann sieht die Frau. ________
(the lady sees the man)
e) Der Mann küsst die Frau. ________
(the man kisses the lady)
f) Den Mann heiratet die Frau. ________
(the lady marries the man)

However, there are certain issues that PI needs to clarify regarding the affective activities. First of all, the 'labelling' of these activities as affective, implies that learners' emotions should be involved during this type of activities. The term however used, is

[18] Please note that the English rendition is not available for learners during treatment, although vocabulary lists may be available.

rather misleading when in fact putting sentences into a chronological order or ranking, matching sentences and or requiring learners to respond on a task based on their ‘world knowledge’ does not necessarily involve learners’ ‘affect’ and/or emotions. A further point of interest is made by Marsden (2004:36), who points out that as the tasks in the affective activities do not require learners to process the meaning of the target features correctly in order to complete the task, any attention to the target feature is incidental. Another possible interpretation could be that affective activities may serve as input flood[19], providing more examples of the target form for the learner. Clearly, these are issues that PI needs to addressed in more detail, as they are important in understanding how PI actually works during input processing; how do learners make correct form-meaning connections and as a result how do they acquire the target form. In the PI literature there is no detailed explanation referring explicitly to the role of referential and affective activities and their necessity in PI. VanPatten (1996, 2002, 2004, 2007, 2009) stresses the centrality of referential activities in assisting the learner in the process of making correct form-meaning connections. Affective activities are provided, according to VanPatten (1996, 2002, 2004, 2007, 2008), in line with ‘communicative teaching’ (1996:64). However he does not explain what he means by communicative teaching nor what this entails for PI. Given the process of acquisition that VanPatten adopts (1996, 2002, 2004:26, 2007, 2009; similar to McLaughlin, 1990), one could assume that affective and referential activities contribute to the restructuring and accommodation process of acquisition. However, this remains vague in PI’s and IP’s theoretical underpinnings, as VanPatten (1996, 2004, 2007, 2009) states that IP and consequently PI are only concerned with the first phase i.e. input processing. Based on this argument, he does not explicitly describe what processes are involved in restructuring and accommodation and/or how PI contributes to these two important processes for language learning. Providing a detailed and clear description of what IP and PI entail and how they are operationalized in SLA can further assist researchers in investigating the role of PI in SLA. The existing vagueness has led many researchers to misinterpret how exactly PI should be operationalized and/or how PI materials should be developed; thus there have been unfruitful attempts to investigate PI effects in SLA

[19] Input flood is another type of Input Enhancement, as discussed above. Wong (2005) defines input flood as a language teaching technique in which the input is flooded with the target form in an attempt to make learners notice the particular target feature. ‘In input flood, the input learners receive is saturated with the form that we hope learners will notice and possibly acquire’ Wong (2005:37).

(see critique on Allen, 2000; DeKeyser and Sokalski, 1996, Salaberry, 1997). Furthermore, results obtained from PI studies cannot be adequately discussed at a theoretical level referring to second language acquisition (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Doughty, 2004; Harrington, 2004).

Another unclear issue in PI is the fact that during structured input (SI) activities, learners are not required to produce the target form. Despite the fact that studies have shown that PI performs as well in production as other types of instruction (which have included production tasks in their treatment packages, see VanPatten and Cadierno, 1993; VanPatten and Wong, 2004; Wong, 2004). The reasons are explained in VanPatten (1996), where he notes that PI is an additional comprehension-based approach and as such it is mainly concerned with learners understanding (comprehension), which precedes acquisition, similar with other comprehension approaches as expressed by Krashen (1982, 1985) and Winitz (1981; see also VanPatten, 1996:83; Krashen and Terrell, 1983; Richards and Rodgers, 1986).

The comprehension approach emphasizes on acquisition of a second language through meaningful input without being provided with opportunities for output, interaction and/or explicit instruction (Verspoor, Lowie, De Bot, 2009). As VanPatten (1996:83) notes the difference between PI and other input-based approaches lies on the fact that PI departs from the concept of just presenting learners with comprehensible input (see Krashen, 1982, 1985; Krashen and Terrell, 1983; Richards and Rodgers, 1986; Winitz, 1981) in that it forces learners to process more form and/or to process correctly the form in the input. Guided from psycholinguistic theory and research its aim in influencing the process of input becoming intake (VanPatten 1996:83; see also White, 1987; Carroll, 2001; VanPatten, 2007, 2009). Based on the PI guidelines, one can assume that the lack of production tasks may be also attributed to the guideline ‘Move from sentences to connected discourse’. According to this guideline, because of working memory capacity limitations, the initial process of input is crucial. ‘Processing time’ limitations do not provide sufficient time for learners to process the form in order to make correct form-meaning connections and at the same time engage in production tasks at the initial stage of input processing (VanPatten, 1996, Terrell 1991). The particular guideline proposes that connected discourse, involving both listening and engaging in conversations, should be introduced at later stages in the lessons. However even if it is assumed that at later stages production is foreseen in PI, there is no reference to the overall aim, form and contribution of production tasks to the processing and acquisition of the target form through PI.

VanPatten (1996, 2002, 2004, 2007, 2009) explains that IP and PI are concerned with the role of input without dismissing the role of output. Despite the fact that VanPatten supports Swain's (1985, 1995, 2005) 'Output Hypothesis' arguing that production is part of the acquisition process (i.e. similar to the 'Input Hypothesis', 'comprehensible output' is also considered to be an important factor promoting language development) he does not explain how output contributes to PI (see criticism from DeKeyser et al. 2002; Harrington, 2004; Carroll, 2004; Collentine, 2004). Clearly, the issue of production in PI requires elaboration, especially, the reasons why production of the target form is not available in the treatment. Finally, the fact that PI achieves equal performance rates to other types of instruction that include production tasks in their treatment packages must be accounted for.

2.3.3 PI studies

In the field of PI there has been extensive research on the FNP. Studies have focused on mainly on Spanish, but there are also a few studies on French, Italian, Japanese and German. The seminal study of VanPatten and Cadierno (1993a, 1993b) has promoted research in the field of SLA and many have a) compared PI with 'traditional instruction' (TI) and other types of instruction (Allen, 2000; Benati, 2001, 2005; Keating and Farley, 2008; Lee and Benati, 2007; VanPatten and Wong, 2004; Wong, 2004; among others), b) investigated the role of EI in PI (Benati, 2004; Farley 2004; Marsden, 2006; Sanz and Morgan-Short; 2004; VanPatten and Oikennon, 1996; Wong, 2004), c) investigated the cumulative effects of PI (Benati and Laval, 2008; Laval, 2008) and d) examined the effects of PI in virtual contexts (Benati and Lee, 2007). Over the past fifteen years the majority of PI studies, regardless of their different research aims and objectives, have been consistent in providing results supporting the basic argument of PI: that PI is effective in improving performance by forcing learners to alter their processing strategies. In the following section, PI studies relevant to the purposes of the present study will be reviewed. Providing an exhaustive review of all PI studies focusing on different IP Principles is beyond the scope of this study. However reference to studies that have explored other IP Principles than the FNP and their effectiveness in SLA will be made when appropriate.

The original VanPatten and Cadierno (1993b)[20] study investigated the FNP with eighty adult English learners of Spanish at second year university level. The target forms were word order and Spanish clitic object pronouns. PI was compared with traditional type of instruction (TI) and no instruction (control group). Participants were divided into three groups, namely, N=27 PI group, N=26 traditional instruction group (TI) and N=27 Control group (C). A two-day treatment was administered after pre-testing. Treatment packages for TI consisted of EI, and explanation of direct object pronouns (full paradigm presentation) and activities taken from a Spanish textbook following TI. Activities for this group moved from oral mechanical to oral meaningful and finally to oral communicative practices. The PI treatment package included EI and SI activities. The control group was exposed to the target form but received no instruction on the target form.

Assessment tasks involved an interpretation and a production task. The interpretation task consisted of ten target sentences five of them of the type $\text{Object}_{\text{Noun}}$-Verb-$\text{Subject}_{\text{Noun}}$ and a set of five target structures of the type $\text{Object}_{\text{Pronoun}}$-Verb-$\text{Subject}_{\text{Noun}}$ sentences and 5 distracters Subject -Verb-Object (SVO) sentences. Participants listened to the sentence and had to choose from a set of pictures the one that corresponded best to the sentence. The production task was a written task consisting of five target items in which participants had to complete the second part of an incomplete sentence based on a set of pictures. Three post-tests were administered after instruction: an immediate post-test, a second one a week after treatment and a third, delayed post-test, four weeks after treatment. Results from repeated measures ANOVA showed that PI outperformed the other two groups in the interpretation task, suggesting the following hierarchy: PI>TI=C. For the production task repeated measures ANOVA showed no difference

[20] The VanPatten and Cadierno (1993a) study has the same test design and format only in a smaller scale in terms of participants and it could be defined as a pilot study of the VanPatten and Cadierno (1993b) study. Given that overall design, administration and results were the same, the review is limited to the seminal VanPatten and Cadierno (1993b) study.

between the PI and TI groups; however both groups outperformed the control group, thus: PI=TI>C. Moreover, raw score analysis showed a U-shaped performance both in the interpretation and production tasks for the PI and TI groups (similar to McLaughlin, 1990). The control group showed improvement from pre- to first post-test then in the second post-test slight improvement and in the delayed post-test performance decreased reaching lower rates than those of the first post-test. In the production task, the control groups' performance was the same after the first post-test. VanPatten and Cadierno (1993b:52) concluded that PI has an effect on the developing system, since a) in the production task there were no significant differences between the PI and TI groups, despite the fact that the PI group did not produce the target form during instruction, counter to the TI group, who received a lot of practice; and b) in the interpretation task there were no significant differences between the TI and Control groups.

Clearly, the interpretation task was biased towards PI, while the production task was biased towards the TI group. It is therefore valid to conclude that the study has provided evidence supporting that PI can improve learners' performance both for tasks that the group has been trained for as well as in unfamiliar tasks. On the other hand, the TI treatment type can improve task performance, when participants are trained (VanPatten and Cadierno, 1993b:51). PI has outperformed groups both in comprehension and production of the target form. However, the production task was an untimed fill in the gap task, and it could be argued that it was not a 'pure production' task that can measure implicit knowledge or provide supporting evidence suggesting that participants have internalised the target form (Doughty, 2003, 2004). Since there was no time limitation for the fill in task it may well be that participants were accessing metalinguistic knowledge. For the interpretation tasks, it could be further argued that responses might have been based on chance, as it does not seem that there were available options such as 'I am not sure' to control for this variable.

Administrating a battery of tasks including those such as grammaticality judgement and/or oral production tasks, might have also been more beneficial in obtaining a better insight into participants' performance and treatment effectiveness in other tasks (Gass and Mackey, 2005). Furthermore, the study claims to be having an effect on the developing system; however the delayed post-test was administered four weeks post-instruction, thus the time elapsed is too short to provide valid and reliable conclusions regarding effects on the developing system (Doughty, 2003, 2004; Lee and Huang, 2008; Mitchell and Myles, 2004; Norris and Ortega, 2000). In the field of SLA it is widely accepted that in order to draw reliable and generalisable conclusions regarding effects of an intervention in the developing system, delayed post-tests should be administered at least 12 weeks post-instruction; whereas the population of the sample size should be at least one hundred for statistical purposes (see Doughty, 2003, 2004; Field, 2009; Gass and Mackey, 2005; Lee and Huang, 2008; Mitchell and Myles, 2004; Mujis, 2005; Norris and Ortega, 2000; among many). Regarding the actual treatment, PI has received criticism referring to its short length, as the mean hours of exposure being one and a half hours, whereas the maximum exposure reaches 2 hours (see Doughty, 2003, 2004; DeKeyser et al. 2002; Mitchell and Myles, 2004; Norris and Ortega, 2000). The latter is indeed a very important factor, one that merits considerable attention and detailed explanation on behalf of PI, as despite the short length of instruction there is considerable evidence supporting changes in the developing system (see PI research). On these grounds, one cannot argue that participants have acquired the target form, although findings seem promising. Moreover, the sample size is not adequate to draw generalizable conclusions. Hence, the original VanPatten and Cadierno (1993) study has provided promising results that are valid for claims made on short-term effects of the method. However, the study fails to link IP propositions and other theoretical frameworks to the obtained results. Findings do not establish when and how participants made correct form-meaning connections and how and why these connections will be maintained in the long-term (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Harrington, 2004).

DeKeyser and Sokalski (1996) have argued that results from VanPatten and Cadierno's (1993b) study are due to the fact that the groups did not receive the same amount both quantitatively and qualitatively of declarative information. The PI group received explanation plus contrasting examples of the clitic pronouns, including the use of the "personal *a"* object marker, as well as examples of Spanish word order (e.g. object pronoun-verb-subject). In contrast, the TI group received explanation on objects and object pronouns while was presented with a single paradigmatic chart including only object pronouns for all person-number combinations (DeKeyser and Sokalski, 1996:

619; see also VanPatten and Cadierno, 1993b: 48). Clearly, DeKeyser and Sokalski (1996) misinterpret the basic underpinnings of PI. It is reasonable that there should be qualitative differences between an approach that focuses attention on meaning (PI) and an approach (TI) that is based on mechanical drills (see also Allen, 2000; VanPatten, 1996, 2004; VanPatten and Fernandez, 2004; VanPatten and Wong, 2004; Wong, 2004). This misinterpretation is evident also in their 'replication' study, in the way they refer to the experimental groups. The TI group is referred to as an output-based group and the PI as an input-practice group.

With this in mind DeKeyser and Sokalski (1996) attempted to replicate the original VanPatten and Cadierno (1993b) study. Based on the hypothesis that the complexity of a structure influences the degree to which input and/or output practice can be effective, they argue that the target structure investigated in the original study, Spanish clitic object pronouns, is a structure easy to produce but difficult to perceive. Thus, they conclude that results have been additionally affected by these two factors (see DeKeyser and Sokalski, 1996:620; see also DeKeyser, 1994). Their argument is based on the notion that Spanish clitic object pronouns area morphologically simple structure for English speakers to produce because 'the structure encodes an obvious agent/ patient relation reflected in the morphological alternations between *yo/me* (I/me), *él/lo* (he/him), and so on' (DeKeyser and Sokalski, 1996:620). However, the particular structure is difficult to perceive for English speakers because of the OVS word order in Spanish, in which case word order tends to override morphology in their interpretation of sentences (DeKeyser and Sokalski, 1996:620; see also Bates and MacWhinney, 1989; MacWhinney, 1987).

In order to balance the described effects, DeKeyser and Sokalski (1996) contrast the Spanish clitic object pronoun with the Spanish conditional structure, which is a structure easy to perceive but difficult to produce, while both experimental groups received the same amount of EI for both target forms. This is a further indication of the misconception regarding the theoretical and practical fundamentals of PI. Indeed the FNP refers to the particular aspect of processing by providing a solution through PI in tackling the issue of word order overriding morphology. PI uses contrasts in order to enhance cue strength, validity and reliability, however it does not use contrasts based on which structure is more salient or easier to produce and/or comprehend (see Allen, 2000; VanPatten, 1996, 2004, 2007, 2009; VanPatten and Fernandez, 2004; VanPatten and Wong, 2004; Wong, 2004). Nonetheless, DeKeyser and Sokalski make an important point: the importance of having the same amount of EI in a comparison study. Perhaps, if IP and PI were described in more detail, such misinterpretations would not occur.

In the DeKeyser and Sokalski's (1996) study, eighty-two university-level participants were divided into three groups, an input; an output; and a control group. A four-day treatment: a pre- test, an immediate test (one day post-instruction) and a delayed post-test (one week post-instruction) were administered. For the immediate post-test, their study revealed a hierarchy for comprehension of the Spanish direct object pronouns, indicating that the PI was significantly different from the output practice group. In turn, the output group was significantly different from the control group. However on the delayed post-test there were no significant differences between the three groups. In terms of production no significant differences between the input and output based groups were reported, although both outperformed the control group. Findings for the Spanish conditional showed that the output practice group was better than the input group for both production and comprehension tasks in the immediate post-test. However, in the delayed post-test, descriptive statistics showed that the input practice group was slightly better for comprehension, while the output group was slightly better for production. These results lead DeKeyser and Sokalski (1996:634-635) to conclude that comprehension and production skills in an L2 are learned as a separate skill. Moreover, they argue that there is a correlation between type of skill practice and type of skill required in a task, while there are indications of transfer between the two skills (based on the fact that the performance of the two experimental groups was significantly better than the control groups, while both groups outperformed each other in the delayed post-tests), a fact that they attribute to transfer via declarative knowledge (see also Anderson, 1993). Nonetheless, according to DeKeyser and Sokalski (1996:636), the results provide information about declarative and non-automatized procedural knowledge[21], due to the short-term application of the delayed post-test.

Clearly, DeKeyser and Sokalski did not carry out a replication nor a 'conceptual replication' study of the original VanPatten and Cadierno (1993) study. Polio and Gass (1997:502) define 'conceptual replication' as any attempt that alters various features of the original study in order to confirm the generalizability or external validity of the original study. Features that have been identified as valid in order to claim replication are a) testing a different population; b) using a different setting; c) using a different testing modality. In DeKeyser and Sokalski (1996) there are many factors that lead to the conclusion that their study provides a comparison of their skills acquisition theory with an input processing theory similar but not identical to what VanPatten and Cadierno suggest (see also Allen, 2000; VanPatten and Fernandez, 2004; VanPatten and Wong, 2004; Wong, 2004). First of all, the formulation of the hypothesis of their study clearly states that they set out to investigate 'Skill Acquisition Theory' (DeKeyser and Sokalski, 1996:623). Secondly, there is no indication of the IP Principle that they have identified as needing to be altered. VanPatten and Cadierno (1993) identified in their study the FNP as the strategy that needs to be altered regarding the Spanish Object pronouns. Assuming that the same principle is valid for the DeKeyser and Sokalski (1996) study, it is not the same for the second target form that they have chosen to investigate, namely the Spanish conditional. Moreover, they clearly state that they have chosen the target forms on the grounds that one is easy to comprehend but difficult to produce and the other form is difficult to perceive but easy to produce. This might be indeed a processing principle that needs to be addressed, however it is not part of the IP

[21] Researchers from the field psychology and psycholinguistics have posited three stages of knowledge development, namely declarative, procedural and automatic (see Anderson, 1982, 1993; Cohen and Squire, 1980; DeKeyser, 1995, 1998; 2007; McLaughlin, 1987, 1990; among many). Declarative knowledge involves "knowing that", for example London is the capital of England, whereas recalling information from declarative memory involves some degree of conscious effort, i.e. information is consciously brought to mind. Procedural knowledge involves "knowing how" to do things. It includes skills, such as "knowing how" to play the piano, while it does not involve conscious (i.e. automatic) thought. For example, we ride a bike with little or no awareness of the skills involved. Once procedural knowledge has been acquired practice gradually leads to automatization of knowledge. For example we drive the car without being consciously aware of the skills involved. However, it should be noted that even highly automatized behaviours are not 100% automatic, as DeKeyser (2007:99) notes.

Principles that VanPatten (1996, 2004, 2007, 2009) describes in his IP model, nor one that VanPatten and Cadierno (1993) have identified (Allen, 2000; VanPatten and Wong, 2004; Wong, 2004). Furthermore, it is a processing principle that is mainly concerned with production, which to some extent contradicts the comprehension approach adopted by VanPatten. More importantly, their study design (with regard to the groups and the task) greatly departs from the original study. Comparing activities and items from both studies, it is evident that both use comprehension activities, however both of them do not use SI activities, which results in violations of the PI guidelines, such as 'keep meaning in focus' and 'teach one thing at a time' and their study ultimately lacks 'treatment fidelity' and therefore replication (Allen, 2000; VanPatten, 1996; VanPatten and Wong, 2004; Wong, 2004).

Nonetheless, DeKeyser and Sokalski's (1996) study has provided evidence regarding the role of skills acquisition in SLA and how this might affect the process of language learning. In addition, they make a very important and valid argument, which should be accounted for in PI and IE studies, namely that administering the delayed post-test four weeks post-instruction still can only provide information about declarative and non-automatized procedural knowledge[22]. Thus, we cannot claim acquisition of a target form when obtained data are gathered in the short-term (Anderson, 1993; DeKeyser, 1995; DeKeyser and Sokalski, 1996; Lee and Huang, 2008; Norris and Ortega, 2000; Mitchell and Myles, 2004).

Salaberry (1997) also attempted to replicate the original VanPatten and Cadierno (1993) study. 33 University students were assigned to three groups, an input, an output and a control group. The target form was Spanish clitic object pronouns as in the original study. Salaberry (1997) provided the same EI for both the input and output based groups, as well as operationalized PI in a different way from the original VanPatten and Cadierno (1993) study. He did not identify the IP Principle that needed to be altered and he did not follow the guidelines in developing PI materials and SI activities. Thus, Salaberry's study cannot be claimed to be a 'conceptual replication' of the original

[22] DeKeyser and Sokalski (1996), as a 'replication study' followed the design of the original VanPatten and Cadierno (1993) study. However, they acknowledge the fact that the time elapsed in the two studies for the delayed post-test cannot provide robust evidence regarding automatized (acquired) knowledge. Clearly, this is an issue that needs to be accounted for both in the field of PI and IE, as well as in other approaches investigating language development.

study (Allen, 2000; Gall, Borg and Gall, 1996; Polio and Gass, 1997; VanPatten and Cadierno, 1993; VanPatten and Fernandez, 2004; VanPatten and Wong, 2004; Wong, 2004).

Salaberry's treatment lasted one and a half hours. A pre-test, an immediate and a delayed post-test (four weeks post treatment) were administered. The assessment tasks included a comprehension and a production task (in accordance with VanPatten and Cadierno, 1993) and a free timed narration task (five-minutes) of a one-minute silent video clip (following Sanz, 1995). Results of the study showed no significant differences between the three groups in the production and the free narration task from the pre-test to the post-test condition. Results from the comprehension task showed no significant differences between the input and output based groups. Again findings of the Salaberry study are inconclusive regarding the possible effectiveness of PI, as the study investigates an input based approach not identical to the one proposed by PI with an output based approach. Moreover, the number of participants (n=33) is too small to make valid claims on effectiveness, acquisition or generalizability.

On the other hand, it should be acknowledged that DeKeyser and Sokalski (1996) and Salaberry (1997) pointed out one important issue in the VanPatten and Cadierno (1993) study, namely the issue of "internal validity", based on the fact that experimental groups did not receive quantitatively the same amount of explicit instruction, which as a variable may have affected the obtained results. Despite the fact that PI does vary both quantitatively and qualitatively from TI (Allen, 2000), their criticism has led to the investigation of the role of EI in PI (Benati, 2001; Culman et. al, 2009; Farley, 2004; Fernandez, 2008; Marsden, 2006, Wong, 2004; VanPatten and Oikennon, 1996).

In agreement with Polio and Gass (1997), Allen (2000) recommended 'treatment fidelity' in conducting a 'conceptual replication' of the VanPatten and Cadierno (1993) study. Allen's (2000) study included a pre-test and three post-tests assessing interpretation and production. The study differs from the original in terms of the target form, as it investigates the French causative with *faire*, it selected an open-ended production test instead of sentence completion, and its sample size of participants was179 (at a university level). The processing problem associated with French causatives with *'faire'* falls under the FNP, as studies have shown that English learners of French rely on word order in decoding the meaning in sentences containing the French causative (Heilenman and McDonald, 1993; McDonald and Heilenman, 1992).For example, in the sentence *Mon profeseur me fait ouvrir le livre* ('My teacher has me open the book'), English learners of French would interpret the sentence as 'My teacher opens the book' (cf. Allen, 2000:73).

Participants were distributed into three groups, a control, a PI and a TI group. Results of the interpretation task revealed no significant differences between the PI and the TI group. The findings from the production task showed that the TI group performed significantly better than the PI group. Regarding the long term effects of instruction, the analysis of the interpretation tasks did not indicate any significant differences between TI and PI, whereas in the production test, the analysis of the third post-test showed that TI was significantly better than PI. The analysis of the second post-test did not indicate any significant differences between the two groups (Allen, 2000:77-78). Allen (2000:80) concluded that the results of VanPatten and Cadierno (1993) 'are not generalizable to the French causative structure'. Allen also pointed out that PI might be 'effective only for certain grammatical structures', since for the French causative both the TI and PI proved equally effective for the interpretation task, whereas PI was shown to be less effective than TI in the production task.

VanPatten and Wong (2004) claimed that results from Allen's (2000) study are due to PI violations, such as 'Event Probability'. The presence of specific lexical items has enabled participants to determine who performed the action without actually processing the target form, as the following extract suggests: *Le professeur fait étudier le verbe "être" à l'élève* (The instructor gets the student to study the verb "to be") (cf. VanPatten and Wong, 2004:101). Given the type of materials provided in Allen's study it is not possible to establish if learners were relying on sentence structure for meaning or on the lexical items providing cues to avoid processing the target form or if they were being encouraged to memorize a pattern with the help of the activities (VanPatten and Fernandez, 2004; VanPatten and Wong 2004; Wong, 2004). According to VanPatten and Wong (2004), the PI treatment package in Allen's (2000) study did not include true SI activities, and therefore learners were not pushed to rely on sentence structure to get meaning (VanPatten and Wong, 2004; Wong, 2004). Furthermore, Allen (2000) did not maintain the third-person singular constant in both the interpretation and production assessment tasks. The third-person singular was limited to the sentences for the interpretation task, while for the production task the first-person singular object pronoun was used. Although the purpose of keeping the third person singular constant in all items for the SI activities is not explained, it seems to be an important element, as all PI studies have maintained parity (VanPatten and Wong, 2004; Wong, 2004).

Allen (2000) also used exclusively causative sentences, which did not push learners to make a distinction between causative and non-causative sentences with *faire*. In PI it is essential to use the target form with a contrasting structure in order for learners to make correct form-meaning connections. IP argues that the contrast forces learners to assign grammatical meaning and semantic roles and that it enhances cue strength, validity and reliability during sentence processing when trying to decode the meaning of the sentence (VanPatten, 2004:15; similar to Bates and MacWhinney, 1989; Carroll, 1991, 2001, 2007; Pienemann, 1984, 1998, 2007). Finally, participants in the PI group were required to process the full verb paradigm and pay attention to all conjugations of the present tense of the verb *faire* rather than focusing on one form according to the PI approach (VanPatten, 1996, 2004). Consequently, the PI group was actually receiving TI (VanPatten and Wong, 2004; Wong, 2004). The TI group processed EI in the same way as the PI should have processed; i.e. instead of moving from mechanical to meaningful to communicative practice, they reviewed a total of 23 sentences before practice began (VanPatten and Wong, 2004; Wong, 2004). Thus, VanPatten and Wong (2004) argue that, in the initial stage, the TI group received the EI that the PI group should have received and that the PI group received the EI that the TI group should have received.

VanPatten and Wong (2004) have attempted to replicate Allen's (2000) study based on the original VanPatten and Cadierno (1993) methodology in designing treatment materials. The obtained results are similar to those of VanPatten and Cadierno (1993), namely the PI group outperformed the TI group in comprehension, while there were no significant differences between groups in the production task. However, the VanPatten and Wong (2004) study is limited to providing data only from the first immediate post-test, as scheduling conflicts did not allow for the delayed post-tests comparable to Allen's (2000) study. Furthermore, the end sample size was too small (N=77) to provide valid and reliable comparable data to those of Allen's (2000) study of 179, as well as overall generalizable conclusions.

Despite the fact that Allen's study cannot be considered as a 'true PI' study because of the violations in the study design, it clearly demonstrates the vagueness and lack of explicit information on what exactly should be included in PI packages; what the differences are between PI and other types of instructions; what constitutes its theoretical framework. Clearly these factors lead to misinterpretations and therefore limited efforts attempting to further explore the possible effects of PI in other languages and other grammatical forms.

Long term effects of PI have been investigated by VanPatten and Fernandez (2004:278), who replicated the original VanPatten and Cadierno study (1993b); while they differed in the administration of the delayed post-test, which measured effects eight months post instruction. Forty-five university level students received PI. The study lacked a control and a comparison group, because of the small number (N=45) of participants (VanPatten and Fernandez, 2004). Moreover, based on findings of previous PI studies that have included a control group which showed no improvement for the control group up to the delayed post-test, they argue that control groups 'appear to be unaffected by test familiarity or any other variable' (VanPatten and Fernandez, 2004:285). However, such claims cannot validate the exclusion of a control group, nor can they enhance the validity and reliability of the study. Control groups are necessary to measure and compare any differences in gains between instruction and non-instruction of the target form (Doughty, 2003; Mitchell and Myles, 2004; Norris and Ortega, 2000).

A pre- test, an immediate test and a delayed post-test, eight months after instruction, were administered in the VanPatten and Fernandez (2004) study, while the treatment lasted for two consecutive days. Repeated measures ANOVA revealed significant differences for the interpretation task in the three conditions. Raw score results showed significant improvement from the pre-to the immediate post-test condition. A decline in scoring was shown from the immediate to the delayed post-test; however scores were significantly better compared to the pre-test performance. The same pattern was revealed for the production task both in terms of scoring and statistical significance in the three test conditions. VanPatten and Fernandez (2004:284) concluded that PI is effective both in the short and long term in terms of altering the FNP regarding OVS and object clitic pronouns. Despite promising results, it is valid to claim that the lack of comparison groups does not allow conclusions regarding the effectiveness of PI compared to other types of instruction, as it might have revealed that the PI was performing better in the three conditions than the other group and/or vice versa.

Now that replication studies of the original VanPatten and Cadierno (1993b) study investigating the effects of 'full PI' to date have been reviewed, I turn to the review of the second seminal study and replication studies in the field of PI, which investigated the role and effectiveness of explicit instruction in PI, namely VanPatten and Oikennon (1996). The experimental groups in this case are divided into groups that receive 'full PI', namely EI+SI, and groups that receive only SI activities.

VanPatten and Oikennon (1996) replicated the original VanPatten and Cadierno (1993) study focusing on the role of EI. Fifty-nine English learners of Spanish were divided into three groups: an SI activities only (n=20), a 'full' PI (n=17) and an EI only (n=22). The target form was the same as in the original VanPatten and Cadierno (1993) study and the focus was also the FNP. The assessment tasks were also the same as in the original study; however the timeline of the study was limited to a pre- and an immediate post-test. Findings of their study showed that from pre to post-test, the SI activities group performed as well as the full PI group, while the EI only group 'flatlined'. VanPatten and Oikennon (1996) argued that the change in learners' knowledge and performance is due to the SI activities and therefore EI is not necessary in PI for it to be effective.

Following the seminal study of VanPatten and Oikennon (1996), researchers have attempted to (partially) replicate the original study, looking at possible effects. Regarding the Italian future tense (Benati, 2001, 2004); the Spanish subjunctive (Farley, 2004) and again Spanish clitic Object Pronouns (Sanz, 2004; Sanz and Morgan-Short, 2004). Apart from Farley (2004), all the above mentioned studies have demonstrated that the use of full PI and the use of SI activities only had equal gains. Some of the studies that are reviewed next also examined the role of feedback with or without EI in PI.

Sanz (2004) investigated the effects of implicit versus explicit computer delivered feedback in PI. Twenty-eight participants were randomly assigned to two groups, an implicit and an explicit feedback group. The processing principle in focus was the FNP, while the target form was OVS sentences. A pre-, immediate post-test design and a two-day treatment were carried out. Both groups received treatment and testing in a CALL environment The explicit group received personalised feedback that was immediate, according to 'the cognitive window of opportunity'[23] (Doughty, 1998, cited in Sanz, 2004:247 and Sanz and Morgan-Short, 2004:55). On the other hand, the implicit group received implicit but still metalinguistic feedback in the form of 'ok' for correct responses or 'Sorry, try again' for incorrect responses. The assessment tasks involved written interpretation, a sentence completion production task and a video retelling task. Results of the study showed that both groups significantly increased their ability to interpret and produce the target form, however no significant differences were reported, as there were no indications of one group outperforming the other. Sanz (2004) concluded that regardless of implicit or explicit feedback, there are no indications in her study that suggest the effectiveness of feedback in PI. For Sanz (2004:253) the need for feedback inclusion, especially in CALL environments, derives from the type of instruction. She argues that for PI feedback is not necessary, since the meaningful tasks draw attention to both form and meaning, can replace feedback and drive the acquisition (Sanz, 2004). However, the fact that, in Sanz's (2004) study, the variables SI and ±feedback were not isolated weakens her claims about feedback not being necessary in PI. The inclusion of a third group receiving only SI activities as well as the inclusion of a control group receiving no feedback and no SI activities could have provided evidence supporting such claims.

[23] The 'cognitive window of opportunity' refers to Lightbown's (1998) and DeKeyser's (1998) proposals claiming that 'even in advance of needing forms for communicative purposes, learners can hold them in mental representation (i.e. memory) for further processing. However, this cognitive ability is limited in the sense that, if no timely opportunity for use arises, the forms will no longer remain in memory' (Doughty and Williams, 1998:5)

In a similar study investigating Spanish clitic object pronouns and word order, Sanz and Morgan-Short (2004) investigated the role of feedback prior to and during a task. Sixty-nine participants were assigned to four groups, namely +Explanation–Feedback, +Explanation+Feedback, -Explanation+Feedback, -Explanation– Feedback. Materials developed were adapted computerised versions of the VanPatten and Cadierno (1993) study. A total of 56 target items were provided through SI activities, while presence or absence of metalinguistic information given prior to the task (Explanation) or during task completion (Feedback) was provided accordingly for each group. Personalised explanation was provided when needed following PI, while explicit negative feedback was immediate, following the 'cognitive window opportunity' (Doughty, 1998; 2001:249).

Assessment tasks involved interpretation and production. Results showed no significant differences between the four groups for interpretation and production, although gains for all groups (including the control group by 31.8%) were reported. Sanz and Morgan-Short (2004:68) concluded that neither EI nor feedback enhances the acquisition process of Spanish word order. However, as they also note, time measures (in terms of reaction time) could have provided a better overview of the groups' processing of the forms and might have also revealed differences between groups. Furthermore, if the software used had provided additional information regarding the time and frequency at which participants accessed feedback, differences between groups might have been more noticeable.

PI studies have also examined the effects of PI versus meaning-output instruction (MOI) exploring the effects of the sub-principle '*Lexical Preference Principle*', which assumes that learners rely on lexical items rather than grammatical form to get the meaning when both encode the same semantic information (VanPatten, 1996, 2004:14). Relevant to the purposes of the present study are the studies that have investigated the effects of the particular sub-principle using textual enhancement in their research design.

Benati (2001) investigated the acquisition of verbal morphology for the Italian future tense of regular verbs. Thirty-nine second semester university level students were randomly assigned to a PI, an output-based type and a control group. The PI group received full PI. The output-based group received EI in terms of presentations of all future forms in a paradigmatic way, where the endings of the verbs were textually enhanced (e.g. bolded and underlined) followed by activities during which learners practiced production of the target form. According to Benati (2001:107), the activities in the PI group were almost all meaningful and communicative, whereas the activities in the output-based treatment contained some mechanical form of practice. The two treatments also differed because one was receptive and the other one was productive. The study followed a pre, post-test design. A two-day instruction lasting for six hours preceded the immediate and delayed (3 weeks post-instruction) post-tests.

Assessment tasks involved one interpretation, one written completion task and an oral limited-response production task. Temporal adverbs were excluded and the verb was never placed at the beginning of the sentence. The task required participants to establish whether the action expressed in the sentence they listened to was taking place in the present or in the future. Scoring for the particular task involved no points for incorrect responses and one point for correct responses. Thus, it is valid to assume that a third option allowing the learners to say ‘I am not sure’ was not available in order to control for guessing. The written production task was a fill in the gap task, where the verb was provided in the infinitive form in brackets. In the oral production task participants were required to re-tell a story in the future tense. Participants received two points for ‘fully correct use’ of the future tense; one point for ‘partially correct use’ of the future tense (wrong spelling but right ending, such as for the first person singular *parlarà* instead of *parlerà* ‘will speak’) and no points for incorrect use (Benati, 2001:109). Vocabulary familiarisation was available before the production tasks, while distracter tasks were conducted in-between the production tasks.

Analysis of results showed no significant differences in the pre-test condition but a hierarchy was revealed in the two post-tests suggesting that the PI outperformed the EI output based group, which in turn outperformed the control group. Analysis of only the production tasks showed no significant differences between the PI and the output based group, however both groups performed significantly differently to the control group. Benati (2001) attributes results to the fact that both groups included meaning oriented instruction, while he argues that 'the output in the output-based group may have served as input for students who were listening to their classmates' responses' (Benati, 2001:116; similar to Farley, 2001 and Spada, 1997). Overall, he claims that the methodological design of the study was a major explanation for such results, because the items provided in the experimental groups were balanced. Interestingly, Benati did not consider the possible role(s) that the applied textual enhancement might have played on the obtained results. However, since both groups received EI with textual enhancement of the target form endings, no valid and reliable conclusions can be drawn. The short-term results show that meaning-based instruction, whether input and/or output based, is effective (Benati, 2001). However, long-term effects cannot be established.

Benati (2004) partially replicated Benati (2001). The two studies differ in the treatment groups, where PI is compared to SI only and EI only experimental groups (Benati, 2004). The materials used are adapted from Benati (2001) accordingly. Thirty-eight university-level participants were randomly assigned to a full PI (n=14); an SI (n=12)[24] and an EI group (n=12). The study maintained the same assessment interpretation and written production tasks as in Benati (2001). A pre, post-test design was followed with the delayed post-test being administered four weeks post-instruction.

Analysis showed that all groups improved in the interpretation tasks. The PI and SI groups improved much more than the EI group but were not significantly different from each other; however they outperformed the EI group both in the immediate and delayed post-test. The same results were obtained in the written production task, while the control group also improved post-treatment. Thus, findings support that the main variable in PI effects is the structured input component (Benati, 2004:216; see also Farley, 2004; Sanz, 2004; Sanz and Morgan-Short, 2004; VanPatten and Oikennon, 1996). Benati (2004:216) argues that the simplicity of the target form compared to previous studies that used more complex forms (such as OVS sentences in Spanish) may have contributed to improvement in performance of the EI group. As in previous PI studies, the findings of this study are not generalizable as the sample size is too small,

[24] The reader is reminded that the difference between full PI and SI lies on the inclusion of explicit instruction. In other words, full PI begins with explicit instruction followed by structured input (SI) activities, whereas SI type of instruction is still part of PI; however instruction begins immediately with referential and affective activities without any prior explicit instruction referring to the target form (see VanPatten, 2004; VanPatten and Oikennon, 1996; Lee and Benati, 2007).

while the delayed post-test cannot provide valid conclusions for long-term effects. Interestingly, there is again in this study (Benati 2004) no reference to the possible effect(s) that the textual input enhancement in the EI phase might have had on the obtained results.

Lee and Benati (2007a) did explore the effects of textually enhanced versus unenhanced SI activities in the acquisition of future tense morphology, partially replicating Benati (2001) and (2004). Twenty first semester undergraduate students were randomly assigned to two groups, an enhanced SI (n=10) and an unenhanced SI group (n=10). The type of enhancement used in the referential activities was oral enhancement of the target form by pronouncing the targeted verb ending more loudly (Lee and Benati, 2007:102). For the affective activities, textual enhancement was applied with the target form verb ending being bolded and underlined (Lee and Benati, 2007a:103). The study adopted a pre- and immediate post-test design measuring interpretation and written production, as in Benati (2001, 2004). Results from the interpretation task showed that both groups made significant gains post-instruction, however no significant differences were reported between groups. The same results were revealed for the written production task. Lee and Benati (2007a:109) concluded that SI enhanced or unenhanced is the main factor contributing towards accurate comprehension and production of the target form (similar to Sanz, 2004; Sanz and Morgan Short, 2004; VanPatten and Oikennon, 1996).

Findings of the Lee and Benati (2007a) study are important since they compare PI to another input based approach namely IE, applying aural and textual enhancement. The study provides evidence that SI can be an effective teaching intervention when combined or not with IE. However, the sample size of the study and the lack of the delayed post-test cannot provide reliable and generalizable conclusions regarding the effectiveness of the proposed teaching intervention. The fact that the enhancement variable was not isolated, i.e. both types of enhancement were applied, namely aural and textual enhancement, does not allow conclusions regarding which of the two types of enhancement is the more effective.

Lee and Benati (2007b), on the other hand have investigated the effectiveness of aurally and written enhanced versus unenhanced SI activities with no instruction in Italian learners of Japanese, partially replicating Lee and Benati (2007). The target form was the Japanese past tense, which for Italian learners of Japanese poses the processing problem of location of the past marker at the end of the sentence (Lee and Benati, 2007:112). Thus, among other processing problems the learners came across the 'Primacy of Content Words Principle: learners process content words in the input before anything else' and sub-principle 'Sentence Location Principle', according to which learners tend to process items in sentence-initial position before those in medial and final position (VanPatten, 2004:14).

Twenty-six Italian adult learners of Japanese were assigned to three groups; a SI activities (n=9); an enhanced SI activities (n=10) and a control group (n=7), all of whom received a two-day treatment. The same type of enhancement (aural and textual bold and underlining) was provided for the enhancement SI group and the same types of materials used in the Lee and Benati (2007) study were adapted in Japanese. A pre, post-test design was used, while a four-hour treatment over two consecutive days was delivered. The delayed post-test was administered one week after instruction. Interpretation and production assessment tasks were applied. Results showed no significant differences between the enhanced and unenhanced groups, however both groups outperformed the control group significantly. Gains were retained from the immediate to the delayed post-test condition for both groups in the two assessment tasks. Lee and Benati (2007:125) concluded that it is the SI activities and not enhancement that accounts for grammatical gains in acquisition (similar to Lee and Benati, 2007; Farley, 2004; Sanz, 2004; Sanz and Morgan-Short, 2004; VanPatten and Oikennon, 1996). However, the small sample size, the timing of the delayed post-test administration, and the failure to isolate the enhancement variable do not allow for valid and generalizable conclusions.

Meta-analyses suggest that studies including explicit rule explanation are more effective teaching interventions than implicit types of instruction (Doughty, 2003; Lee and Huang, 2008; Norris and Ortega, 2000). The role of explicit instruction and feedback has been widely investigated in the field of PI. Studies have shown that PI is equally effective, whether or not it includes EI (Benati, 2001, 2004; Sanz, 2004; Sanz and Morgan-Short, 2004; VanPatten and Oikennon, 1996). Studies investigating the role of feedback have been, on the other hand, inconclusive regarding the role of feedback in PI (Sanz, 2004; Sanz and Morgan-Short, 2004). However, recent studies investigating the role of EI and feedback in PI suggest that groups receiving 'full PI' versions (EI and feedback) establish correct form-meaning connections earlier and have a greater percentage of accuracy than groups that receive only SI activities.

Fernandez (2008) investigated the effects of EI and 'implicit feedback' (following Sanz, 2004) in PI. Assuming that EI can have a beneficial effect on language learning, she investigated whether EI encourages more accurate processing than absence of EI in materials; whether EI promotes faster processing of the target form (in terms of response time in providing an answer); and whether EI promotes and maintains correct input processing. The criterion for correct processing was established as the point when learners provided correct responses for at least three target forms and one distracter in a row ('criterion analysis', Fernandez, 2008:285). Materials were computerised in order to track response time and accuracy, and one group received full PI and the other only SI activities. Fernandez (2008) conducted two experiments, one with Spanish OVS word order clitics and a second with Spanish subjunctive.

The treatment materials used for the Spanish OVS word order were the same as in the VanPatten and Cadierno (1993) and VanPatten and Oikennon (1996) studies in terms of the EI and SI activities. However, only referential activities were included because they could be scored[25]. A total of 30 referential types of items for the first target form, 20 target OVS and 10 SVO/SOV distracters were delivered through e-prime. Feedback was available on the screen in the form of 'correct' or 'incorrect' for both groups after each response. 84 participants were distributed into two groups, namely 42 in the full PI and 42 in the SI (referential) group. Results showed for the Spanish direct objects and word order that there were no significant differences between the explicit and non- explicit instruction groups in the three measures. A similar number of participants from both groups (26/42 for PI and 21/42 in the SI group) reached the criterion. However,

[25] Affective activities do not have correct or incorrect responses, as learners provide what applies to them personally.

hypotheses about trials to criterion, accuracy and faster processing were not supported, as there were no statistically significant differences between the two groups.

In the second experiment, hypotheses, participants and study design were the same as in the first one, the only difference being the target form, namely Spanish subjunctive. Materials in the second experiment were taken from Farley (2000) and were transcribed into computerised materials. As in the first experiment, there were 30 items (Fernandez, 2008:294). Findings of the second experiment showed that for the Spanish subjunctive, EI did make a difference, as significant differences confirmed that the number of participants that reached the criterion was proportionally and statistically significantly different in the full PI (32/42 full PI vs. 21/42 for the SI group). The full PI group took significantly fewer trials to reach the criterion, responded significantly faster than the SI group and maintained accuracy after having reached the criterion in comparison to the SI group (Fernandez, 2008).

Fernandez (2008:297) attributes the differences in results between the two experiments primarily to the nature of the target forms. In the first experiment, the target form (OVS and word order) required a more demanding type of processing, as learners had to interpret the meaning of the sentence in order to match the sentence with the corresponding picture. For the Spanish subjunctive, participants had to attend to the verb inflection and match it with the corresponding picture. Fernandez (2008:298) thus concludes that 'EI seems to be beneficial when the task is to notice and process a single form, but it does not seem to play any role when the task is to assign different grammatical roles in sentences'. This might be a possible, though not generalizable, explanation. One could claim that learners can notice and process a single form with OVS sentences as well. For instance in German masculine OVS sentences, learners can process only one of the two nouns in a given sentence (correctly or incorrectly) instead of processing and assigning roles to items in the whole sentence, e.g. the can process the first noun of the sentence and decide who performs the action without processing the whole sentence (VanPatten, 1996, 2004, 2007, 2009; see also Culman et al. 2009). Interestingly, Fernandez (2008) emphasizes the role of L1 transfer in her explanation of the obtained results. She argues that English learners are used to processing redundant forms in their L1, such as 'he walks' or 'he walked yesterday' (Fernandez, 2008:298).

Therefore in the case of the subjunctive, they do not have to build a new strategy, which is not the case for word order, as English follows a strict SVO structure and therefore learners have to look in the sentence for determiner forms in order to establish subject versus objects and verb forms in order to correctly assign roles (ibid). Indeed, the suggested role of L1 transfer strategies and effects in acquiring a language through PI is a valid claim, one that merits detailed explanation first at a theoretical level (IP) and investigation at a practical level (PI).

Task familiarization might have been an additional factor for the results obtained. Although EI seems to have helped participants notice information that they might not have noticed or might have taken more time to notice, it is obvious that participants did not start instantaneous correct matching of the target forms in both experiments (Fernandez, 2008:298). Task familiarization can be indeed one factor, however individual differences might also be another (Benati, 2001; Gass and Mackey, 2005; Skehan, 1998).

Nonetheless, Fernandez (2008) raises very important issues that IP and PI should address. The results regarding the possible contribution towards earlier, faster and more accurate processing of the target form when full PI is applied are very important, especially when considering the criticism that the field has received in terms of the role, the form and necessity of EI in PI (see DeKeyser et al. 2002; DeKeyser and Sokalski, 1996; Salaberry, 1997). Yet, Fernandez's study fails to provide data regarding long-term effects as delayed post-tests were not applied. It is true that the aim of the study was to observe learners' behaviour during input processing; however generalizable data are necessary in order to make claims about the role and effectiveness of EI in language learning. In addition, a larger sample size could support such claims. Finally, as in most PI studies, there is no analysis of the role of the 'implicit feedback' provided in the study, although participants received feedback (i.e. 'correct' and/or 'incorrect'). PI and IP need to address the role of such feedback. Does it assist in hypothesis formation and testing? Does it promote correct form-meaning connections?

The Culman et al. (2009) study, replicating Fernandez (2008), differed with regard to the target language, which was German, and the target form, which was accusative case marking on articles and word order. The processing strategy in focus was the FNP. Fifty-nine first (n=31) and third semester (n=28) English learners of German participated in the study. Intact classes were used in the study, where participants were divided into four groups: PI first semester (n=16); SI first semester (n=15); PI third semester (n=14); and SI third semester (n=14). The design of the study followed the same pattern as Fernandez (2008), and online materials using e-prime included 30 target items; 10 sentences in the canonical (SVO) word order with at least one masculine noun (Type 1); 10 sentences of inverted (OVS) word order including at least one masculine noun (Type 2); and 10 distracter sentences that did not include masculine nouns (Type 3), all illustrated with examples next page (Figure 2.6):

Type 1: ***Die Frau ruft den Mann.***
The-NOM woman calls the-ACC man
(the man calls the woman.)

Type 2: ***Den Mann ruft die Frau.***
The-ACC man calls the-NOM woman
(the woman calls the man)

Type 3: ***Die Frau ruft das Mädchen.***
The woman-NOM/ACC calls the/NOM /ACC girl
(the woman calls the girl/the girl calls the woman)

Figure 2.6: Types of sentences applied in Culman et al. 2009 study
(Source: Culman et al. 2009:24)

Participants listened to the sentences and were required to choose between two pictures, the one that corresponded best to the sentence. Post treatment, they were asked to complete a questionnaire, requiring stating the rule of German inverted OVS sentences, the processing of the German accusative case, and to report if they had learned any ‘tricks’ they had applied as processing strategies when working with the materials. The scoring procedure was different from the one used in Fernandez’s (2008) study. Data analysis focused on four item sets distributed across the 30 items (Set 1: 1–3 items; Set 2: 8–10 items; Set 3: 15–17 items; and Set 4: 28–30 items). Each of these sets consisted of three items, one SVO followed by an OVS and in turn followed by a SVO item. Each set was scored on a scale of0–3 in which each item was weighted equally (Culman et al. 2009). Results showed that the ‘full’ PI group correctly processed sentences earlier than the SI group, suggesting that EI is the variable that speeds up processing and accuracy (Culman et al. 2009). Interestingly, in the questionnaires, twenty out of the thirty participants of the PI group but only three out of the twenty nine participants of the SI group reported using the definite article as a cue to determine the agent of the sentence.

Comparing the results of the two studies (Culman et al. 2009 with Fernandez, 2008), researchers concluded that the contradictory results derive from the complexity of the structure in Spanish, as also previously discussed. Culman et al. (2009:28) assume that despite the fact Fernandez (2008) focused only on one third person singular Spanish object pronoun, if learners were required to have some knowledge of all four possible object pronouns then the target form would be ‘a bit too much for average learners to keep in working memory’ (Culman et al. 2009:21). However, the particular assumption is not adequately explained in terms of why learners should have knowledge of the other possible pronouns. In IP and PI terms, it is rather a contradictory argument, as materials should be developed in such a way that no processing of other forms is required at the same time. This means that if there are four object pronouns in Spanish, materials should be designed so that learners are forced to process only the target form; thus other forms should not be a prerequisite to complete the SI activities. In the IP theoretical framework, there are no guidelines concerning when a target form should be more appropriately applied through PI; or when a IP Principle presented through PI will be more or less effective. Moreover, the research design of the Fernandez (2008) study was based on the original VanPatten and Cadierno study (1993) since it was the particular form, the particular items and the same level and age of students that showed that PI was effective in processing and acquiring the specific target form. Perhaps the fact that affective activities were not included in Fernandez (2008) and Culman et al. (2009) had an effect; however further investigation is necessary to support such a claim, especially where affective activities do not involve ‘correct’ and ‘incorrect’ responses but elicit learners’ opinions.

Overall the Culman et al. (2009) study showed that EI is more beneficial when used in conjunction with SI activities. On the other hand, the lack of a delayed post-test does not allow for conclusion regarding long term effects. Results contradict Sanz (2004), Sanz and Morgan Short (2004), the original VanPatten and Oikennon (1996) and partially[26] Fernandez, (2008). However, all studies mentioned are limited in their sample sizes and lack of a delayed post-test. Therefore valid and generalizable conclusions cannot be drawn (see also critique from Allen, 2000; Gass and Mackey, 2005; Mitchell and Myles, 2004; Norris and Ortega, 2000).

2.3.4 Conclusion for PI studies and present study's contributions towards PI

The PI studies reviewed above investigated three main points: the effects of the 'FNP'; the role of EI and feedback, in terms of applying 'full PI' versus SI activities and/or implicit/explicit feedback and the effects of enhanced versus unenhanced full PI and/or SI only. Findings of the studies have shown that the application of 'full PI' (VanPatten and Cadierno, 1993; VanPatten and Fernandez, 2004; Wong, 2004) vs. SI activities only (Farley, 2004; Sanz, 2004; Sanz and Morgan Short, 2004; VanPatten and Oikennon, 1996) as well as the application of enhanced vs. unenhanced versions of either 'full PI' and/or SI activities only (Benati, 2001, 2004; Lee and Benati, 2007, 2007b) are equally effective. Recent studies have also shown that EI with the combination of implicit feedback is more effective in affecting input processing (in terms of promoting speedier and more accurate processing of the target form) than the sole application of SI activities (Culman et al. 2009; Fernandez, 2008; for further details regarding these studies please refer to section 3.1.2.1). Based on the latest findings, the present study adopts the full PI version, while exploring possible effects of coloured typographical/textual enhanced vs. unenhanced PI. The novelty of the present study is to focus on the FNP and the acquisition of German word order and case marking.

The review of IP studies has shown that attempts to replicate original studies have not been successful (see critique on Allen, 2000; DeKeyser and Sokalski, 1996; Salaberry, 1997). Criticism has mainly focused on the fact that IP and PI are vague and lack adequate theoretical explanation, especially in terms of what should be included in a PI treatment, the aims and objectives and when and how form-meaning connections are established, i.e. are learners making form-meaning connections during and/or after PI? How can PI findings be explained based on PI and IP theories? So far, no PI study has provided an answer to these thorny issues (Carroll, 2004:297; Collentine, 2004:172;

[26]"Partially" refers here to the second study that Fernandez (2008) conducted and which does not contradict Culman et al (2009).

DeKeyser et al. 2002; Doughty, 2003, 2004; Harrington, 2004; Lee, 2004). The present study, although it is not a replication study, aims at addressing these issues. Findings of the present study will be analysed and discussed based on the theoretical frameworks and models reviewed in this chapter.

2.4 Language processing frameworks and second language acquisition Introduction

The focus of this chapter is to provide a review of frameworks focusing on the role of input processing while also explaining the processes involved in second language acquisition. Because of the processing view of language acquisition that both IE and PI adopt, the majority of the frameworks that will be presented derive mainly from cognitivist, emergentist and processing-cognitivist theoretical perspectives. However reference to other approaches, such as Universal Grammar 'UG', informing IE and PI, as well as the frameworks under review will be made. An exhaustive review of all available theoretical frameworks and models of SLA is, however, beyond the scope of the present study.

In this section McLaughlin's (1987, 1990) 'Restructuring' approach followed by the 'Competition Model' (Bates and MacWhinney, 1989) and Pienemann's (1984, 1998, 2007) 'Processability Theory' will be reviewed. In the second part of this section, a modular approach of language processing will be presented with the review of two interrelated albeit different approaches of language processing, namely; Carroll's (2001, 2007) 'Autonomous Induction Theory'; and Sharwood Smith and Truscott's (2004, 2005 and in prep) 'Modular Online Growth and Use of Language' (MOGUL).Other prominent frameworks, such as Jackendoff's (1987, 1997, 2002, 2007) 'Parallel Architecture' and Baars' (1988, 2007) 'Global Workspace Theory' are not reviewed in the present study; however, reference to these frameworks will be made where necessary, as some of them inform frameworks relevant to the purposes of the present study.

2.4.1 Restructuring

McLaughlin (1987) adopts a cognitive psychological approach to second language learning, emphasising the importance of ‘automatic’ vs. ‘controlled’ processes and restructuring. He assumes that the human mind has a limited capacity for processing information and that second language learning involves a shift from controlled to automatic processing (see also Anderson, 1985). Language Acquisition is viewed as a complex cognitive skill, because various aspects must be practiced and integrated into fluent performance. McLaughlin argues that at the initial stage, complex cognitive skills are time-consuming and require attention, which is why they are defined as controlled processes. For instance, when we start learning how to drive a car, a great deal of attention is required in order to learn how to use the pedals, e.g. engage the clutch with the gas pedal and change gears. Through practice, sub-skills become automatic and controlled processes are free to be allocated to higher levels of processing, i.e. we can drive a car and carry a conversation at the same time. However, if we drive on an icy road, controlled processing comes into play again and it is difficult to keep up the conversation (McLaughlin, 1990: 114). Hence, a skill must be practiced again and again until it is over-learned and becomes acquired, which in turn leads to freeing up controlled processes in order to acquire new learned knowledge, as control processing lays down the ‘stepping stones’ for automatic processing (Shiffrin and Schneider, 1977).

Practice for McLaughlin can be effective in two different ways. Firstly, it can lead to improvements in performance as sub-skills become automated, as in the driving example. Secondly, improvement in performance can lead to ‘restructuring’, i.e. structural change (see McLaughlin, 1987, 1990:117). ‘Restructuring’, from an information processing perspective, is a cognitive process ‘in which the components of a task are coordinated, integrated or reorganised into new units, thereby allowing the procedure involving old components to be replaced by a more efficient procedure involving new components’ (Cheng, 1985; McLaughlin, 1990:118).

According to McLaughlin (1987, 1990), acquisition is not seen as a linear or cumulative process. Restructuring involves three stages. In the first stage learners of English can for example be using the English irregular past form ‘went’. In the second phase ‘organisation-oriented-procedures’ take place, as a result of the learners’ attempt to simplify, unify and control the internal representation. In this phase a developmental shift may occur in performance, as learners go through a transition phase. This means that learners may go through a stage where instead of supplying *went* for the English irregular past form, they supply *goed* based on acquired rules for the formulation of the regular forms. In time (third stage) *goed* will be replaced with the initial form *went* (see study by Wode, Bahns, Bodey, and Frank, 1978). This example of morphological development shows that learning is not a linear or a cumulative process. Rather it shows that development is U-shaped, where U could potentially symbolise the shape of the learning curve taking place, i.e. initial appearance of the correct irregular verb form (*went*), subsequently regularised to (*goed*) and reappearance of the correct form (*went*) (Cheng, 1985; McLaughlin, 1987, 1990; McLaughlin and Heredia, 1996; Wode, Bahns, Bodey, and Frank, 1978). SLA can be therefore defined by ‘backsliding and loss of forms that seemingly were mastered’ (McLaughlin, 1990:121; see also Lightbown, 1985). New forms may therefore be responsible for a decline in performance, as their appearance causes restructuring to the whole system even to forms that have been mastered (Lightbown, 1985; McLaughlin, 1990; McLaughlin and Heredia, 1996).

Ellis (1985) describes syntactic restructuring as a transitional shift that occurs between two stages, when processing form-meaning mappings. The initial stage involves a phase where learners formulate hypotheses that may or may not correspond to the target language. Learners may therefore use two or more forms freely, e.g. *no* and *don’t* both for the indicative and imperative. In a second phase, learners go through an ‘economy principle’ in an attempt to maximise linguistic resources. In this phase, redundant forms are eliminated. For example either *no* or *don’t* will be used. McLaughlin (1990) argues that, in the initial stage, there is ‘non-systemic variation’ because forms are assimilated and not yet integrated into the learner’s system. On the contrary, in the second phase, forms are accommodated by the restructuring of the existing form-function system. In this phase, new forms are given their own meanings to perform because of systemic variation. In the third phase, ‘learners restructure their knowledge until they sort out form-function relationships’. However, it is not always the case that all learners reach this third stage (McLaughlin, 1990:121). The findings of Meisel, Clahsen and

Pienemann's (1981) study on the acquisition of German word order provide supporting evidence for these three stages of restructuring. The results indicated that learners went through a temporary stage where they omitted object noun phrases, or left out categories that should be inverted such as subjects and verbs. Contrary to Pienemann's (1984, 1998, 2007) interpretations (see section 2.1.3 for a detailed discussion), McLaughlin (1987, 1990) attributes this backsliding to restructuring. Many studies in first language acquisition also refer to lexical, syntactic and comprehension restructuring (among many Keil, 1983; Keil and Caroll, 1980; Slobin, 1987). However, these studies will not be dealt with in the present study, which focuses on the effects of restructuring in second language acquisition.

McLaughlin (1987, 1990) discusses the strategic shifts that learners develop when learning a second language. Learners begin by developing a strategy that involves memorizing formulas and chunks, which they hardly understand and which contain complex syntax and vocabulary (Ellis, 1996a, 1996b; Myles et al. 1998, 1999; Wong Fillmore, 1976; Wray, 2002)[27].

Restructuring is a very important concept used to explain the process of first and second language acquisition. Controlled and automatic processes provide a coherent interpretation of the processes involved in language acquisition. Restructuring as a concept has been rather influential in the formulation of the Input Processing approach proposed by VanPatten (1996, 2004, 2007, 2009). For the purposes of the present study, I argue that, despite the fact that restructuring is a cognitive psychological approach that can provide a coherent interpretation of second language processes such as learning and acquisition, it fails to provide precise information regarding the initial stages of the process, e.g. how we come to process input; what the processes involved are; what the role and impact on language learning of other psycholinguistic concepts is, such as attention and awareness.

[27] McLaughlin (1990) argues that once a new language has been acquired, subsequent language learning is facilitated, supporting the idea that positive L1 transfer can be carried out in the process. Moreover, empirical evidence supports the idea that multilingual second language (i.e. L3) learners develop the ability to shift strategies and can notice structural similarities and differences between languages. Thus they may progress faster than non-multilinguals (McLaughlin, 1990; see also studies Morgan and Newport, 1981; Nation and McLaughlin, 1986; Nayak, Hansen, Krueger and McLaughlin, 1990; Ramsey, 1980).

In the following section another model that partially informs Input Processing is reviewed, namely the 'Competition Model' (Bates and MacWhinney, 1982, 1987; MacWhinney, 1987, 2001, 2002; MacWhinney and Bates, 1989; see also VanPatten, 1996; 2004, 2007, 2009).

2.4.2 Competition Model (CM)

The 'Competition Model' (CM) proposed by Bates and MacWhinney (1982, 1987; MacWhinney, 1987, 2001, 2002; MacWhinney and Bates, 1989), adopts a lexical functionalist approach, and applies it to both first and second language acquisition. In its original instantiation, the CM was proposed as a theory of cross-linguistic sentence processing, arguing that people interpret the meaning of a sentence by taking into account various linguistic cues contained in the sentence context, such as word order, morphology, and semantic characteristics (e.g., animacy). According to the weight of cues, which are computed inductively at a sentence level, meaning is interpret based on probabilistic values, i.e. the higher probabilistic value a cue receives the more chances it has to be chosen in interpreting meaning. For the Competition Model the weight of cues differs between languages. Thus, second language learning is based on learning the cues that are important for the second language in order to successfully interpret sentences in the given language (Bates and MacWhinney, 1982, 1987; MacWhinney, 1987, 2001, 2002; MacWhinney and Bates, 1989; for further details please refer to the next paragraphs). By definition the CM adopts an emergentist approach to second language acquisition. Although the theoretical underpinnings of the CM support the idea of innate mechanisms being responsible for language learning it is not assumed that there is a special mental organ consisting of predetermined linguistic properties and universal grammar (i.e. a Language Acquisition Device, LAD, Chomsky, 1972; see also Carroll, 2001, 2007). Instead, Bates and MacWhinney (1982) argue that language acquisition is governed by general cognitive mechanisms and is dependent upon experience with language and influence of the environment. In the next paragraphs these processes and mechanisms will be explained in more detail.

The major theoretical underpinning of the CM is that 'mental processing is competitive' (Bates and MacWhinney, 1987:3). In the CM a language processor consists of three parts, namely a segmenter, a lexicalizer and a parser. The ‘segmenter’ segments the noise entering the processor, while the lexicalizer activates candidate lexical items which, in turn, activate role expectations. Once a role expectation is activated, the processor checks if there is a match in the currently lexicalized items and clusters, in which case a candidate’s ‘valence[28]’ attachment is formed. However, spreading activation can occur, that is when several competing attachments are formed. In this case, if the strength of one of the competing attachments becomes overwhelming, it wins the competition. Although MacWhinney (1987:9) notes that, in many cases, the final decision between competing attachments is not made until the end of the sentence or clause.

The functionalist belief of the CM is explicitly stated: the ‘forms of natural languages are created, governed, constrained, acquired and used in the service of communicative functions’ (MacWhinney, Bates and Kliegl, 1984:128) and lexical knowledge is emphasized as the main controller of parsing, processing and acquisition. This assumption is compatible with the dominant view held by research on sentence processing, which emphasizes the role of lexical knowledge in language processing (Ellis, 1998, 2002; Harrington, 2001; Lakoff, 1987; Massaro, 1987; Rosch, 1977; Whorf, 1956; see also MacWhinney, 1989). In other words, in the CM, sentence processing is controlled by lexical items (Givón, 1979; MacWhinney, 1989).

The underlying assumption of the CM is that humans must develop form-function mappings for language use and acquisition. In the CM, surface forms such as word order patterns, lexical-semantic animacy, morphological markings and prosody (i.e., contrastive stress), are termed ‘cues’ (MacWhinney, 1982, 1992). Cues are any piece of information that can be used by listeners and speakers to determine the relationship between form and meaning (MacWhinney and Bates, 1989). In English, when we hear the sentence ‘the woman kisses the man’ we interpret the woman as the agent of the sentence by using the preverbal position, animacy, agreement and initial position cue.

---

[28]Valence refers to the number of arguments controlled by a verbal predicate, i.e. arguments can include subject and objects of a verb. It is related to verb transitivity, however it is not the same, as transitivity accounts for only object arguments of the verbal predicate. The linguistic meaning of valence derives from the definition of valency in chemistry, a metaphor provided by Tesniére (1959). Similarly, MacWhinney (1987) and Bates and MacWhinney (1989) propose ‘valence bridges’ in determining the meaning and the functions of the words involved in construction.

At an early stage of acquisition all ‘cues’ have equal validity, however usage strengthens the weight of cues, resulting in different types of ‘cue strength’ and ‘cue validity’. ‘Cue validity’ refers to the degree to which a cue helps in making correct form-meaning interpretations during on-line processing. Bates and MacWhinney (1989) argue that ‘cue validity’ determines cue strength, since it is the product of ‘cue availability’ and ‘cue reliability’. Cue availability refers to the frequency of a cue in the input, while cue reliability denotes the consistency of mapping a particular form. In English, we have cue availability in determining the agent of the sentence referring to the preverbal position cue, as English has a strictly SVO word order. On the contrary, German does not have a strict SVO word order. Therefore the preverbal position cue is not a reliable cue, when processing German sentences, as it does not always lead to correct conclusions. In German both ‘Der Mann küsst die Frau’ (The man kisses the woman) and ‘Den Mann küsst die Frau’ (The woman kisses the man) are possible. An English learner of German relying on the preverbal positional cue would interpret ‘the man’ as the agent in both sentences. In order to correctly decode the meaning, the learner should instead use cue reliability, i.e. case marking, to establish that in the first example ‘the man’ is the subject while in the second, ‘the man’ is the object.In a nutshell, for the CM language acquisition is a process consisting of a series of competitive cognitive processes activated in a rich and stimulating environment when analog signal, i.e. language is present. The weights of the competing representations are computed and adjusted based on the learner's experience with the target language. Therefore learner’s extensive exposure to the target language increases understanding of the meaning of sentences in the target language and leads to acquisition (see Bates and MacWhinney, 1989; MacWhinney, 1987, 1989; MacWhinney and Bates, 1989).

When two cues are in conflict for the same role, the CM proposes ‘conflict validity’ in order to resolve the conflict and assign correct roles and make correct mappings. For instance, in English in a sentence like ‘the ball hit the child’ the preverbal position cue indicates the ball as the agent of the sentence. ‘Cue animacy’ determines the child as the agent of the sentence. Therefore the two cues, ‘the ball’ and ‘the child’, are in competition for the agent role. The solution will rely solely on cue validity. This means that a cue high in conflict validity is usually the one that is maximally reliable because it entails more weight than a cue high in overall validity (McDonald, 1986, 1987). This implies that conflict validity is of greater significance when looking at adults' second language processing. According to the CM, children, when strengthening cues in their systems, are more influenced by overall validity because they have not been exposed to many conflict sentences, (MacWhinney, 1987). A valid assumption for the role and essence of conflict validity in the CM is that it provides an understanding of how rare and/or non existing phenomena in the L1 can be processed and acquired when available in the L2, for instance English learners of German acquiring case marking.

Capacity limitations in attention and working memory are one of the two reasons why a cue can be harder and costlier to detect. For the CM this limitation is referred to as a 'perceivability limitation' denoting the degree of difficulty in detecting a cue (e.g. the accusative *–t* in Hungarian after consonants), which can lead in delayed initial acquisition. A second factor can be the cue itself, as some cues are inherently harder to detect, such as case marking in German, which is referred to in the CM as an 'assignability limitation[29]'. The outcome of the two processing limitations is 'cue cost' (Bates and MacWhinney, 1989). The underlying assumption is that cues become less reliable when they are difficult and costly to process. With this in mind, the model proposes that some cues are considered highly assignable because they demand less information processing, while others have low or limited assignability (Kail, 1989).

Findings of studies investigating the CM have shown that languages differ with regard to cues, both quantitatively, i.e. the degree of reliance, and qualitatively i.e. types of cues (Bates et al. 1999). For example, in cue reliance, adult English speakers rely heavily on the preverbal position (McDonald, 1987), whereas German speakers rely primarily on case marking, followed by animacy, agreement and word order (MacWhinney, Bates and Kliegl, 1984). It should be noted that materials used to test the CM have included both simple grammatical and ungrammatical combinations at sentence level, which has been largely criticized due to the use of ungrammatical sentences and their implications on the learning process, i.e. participants may in this way learn ungrammatical sentences (although there have been exceptions see Bates et al. 1999; MacWhinney and Pléh, 1988; Schelstraete and Degand, 1998). Critics have

[29] Assignability also arises from a limited working memory system. According to the CM, cues have different kinds of processing demands. Some cues demand less information processing and therefore are highly assignable, while others are more demanding and thereby low in assignability (Kail, 1989).

also questioned whether grammaticality judgement at sentence level mirrors actual communicative situations where it has been argued that learners are prompted to develop a strategy to decipher meaning when processing simple sentences at a communicative level (Gibson, 1992; McLaughlin and Harrington, 1989). Processing Instruction has also received the same criticism, since its operationalization echoes the propositions of the CM (DeKeyser et al. 2002; Harrington, 2004).

VanPatten (2004:24) points out that the CM cannot provide explanations about non-conflict situations in sentences such as *'monkey-bite-baboon'*. According to VanPatten, in these sentences learners would assign the first noun, *monkey*, the subject role regardless of their first language. He therefore argues that 'there is something much more primitive and default about word order (the first noun as the subject) and that differences only surface when we deliberately put conflict into sentences' (VanPatten, 2004:24). For the purposes of the present study I argue that the CM could provide a theoretical explanation about the role of cues developed through PI and/or IE at a sentence level and their impact on language processing and acquisition. Furthermore, the role of attention, which is central for both PI and IE could be linked at a theoretical and practical level explaining the workings of the two approaches and their impact on second language development. However, the CM is limited and cannot provide sufficient interpretations regarding what form-meaning connections learners make and why those than others.

2.4.3 Processability Theory (PT)

Pienemann's Processability Theory, PT, (1984, 1998, 2005, 2007:137) is a theory of second language development. The underlying assumption that PT makes is that at any stage of language development, learners can only produce and/or comprehend L2 linguistic forms that the current state of the language processor can handle. In other words, PT assumes 'developmental readiness' as prerequisite for processing and acquisition. According to Pienemann (2007:137) the architecture of language processing that PT adopts is based on Levelt's[30] (1989) theory of speaking, which overlaps to some extent with Garret (1976, 1980, 1982) and the computational model of

[30] Reviewing Levelt's (1998) theory of speaking is beyond the scope of the present study, as the present study adopts an input-based approach focused on written input.

Kempen and Hoenkamp (1987). PT hypothesises that processing operations are automatic and implicit, while processing is incremental. This suggests that learners start processing of the required knowledge implicitly, however stepwise, starting with simpler processing procedures that are prerequisite for more complex ones, a principle derived from Kempen and Hoenkamp's (1987) model of incremental grammatical processing. Furthermore, grammatical processing is assumed to have access to a temporary memory store that can hold grammatical information, while the output of the processor is linear (Pienemann, 1982, 1998, 2005, 2007).

PT predicts a hierarchy of six stages in language development. These stages are universal and apply to all languages. However, language learners have to overcome first a processing challenge, in terms of learning to exchange grammatical information within a sentence. The language processor in this case will check if the new entry contains the same grammatical information and will match the features that correspond to each other. This process is referred to as 'feature unification' and is constrained by Lexical Functional Grammar (LFG) as proposed by Bresnan (1982). For example, language users have to ensure that a verb and its subject have the same number feature, or that the noun and its article have the same case feature, depending on the language. For instance, the sentence 'John read a books' is ungrammatical because *read* and *John* do not have the same person and number feature (third person singular) and *a* and *books* do not share the same number feature. If learners have not yet developed a fully functioning sentence procedure, these mismatches will not be detected. Therefore step-by step processing is required, as each step requires processing prerequisites that are developed in previous stages. Learners cannot skip stages, while processing of each structure must be successful in order to able to continue. Thus, the sequence of stages is hypothesized to be *'cumulative'* (Pienemann, 1984, 1998, 2005, 2007).

Initially, learners are able to map conceptual structures onto individual words and fixed phrases ('lemma access'). Once grammatical categories have been assigned to lexical items, then morphological markers can be produced. However, grammatical information cannot at this stage be exchanged. Therefore at this stage learners rely on strictly serial word order ('canonical SVO order'). At the next stage sharing of information is available between a Head and its modifiers. Once phrasal procedures are available, ‘Appointment Rules’ and ‘S-procedure’ can follow. This means that phrases can be assembled into sentences, while the function of each phrase is clear (i.e. the subject, verb of the S). Once this syntactic information is available at a sentence level, then subordinate clauses can develop. These stages predict the developmental sequences for acquisition (Pienemann, 1998:83-86; see also Mitchell and Myles, 2004:113-114). Referring to the acquisition of German[31] word order, the stages of L2 development that learners follow are predicted in the next page (Figure 2.7):

Stage	Formation	Example
1	Lemma access	Kinder (children)
2	Canonical Order- SVO (stage x)	Ich sehe die Kinder. (I see the children)
3	ADV Preposing (stage x+1)	Da sehe ich die Kinder. (There I see the children)
4	Verb SEP (stage x+2)	Ich habe die Kinder gesehen. (I have seen the children)
5	INV (stage x+3)	Gestern habe ich die Kinder gesehen. (Yesterday I have seen the children)
6	Verb Final (V-END) (stage x+4)	Ich habe gesagt, dass ich die Kinder gesehen habe. (I have said that I have seen the children)

Figure 2.7: Stages of acquisition in L2 German
(Pienemann, 1998)

PT also provides predictions regarding the relationship between functional and constituent structure through the ‘TOPIC Hypothesis’ (Pienemann et al. 2005). According to the ‘TOPIC Hypothesis', at initial stages of acquisition learners will not differentiate between SUBJECT and other grammatical functions in sentence-initial position (e.g. TOPIC). In LFG, TOPIC is a grammatical function. Therefore based on LFG, in a sentence such as ‘Den Mann küsst die Frau’ (The man-Obj. kisses-V the woman-Subj.), ‘Den Mann’ has in this case two functions, OBJECT and TOPIC (see

[31] PT stages for acquisition accounts for all languages, however, I am focusing on German, as the present study investigates the acquisition of German word order by English learners.

Pienemann, 2006, 2007 on SOV L1 influence and an SOV stage). PT hypothesises that learners go through three stages to reach ‘topicalization’, a process that allows learners to add a constituent before the subject position, while it allows them to differentiate between TOPIC and SUBJECT vs. TOPIC and OBJECT. At the initial stage, TOPIC and SUBJECT are not differentiated, e.g. ‘Der Mann sieht das Kind’ (the man-TOPIC and SUBJECT sees the child-OBJECT). At the second stage TOPIC is assigned to ‘noncore functions’, i.e. functions that do not relate arguments listed in the lexical entries of verbs such as adverbs. For instance, ‘Gestern sah der Mann das Kind’ (Yesterday-ADJUNCT saw the man-SUBJECT the child-OBJECT). Thus, TOPIC is differentiated from SUBJECT. In the third stage the TOPIC function is assigned to core functions such as OBJECTS, e.g. ‘Den Mann sieht das Kind’ (the man-TOPIC and OBJECT sees the child-SUBJECT) (see also Keβler, 2008:25-26; Pienemann et al. 2005).

PT provides a well-defined description of developmental sequences and problems that learners come across while acquiring a second language, and therefore PT could explain the developmental stages that English learners of German go through in the acquisition of case marking and word order[see also Vainikka and Young-Scholten (1995, 1998) on the role of ‘triggers’, which are defined as elements that cause re-organisation in the grammar (1995:6), where morphology is claimed to act as the trigger for acquisition of functional projections both in the L1 and L2 (see also White, 2003). However, instruction in PT is constrained by the developmental stages predicted; therefore instruction can only be effective if stages are not skipped (see ‘Teachability Hypothesis’ Pienemann, 1984). Thus, for PT, practice does not necessary make perfect in language learning. It therefore seems that PT has little to say about the aims, objectives and overall design methodology applied in the present study (for further details see Chapter 3).

In the present study it is necessary to account for how input and the way it is presented to learners can have an impact on language processing and acquisition. I therefore turn to reviewing modularity and acquisition focusing on the theoretical frameworks proposed by Carroll (2001, 2007) and Sharwood Smith and Truscott (2004, 2005 and in prep).

2.4.4 Modularity and acquisition

2.4.4.1 Autonomous Induction Theory

Carroll (1999, 2001, 2007) formulated a set of proposals which she called Autonomous Induction Theory (AIT). This theory integrates UG properties with modular processing systems (similar to Jackendoff, 1987, 1997, 2002, 2007). There are four main points that AIT attempts to address as a theory of SLA: linguistic competence, restructuring, processing, and learning (Carroll, 2001:39).

AIT is an adaptation of Jackendoff's (1987, 1997, 2002, 2007) Parallel Architecture and Holland et al.'s (1986) induction model. Based on Jackendoff's (1987, 1997, 2002, 2007) proposal, there are three autonomous representational systems, i.e. phonological, syntactical and conceptual structures, which are linked by integrative and correspondence processors, operating bottom-up or top-down between each system both in an autonomous and parallel way. The integrative processors entail primitive representations while they build structural representations. Therefore input for the integrative processors is any sort of symbols that can be processed. In other words, the integrative processors process the 'transduced stimuli' and produce some sort of 'input' for each of the three modules, i.e. phonological, morphosyntactic and conceptual structures. The correspondence processors link the information available in the integrative processors, as they are working in an autonomous way, both 'bottom up' and 'top down'. Carroll (2002:236-237) summarises bottom up and top down processing as follows (Figure 2.8):

- **Bottom-up correspondence processors:**

a) transduction of sound wave into acoustic information (via peripheral and central auditory analysis);
b) mapping of available acoustic information into phonological format;
c) mapping of available phonological structure into morphosyntactic format;
d) mapping of available morphosyntactic structure into conceptual format.

- **Top-down correspondence processors:**

a) mapping of available conceptual structures into morphosyntactic format;

- **Integrative processors:**

a) mapping of newly available phonological information into unified phonological structure;
b) mapping of newly available morphosyntactic information into a unified morphosyntactic structure;
c) mapping of newly available conceptual information into a unified conceptual structure.

Figure 2.8: Bottom up and Top down processing in Autonomous Induction Theory
(Adapted from Carroll, 2002: 236-237)

The aim of AIT is to explain a learner's linguistic competence based on psychological mechanisms. AIT presents changes in linguistic competence as changes in the mental grammar (similar to Jackendoff, 1987, 1997, 2002, 2007 and Sharwood Smith and Truscott, 2004, 2005 and in prep). These changes can occur through activity taking place in the Language Acquisition Device (LAD), which entails predetermined linguistic properties (Chomsky, 1972). The role of UG in AIT is limited to explaining how learners possess knowledge of essential grammatical components. Thus for AIT, UG provides primitives that are available implicitly through operations in the LAD. While for Carroll (2001:112), UG Principles and Parameters are 'a metaphor that has outlived its usefulness', as they fail to explain the restructuring of knowledge in adult SLA, the LAD can, through 'feature unification', (re)combine features or structures to create categories as well as equate two categories, i.e. unify a phonological with a morphosyntactic representation of a word (Carroll, 2007:155). Restructuring of grammar takes place in the LAD (Carroll, 2007:155).

There are three distinct types of input in AIT, namely primary linguistic data or stimuli, input to processing mechanisms, and input to the LAD (Carroll, 1999, 2001, 2002, 2007). Primary linguistic data or stimuli first enter the perceptual system, as acoustic strings and/or as written stimuli. These 'stimuli' are processed, i.e. encoded ('transduced') into representations, for instance sound waves are 'transduced' into phonetic representations. These 'stimuli' that have undergone this process are termed 'transduced stimuli' because of the transformation that has occurred, while they serve as input to the processing mechanisms, i.e. in this case the sound waves that have been transduced into a phonetic representation serve as input for the phonological processor. In turn, the activation of the lexical entry[32] from the mental lexicon that has begun with the activation of the phonetic representation activates the morphosyntactic and the semantic processors for further processing in order to build and/or assign the appropriate features in a modular way. 'Transduced stimuli' are viewed in AIT as intake for each of the processing mechanisms. For instance, when a German lexical entry

[32] Following Jackendoff (2002), a lexical entry for a word can be seen in two alternative but compatible ways: 1) as a 'rule' associating three separate and independent types of structure, a phonological structure (PS) a syntactic structure (SS) and a conceptual structure (CS); and 2) as a chain of structures that have been activated and put in correspondence with one another (PS<=>SS<=>CS); whereas each structure is processed independently within its own processing unit (module) according to the principles of that module (see also SST, in prep and http://www.msharwood.pwp.blueyonder.co.uk/mogul/).

becomes activated, *Katze* ‘cat’, the morphosyntactical processor informs the parser of its morphosyntactical features, such as NOUN, ANIMATE, FEMININE and SINGULAR. These formal features are available in the morphosyntactic system through UG, more specifically through the LAD[33]. If the learners do not have the feature for German nouns belonging to gender classes in their grammar, then LAD needs to restructure this representation in order to contain this information. In this case, some ‘input to the LAD’ will be necessary in order to trigger the restructuring process; for instance the determiner *die*, indicating the feminine gender in German (Carroll, 1999, 2001, 2002, 2007). Input in AIT is a multifaceted concept that stimulates different types of processing, i.e. input processing for the processing mechanisms, input processing for comprehension, input processing for the LAD, input processing for acquisition. That is why Carroll (1999, 2001, 2002, 2007) argues that the term ‘input’, as used in the SLA literature, should be abandoned because it is too vague.

Following Holland et al. (1986), Carroll (2001:120) also defines ‘induction, as a kind of learning, occurring in a wide variety of cognitive domains’, though she departs greatly from the original theory, as she follows Jackendoff’s propositions on autonomous and modular operations. She also distinguishes induction from inductive reasoning, which is located in the conceptual system and associated with problem-solving, hypothesis-formation and testing. According to Carroll (2001, 2002, 2007), when parsing fails, inductive learning, that is ‘i-learning’, begins. ‘i-learning’ is an on-line process that can change perceptual and parsing procedures such as detection, storage and recognition, suggesting that it can both encode and alter parsing procedures in order to build a representation, where ‘i-learning’ depends on working and long-term memory. However, this process is limited to creating a new representation which will be minimally different from “parent” representations. Following the ‘Competition Model’ (Bates and MacWhinney, 1989), Carroll (2001, 2002) argues that the novel structure will compete with the ‘parent’ structures, where the best representation will win (Carroll, 2001:203, 2002). ‘i-learning’ is therefore for Carroll (1999, 2001, 2002, 2007) synonymous with acquisition. When looking at L2 acquisition, L1 transfer has a

[33] Carroll (2001: 208) argues that ‘UG is part of the LAD in precisely the sense that it provides the basic representational system for the pre-linguistic child who neither induces the primitives of a grammatical representational system from linguistic stimuli nor maps them in a one-to-one correspondence from an initial innate conceptual system’.

prominent role, as processes are 'attuned to L1 properties' and therefore will automatically attempt to process the L2 stimuli, which will in turn lead to 'I-learning'.

Carroll (1999, 2001, 2002, 2007) also addresses the issue of linguistic competence versus learned knowledge. Competence refers to what Chomsky (1965) describes as an idealised linguistic knowledge of a language, of which the speaker is not necessarily aware. However, she differs from Chomsky (1965) in her use of the term 'I-language', as for her, 'I-language' also includes metalinguistic information encoded in the conceptual system. However, she does not specify how and where this knowledge is represented, nor does she address whether metalinguistic knowledge is projected into conscious awareness (widely accepted by Anderson, 1983, 1992; DeKeyser, 1998, 2003; Ellis, 2005; Ellis and Larsen-Freeman, 2006; Paradis, 2009; Sharwood Smith, 1991; 1993; Sharwood Smith and Truscott, 2004 and in prep; Truscott, 1998; see also Carroll, 2001:24).

Carroll hypothesises that acquisition is failure driven, like other generative researchers (e.g. Schwartz, 1993). When a mismatch occurs between input and the current activated representation, parsing fails and the LAD is activated. Then implicit processes can either re-structure I-language and can drive i-learning (Carroll, 1999, 2001, 2002, 2007). Acquisition occurs incidentally, as it is a 'byproduct of learners' interacting with language in some kind of setting' (Carroll, 2007:169). Processing either in the parsing processors or in the LAD is automatic, modular and implicit without conscious awareness. 'Noticing' is defined by Carroll following Jackendoff (1987, 1997, 2002, 2007) as conscious awareness and is a 'by-product of the processing of phonological representations' (Carroll, 2007:164). In other words, Carroll (1999, 2001, 2002, 2007) assumes that phonetic, morphosyntactical and semantic systems are inaccessible to conscious awareness, thus we can only notice articulated speech.

According to Carroll (2001, 2007), child language acquisition differs from SLA with regard to the 'logical problem' (Chomsky, 1981). In child acquisition, the logical problem is that children start acquiring a first language with a less developed representational system for the phonological, syntactic and conceptual representations than adults. Second language learners already have an established representational system in their L1. Therefore they have an already mature system, which is a 'product of '"innate" capacities and acquired information' and have already acquired the ability to parse L2 speech stream (Carroll, 2001:209). Hence, for Carroll (2001, 2007), SLA is a matter of empirical development, whereas 'access' to UG is not necessary for SLA but is implicated in an indirect way in order to explain how second language learners have already set representational systems when learning a second language. Carroll also explains the role of instruction in SLA for AIT. She argues that neither explicit instruction nor practice will be effective in changing the learners' mental representation 'unless the instruction causes changes in processing of the primary linguistic data' (Carroll, 2007:170). This effect is attributed to externally induced manipulation of input that can lead to failure in processing. She views VanPatten's (1993, 1996, 2004, 2007) Processing Instruction (PI) as an effective intervention in triggering parsing failure. On the contrary, Input Enhancement (IE) does not induce parsing failure.

To conclude, it seems that Carroll's AIT can provide a better explanation for teaching interventions that include activities requiring speech parsing. She argues that external acoustic stimuli are processed strictly bottom-up and like Schwartz (1993) that information from the conceptual system cannot interact with grammatical information. Language processing and acquisition occur only implicitly and learners only become consciously aware of phonological representations in speech production. Based on these grounds, it seems that AIT could be a suitable candidate to account for PI and IE, if treatment materials include oral input as well. However, as is discussed in section 3.4.1 the present study focuses only on written input, due to the nature of the target form and the restrictions provided from PI guidelines in developing treatment materials. Furthermore, the overlap between conscious and unconscious processes is not clearly defined, despite the fact that it is hypothesised that inductive learning can be both implicit and explicit. Moreover, although in AIT it is argued that L1 transfer plays a role in L2 acquisition, it is not sufficiently explained how transfer affects the process nor how variability occurs. On the other hand, it should be acknowledged that Carroll provides a precise description of what constitutes input and acquisition and how the two interact. Although she argues that instruction can have a limited effect on L2 acquisition, she views PI as the strongest candidate compared to other types of instruction in inducing parsing failure while processing the input. On the other hand, the extent to which IE can induce parsing failure through manipulation of the input remains unclear. Overall, it seems that AIT can potentially explain the processes involved when parsing input with PI and IE. However, interpretations will be limited, as there can be no elaboration on how the underlying psycholinguistic concept for the two teaching interventions, namely attention, affects second language processing and acquisition. Modular Online Growth and Use of Language (MOGUL) reviewed in the next section can provide a better framework for accounting for the role of attention.

2.4.4.2 Modular Online Growth and Use of Language (MOGUL)

Modular Online Growth and Use of Language (MOGUL) proposed by Sharwood Smith and Truscott (2004, 2005 and in prep, henceforth SST) is a language processing framework which emerged primarily from ideas developed by Fodor (1981, 1983) and Jackendoff (1987, 1997, 2002, 2007) in terms of integrating UG representations with a theory of real-time processing. MOGUL also adopts ideas developed by connectionist approaches such as the 'Competition Model' (Bates and MacWhinney, 1987), while it adopts Baars' (1988) ideas of the 'Global Workspace Theory' in relating the role of consciousness to language learning and acquisition. With 'Acquisition by Processing Theory' (APT), MOGUL provides an adaptation of the previously mentioned theories and models in order to explain language learning, also addressing the role of L1 transfer in SLA. According to SST (2004:1), MOGUL aims to build'a cross-disciplinary platform which can bring together research on linguistic structure and research on general cognition, all framed within a real-time processing perspective, and which, in the process, can generate new insights of its own'.

SST, like Carroll, see language as special, operating in a different way from other types of information processing and consisting of separate language sub-modules, namely the phonological (PS), (morpho) syntactic (SS) and conceptual structures (CS), processors and interfaces (similar to Chomsky, 1995; Jackendoff, 1987, 1997, 2002, 2007). Based on Jackendoff's (1987, 1997, 2002, 2007) propositions, the modules operate in their 'own unique code' and are autonomous but their basic internal structure is the same for all modules. They consist of an integrative processor unique to each module, interface processors that enable cross-matching with modules in adjacent modules and a memory store (SST, 2004, 2005 and in prep). The integrative processor processes elements that appear in the blackboard, which is roughly equivalent to working memory (WM). Each module comprises both an information store (long-term memory-LTM) and a blackboard (WM). The term 'blackboard' proposed by Jackendoff, is used as a metaphor to describe WM because it is used to temporarily write information during processing (SST, 2004:2). SST differ from Jackendoff (1987, 1997, 2002, 2007) in terms of not conceiving WM as a separate 'blackboard' or an alternative construct of long term-memory (LTM), following Cowan's (1993, 2001) model. They assume that WM is "a transient pattern of activation of elements within long-term memory stores". In other words, the blackboard is the most highly activated layer of long-term memory (SST, 2005:232 and in prep: 39). This means that when a stimulus activates elements in the long term memory (LTM), items can be activated and elevated to varying degrees until one item reaches the most highly activated layer of LTM, the blackboard, where the item is further processed (Cowan, 1993, 2001; Ruchkin, Grafman, Cameron and Berndt, 2003; Miyake and Shah, 1999; SST, 2005 and in prep; please refer to next paragraphs for more details).

Following Jackendoff (1987, 1997, 2002, 2007), there is a parallel domain language specific module, which consists of the phonological (PS) and morphosyntactical (SS) systems and a conceptual system (CS), which entail innate 'primitives' in terms of chunks, i.e. phonological features for the PS, syntactical features for the SS, and chunks of conceptual knowledge connected to a SS and PS for the CS. According to SST (in prep:25), the removal of the CS from the language module is based on propositions supporting that lexical meaning is acquired more explicitly from other aspects of lexical knowledge (see also N. Ellis, 1994). Structures can work both in isolation or can communicate information with each other through the interfaces in a modular way, as Figure 2.9 below depicts:

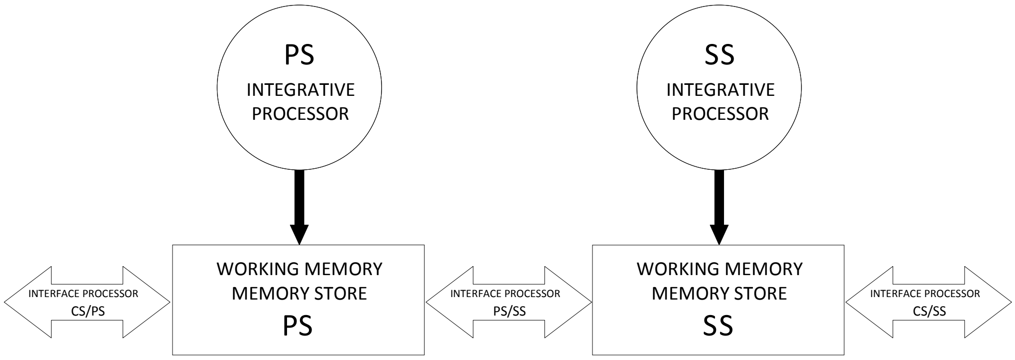


Figure 2.9: MOGUL memory stores, processors and interfaces
(adapted from http://www.hw.ac.uk/langWWW/mogul/)

The language module, otherwise referred to as the core language system, is considered in MOGUL as a 'blind spot' as we cannot become consciously aware of the processing of phonological and morphosyntactical structures. On the contrary, the conceptual store can process both linguistic and non-linguistic information, which can be communicated through the interfaces with the language module. Additionally, information processing in the conceptual structure (CS) can reach high levels of consciousness, thus we can become consciously aware of CS.For SST (2004 and in prep) the CS is located outside the language module because of the innate universal properties the latter entails (SST, 2004:4; see also Jackendoff, 1990). However, all modules are also connected to the non-linguistic perception system through a 'composite blackboard' of extralinguistic modules, i.e. auditory, kinaesthetic, olfactory, gustatory and visual processors, for perceptual processing and representations of sensory input, called perceptual output structures (POpS). POpS serve as the basis for higher level processing and are therefore accessible to consciousness because of their rich interconnectivity. As Figure 2.10 also illustrates, MOGUL is bi-directional, i.e. it accounts for both comprehension and production (SST in prep; a detailed presentation and review is provided in section 2.4.4.2.1):

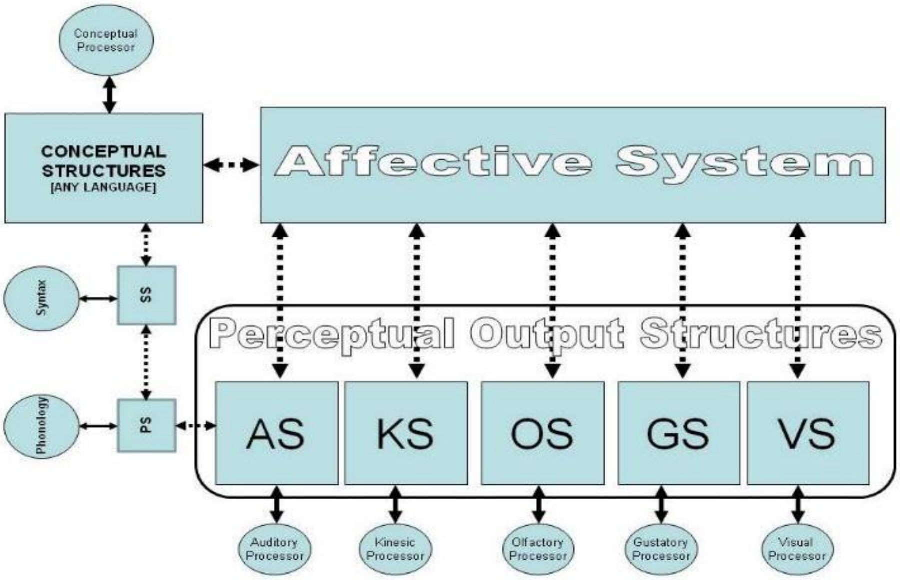


Figure 2.10: MOGUL architecture in a nutshell
(Sharwood Smith and Truscott, available at:
http://www.hw.ac.uk/langWWW/mogul/index.htm, cited with permission).

According to MOGUL, in line with Jackendoff, the three structures, i.e. PS, SS and CS, build the 'lexicon' of a language, which is defined as the 'metalinguistic notionuseful for an analytic understanding of language' (SS, available at: http://www.msharwood.pwp.blueyonder.co.uk/mogul/; see also SST in prep). This means that each of the three modules possesses a unique memory store (sub-lexicon), in which elements are stored either as independent elements or in various combinations that are unique for each language and each language user, as they are formed during exposure to the particular language (ibid).For instance, the phonological lexicon contains only phonological structures, while the syntactic lexicon only morphosyntactic structures. In turn, structures can be stored either as independent items, a PS representation, or in various combinations, i.e. as a chain of structures PS<=>SS<=>CS that have been formed during exposure to language. For instance, the word 'car' might be available as written input, entering our system through POpS (visual structure). If the visual processor processes the written representation and provides an index for the word 'car', then this representation is considered as the visual structure's intake, which can be communicated to other structures through the interface processors in order to get further processed and formulate a chain car<=>noun, non-animate, singular<=>/kar/, <=>[kar] (SST, 2004 and in prep, Truscott and Sharwood Smith- henceforth TSS, 2004:3,4, in prep: 10). The result of this processing activity is what 'Acquisition by Processing Theory' (APT) refers to as 'indexing' and 'co-indexing'. Indexing 'has the effect of matching up a representation in different modules', while co-indexing refers to the activity of a representational chain formulation. Co-indexing is not limited in indexing only a phonological representation in the phonological structure, as it can also include, for instance a chain of structures, i.e. PS<=>SS<=>CS matched up with affective and/or POpS structures, while it can occur with or without any awareness at the level of understanding or even noticing. However, if co-indexing occurs with a very highly active perceptual representation it can become a conscious representation and therefore can have more chances (although there is no guarantee) of succeeding in becoming a form-meaning connection (SST in prep, cited with permission). This means that if during co-indexing a car is actually present at the situation (thus forming a visual-POpS structure) and someone is pointing perhaps at the car while co-indexing, i.e. matching up of phonological, syntactical and conceptual structures, the presence of the POpS structure (car), might lead the learner in becoming conscious aware of the representation. In turn, it can assist the learner in succeeding in formulating a form-meaning connection, i.e. identify that the object the person is pointing at is called in English [kar], is a noun and is actually a car.

Input is therefore defined in MOGUL as a 'perceptual representation of spoken or written language' (TSS, forthcoming: 14), suggesting that input is acoustic and/or visual structures which make it into the system through the module(s), (similar to Carroll (1999, 2001; Jackendoff, 1987, 2002, 2007; TSS, forthcoming). Intake, on the other hand, is a more 'abstract entity' which refers to whether 'input representations can or cannot be processed by the given processing modules' (TSS, forthcoming: 15). This means that input is seen in MOGUL as just a representation that attempts to enter our system, which might not be processed any further. If the particular representation is processed through one of the structures leading to co-activation of other structures, then this is referred to as intake. Hence, in MOGUL each module has its own input and intake. Clearly, input and intake are interrelated processes, since what serves as input for one of the processors might lead as intake for another processor (similar to Carroll, 1999, 2001).

Activation levels and competition play a significant role in MOGUL. Each structure has resting, current, high or low levels of activation. As soon as a representation enters one of the processors, structures that have high activation levels are the ones that have a greater chance to win the competition during processing. The ‘resting level’ is the starting point for each item and it determines if and how quickly an item will become available for processing, and/or incorporated into a representation. An item’s ‘resting level’ is therefore determined by its past use. The speed at which the item becomes available for processing is also determined by its ‘resting level’ (SST in prep: 46). An item’s resting level might be not the same as originally, as it might have different resting levels depending on the frequency it becomes activated and selected. It could therefore rise or fall and it might not be the same each time for the same item (see Fig.8 for an illustration). On the other hand, the current activation level of an item determines its availability for inclusion in the current processing activity. SST (2005:234) argue that ‘an item’s current activation level is its resting level plus any additional activation it has received from a processor during the current processing’. The two differ in the sense that an item’s resting level usually refers to LTM storage, while the current level of activation refers to the item’s elevation from its resting level towards the upper level of the memory store, namely WM/blackboard. When an item’s current level is activated then spreading activation also occurs (similar to Levelt, 1999; Levelt and Roelofs and Meyer, 1999), meaning that if the PS structure is activated then the SS and CS can also be activated. Spreading activation can also refer to the activation within each of the sub-modules, as interfaces trigger stimulation and try to match the information with any available representation from the LTM store of the sub-module that is activated, as Figure 2.11 below depicts:

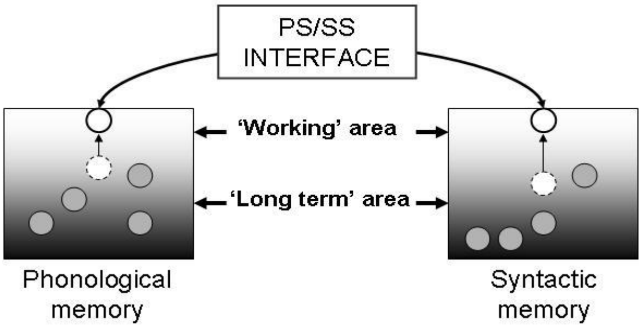


Figure 2.11: Memory stores and activation levels in MOGUL[34]
(Sharwood Smith and Truscott in prep: 38, cited with permission)

It is clear that MOGUL supports connectionist theories, given the fact that processors operate based on stimulation, activation levels and competition. SST (in prep) argue that the difference between MOGUL and other connectionist approaches is thatin MOGUL computation[35] is involved in a representation (similar to Pinker and Ullman, 2002).This means that frequency of a structure alone that is without registration in some part of the parser (i.e. if a structure is not recognised frequently by one of the possible structures, the PS, SS, CS, POpS and/or AfS) will not have an impact on language development (SST, 2005 and in prep). In other words, the difference between MOGUL and other connectionist approaches is that for MOGUL processing is not just about retrieval of a stored representation but about processing involving the development of a 'rule' via co-indexing (SST, 2005, in prep; similar to Pinker and Ullman, 2002; different from Bates and MacWhinney, 1987). Thus, learning occurs when an indexed item hits an empty node and thus creates a new item. Acquisition or, in MOGUL terms 'growth of language', is the 'lingering effect of processing' (SST, in prep: 59). Thus, acquisition in

[34]Note that the circles in the model depict the different structures that can be activated once processing of a new representation begins. Depending on the activation levels that one existing structure holds the structure rises until it wins the competition and reaches the current level of activation. This is the reason why the circles are not placed all in the same level, as well as by comparison the number of circles between the phonological and syntactical memory store as depicted in the figure do not contain the same number of circles; suggesting that each structure may activate different number of structures and in different elevation stage. A fact that further shows how each structure operates not only autonomously but also in a modular way through the interfaces(similar to Cowan, 1993, 2001; Ruchkin, Grafman, Cameron and Berndt, 2003; Miyake and Shah, 1999; SST, 2004, 2005 and in prep).
[35]Computation refers to the view of language in terms of 'words and rules' as discussed mainly by Pinker and others (e.g. Pinker, 1998; Pinker and Ullman, 2002). 'Rules' are the product of computation and refer to the workings of the processors, while 'words' are the stored items, i.e. the representations in the stores (see SST in prep).

MOGUL is not failure driven, as argued by Carroll (2001) but is viewed as an automatic procedure, since the 'parser is constructed so as to automatically acquire (even if the effects are often fleeting)' (SST, 2005:233, see also SST, 2004 and in prep).

The role of L1 transfer is treated in MOGUL in the context of 'full transfer/full access' (SST, 2004:15). SST argue that L2 knowledge develops with L1 knowledge, however possesses different activation levels. In time, sufficient input, metalinguistic information and output can increase the activation levels of L2 features, thus increasing their chances of winning the competition against L1 features (SST, in prep: 69).

MOGUL offers an interdisciplinary approach for a potentially coherent transition theory for SLA; however empirical evidence is necessary, as it remains at a theoretical stage. SST's (2004, 2005 and in prep) account of competition based on activation levels and the role of consciousness provides a theoretical framework to explain language learning and acquisition. This framework can be adapted to provide the theoretical basis for applied models exploring the role of instruction in SLA, such as IE and PI.

2.4.4.2.1 Consciousness and Acquisition according to Modular Online Growth and Use of Language (MOGUL)

In MOGUL it is argued that learners can process structures both with or without conscious awareness. Structures that are processed in the language module are not consciously processed by the learner (SST in prep). When, for example, the learner hears the word 'dog', he/she is not conscious that the syntactic processor is processing that the particular word is a noun, animate and according to the sentence it can be the agent or the recipient of that sentence. In turn, when we hear the word 'dog' we cannot become conscious of the processing that occurs in the acoustic processor that turns the sound into a string generating the word 'dog'. On the contrary, we can become conscious when information is processed in the perceptual output structures (POpS) (SST in prep). POpS refer to information entering through our sensory system and 'are the representations that are accessible to consciousness and serve as the basis for higher level of processing and action' (SST in prep: 94). SST argues that POpS can become accessible to consciousness because of their rich interconnectivity and synchronization with other structures, i.e. the morphosyntactical, phonological and conceptual processors. Therefore POpS can provide additional sources of activation for the representations available through the sensory system (ibid).

As Figure 2.12 illustrates, the language module and the conceptual module (CS) are connected via interfaces to the non-linguistic perceptional system through a 'blackboard' of extralinguistic modules for perceptual processing and representations of sensory input, namely POpS (SST in prep). There are three characteristics attributed to POpS: a distinct store for each of the sensory modalities; activation level variability; and a strong tendency for synchronization (p.100), while all are strongly connected to one another (p.178). Truscott and Sharwood Smith (forthcoming: 11) argue that 'It is possible that the output of each of the sensory modules, especially vision and audition, has inherently high activation levels, beyond the interconnectedness of POpS. MOGUL also includes an affective system (AfS), which involves emotions that can be stimulated when processing information, as illustrated in the next page (Figure 2.12):

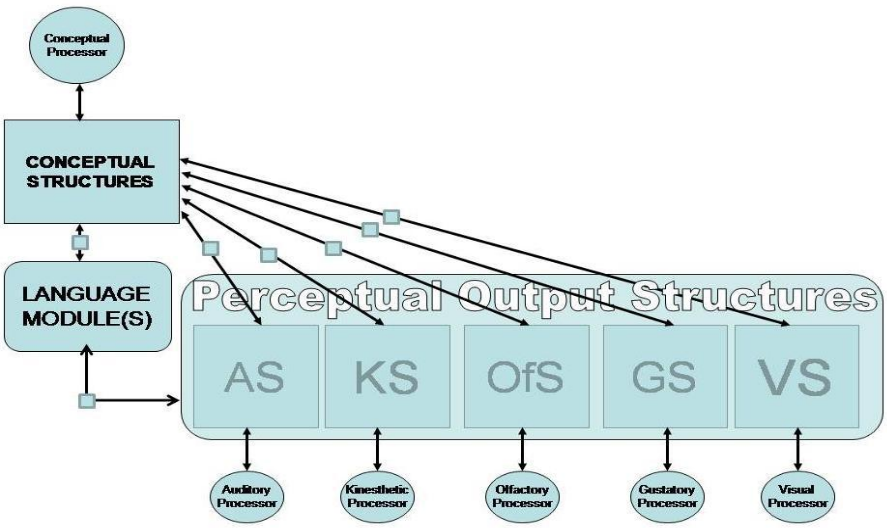


Figure 2.12: Interconnectivity of POpS in MOGUL[36]
(Sharwood Smith and Truscott, in prep, cited with permission)

[36] Please note that in Figure 2.12 the affective structure (AfS) is missing and that a full figure depicting MOGUL is available in p.95. The figure in this page (Fig. 2.12) focuses on depicting the interconnectivity of perceptual output structures (POpS) in order to make clear why POpS are considered in MOGUL highly linked with higher levels of processing and metalinguistic knowledge.

According to SST (in prep), a POpS representation should not be viewed as a sole product of sensory processing but as an output product of linguistic processing. Given that POpS are connected to other processing units (e.g. CS, PS, SS, etc) processing in the POpS is influenced by the activity taking place in these units and therefore processing of a representation takes place in the same way, as would processing occur in one of the other structures, e.g. the language module. Hence, in the same line, a structure's activation level is crucial for it to be stimulated and also to win the competition from other competing stimuli in order to get further processed in the POpS blackboard. SST (in prep) argue that a representation's current level can be raised at a higher level of activation than other representations that are processed in the language module and/or the conceptual structure, if information enters through POpS. It is the exceptionally high level of activation that POpS can achieve that makes them unique in being able to trigger conscious awareness therefore providing an advantage for SLA (SST in prep). Hence, if the representation enters our language learning system through our visual system we can become conscious of the particular structure, although we will not become conscious of the syntactic processing of that structure. The particular representation will have higher levels of activation if there is co-activation and synchronization with other structures. This means that the representation will have to be processed in the morphosyntactical and in the conceptual structure in a modular way in order to reach the exceptionally high levels of activation for the learner to become conscious of the target form. In turn, this implies that the target structure will have to beat other competing structures that might be activated once the information enters the system. In addition, it should generate spreading activation in order for the rest of the necessary structures to be activated (similar to Baars' 1988 'Global Workspace Theory').

SST point out that there is no guarantee that if the form enters through the visual structure it will definitely be processed, echoing Sharwood Smith (1991, 1993). For SLA, this suggests that no matter how well an instructor has planned the teaching materials in order to generate processing through POpS, as described above, there is no guarantee that each learner will perceive the target structure and that the structure will be successfully processed. POpS are highly linked in MOGUL with metalinguistic knowledge. Metalinguistic knowledge in MOGUL is considered to be encoded in the conceptual structure and therefore differs from 'linguistic knowledge' (similar to Carroll, 2001). According to SST (in prep) knowledge about language (metalinguistic knowledge) can be raised to awareness through POpS since they can adjust the input that learners receive in order to promote growth of language (see also http://www.msharwood.pwp.blueyonder.co.uk/mogul/). Promoting metalinguistic knowledge can occur in MOGUL either through POpS or the CS. In the case of POpS, metalinguistic knowledge alters a perceptual representation of the input in order for information to be also processed in the language module. Metalinguistic knowledge from the CS makes linguistic input more comprehensible and therefore more likely to encourage growth of the language module, thus leading to language development (see SST, in prep: 215). In this case, MOGUL gives a significant role to instruction, if metalinguistic knowledge provided through instruction can successfully aid in the construction of representations that accurately capture the information, by enhancing the reception of the input and coaxing the learner to make correct form-meaning connections (ibid:216). PI is seen as the best type of instruction to promote language development in this way (SST, in prep: 217).

2.5 Motivation, limitations, research questions and hypotheses of the present study

The present study's rationale stems from the fact that input manipulation results in drawing learners' attention to specific items of the input in order to help the process of input becoming intake and thus getting further processed through the language learning mechanisms (Schmidt, 1990, 1993, 1995, 2001; Sharwood Smith, 1991, 1993; Tomlin and Villa, 1994; VanPatten, 1996, 2004, 2007, 2009). As we have seen, this process has been shown to have an impact in the short-term. However, gains in the long term remain to be empirically verified in order to validate claims that IE and PI are effective teaching interventions in SLA (among many studies see Doughty, 1991; Jourdenais et al. 1995; Leeman et al. 1995; VanPatten and Cadierno, 1993; VanPatten and Fernandez, 2004). However, in this quest, theoretical issues arise such as: What is the impact of noticing and how can its impact be measured when applying IE, PI or a combination of the two (namely +IE+PI)? What is the attention drawing factor's impact (that both PI and IE claim to promote) on SLA; in what way do they differ; and in what way can they be more effective, i.e. solely and/or in combination? Can PI and/or IE provide valid, reliable and generalizable conclusions regarding effective acquisition of a target form and under what circumstances is acquisition of a target form through these two types of intervention possible? What is the impact of the three types of proposed intervention on the language learning mechanisms? Can they have an impact on input processing and acquisition and if they can, in what way?

In the present study, some of these questions will be answered and will be elaborated further based on the research findings and theoretical explanations provided by the existing frameworks and models that have been reviewed in this chapter.

The focus of the present study is to investigate to what extent IE, PI and/or a combination of the two approaches can be effective in helping English learners acquire word order and case marking in German. Will there be any short and long term effects of these types of instruction? Given the fact that the present study is motivated in investigating the effectiveness of specific types of teaching interventions, a novel structure for the participants has been selected. Furthermore, participants in the present study are secondary school beginner learners of German, i.e. aged 12-14 years in their second year of learning German (for further details please refer to section 3.2 in Chapter 3). In this way it can be controlled that there is no prior knowledge of the target form and therefore valid and reliable conclusions can be drawn regarding or not the effectiveness of the chosen types of teaching interventions; as well as the interference of any ‘meta-linguistic baggage’[37]that participants may carry can be (to the extent possible) avoided (see Gass and Mackey, 2005; Vainikka and Young-Scholten, 2002, 2007b). In addition, the particular target structure has been selected as it fulfils PI’s criteria in terms of identifying a processing Input Processing (IP) Principle that needs to be altered and guidelines in developing treatment and testing materials (see section 2.3 and subsections; see also Chapter 3 for further details regarding the choice of the target form).

For the purposes of the present study, the following research questions will be therefore investigated:

[37]According to Vainikka and Young-Scholten (2002, 2007b), adults’ ‘metalinguistic knowledge can interfere with UG based unconscious acquisition mechanisms hindering second language development, as it can generate competition between general cognitive and linguistic mechanisms (see Young-Scholten and Strom, 2006:52; see also Felix, 1985; Ioup et al, 1994).

1. Is instruction using colour as Typographical/Textual Input Enhancement (+IE-PI) of the target linguistic form more effective than Processing Instruction (-IE+PI) and/or no instruction (-IE-PI) in the short and long term?

2. Does instruction using Processing Instruction (-IE+PI) prove to be more effective than Colour Input Enhancement (+IE-PI) and/or no instruction (-IE-PI) in the short and long term?

3. Is the combined instructional method (+IE+PI) more effective than the application of each technique separately and/or no instruction in the short and long term?

4. Based on previous findings from studies conducted both in the field of IE and PI as well as meta-analyses comparing in this chapter the two types of teaching intervention, the hypotheses of the present study are as follows:

Hypothesis 1: The combined teaching intervention, being more elaborate, obtrusive and explicit than PI and IE will be a more effective teaching intervention.

Hypothesis 2: PI, being elaborate, more obtrusive and more explicit than IE will be a more effective teaching intervention than IE.

Hypothesis 3: IE, will be a more effective teaching intervention than no instruction.

Hypothesis 4a: Effects will be found in the short term.

Hypothesis 4b: Effects will be found in the long term.

The present study adopts a cognitive processing approach to language learning and acquisition, thus the findings of the present study will be interpreted accordingly.

3 Methodology

Introduction

This chapter provides a detailed presentation of the overall design, aims and objectives of the present study and is divided into four parts. In the first part IE and PI studies will be evaluated, while reference to the present study's contribution to the field of IE and PI will be provided. The purpose of this section is to account for the design adopted in the present study. The second part will present how the theoretical underpinnings of PI and IE have been operationalized and combined in the development of the treatment and testing materials for the three groups, and will also include information about the control group. The third part will provide information about the participants of the present study, while the fourth part will describe the data analysis.

Research questions and hypotheses have been presented in the previous chapter, however for ease of reading they are repeated below:

Research questions of this study:

1. Is instruction using colour as Typographical/Textual Input Enhancement of the target linguistic form more effective than Processing Instruction and/or no instruction in the short and long term?

2. Does instruction using Processing Instruction prove to be more effective than Colour Input Enhancement and/or no instruction in the short and long term?

3. Is the combined instructional method more effective than the application of each technique separately and/or no instruction in the short and long term?

3.1 Study Design

3.1.1 Target form

The target form of the present study is the German Object Verb Subject (OVS) structure and the marking of accusative case on the definite article. German's canonical surface sentence structure is Subject Verb Object (SVO) word order, as in (5).

(5) *Der Mann küsst die Frau.*
The (Nom) man –SUB kisses the woman-OBJ (The man kisses the woman.)

However, in German, the subject, object, and indirect objects can move freely within a sentence as long as the verb stays in second position (in the case of a main clause). A grammatical OVS sentence structure appears in (6).

(6) *Den Mann küsst die Frau.*
The (Acc) man-OBJ kisses the woman-SUB (The woman kisses the man.)

The only cue that German speakers have when interpreting sentences like (5) is that subjects/objects can be identified by their definite article endings, if the object is masculine because the definite/indefinite article has a different form from the nominative. As seen in Table 3.1, the endings of the definite articles in German can change on the basis of both gender and case:

Table 3.1: German Definite Article System

Case	Masculine	Feminine	Neuter	Plural
Subject (Nominative)	der	die	das	die
Object (Accusative)	den	die	das	die
Indirect Object (Dative)	dem	der	dem	den
Possessive (Genitive)	des	der	des	der

An additional cue for German speakers when interpreting OVS sentences is available from stress intonation when sentences are produced orally. At a discourse level, stress intonation is added at the Object of the structure, which is placed at the beginning of the sentence indicating the differences in pragmatics, as well as acting as an additional cue for context and topicalisation[38]. English learners do not have this sort of variable word order or case marking on determiners in their L1, nor do they have grammatical gender, as German nouns (Feminine, Masculine or Neuter) have. According to the First Noun Principle (FNP) of IP's theoretical framework, it would be predicted that English learners, who do not have these features available in their L1, will initially process the first noun in every German sentence they encounter as the subject or agent. For instance, when the learner hears or sees example (5) *den Mann küsst die Frau* 'the woman kisses the man', the learner will interpret the first noun as the subject and process it as 'the man kisses the woman, instead of taking the cue from the definite article ' *den',* which signals a direct object.

According to PI theoretical underpinnings and guidelines, the instructor should design materials that 'teach one thing at a time'. Moreover, if the target processing principle is the FNP and the target structure is OVS sentences, treatment materials must include both OVS and SVO structures in order to force learners to make correct form-meaning connections. With this in mind the focus of the present study was singular masculine OVS structures and materials also included SVO structures. The nominative and accusative singular masculine cases were chosen, as many verbs in German require an accusative object noun. The chosen target structure was also more appropriate for the beginner levels, which were tested, than the dative and genitive cases. The masculine determiner was chosen, since the ending of the determiner is not the same in the nominative and the accusative case (whereas it is the same form for both nominative and accusative of the feminine and neuter determiners), and therefore it would indicate if and when learners are establishing correct form-meaning connections, which could not have been possible with the feminine and neuter determiners.

[38] Topicalisation refers to a process, where the learner is able to assign a constituent in sentence-initial position other than the Subject introducing new information to the discourse (see Pienemann, 2007).

3.1.2 Materials

Norris and Ortega (2000:486) argue that the impact of instruction may 'be directly associated with the type of [test] response required from learners'. VanPatten and Sanz (1995) suggest that different modalities and task types should be used in order to assess and evaluate impacts of instruction. For the purposes of this study, tasks included both in the treatment and in the testing materials comprised a range of tasks (i.e. multiple choice, true or false, putting sentences in chronological or logical order, and fill-in- the-gap). However, the focus was on written tasks, in line with the majority of IE and PI studies. Oral tasks could have been included but only in the testing materials, since in both IE and PI, production tasks are excluded during treatment activities. Due to the nature of the target form, listening tasks were also excluded. This decision was based on the fact that OVS sentences in German when produced orally contain stress intonation to add emphasis, acting as an additional cue for participants. Therefore the use of oral tasks would have had to become artificial by eliminating stress intonation that would promote 'world knowledge', as suggested by Bachman and Palmer (1996), and that would violate the 'Event Probability's guideline, as proposed by the IP framework (see VanPatten 1996, 2004). Moreover, the variables written and oral enhancement would not be isolated, (as in Benati 2001, 2004; Lee and Benati, 2007a, 2007b), if treatment materials contained both types of enhancement. Therefore valid conclusions could not be drawn regarding effectiveness of type of IE in the acquisition of word order and case marking in German.

One of the differences between the two teaching interventions (PI and IE) is the fact that in PI there is a well-defined treatment outline with specific guidelines in developing materials and types of activities, whereas in IE there are no set guidelines suggesting which type of TTIE and what type of activity should be included in the treatment materials (see also Sharwood Smith and Trenkic, 2001). The review of the literature on TTIE in Chapter 2 concluded that the majority of studies included free-recall, reading comprehension, and true-false tasks. Similarly, PI studies included comprehension and interpretation tasks, excluding production tasks from the treatment. With this in mind and in order to control for the variable exposure to input it was decided to include only comprehension and interpretation tasks in the treatment materials for all four groups in the present study. In line with IE and PI research, the testing materials include a written production fill in the gap task.

The majority of PI and to some extent IE studies have assessed the effectiveness of PI and IE over ‘traditional’ types of instruction. These studies have been criticised on the basis that comparisons should be drawn from comparable input-based instructional approaches (Collentine, 2004; DeKeyser et al. 2002; Norris and Ortega, 2000 as discussed in Chapter 2). Based on the latter, the present study explores the effectiveness of three inter-related input based approaches, namely IE PI and a combination of IE+PI.

Four on-line treatment packages were developed for the four different groups: One treatment package for the PI group, another for the IE group, a third for the Combined group and a fourth for the Control group. A fifth-online package was also developed containing the testing materials. Since an intact class would have to be divided into four groups, computerised materials enabled each participant to have access to the specific material allocated to his/her group during normal teaching hours and at the same time as the rest of the participants in the class. This therefore eliminated the need for four classes and four teachers (Alanen, 1995; Jourdenais et al. 1995; Sanz, 2004; Sanz and Morgan-Short, 2004). Gass and Mackey (2005) argue that the use of on-line computerised materials could contribute towards controlling for affective factors, such as learners’ familiarisation and relationship with the teachers, shyness, stress, (de)motivation and unfamiliarity with the ‘new’ teacher, which may result in limited input and thus affect the outcome of the study. Taking into consideration the factors mentioned above, the development of computerised materials enhanced the validity and reliability of the study by providing not only a more ‘neutral’, but also a more controllable and comparable learning environment (see Alanen, 1995; Sanz, 2004; Sanz and Morgan-Short, 2004).

Think aloud protocols (TAP) have been largely applied in TTIE studies investigating the role of noticing through input manipulation in SLA (Alanen, 1995; Jourdenais et al. 1995; Leow, 1997, 2001, 2003; among many). The majority of TTIE studies based on the analysis of TAP have argued that they provide on-line data that can give an insight into conscious processing of the target form (Alanen, 1995; Jourdenais et al. 1995;

Leow, 1997, 2001, 2003, 2006). However, it is widely accepted that the application of TAP have possible pitfalls, such as 'positive and negative reactivity' (see also Bowles, 2008; Bowles and Leow, 2005; Ericsson and Simon, 1993; Godfroid, Housen and Boers, 2010; Leow and Morgan-Short, 2004; Sachs and Polio, 2007; Sanz et al. 2009; as discussed in Chapter 2). The fact that, in the present study, participants were not adult learners raises an issue regarding their ability to verbalise (non-) metalinguistic knowledge. Moreover, studies have shown that concurrent verbalisation increases the time required to perform the task (Bowles, 2008; Bowles and Leow, 2005; Leow and Morgan-Short, 2004; Sachs and Polio, 2007; Sanz et al. 2009). Furthermore, the language processing frameworks reviewed in the previous section argue that there are processes that we cannot become conscious of, such as the processing of morpho-syntax (see Carroll, 1999; 2001; Jackendoff, 1987, 2002, 2007; Sharwood Smith and Truscott, 2004, 2005 and in prep). More importantly, the present study's aims and objectives were not to investigate noticing per se. For all these reasons, TAP were not carried out in the present study.

3.1.2.1 Role and use of EI and feedback in the present study

Based on the findings of the latest studies (for details see section 2.3.3 in Chapter 2), the present study adopted the full PI paradigm including EI, which seems to promote speedier and more accurate correct form-meaning connections (Culman et al. 2009; Fernandez, 2008). To control for the 'internal validity' variable, the same amount of EI in terms of types and tokens was available for all three treatment packages; however, information was adjusted for each treatment package. Participants of the three groups (+IE+PI, +IE-PI, -IE+PI) received, at the beginning of their instructional packages the same basic brief explicit rule explanation of the target form followed by examples in German and their English translation. The explicit information for all three groups covered two main points: (1) the German nominative and accusative case markings on articles; (2) SVO versus inverted OVS word order in German with examples. However, the materials for the IE and the Combined groups differed from the PI group in that the target form was typographically enhanced with colours. This means that the FNP information was available for all three groups, while it was enhanced for the IE and the Combined groups. In this way, grammatical input was controlled for all three groups while ‘internal validity’ was also addressed (see DeKeyser and Sokalski, 1996; Salaberry, 1997). The following extracts (Figure 3.1 and Figure 3.2) included in the treatment materials for each of the three groups provide an illustration of the explicit instruction they received:

> The meaning is the same but the word order is not.
>
> If you see or hear ‘den’ in front of the noun, this means that the noun is not performing the action.
>
> Be careful to notice the article before the nouns. The noun that comes first in the sentence is not always the one that performs the action. You should pay attention to the case in order to establish who is doing what.

Figure 3.1: Extract from the PI group’s explicit instruction

> **The meaning is the same but the word order is not!**
>
> If you see or hear ‘**den**’ in front of the noun, this means that the noun is **not performing the action**.
>
> **BE CAREFUL** to notice the article before the nouns. The noun that comes first in the sentence is not always the one that performs the action. You should pay attention to the case in order to establish who is doing what.

Figure 3.2: Extract from the IE and Combined groups’ explicit instruction

Access to explicit instruction and feedback was limited at the beginning of each treatment and before the treatment activities. Feedback in this phase was provided in the form of a reminder, while it differed qualitatively between groups (see Figure 3.3 and Figure 3.4). Depriving learners of feedback and explicit instruction was essential in order to avoid familiarisation and priming effects, as the same test version was applied in the three testing conditions. Furthermore, the aim of the present study was to investigate if and to what extent the three teaching interventions were effective in the acquisition of German word order and case marking. Therefore investigating the role of EI and feedback in PI was beyond the scope of the present study.

In line with the PI framework there were reminders after the first and last referential activity available in the three treatments, as illustrated in Figure 3.3. However, it should be noted that reminders differed qualitatively between groups. For the PI group the reminder explicitly referred to the form and specifically required processing of the target form, while for the IE group, it explicitly referred to the colour enhancement and specifically required assigning forms a colour according to the coding. For the combined group, the reminder was the same as for the PI group, while the target form was typographically/textually enhanced using the same colour coding as for the IE group (Figure 3.3):

Who performs the action in the next example? Den Lehrer fragt der Student. a. den Lehrer   b. der Student The correct answer is der Student.

Figure 3.3: Extract of the reminders received in the PI treatment

In the IE group, in order to balance and control for the variable input exposure, participants also received reminders after the first and last activity of their treatment package. However, attention was focused on assigning the correct colour for each case, nominative and accusative of the definite articles, according to the colour coding available, as the extract below illustrates:

Which is the appropriate colour for the definite articles in the next example? Den Lehrer fragt der Student. a. **Den** Lehrer fragt **der** Student. b. **Den** Lehrer fragt **der** Student. c. I am not sure The correct answer is a.

Figure 3.4: Extract of the IE Group's reminder

The Combined group received the exact same reminder as the PI group. However the target forms were enhanced following the predetermined colour coding (Figure 3.5 next page):

Who performs the action in the next example?

Den Lehrer fragt **der** Student.

a. **den** Lehrer b. **der** Student

The correct answer is **der**Student.

Figure 3.5: Extract for the Combined Group's reminder

The instructional materials used simple grammar and vocabulary in view of the beginner level of the learners. However, it should be noted that it was not possible to avoid incorporating unknown vocabulary as schools use different textbooks[39] and teachers often provide personal materials. Therefore, vocabulary hard copies were given to each participant containing translations of the possible unknown words[40]. Participants were also given permission to ask the researcher for translation in case of unknown words not contained in the hard copy (for further details please refer to appendix).

All treatment and testing materials were checked by native speakers of German.

In the following subsection, an outline of the treatment materials for each group is presented.

3.1.2.2 Treatment Materials for the PI Group

Participants of the PI group received brief explicit instructions at the beginning of each treatment (section 3.1.2.1). SI activities followed and were split into referential and affective activities. Referential activities required the learners to attend to the meaning in order to decode the target form. Activities of this type were mainly interpretation activities; the target form was provided in German and four possible responses were available to the participant. Each sentence contained two nouns, one in the nominative and one in the accusative, i.e. singular OVS sentences with masculine and feminine; masculine and masculine; masculine and neuter; feminine and feminine; feminine and

[39] The 'Focus Deutsch' and 'Klasse 1' textbooks were consulted during the preparation of materials; however in some cases it was not possible to avoid incorporating on familiar vocabulary.
[40] In the case of nouns the gender of the noun in German was also provided in the form of (m) to indicate that the noun in German is masculine, (f) for feminine and (n) for neuter nouns.

neuter; neuter and neuter. The first option interpreted the first noun as the subject (SVO option) of the sentence, while the second option interpreted the first noun as the object (OVS option) of the sentence. Option c stated that both options a and b were possible interpretations of the German sentence, thus suggesting that both the accusative and the nominative nouns could be the subjects of the sentence. In order to avoid guessing, the fourth option gave the option of choosing ‘I am not sure’ (see example below Figure 3.6, for ‘The ant sees the beetle’).

A. Choose which sentence corresponds to the German sentence:

2. Den Käfer sieht die Ameise.

- a) The beetle sees the ant.
- b) The ant sees the beetle.
- c) Both a and b options correspond to the German sentence.
- d) I am not sure.

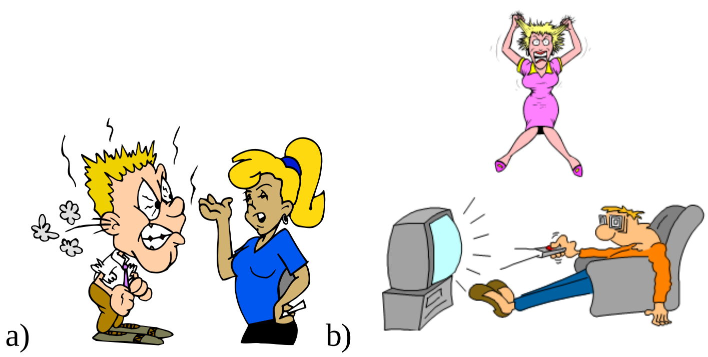
Figure 3.6: Referential activity for the PI group- example from the treatment

In this example the target form is Masculine OVS; ‘den Käfer’ (the-Masc.Acc. beetle) ‘sieht’ (sees) ‘die Ameise’ (the-Fem. Nom. ant), meaning that option b is the correct response (the ant sees the beetle).

Another example of a referential activity available for the PI treatment is provided below, where participants were provided with the target structure and they were asked to choose which picture corresponded to the sentence. Having the option ‘I am not sure’ available as a means to avoid guessing (see example 1c below Figure 3.7):

Choose which picture corresponds to the sentence. If you are not sure choose c) I am not sure:

Den Mann ärgert die Frau



- a
- b
- c I am not sure.

Figure 3.7: Referential Activity for the PI group-example from the treatment

‘Den Mann’ (the-Masc. Acc, man) ärgert (makes angry) ’die Frau’ (the-Fem. Nom. woman), ’the womanmakes the man angry’, thus picture a) corresponds to the German sentence.

Affective activities were also available during treatment in order to provide more examples of the target linguistic form. Here any attention to form would have been incidental to the task (see Chapter 2 section 2.3.2), as the extract below illustrates:

C. Put the following into a chronological order so that a story can be told .

a) Den Vater pflegt die Mutter.

b) Den Dieb sieht der Vater im Zimmer.

c) Der Vater ruft sofort die Polizei an.

d) Den Vater schlägt der Dieb und geht weg.

e) Den Vater weckt die Mutter.

Figure 3.8: Affective Activity for the PI group-example from the treatment

In this case there are various ways of putting the given sentences in a chronological order to tell a story. As this is an affective activity, this is not an issue, as according to the PI framework these types of activities are included in the treatment in order to enhance PI as a communicative type of instruction (basically they provide more examples of the target form, similar to input flood). As a result, attention to the form is incidental (for details please refer to 2.3.2 section)[41].

3.1.2.3 Treatment Materials for the IE Group

The treatment package for the TTIE Group also contained a brief explicit instruction about word order, cases and case marking in German (as described in section 3.1.2.1), followed by treatment activities.

In the IE framework, there are no set specifications suggesting specific tasks for inclusion in the treatment. However, a review of TTIE studies concluded that in most studies, among other types of tasks, the consensus was to include reading

[41] However, as discussed in Chapter 2 (see section 2.3.2) it still remains an open issue in PI why these type of activities are defined as ‘affective’, since the focus of the task is still on the target form and participants have to complete the task irrespective of their personal emotions.

comprehension tasks, while production tasks were used only in testing materials (Alanen, 1995; Doughty, 1991; Jourdenais et al. 1995; Leeman et al. 1995; Leow, 1997, 2001, 2003; Shook, 1994; Overstreet, 1998; White, 1998; Wong, 2003). Therefore it could be concluded that PI and IE studies have a common treatment design in including comprehension tasks, while production tasks are applied as testing materials. In the present study, the tasks included for the IE group were reading comprehension tasks, in line with most of the TTIE studies. Additionally, true/false tasks, which measured comprehension, were also included in the treatment package. The two types of comprehension tasks of the IE group corresponded to the referential and affective activities of the PI group's tasks. In this way, all packages (PI, IE and Combined) included the same amount of types and tokens. However, the text itself used in the IE treatment package was a non-authentic text. The reason for not including an authentic text was because the chosen target structure is not a common structure in authentic texts, although it is often used in oral speech to add emphasis. To enable a wide range of target forms, writing a non-authentic text was therefore necessary. It also allowed better control of the variable exposure to input for all three groups, as the target sentences that were included in the reading comprehension texts were then isolated and turned into SI activities. However, as the lack of authentic texts may have increased artificiality, texts were checked by native speakers of German.

The aim for the IE group was that learners' attention was drawn to the target linguistic form through colour enhancement. However, the reading comprehension activity (containing the colour TTIE) requires general understanding of the text. Hence, the IE group could be described as a less explicit and less obtrusive type of instruction (as described in Doughty, 2003; Doughty and Williams, 1998; Norris and Ortega, 2000) than the PI and the Combined groups. The following extract of a reading comprehension task (Figure 3.9 next page) allows direct comparison with the previous extract of a referential activity from the PI framework (Figure 3.6).

Die Ameise

Die kleine Ameise möchte die Welt kennen lernen und so geht sie von zu Hause weg. Auf ihrer Reise trifft die Ameise zuerst die Grille. Sie erzählt von ihrem Traum und lädt sie ein aber sie ist faul und geht nicht mit. Den Käfer sieht die Ameise danach. Er legt sich hin und ist sehr müde. Plötzlich kommt die Spinne an. 'Pass auf!' sagt der Käfer, und schnell packt der Käfer die Ameise und fliegt. Den Käfer belohnt der Schmetterling. 'Bravo!' sagt der Schmetterling. 'Die Spinne will die Ameise fressen! Und jetzt kannst du deine Reise weiter machen. Komm mit, ich zeige euch was Schönes!' Aber, wenn sie da sind, frisst den Schmetterling der Bär. Und plötzlich beißt den Bär der Panther. 'Schnell, schnell, fliegen wir weit von hier, sagt die Ameise'. 'Mein lieber Gott!, sagt die Ameise, ' die Welt ist nicht so sicher!'. Aber der Käfer ist müde und kann nicht mehr fliegen. Plötzlich sieht der Käfer das Känguruh und bittet es sie mitzunehmen. Das Känguruh nimmt beiden mit. Den Löwe besucht das Känguruh. Dort sind viele Tiere, da er Gerburtstag hat. Dort lernt die Ameise das Zebra kennen und hört den Hirsch spannende Geschichten erzählen, aber die Ameise vermisst schon die Familie. Der Gepard versteht die Ameise und entscheidet sich die Ameise zurück zu der Familie zu bringen. Und so ist die Ameise wieder zurück bei der Familie. 'Vielen Dank', sagt die Ameise. Den Gepard verlässt die Ameise.

A. Choose the correct answer. If you are not sure choose I am not sure.

2. Who did the ant first meet?

- a) den Käfer
- b) die Grille
- c) Ich bin nicht sicher

Figure 3.9: Reading Comprehension Activity - Extract from the IE group.

The following extract from the true/false comprehension task used in the IE group (Figure 3.10) provides a direct comparison with the PI referential activity (Figure 3.7) provided in the previous section (section 3.1.2.2).

Choose a) if the picture corresponds to the sentence, b) if it does not, and c) if you are not sure:

Den Mann ärgert **die** Frau[42].

- a) yes
- b) no
- c) I am not sure

Figure 3.10: True or False Activity extract from the IE group's treatment package

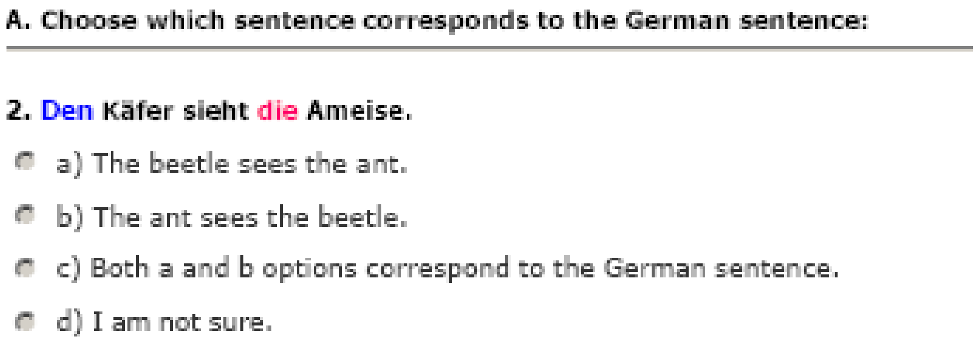
The treatment tasks for each group were controlled for the amount of types and tokens but also for the visual information (i.e. pictures) in each package. The true/false task

[42] ‚Den Mann' (the man-Masc. Acc. Obj.) ärgert (Verb-makes angry) 'die Frau' (the lady-Fem. Nom. Subj),'the lady makes the man angry', thus the correct answer is a: yes, the picture corresponds to the German sentence.

corresponded to the PI and the Combined groups' referential activity. The difference between the IE and the PI group treatments in this particular task is the enhancement of the target form. In the IE group the true or false task requires participants to decide if the given sentence corresponds to the given picture. In the case of the PI and the Combined groups, the task includes two pictures and one sentence and participants are required to decide if the depicted situation, as shown in Fig. 11, is an SVO or an OVS situation. It is valid to conclude that the PI and Combined groups' participants are forced to make (correct) form-meaning connections, whereas for the IE group's participants, the coloured typographical enhancement is more implicit, as the task only requires general comprehension. The IE group's treatment package is therefore less explicit and less obtrusive.

3.1.2.4 Treatment Materials for the Combined Group

The treatment for the Combined (+IE+PI) group was a combination of activities included in the PI and IE treatment packages. The variable exposure to input was controlled for the amount of treatment tasks included in the treatment package. In order to comply with the research design, the target form was typographically enhanced using the same colour coding as in the IE group, while some of the IE group's true/false tasks were incorporated in the Combined group's treatment package. The remaining treatment tasks of the combined group included the same referential and affective activities as in the PI group, the only difference being the colour enhancement of the target form (see subsections 3.1.2.2 and 3.1.2.3 and/or Appendix B). In this way, the amount of types, tokens and pictures included was exactly the same as in the PI group, while the same coloured Typographical/Textual enhancement applied in the IE group was adopted (illustrated below Figure 3.11):


A. Choose which sentence corresponds to the German sentence:

2. Den Käfer sieht die Ameise.

- a) The beetle sees the ant.
- b) The ant sees the beetle.
- c) Both a and b options correspond to the German sentence.
- d) I am not sure.

Figure 3.11: Combined Group's Referential Activity with Colour Input Enhancement-example from the treatment

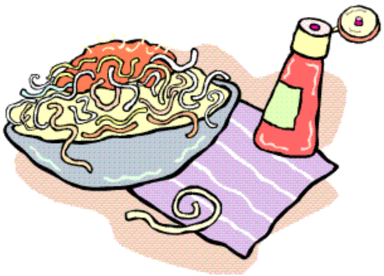
The coloured typographically enhanced referential and affective activities for the Combined group can easily be matched with the reading comprehension and/or the true/false tasks from the IE package (section 3.1.2.3). Reading comprehension in this treatment package was not included because the aim of the present study is to investigate if the combination of PI with IE is more effective than the sole application of each teaching intervention. Moreover, there was an attempt to use IE in a more 'explicit' and 'obtrusive' way (as described in Doughty, 2003and Doughty and Williams, 1998) given the fact that meta-analyses (Doughty, 2003; Lee and Huang, 2008; Norris and Ortega, 2000) have shown that, as a teaching intervention, explicit and obtrusive type of enhancement is more effective than implicit and less obtrusive enhancement.

3.1.2.5 Treatment Materials for the Control Group

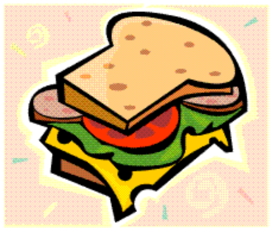
As far as the control group is concerned, participants did not receive explicit instruction or treatment tasks focusing on the target form. The instructional package contained reading comprehension, matching and ranking activities, focused either on verbs and/or vocabulary (see example below and for further details please refer to the Appendix B):

Was essen die Leute gern?

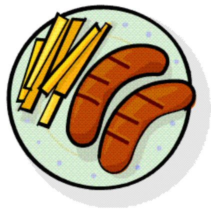
a.



b.



c.



d.

Peter: Ich esse gern Pommes Frites mit Wurst.

Klaudia: Ich mag Brot mit Käse, Schinken und Salat.

Heike: Ich mag Nudeln mit Ketchup.

Thomas: Ich mag Eis.

Figure 3.12: Treatment materials for the Control group

Including a control group was essential, as it would allow direct comparisons regarding the effectiveness of the instructional frameworks tested. The control group would act as the baseline for comparisons among the three instructional treatments, since instruction of the target structure was not included in its treatment package. On the other hand, it is anticipated that the control group will show improvement from the pre- to the immediate and delayed post-test condition in its overall performance, since its instructional package included reading comprehension and true/false on-line tasks. Failure of one of the three instructional groups to outperform the control group will be interpreted as non-effectiveness of the instructional technique.

3.1.3 Testing Materials

Testing materials were also administered online. They comprised four different tasks, namely an error correction task, a comprehension task, a production task and an interpretation task. In total all four tasks included 44 target forms (masculine OVS sentences) and 33 distracters. Originally there was an equal number of distracters and target forms; however, after piloting, a number of distracters were removed, in order for participants to have sufficient time to complete the tasks. One version of the same test was used for the three conditions (pre-, immediate and delayed post-test); participants did not receive any feedback, in terms of either which responses were correct or incorrect, nor a score for each task and/or total at the end of the test[43]. There were a number of distracters in order to distract participants from the target form, avoid familiarisation and avoid memorisation of tasks and/or responses. Moreover, items used in the treatment were excluded from the testing materials, thus controlling for the 'task familiarity' and 'training from teaching materials and instruction' (Bachman and Palmer, 1996; Skehan, 1998). There was a gap of one week between the administration of the pre-test and the immediate post-test, as well as a gap of 12 to 14 weeks[44] between the administration of the immediate post-test and the delayed post-test. However, familiarity with the test, particularly for the immediate post-test condition, may have played a role in performance, especially for the three experimental groups (PI, IE and the Combined groups), since instruction was focused. However, it could also be argued that the Control group might show improvement because of task familiarisation (Skehan, 1998). Therefore having the same test version could provide a good baseline to draw valid and reliable comparisons in terms of performance by looking at variation prior to, during and after treatment. As in the case of the treatment, participants received a hard copy with unknown vocabulary and could also ask the researcher about unknown vocabulary. In the following sub-section, the testing materials and tasks are presented in more detail.

[43] After completion of the delayed post-test the researcher did provide feedback and did answer questions relating to the testing and treatment items.

[44] In some schools despite original planning of the delayed post-test to be available 12 weeks after the administration of the immediate post-test, unforeseen circumstances, such as technical difficulties, half-term school holidays and relocation of the school postponed the delayed post-test. Therefore in three cases the delayed post-test was administered 14 weeks post instruction.

The general trend emerging from TTIE and PI studies meta-analyses, as discussed in Chapter 2, shows that explicit instruction has a more marked effect than implicit instruction, meaning that there is no pedagogical effect but simply the effect was more obtrusive (Doughty, 1991; Doughty, 2003; Lee and Huang, 2008; Norris and Ortega, 2000). This is in line with the fact that many of the treatment and testing tasks applied in the studies require the use of metalinguistic knowledge, e.g. knowledge about language, rather than the knowledge of language (Doughty, 2003; Schwartz, 1993; Truscott 1998). Despite the fact that metalinguistic knowledge has been defined by many as 'pseudoknowledge' which cannot become linguistic competence (Doughty, 2003, 2004; Lightbown et al. 1980; Truscott, 1998), it is also agreed that it is necessary, as it can focus learners' attention and generate conscious processing of items that need to be learned/acquired (Paradis, 2004; Schmidt, 1990, 1993, 1995, 2001; Doughty, 2003). The research pointed to the conclusion that a better insight into the processes involved in language learning could be obtained, if a battery of tasks was used to measure performance. The present study aimed to address this point. In this way a better understanding of the effectiveness of teaching interventions can be reached (Doughty, 1993; Paradis, 2004; Sheen, 2005).

3.1.3.1 Error-Correction Task

A three-step error-correction task was developed, avoiding world knowledge (Bachman and Palmer, 1996) and Event Probabilities (VanPatten, 1996, 2004) cues that could enable learners to avoid processing the target form when decoding the message (e.g. the ball kicked the boy)[45].

Participants were provided with a picture and a German sentence and they were initially asked to decide if the given sentence corresponded to the given picture. In this part of the task, participants were provided with three options to choose from: 'Correct', 'Incorrect' and 'I am not sure'. For all items of the error correction task the 'Incorrect' option was the correct response, as all items depicted an OVS sentence, while the

[45]It should be noted that the variable animacy has been controlled for, by including in each paradigm/item either pairs of animate nouns or pairs of inanimate nouns.

sentence provided describing the depicted situation was an SVO sentence. Participants were therefore required to identify the mismatch between the depicted (OVS) and the described (SVO) situations in order to proceed in the next step and match the picture and the sentence providing an OVS sentence. Therefore the rationale of the design of this task was that learners had not learned/acquired the target form and/or that the teaching intervention was not effective in acquiring the target form in the event that they had chosen ‘Correct’ and/or ‘I am not sure’. The option ‘I am not sure’ was provided in each of the three steps in order to avoid guessing. Whenever the participants chose this option in any of the three steps of the error-correction task, they were directed to the next item. The option ‘Correct’ was also available in each first step of the error-correction task. If participants chose this option, they were also directed to the next item. However, if they chose ‘Incorrect’, which was for all the testing items the correct option, then they were directed to the second step of the exercise. In this second step, a table containing all the words of the given sentence was available, while participants were requested to choose the incorrect word. Depending on the word they had chosen, they were directed to another web-page. In this page they were presented with the picture and the sentence containing a gap instead of the word they had previously chosen. In this third step, participants were required to produce the word that best described the picture. It should be noted that the option ‘I am not sure’ was also available in this last step. The extract next page (Figure 3.13) illustrates the three steps required to complete the error correction task:

Step 1

Select if the sentence below is ‘Correct’ or ‘Incorrect’ according to the picture. If you are not sure select ‘I am not sure’.

Der Sohn verfolgt die Mutter.

- Correct
- Incorrect
- I am not sure

Step 2

Only if they chose ‘Incorrect’

	Der	Sohn	verfolgt	die	Mutter	I am not sure
A Select Incorrect						

Step 3

Only if they chose ‘Der’ as Incorrect

Please type in the correct form:

_____ Sohn verfolgt die Mutter.

- I am not sure
- Correct Form:

Figure 3.13: Error Correction Task example

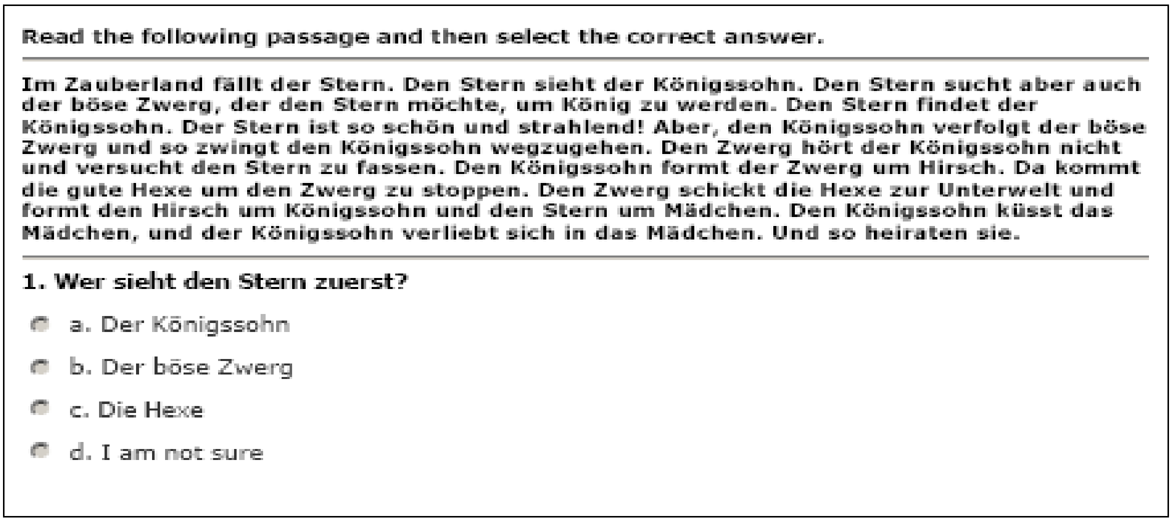
The task design of the error-correction task aimed at controlling for the variable guessing. Therefore, apart from the ‘I am not sure’ option available in every step, participants were directed to a specific webpage, depending on their response. In this way, guessing and/or additional cues that might have assisted participants in the choice and production of forms were avoided.

Judging the grammaticality, correcting and producing the word that would best describe the picture was a novel task not included in the treatment, thus controlling for the variables task familiarity and training from the treatment materials and tasks. In the literature, the information that the error correction task can provide is considered a vexed issue. The debate concentrates on whether the particular task provides evidence of metalinguistic vs. implicit knowledge (Bialystok 1994, 2002; Larsen Freeman 1991, 2003; Mandell 1999; Pica, 1994; Sorace 2003; Truscott, 1999, 2007). Stepping into this debate is beyond the scope of the present study. However, it could be argued that developing a timed error correction task could enhance claims about tapping into implicit knowledge. The fact that participants had adequate time to respond could support claims that the data obtained regarded participants' metalinguistic knowledge. On the other hand, one could also argue that participants' responses were based on factors such as 'it feels right', as they were able to provide a correct response in judging grammaticality of the sentence, while in the next steps they could choose the option 'I am not sure' (Ellis, 2005; Gass and Mackey, 2005; Wong, 2000).

The error correction task comprised 15 target forms and 15 distracters all of which were incorrect (developed according to the three step paradigm illustrated above). Participants received one point, if they provided the correct response in all three steps, or zero points if they chose 'correct' or 'I am not sure' in any of the three steps involved in the error correction task (Wong, 2000). Up to now, the error correction task has not been applied to measure for PI effects. Despite the fact that the particular task has not been part of any treatment package in the present study, and therefore could be interpreted as an 'independent' task to measure for IE, PI and/or combined instructional effects, one could argue that the first two steps of the error-correction task resemble the picture and sentence matching tasks available in all three treatment packages (referential activities for the PI and Combined groups and true or false tasks for the IE group). However, it was beyond the scope of the present study to investigate if there were effects of task familiarity and training from teaching materials and instruction as proposed by Bachman and Palmer (1996) and Skehan (1998).

3.1.3.2 Comprehension Task

A reading comprehension followed the error correction task. A 10 line non-authentic reading comprehension text included ten target masculine OVS sentences. Participants were asked to read the text and answer seven multiple choice questions, each of them comprising four possible options, the last option always being ‘I am not sure’ in order to control for guessing, as the example below (Figure 3.14) illustrates:



Read the following passage and then select the correct answer.

Im Zauberland fällt der Stern. Den Stern sieht der Königssohn. Den Stern sucht aber auch der böse Zwerg, der den Stern möchte, um König zu werden. Den Stern findet der Königssohn. Der Stern ist so schön und strahlend! Aber, den Königssohn verfolgt der böse Zwerg und so zwingt den Königssohn wegzugehen. Den Zwerg hört der Königssohn nicht und versucht den Stern zu fassen. Den Königssohn formt der Zwerg um Hirsch. Da kommt die gute Hexe um den Zwerg zu stoppen. Den Zwerg schickt die Hexe zur Unterwelt und formt den Hirsch um Königssohn und den Stern um Mädchen. Den Königssohn küsst das Mädchen, und der Königssohn verliebt sich in das Mädchen. Und so heiraten sie.

1. Wer sieht den Stern zuerst?

a. Der Königssohn

b. Der böse Zwerg

c. Die Hexe

d. I am not sure

Figure 3.14: Reading Comprehension Task, extract from the testing materials

The particular task was biased towards the IE and Control groups, since their treatment activities included reading comprehension tasks. Therefore it was anticipated that the participants of the IE and Control groups would outperform participants from the PI and the Combined groups, as task familiarity and training from teaching materials and instruction would have an effect on their performance (Bachman and Palmer, 1996; Skehan, 1998). However, it remained to be empirically shown if familiarity had an effect on performance on the reading comprehension task.

The reasons for including non-authentic texts were discussed in section 3.1.2.3. As in the treatment materials, in order to control for artificiality, the reading comprehension text included in the test was also checked by native speakers of German. The task included seven items, and participants received two points[46] for each correct response and zero points for an incorrect response.

3.1.3.3 Fill-in-the-gap task

The third task, a fill-in-the-gap task, followed the reading comprehension task. In this task, participants were presented with a picture and a German sentence containing a gap

[46] In order to balance the score received in all four tasks, it was decided to give double points in the reading comprehension and the interpretation tasks, as these two tasks comprised seven and six target items respectively compared to the other two tasks, which comprised of 15 and 16 target forms.

(always at the beginning of the sentence). They were asked to fill the gap with the target form. The option 'I am not sure' was also available in order to control for guessing:

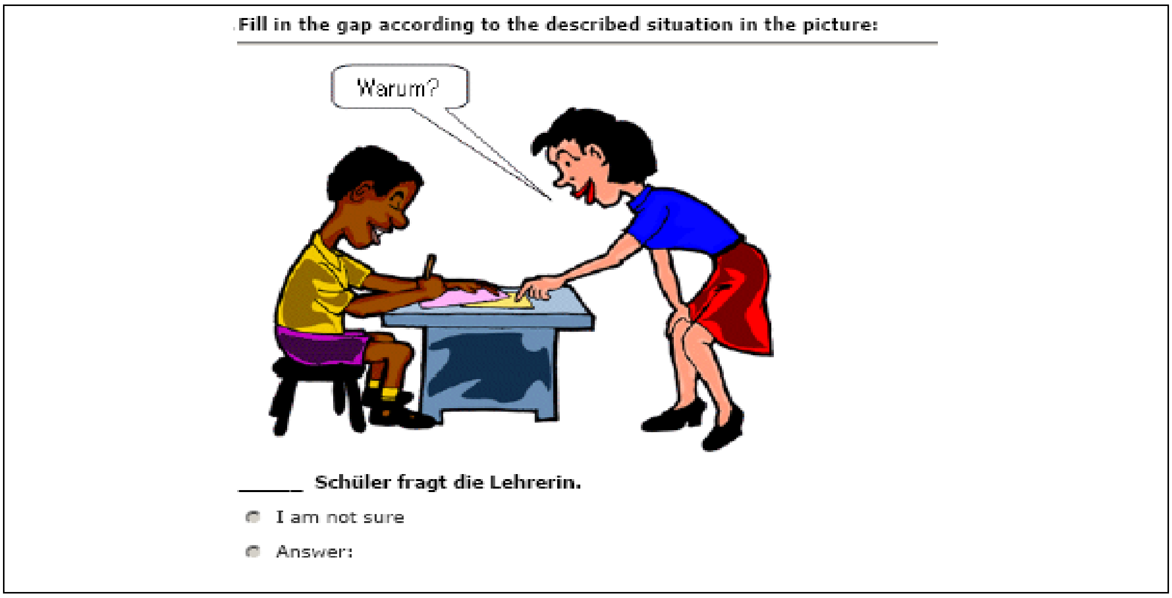


Figure 3.15: Fill in the gap task-extract from the testing materials

Altogether the task comprised thirty terms, sixteen target items and fourteen distracters. Only target items were scored. Participants received one point for each correct response and zero points for incorrect responses. It could be argued that the particular task forces learners to focus on the form and pushes them to call upon their metalinguistic knowledge. In contrast to oral production tasks that tap into implicit knowledge, this written fill in the gap task focuses on language in a rather artificial way. Participants could fill in the task based on the exercises they have previously completed, or because 'it feels right', and/or because they might assume that all sentences should start with 'den'. The fill-in-the-gap task was, however, an independent task, as it was not included in any treatment package.

As discussed in Chapter 2, both PI and IE studies use fill-in-the-gap production tasks to measure for effects of the instructional treatments in the acquisition of a target form (Alanen, 1995; Doughty, 1991; Leow, 2001; VanPatten and Cadierno, 1993; VanPatten and Wong, 2004; Wong, 2004). In fact, one of the claims arguing for the effectiveness of PI is that participants are able to produce the target form, even though it is not included in the treatment, just as successfully as participants that are trained to produce the target form during instruction. Similarly, IE studies use written production tasks in order to assess and evaluate learners' acquisition of the target form (see also 2.2.3 and

2.3.3 sections for further details). However, it should be noted that both PI and IE studies have been criticised for using written fill in the gap production tasks rather than applying tasks that tap implicit knowledge, such as oral production tasks (see Doughty, 2003, 2004). With hindsight, since the use of oral data was not possible due to the nature of the target form (see 3.1.1 and 3.1.2), a timed-constraint fill in the gap task might have been a better choice in collecting data that could provide an insight of implicit knowledge. The present study is therefore limited in providing a better understanding on which instructional method, namely PI, IE and/or the Combined instructional method(s) are effective teaching intervention(s) in acquiring the production of a structure, even when instruction does not promote written production of the form.

3.1.3.4 Interpretation Task

The fourth task was an interpretation task comprising ten items, six target forms and four distracters. The format of the task was the same as in the PI and Combined groups' treatment. It contained a German sentence and two possible English interpretations, the first interpreting the sentence as an SVO and the second as an OVS structure; a third option claiming that the aforementioned options a) and b) are possible interpretations, as well as a fourth option 'I am not sure'. The third and fourth options were included to control for guessing. This task was biased towards the PI and the Combined groups, as it was included in their treatment activities, hence it was anticipated that these two groups would outperform the IE and the Control groups (see Figure 3.16).

Choose which sentence corresponds better to the German sentence.

1. Den Opa fragt das Kind.

- a. The grandfather asks the child.
- b. The child asks the grandfather.
- c. Both a and b options correspond to the German sentence.
- d. I am not sure.

Figure 3.16: Interpretation Task, extract from the testing materials

Participants received two points for each correct answer and zero points for any of the three incorrect options.

3.1.3.5 Summary and Conclusions (Treatment and Testing materials)

Groups	IE Group	PI Group	Combined Group	Control Group
Explicit Information	Same amount of Explicit Instruction (EI) with the other 2 groups, adapted to the IE theoretical framework	Same amount of Explicit Instruction (EI) with the other 2 groups, adapted to the PI theoretical framework	Same amount of Explicit Instruction (EI) with the other 2 groups, adapted to the IE and PI theoretical framework	No instruction
Treatment tasks	Reading comprehension and true/false tasks with the target form being typographically enhanced with colours only comprehension and interpretation tasks, all available online	Structured Input activities (referential and affective activities focusing on comprehension and interpretation of the target form), all available online.	Structured Input activities with the target form being typographically enhanced with colours focusing on comprehension and interpretation of the target form), all available online.	Fill in the gap, and reading comprehension tasks focusing on vocabulary, all available online.
All groups				
Testing Tasks	Same one online version for the three conditions (pre-, immediate-, delayed-post tests for all four groups including four tasks, two dependent tasks, i.e. a reading comprehension and an interpretation task and two independent a fill in the gap and an error correction task.			
Testing Items	44 target forms and 33 distracters, while examples available in the treatment were excluded from the testing materials.			
Feedback	No feedback regarding if a response was correct/incorrect; and/or regarding score was available during the treatment and testing conditions for all four groups. After completion of the delayed post-test the researcher provided feedback when requested.			
Control for guessing	Both treatment and testing materials contained ‘I am not sure’ option in order to control for guessing.			
Control for artificiality	Both treatment and testing materials were checked by native speakers of German in order to control for errors and artificiality of language.			
Unknown vocabulary	Hard copies of unknown vocabulary were available both during treatment and testing and the researcher was present to respond to questions about unknown words not included in the hard copy.			

3.2 Participants

PI and IE studies have focused mainly on adult SLA and very little research has been carried out on subjects in secondary schools. Since participants in the present study are secondary school learners, information about the effectiveness of instruction for this particular age group will be provided. Initially, 156 English learners of German in their second year tested roughly after 100 hours of learning German participated in the study; however the final number of participants was 131, since some participants did not attend all four sessions. In some of the secondary schools, German was available from Year 7, in some from Year 8, while in others from Year 9. As a result, participants in the study ranged from Year 8-Year 10 and were aged 12-14. It should be noted that according to the British educational system, learners are allocated to groups (top, middle and bottom sets) according to their abilities. Overall, there are three higher-set and three lower-set classes participating in the study. However, it should be noted that teachers reported that participants in all six classes were of mixed abilities.

There is also variability with regard to the teaching hours per week in the participating classes. One top set and one bottom set class have three hours of German per week, while another bottom set only have two hours per week. Another top set class has three hours per fortnight, while a top and bottom set class have four hours per fortnight. In addition, course books and teaching materials used in the classes vary. One of the participating classes also had access to German Language web pages once a week, mostly with interactive language-oriented games, while the remaining six participating classes used the same or similar web pages or software accompanying the course book, though not on a regular basis.

Most of the schools involved in the study participated in school and student exchange programmes with German schools, and/or organised school trips to Germany. However, participants in this study had not yet had the opportunity to participate in these activities.

Participants took an online placement test available from the Goethe Institut (http://www.goethe.de/cgi-bin/einstufungstest/einstufungstest.pl)[47], in order to have an independent assessment of their proficiency. The placement test revealed that the majority of participants were at the A2 level; however, two learners were placed on level A1. The range of scores (there is a range in scores for level A1 and A2) confirmed that participants in each class were of mixed abilities.[48]

As required for ethics approval, the researcher had undergone a Criminal Records Bureau (CRB) check before approaching the schools, and written informed consent was obtained following British Association for Applied Linguistics (BAAL) and Newcastle University ethical guidelines. The present study was also assessed and approved by Newcastle University Ethics Committee following the University's ethical codes of research, prior to its realisation. The participants were reassured that their participation was anonymous and that the study was not in any way evaluating their school's performance or grades. The researcher also attempted to overcome affective factors that might have affected the study, such as stress, anxiety and loss of motivation by using online materials both for the treatment and testing materials (Gass and Mackey, 2005).

3.2.1 Schools

The study was carried out in five different secondary schools in the North East of England. A sixth school was also involved in the pilot study. All six schools were mixed comprehensive 11-16 schools. Two participating schools are also Language Specialist Colleges (i.e. schools with special status and extra funding awarded on the basis of a range of language teaching and learning initiatives and successes).

---

[47] Goethe Institut is the German equivalent of the British Council. It runs courses on German language and culture and is acknowledged worldwide for running standardized exams in the German language. The online test rated students' level according to the Common European Framework of Reference for Languages (CEFR).

[48]

A1	A2
Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

(Source: http://www.goethe.de/ins/de/ler/kst/enindex.htm)

Consent for the study was given by the schools' senior management, namely the Headteacher and the Head of Modern Languages/ Head of German Department. In addition, signed consent was sought from both the students and parents/guardians wishing to participate in the study on a voluntary basis. All students agreed to participate in the study.

The five schools participating in the main study provided six classes for the investigation. Classes were used intact in each school. Gall, Borg, and Gall (1996:490) suggest that the internal validity of an experiment in which intact groups are used can be maintained if there is more than one class per experimental treatment and the classes are randomly assigned to the experimental and control groups. Therefore, in order to enhance internal validity, participants in each class were divided randomly into the four groups (-IE+PI, +IE-PI, +IE+PI, -IE-PI).

3.2.2 Teachers

After obtaining consent from the schools, teachers were informed by the researcher about the broad aims of the study, namely an investigation of the effectiveness of different instructional methods in acquiring German grammar. In order to control for any external and internal factors that may have influenced the study, although teachers were informed of the general focus of the study, they were not told about the precise grammatical phenomenon under investigation. Moreover, they did not have access to the intervention and testing materials used in the study until after data collection was completed. As teachers would be present during the actual study and therefore would be able to find out which grammatical phenomenon was investigated, it was agreed that they would not provide any assistance to participants during the whole duration of the study. To my knowledge, teachers did not attempt in any way to provide assistance during the course of the study. A timeline for the whole study was agreed with each school. Furthermore, as access to computers and the internet were essential for the aims of the study, it was ensured that participants would have access to computers and the internet in all four sessions of the study. Further communication was established with the schools' technicians in order to ensure that participants would be able to access the testing and teaching materials.

Prior to the study, teachers provided information about the classes participating in the study, in terms of participants' level, hours of German per week, school's curriculum and extra-curricular activities (i.e. participation in student exchange programs, visits to Germany), course book and general information about their teaching intervention (see section 3.5 for further details).

The presence of the teachers during the study was essential. The role of the teachers in the study mainly involved registration and discipline of the students, as stipulated by school regulations.

3.2.3 Timeline of the study

The timeline of the study was designed to test for short and long term effects of the instructional methods. In the first week of the study, participants took the pre-test in order to test and control for prior knowledge of the target form. In the second week of the study a one and a half hour treatment was followed by the immediate post-test in order to measure short term effects of the instructional treatments. Twelve to fourteen weeks after the administration of the immediate post-test, the delayed post test was administered in order to measure the long term effects of the instructional packages. The administration of the delayed post test in this timeframe was motivated by criticism of the majority of PI and IE studies for administrating delayed post-tests three to four weeks after the treatment, which is too soon after treatment to provide generalisable and reliable conclusions on effectiveness of the teaching intervention and acquisition of the target form (see Mitchell and Myles, 2004; Norris and Ortega, 2000; see also 2.2.3 and 2.3.3 sections). As one of the research questions of this study is to measure if long term effects of the applied methods can be claimed, the timeline of the delayed post-test was set initially at 12 weeks after the second treatment and immediate post- test administration. However, in some schools, despite careful organisation of testing, delays of 2 weeks occurred due to technical difficulties (in terms of computer maintenance), half-term school breaks, or relocation of the school's facilities. The table (Table 3.2) in the next page summarises the timeline of the study:

Table 3.2: Overview of the timescale of the study

Pre-test	Treatment 1 1 hour	Treatment 2 (half an hour) + Immediate post-test (IPT) half an hour	Delayed post-test (DPT)
Week 1	Week 2	Week 2	12 -14 weeks after treatment and IPT

3.2.4 Questionnaire

The majority of PI and TTIE studies have investigated adult SLA. However, they have not accounted for effects of multilingualism, i.e. the fact that some participants might speak and/or learn more than one second language. Evidence supports the proposal that that multilingual second language learners may progress quicker than non-multilinguals (Morgan and Newport, 1981; Ramsey, 1980; see also studies McLaughlin, 1990; Nation and McLaughlin, 1986; Nayak, Hansen, Krueger and McLaughlin, 1990). Review of language models and theories has also shown that younger learners acquire a second language in different ways than adults (see Carroll, 1999, 2001, 2007; MacWhinney and Bates, 1989; McLaughlin, 1990; Pienemann, 1984, 1998, 2000; Sharwood Smith and Truscott, in prep, among many). Investigating the role of TTIE, PI and/or the combination of the two in younger SLA learners could provide further insight into the processes involved in language learning at the initial stage, i.e. when the learner first starts acquiring a foreign language, especially when language learning starts at younger ages (e.g. many countries introduce a second language in their schools' curricula in elementary schools).

Upon completion of the first treatment, participants were asked to fill in a questionnaire (for further details please refer to Appendix B), in order to document background variables, such as L1, exposure to the L2, years of learning German, and if they had received any German teaching outside the classroom. In the second part of the questionnaire, participants were asked to rate the activities (using a Likkert scale) provided in their group. They were also asked to provide any negative and positive feedback regarding the activities of their treatment group. The questionnaire also included questions that could provide qualitative evidence as to whether the instructional method focused learners' attention on noticing and recalling the linguistic properties of the target form and in what way, e.g. 'Did you notice how definite articles are used in German? If yes, please explain briefly'. The aim of including open-ended items in the questionnaire was to obtain insightful information about the thoughts and beliefs of the learners (see Mackey and Gass, 2005).

3.2.5 Pilot study

A pilot study took place in a sixth school in the North East of England. Twelve English learners of German in their second year participated in the pilot study and were divided into four groups (+IE+PI, -IE+PI, +IE-PI and –IE-PI). The participating class was a top set class with similar characteristics to the ones described in subsections 3.2 and 3.2.1 in terms of pass rates, curriculum, textbooks and software used, etc.

Both the treatments and testing materials were piloted. The pilot study provided useful feedback about the logistics of the study, such as the creation of additional email accounts and links in case participants accidentally accessed material for the groups they had not been allocated to. Moreover, the number of distracters in the testing package was cut down to 33 instead of 44 in order to ensure that participants would have sufficient time to complete the exercises especially in the third session where they would have to complete both the treatment and the immediate post-test. Finally, feedback received from the pilot led to the refinement of the treatment and testing materials. For example some of the pictures used had to be replaced and/or amended in order for them to be clearer.

3.3 Data analysis

A range of statistical tests was carried out to analyse the data, looking at differences between test scores for the four groups. The statistical analysis and evaluation was checked by a statistician, ensuring the validity and reliability of the analysis. The statistical procedures carried out are common in the social sciences as well as in IE and PI research. However, it should be noted that a detailed mathematical discussion of the statistical procedures applied in this study is beyond the scope of this thesis. In the following sub-sections the procedure applied in the statistical analysis will be described and justified, discussing the following issues: eliminating data, choosing parametric or non-parametric tests, testing the normality and homogeneity of distributions as well as repeated measures of analysis of a variance (repeated measures ANOVA).

3.3.1 Scoring

Previous PI studies have eliminated learners on the grounds that their scores were considered to be outliers, as their pre-test results were 'too high to show improvement' (VanPatten, 2004; studies in PI see section 2.3.3). However, there is no consensus among the PI studies about the cut off point for exclusion. In the original VanPatten and Cadierno (1993) study (which many have replicated), there were no exclusions (Salaberry, 1998). For the present study, there are no strong arguments suggesting the exclusion of data or learners, since all learners were at beginner level. Furthermore, statistical tests were carried out and provided information regarding outliers. In the event that outliers exist, non parametric tests were carried out (see also Muijs, 2004; Field, 2009).

The actual scoring procedure was described in detail for each of the testing tasks (see sections 3.1.3.1-3.1.3.4). In total, the maximum score that could be achieved was fifty-seven points: fifteen in the error-correction task, fourteen in the comprehension task, sixteen in the production task and twelve in the interpretation task. Both PI and IE studies do not use acquisition criteria [49] (as proposed by Andersen 1978; Bahns 1983; Dulay and Burt 1974; Ellis 1988; Pallotti, 2007; Pienemann, 1998; Vainikka and Young-Scholten, 1994; see also section 5.5 for further details) for scoring, when claiming acquisition of the target form. The present study followed the paradigm of the IE and PI studies.

[49]Test scoring is used as a means to measure acquisition. Researchers have proposed a wide range of percentage scores (60-90%), which are assumed to be establishing the threshold indicating acquisition of the target structure. These percentage scores have been defined in the literature as 'acquisition criteria' (see Andersen 1978; Bahns 1983; Dulay and Burt 1974; Ellis 1988; Pallotti, 2007; Pienemann, 1998; Vainikka and Young-Scholten, 1994).

Collected data were analysed using the Statistical Package for the Social Sciences (SPSS).

3.3.2 Eliminating data

Data was excluded in the analysis when participants did not attend one of the four sessions and/or did not complete the five package materials (two treatment and three testing materials)[50]. The reasons for excluding participants were due to the fact that the study is set to investigate the effectiveness of the instructional treatments available. In addition, the use of repeated measures ANOVA to analyse data requires comparing and measuring performance of the same participants in the three conditions (pre-, immediate and delayed post- test).

The final number of participants was therefore 131.

3.3.3 Parametric vs. non parametric tests

According to the literature review, there is a preference in PI and IE studies for the use of parametric tests to analyse the data, even when datasets do not meet the criteria to run parametric tests (see criticism Doughty, 2003; Lee and Huang, 2008; Norris and Ortega, 2000). In general, parametric tests are considered to be more robust than non-parametric tests[51]. However, there is at least one non-parametric equivalent for each parametric general type of test (see Bryman and Cramer 1997; Field 2009; Muijs 2004). If parametric tests are to be used, sample size, distribution and homogeneity of variance need to be controlled for. With regard to the sample size and the use of parametric tests, Bryman and Cramer (1997:118) suggest a sample size of 15 participants and above, whereas Diamond and Jeffries (2001) suggest a sample size of 10 or more if the

[50]There were limited cases were participants attended the sessions and completed the second treatment, however they did not complete the immediate post-test for various reasons (sickness, other obligations, such as preparing and/or attending school events; one to one tutoring). These participants were excluded from the study.

[51]The difference between the two lies in the fact that parametric tests give information about the size of the differences between scores by using the mean, standard deviation and variances of the group scores and are therefore more likely to detect a significant difference where there is one. On the other hand, non-parametric tests rank the outcome variable from low to high and then analyze the ranks and are considered assumption-free tests because they make fewer assumptions about the type of data on which they can be used, and are therefore considered as less powerful than parametric tests (Field, 2009:540).

population is normally distributed; Goodwin and Goodwin (1996:41) suggest that for every ‘predictor’ 10 subjects are necessary; while Woods, Fletcher and Hughes (1989) suggest a sample size of 25 or more in case the population is skewed. The present study fulfils the criteria for carrying out parametric tests, as the end number of participants is N=131 divided into four groups, +IE+PI n=34; -IE+PI n=33; +IE-PI n=32; and –IE-PI n=32; and predictors are two, namely time of test and type of instruction.

The other two criteria that need to be met in order to carry out parametric and non-parametric tests are the normality test of distribution which checks that the sample is normally distributed prior to the intervention; and the homogeneity of variance test, which verifies if groups were equally distributed prior to the intervention. In the present study the dataset did not always pass the normality and homogeneity of variance tests. The normality tests carried out for the pre-, immediate and delayed post tests revealed that in more than one group in all three conditions the criterion of normally distributed sample size had not been met, since significance should be at $p \geq .05$.For uniformity and reliability reasons non-parametric tests were carried out.

Repeated Measures ANOVA (RMA) is a parametric test type used for running comparisons when three ‘repeated’ measures are available for each case. The fact that the same participants took the pre-, immediate- and delayed-post-test allowed for RMA analysis that can provide information about differences in performance from Time 1 (pre-test) to Time 2 (immediate post-test) and Time 3 (delayed post-test).

3.3.4 Summary of research design and methodology

- The timeline of the study was set as follows: Week 1-Pre-test; Week 2-treatments and immediate post-test; Week 12-14-delayed post test
- Including the pilot study, six schools in the North East of England participated in this study. There were 7 classes, totalling 131 students.
- A placement test available from Goethe Institut online and a qualitative questionnaire were two independent measures to assess the level of participants and control for the variable exposure to input.
- The Questionnaire also allowed participants to assess the treatment materials.

- The maximum scoring that could be achieved was 57 points; 15 for the error correction based on three steps; 14 for the comprehension; 16 for the fill in the gap task and 12 for the interpretation task.
- Data was eliminated if one participant failed to attend one of the four sessions or complete the five available packages for each group.
- Normality and homogeneity tests revealed the need to carry out non-parametric tests.
- Repeated Measures ANOVA parametric were carried out in order to compare and report differences if any in the three conditions, pre-, immediate- and delayed post-test.

3.3.5 Summary and Conclusions: Original features of this study and contribution to PI and IE research

- The time between the immediate and delayed post-test was longer (12-14 weeks post instruction) than in many PI and IE studies, where most delayed post-tests take place in week 4 of the study design.
- To the best of my knowledge, this is the second study conducted on German word order and case marking for PI and the first one on young learners of German for both PI and IE (Culman et al. 2009). Therefore the present study adds to the body of research on PI and IE approaches, while it also provides a better understanding of the acquisition of German word order and case marking in young English learners of German.
- All three groups had the same amount of explicit grammar instruction, in order to control for the variable exposure to grammatical input.
- All three groups had the same amount of types and tokens, contrary to other PI studies that have failed to achieve this criterion.
- The number of participants (*N*=131) was much higher than in most PI and IE studies.
- Prior to the teaching intervention, an online independent placement test established the participants' L2 proficiency (available from the Goethe Institut at http://www.goethe.de/cgi-bin/einstufungstest/einstufungstest.pl). A battery of testing tasks measured the participants' performance post-instruction.

- To the best of my knowledge, this is the only study that has isolated and combined the variables coloured typographical/textual IE and PI. In this way, the present study aims to provide a better understanding of the effectiveness of PI, IE or a combination of the two in L2 instruction in the acquisition of German word order and case marking with English learners of German. It is hoped that a greater understanding will be reached regarding the effectiveness of explicit and/or more elaborate and obtrusive vs. implicit and/or less elaborate and obtrusive types of instruction.

4 Results

Introduction

In this section, the results of the present study will be reported. Let us first summarise the information presented in Chapter 3. The total number of participants was 131 and participants were allocated into four groups, namely a PI (N=33); an IE (N=32), a Combined (N=34) and a Control group (N=32). A battery of testing tasks was deployed, including an error-correction task; a reading comprehension; a production; and an interpretation task. The same test version was available in all three conditions. After the pre-test, a two-day treatment was provided followed by the immediate post-test. Participants received the delayed post-test 12 to 14 weeks post-instruction. The maximum score that participants could achieve was fifty-seven points. Participants also received a hard copy with unknown vocabulary (please refer to Appendix B).

Analysis of results was conducted with the Statistical Package for the Social Sciences (SPSS). Statistical tests were carried out to explore normality of distribution and homogeneity of dataset, as well as the presence of outliers in the dataset. Shapiro-Wilk tests showed a significantly non-normal distribution, while Levene's homogeneity of variance tests showed that groups were not equally distributed across a number of variables ($p<.05$). In addition, the boxplot analysis revealed a number of outliers. For these reasons, statistical analysis was carried out used non-parametric tests. Kruskal-Wallis tests were carried out to measure whether there were statistically significant differences across groups. Paired Mann-Whitney tests were used to test for significant differences between scores.

Results will be analysed and reported in this chapter firstly by reporting the overall groups' performance in the three testing conditions. Secondly, each of the four testing tasks, i.e. error-correction; comprehension; fill in the gap; and interpretation tasks, will be analysed and reported. Analysis will take place in a specific order. It will begin with the presentation of descriptive data, in order to provide a better overview of the findings, followed by the Kruskal-Wallis and Mann-Whitney non-parametric tests, which can provide an in-depth understanding about the possible differences in the dataset. Finally, results will also be analysed and reported with Repeated Measures

ANOVA (RMA) tests, in order to provide a better understanding about groups' performance in the three testing conditions and in each of the testing tasks.

For ease of reference the hypotheses of the present study, as presented in the literature review chapter, are provided below:

Hypothesis 1: The combined teaching intervention, being more elaborate, obtrusive and explicit than PI and IE, will be a more effective teaching intervention.

Hypothesis 2: PI, being more elaborate, obtrusive and explicit than IE, will be a more effective teaching intervention.

Hypothesis 3: IE will be a more effective teaching intervention than no instruction, based on IE studies and meta-analyses conducted up to now.

Hypothesis 4: Effects will be found in the short and the long term.

4.1 Overall groups' performance

The battery of tasks used in the present study included an error-correction, a reading comprehension, a fill in the gap (written production) and an interpretation task. The same test version was used in the three testing conditions and participants could score a maximum total of 57 points.

In the pre-test condition, descriptive analysis of results showed that groups did not vary in their overall performance. Raw scores ranged in the four groups from zero to eight points out of the fifty seven they could achieve. According to the median scores, the majority of participants scored four out of the fifty seven points, while standard deviation (SD) showed that there was great variability between and within groups, as Table 4.1 summarises:

Table 4.1: Pre-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	8	4.00	2.256
PI	33	0	8	4.24	2.437
IE	32	0	8	3.94	2.663
Control	32	0	8	4.13	2.826

Further statistical tests showed that in the pre-test condition, groups were not significantly different. The Kruskal-Wallis test showed that there were no statistical significant differences (p=.974) across groups. Between groups' analysis with Mann-Whitney paired tests showed also that groups were not statistically significantly different (p>.05) from each other. Findings from this analysis are in line with the placement tests measuring participants ‘proficiency in the L2. Out of the 131 participants only two were placed at level A1, while the rest were placed at level A2. However, it should be noted that despite the fact that the majority of the students were placed at the A2 level there were variations in participants’ actual scoring ranging ±3 points[52], confirming that groups were not homogenous, as described also by the teachers during the interviews.

Based on the analysis of results as well as the information obtained from the teachers and the questionnaire, it is valid to conclude that participants did not have prior knowledge of the target form. The range in the scoring both in the pre-test condition as well as in the placement test further indicates that there is great variability between participants. This finding comes in line with the fact that the dataset includes participants from six different schools; three top and three bottom-set groups. Thus, individual differences may have affected the dataset.

In the immediate post-test, raw score data revealed that the Combined group post-instruction outperformed all three groups. Out of the 57 total points that one could achieve in the Combined group the score ranged from 0 to 53 points; in the PI from 4 to 47 points; in the IE from 0 to15.47 and in the Control group from 0 to 12 points. Based on the mean score performance, a hierarchy is obtained suggesting that the Combined group outperformed the PI group, which in turn outperformed the IE group, which outperformed the Control group (Combined>PI>IE>Control). However, SD deviation also shows a wide range in the immediate post-test suggesting that there is great variability within groups (Table 4.2).

[52] The placement test comprised of a cloze-test with thirty multiple choice items. The maximum score that participants could achieve was 30 points. Based on their scoring an automated message was generated placing students at the appropriate language level. Participants were asked to note on the questionnaire-sheet both the score and the level they have achieved. It was clear that participants were placed at the same level even though their scores ranged with a difference reaching ±3 points.

Table 4.2: Immediate post-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	53	34.18	12.423
PI	33	4	47	24.79	13.014
IE	32	0	37	15.47	10.491
Control	32	0	12	5.75	3.510

The Kruskal-Wallis revealed highly significant differences ($p=.001$) across groups post-treatment. Paired comparison tests using Mann-Whitney showed that the Combined group was statistically significantly different ($p \leq .003$) from the PI, IE and the Control groups. In turn, paired comparisons between the PI, the IE and the Control groups showed statistically significant differences between groups ($p \leq .007$). Significant differences ($p=.001$) were also revealed from paired comparisons between the IE and the Control groups.

In the delayed post-test, condition analysis showed that gains made post-treatment were maintained. However, based on the mean score data, gains were at a lower rate than in the immediate post-test condition. Nonetheless, the hierarchy obtained in the immediate post-test condition is also maintained in the delayed post-test, as Table 4.3 summarises:

Table 4.3: Delayed post-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	56	18.74	13.415
PI	33	0	48	13.85	11.568
IE	32	1	29	9.72	5.990
Control	32	0	12	5.25	3.583

The Kruskal-Wallis statistical analysis showed significant differences ($p=.001$) between groups twelve to fourteen weeks post-instruction. Paired Mann-Whitney comparisons between the Combined and PI groups showed non-significant differences ($p=0.66$). However, significant differences were found from comparisons between the Combined and the IE and the Combined and the Control groups ($p \leq .004$). Paired Comparisons between the PI and the IE groups also showed non-significant differences ($p=.305$). Significant differences ($p=.001$) were reported for comparisons between the PI and the Control and the IE and the Control groups.

Overall, Repeated Measures ANOVA (RMA) showed that there were significant differences (p=.001) across groups from the pre- to the immediate post-test and from the immediate to the delayed post-test. Post-hoc Scheffe and Bonferonni tests comparing paired groups performance in the pre- to the immediate and in the immediate to the delayed post-test revealed that groups were significantly different (p<.05). Figure 4.1 provides an illustration of groups' performance based on RMA in a summarised form. The blue line corresponds to the pre-test, the green to the immediate post-test, while the yellow line corresponds to the delayed post-test. The horizontal axis depicts groups, while the vertical axis depicts raw scores.

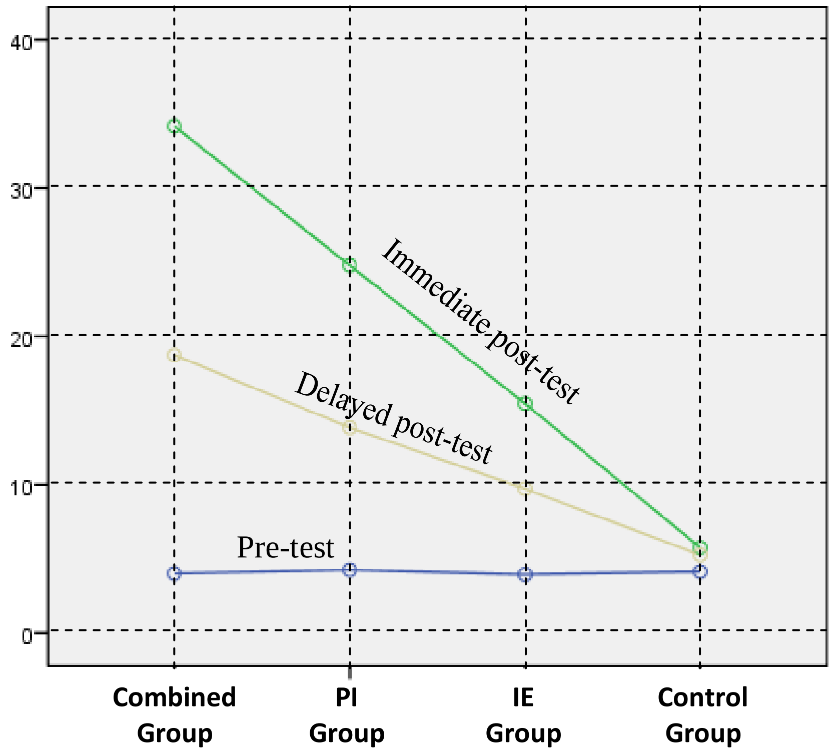


Figure 4.1: Overall groups' performance in the three testing conditions

To sum up, as Figure 4.1 illustrates, prior to instruction, groups did not significantly vary in performance across and between groups, as raw scores are low, indicating that the target structure was a novel structure for participants. Evidently, instruction did play a role in affecting performance, as raw scores increased significantly in the immediate post-test. An effect is also observed regarding the type of instruction, as raw scores vary across groups. Gains made due to instruction are maintained in the delayed post-test after 12-14 weeks, however at a lower level, where the hierarchy generated in the immediate post-test condition is maintained. Furthermore, the reader is reminded that between the immediate and delayed post-test participants, to the best of knowledge, had no instruction and/or exposure to the target form. As discussed in Chapter 3, the researcher and the participating teachers had agreed that students would not receive any sort of feedback, instruction and/or exposure of the target structure until after the delayed post-test. However, it should be noted that based on the questionnaire administered after the immediate post-test there were indications based on participants' responses suggesting a preference on coloured enhanced and non enhanced PI materials, i.e. short sentences with pictures with or without colour enhancement, instead of IE materials, involving by comparison long reading comprehension texts with textual colour enhancement.

4.2 Error-Correction Task Results

The error correction task comprised three steps. Participants received one point when providing the correct response in all three steps of the task. The reader is reminded that in every step the option 'I am not sure' was available in order to control for guessing, which was scored zero points, as was the case for incorrect responses.

Pre-test raw score data show that groups were not able to successfully complete the task, verifying that the target form was a novel task for participants within and between groups (Table 4.4):

Table 4.4: Pre-test raw scores from the error-correction task

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	0	.00	.000
PI	33	0	0	.00	.000
IE	32	0	0	.00	.000
Control	32	0	0	.00	.000

The error correction task was an independent task in the battery of tests, as it was not included in any treatment package of the four groups. Statistical analysis measuring performance across groups with the Kruskal-Wallis test and between group comparisons carried out with Mann-Whitney tests showed that groups did not significantly differ in performance in the pre-test condition ($p>.05$).

The factors instruction and type of instruction affected the dataset, as the analysis of the immediate post-test raw scores showed that groups significantly improved. Based on the mean score performance, the Combined group outperformed the three groups, while a hierarchy is obtained suggesting that the Combined group outperformed the PI, which outperformed the IE group, which in turn outperformed the Control group (Combined>PI>IE>C). However, SD shows that there was great variability within groups, as Table 4.5summarises:

Table 4.5: Immediate post-test raw scores from the error-correction task

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	15	10.09	3.988
PI	33	0	14	7.27	4.125
IE	32	0	11	3.44	2.884
Control	32	0	2	.44	.669

Across groups statistical analysis with the Kruskal-Wallis test showed statistically significant differences ($p=.001$). Between groups' analysis using the Mann-Whitney test showed that the Combined group was significantly different ($p\leq.007$) from the other three groups. Paired comparisons between the PI group and the IE and Control groups showed statistically significant differences ($p=.001$); while comparisons between the IE and the Control groups showed that groups were significantly different ($p=.001$).

In the delayed post-test, condition analysis indicated that gains were maintained, although at a lower rate. Variability between groups is evident based on the SD data. Despite the fact that overall, the Combined group outperformed the three groups, the mean scores showed that in the delayed post-test, the Combined and PI groups did not vary significantly, as Table 4.6 illustrates:

Table 4.6: Delayed post-test raw scores from the error-correction task

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	14	4.59	3.978
PI	33	0	13	3.85	3.203
IE	32	0	8	1.75	2.064
Control	32	0	0	.00	.000

Statistical analysis with the Kruskal-Wallis test showed that groups were significantly different ($p=.001$) in the delayed post-test. Paired group comparison with Mann-Whitney tests showed that the Combined and PI groups were not significantly different ($p=.587$). However, significant differences are reported between the Combined and the IE groups as well as between the Combined and Control groups ($p=.001$). In turn, the PI group was significantly different ($p\leq.002$) from the IE and Control groups. Paired comparisons between the IE and the Control group also showed significant differences ($p=.001$). The results provide support to the fact that instruction and more specifically type of instruction does play a role. It should be noted that during the 12 to 14 weeks of gap between the immediate and delayed post-test conditions participants had to the best of my knowledge no instruction of and/or exposure to the target form. Findings of the error-correction task seem to suggest that both the Combined and the PI types of instruction are more effective than IE and/or no instruction. On the other hand, standard deviation shows that within groups there is great range of individual variation, which might be the outcome of learners' individual variation in terms of sensitivity to different types of textual enhancement. A further assumption could be the factor gender in terms of sensitivity to the colour textual enhancement.

Overall, RMA showed that groups were significantly different ($p=.001$) post-instruction in the immediate and delayed post-test conditions. Post-hoc Scheffe and Bonferonni tests showed that overall groups were significantly different ($p\leq.05$). Figure 4.2 in the next page provides an overview of the groups' performance in the three testing conditions. The blue line corresponds to the pre-test, the green to the immediate post-test, while the yellow line corresponds to the delayed post-test. The horizontal axis depicts groups and the vertical axis depicts raw scores.

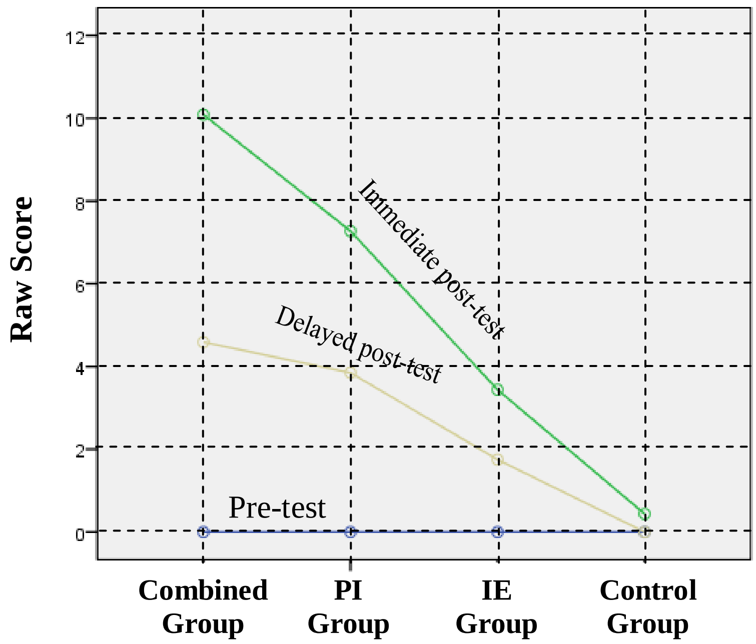


Figure 4.2: Error Correction Task performance of the four groups in the three testing conditions

Based on Figure 4.2, raw score performance showed that groups were equal prior to instruction. The low raw scores indicate that the target structure was a novel structure for participants across and between groups. Immediate post-test performance revealed that score performance significantly improved for the Combined, PI and IE groups, while there was some minor improvement for the Control group. Thus, the groups' performance in the immediate post-test condition suggests that instruction did play a role in improving performance. The variations in scores across groups also show that the type of instruction is an important factor affecting performance. Gains made due to instruction are maintained in the delayed post-test after 12-14 weeks post-instruction, though at a lower rate, while the hierarchy generated in the immediate post-test condition is maintained (Combined>PI>IE>Control).

4.3 Reading Comprehension Results

The second task in the battery of tests was a ten-line reading comprehension text, where multiple choice responses included four options, one being 'I am not sure'. The task included seven target items. Participants received zero points for incorrect and 'I am not sure' responses and two points for each correct response[53]. The maximum score that could be achieved was therefore fourteen points.

Pre-test raw mean scores showed that groups did not vary significantly in performance prior to the teaching interventions. Scoring in the Combined, PI and Control group ranged from 0-8 points, while the mean score for the three groups was almost four points. However, SD indicates that there is great variability within groups, as Table 4.7 below summarises:

Table 4.7: Pre-test raw scores from the reading comprehension task

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	8	3.71	2.419
PI	33	0	8	3.88	2.595
IE	32	0	8	3.63	2.709
Control	32	0	8	3.38	2.848

Statistical analysis using the Kruskal-Wallis test showed that in the pre-test condition groups were not significantly different ($p=.864$). Paired group comparisons using the Mann-Whitney test also showed that groups were not significantly different ($p\leq.05$).

It was anticipated that the IE and the Control groups would outperform the PI and the Combined groups in the reading comprehension task post-instruction. This type of task was included in both groups' treatment materials and therefore it was expected that it would affect the participants' performance. However, raw scores analysis from the immediate post-test showed that all four groups improved post-instruction. Based on the mean scores, it is evident that the Combined group outperformed all groups, while the PI and IE groups did not vary; however both outperformed the Control group. Nonetheless, in the immediate post-test, there was also SD above one SD, suggesting great variability and range in scoring within all four groups (see Table 4.8 next page):

[53]The reader is reminded that due to the imbalance of the final tokens in the four testing tasks (15 target tokens for the error correction task, seven for the reading comprehension, 16 for the fill in the gap and six for the interpretation task), it was decided to give two points for each correct response in the reading comprehension and the interpretation tasks in order to balance the score for the analysis of groups' performance. In this way the analysis would not be based mainly on groups' performance in the two out of the four testing tasks.

Table 4.8: Immediate post-test raw scores from the reading comprehension task

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	14	7.00	3.124
PI	33	0	12	5.45	3.545
IE	32	0	12	5.56	3.242
Control	32	0	12	4.44	2.994

The Kruskal-Wallis test showed that groups were statistically significantly different ($p=.011$) post-instruction. Paired Mann-Whitney group comparisons between the Combined and PI groups revealed that groups were not significantly different ($p=.070$). However, comparisons between the Combined and the IE groups and between the Combined and the Control groups revealed that groups were significantly different ($p\leq.047$).

In the delayed post-test, the three groups maintained their gains, though at a lower rate, while the Control group showed minimal improvement. Mean scores performance therefore suggests the following hierarchy for the reading comprehension delayed post-test: the Combined group outperforms the Control group, which in turn outperforms the PI and IE groups, which perform equally (Combined>C>PI=IE), as Table 4.9 below shows:

Table 4.9: Delayed post-test raw scores from the reading comprehension task

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	14	5.47	3.760
PI	33	0	10	4.30	3.468
IE	32	0	10	4.19	2.799
Control	32	0	10	4.75	3.473

Statistical analysis with the Kruskal-Wallis showed that groups were not significantly different ($p=.520$) in the delayed post-test condition. Paired group comparisons between the Combined group and the other three groups demonstrated non-significance ($p\geq.05$). Similarly, comparisons between the PI and the IE and Control groups; as well as the IE and the Control groups showed non-significant differences ($p\geq.05$).

Overall, RMA tests showed that there was a positive interaction for comprehension from the pre- to the immediate condition and from the immediate to the delayed post-test condition ($p \leq .05$); however it was not significant in relation to groups ($p \geq .05$). Post-hoc Scheffe and Bonferonni tests demonstrated also non-significant differences ($p \geq .05$) between groups in the three testing conditions. Figure 4.3 provides a summarised overview of groups' performance in the reading comprehension task in the three testing conditions. The blue line corresponds to the pre-test, the green to the immediate post-test, while the yellow line corresponds to the delayed post-test. The horizontal axis depicts groups, while the vertical axis depicts raw scores.

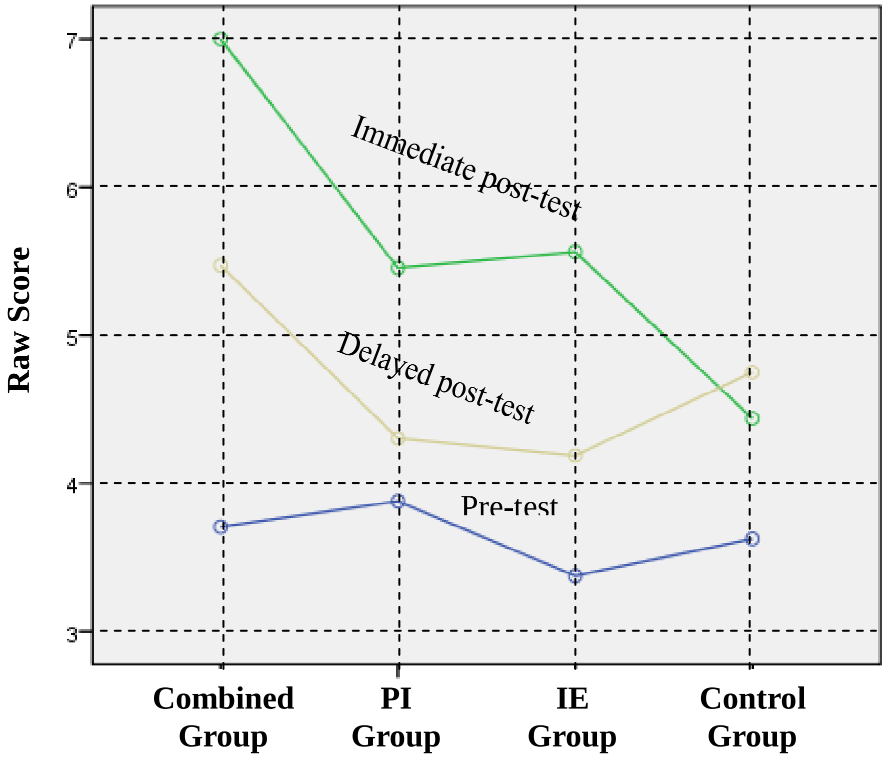


Figure 4.3: Comprehension Task- Repeated Measures ANOVA

Clearly groups' performance prior to instruction shows that between groups scores fluctuated suggesting a hierarchy, namely that the PI group outperformed the Combined group, which in turn outperformed the Control group, which outperformed the IE group (PI>Combined>Control>IE). Post-instruction and in the immediate post-test condition groups improved significantly, while the obtained hierarchy differs from the one obtained in the pre-test condition. The Combined group outperformed the IE group, which in turn outperformed the PI group, which outperformed the Control group

(Combined>IE>PI>Control). In the delayed post-test condition, groups' performance differs compared to the pre and the immediate post-test condition. The obtained hierarchy shows that the Combined group outperformed the Control group, which outperformed the PI group, which in turn outperformed the IE group (Combined>Control>PI>IE). Despite the fact that statistical analysis showed that groups were significantly different only in the immediate post-test condition, it is clear that post-instruction, the Combined group outperformed the other three groups in the two conditions. Gains and hierarchies are not constant in the two conditions for the other three groups, while the Control group's performance improves in the delayed post-test outperforming the PI and IE groups.

4.4 Fill-in-the-gap results

The fill-in-the-gap written production task was the third task in the sequence of the battery of tests. In total there were 16 target items. Participants were presented with a picture and a sentence with the gap at the beginning and were required to fill-in-the-gap. The option 'I am not sure' was also provided in order to control for guessing. The maximum score that could be achieved was sixteen points, one point for each correct response.

Descriptive analysis of results revealed that participants from all four groups had no prior knowledge of the target form, as the target structure was not produced in the pre-test condition in any of the four groups (Table 4.10):

Table 4.10: Fill in the gap pre-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	0	.00	.000
PI	33	0	0	.00	.000
IE	32	0	0	.00	.000
Control	32	0	0	.00	.000

As all scores were zero, statistical analysis, using the Kruskal-Wallis test, showed that groups were not significantly different ($p=1.00$) prior to instruction. Paired group comparisons also indicated no significant differences ($p=1.00$) between groups' performance in the pre-test.

One of the assumptions of the present study was that the experimental groups would improve post instruction. Descriptive analysis of results verified the original assumption. Raw mean score performance shows that all three groups improved post-instruction. Based on the mean scores, a hierarchy is obtained showing that the Combined group outperformed the other groups. In turn the PI group outperformed the IE group, which in turn outperformed the Control group (Combined>PI>IE>Control), who showed no improvement at all. However, SD shows that there is yet again dispersion from the average within groups, as Table 4.11 shows:

Table 4.11: Fill in the gap immediate post-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	15	7.91	4.699
PI	33	0	15	4.30	4.779
IE	32	0	15	2.66	4.770
Control	32	0	0	.00	.000

The Kruskal-Wallis statistical analysis showed that groups were significantly different post instruction ($p=.001$). Paired group comparisons between the Combined and the PI, IE and Control groups showed significant differences between groups ($p\leq.004$). Paired comparisons between the PI and IE and Control groups also revealed significant differences ($p\leq.028$) between groups. Paired comparisons between the IE and the Control groups indicated as well that groups were significantly different ($p=.001$) post-instruction.

Delayed post-test performance based on descriptive analysis of results showed that groups maintained gains, however at a lower rate than in the immediate post-test condition. The hierarchy obtained in the immediate post-test condition was also maintained in the delayed post-test condition (Combined>PI>IE>Control). However, dispersion is maintained based on SD, as Table 4.12 in the next page shows:

Table 4.12: Fill in the gap delayed post-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	16	3.50	5.047
PI	33	0	15	1.76	4.131
IE	32	0	7	.97	2.265
Control	32	0	0	.00	.000

The Kruskal-Wallis statistical analysis showed that groups were still significantly different (p=.001) twelve to fourteen weeks post instruction. Paired group comparisons between the Combined and the PI, the IE and the Control groups showed significant differences (p≤.017). However, comparisons between the PI and the IE groups showed no significant differences (p=.679) between the two groups; though significant differences (p=.006) were reported between the PI and the Control groups. The IE and the Control groups were also significantly different (p=.011).

Overall performance of groups in the three testing conditions using RMA showed that there were significant differences from the pre- to the immediate post-test condition, as well as from the immediate to the delayed post-test condition (p=.001). Post hoc Bonferonni and Scheffe tests comparing paired groups' performance in the three conditions showed significant differences between the Combined and the PI, IE and Control groups (p≤.010). On the other hand, post-hoc comparisons between the PI and IE groups showed no statistically significantly differences (p≤.752). Comparisons between the PI and the Control groups showed significant differences (p=.003) in the three testing conditions. Non-significant differences (p≤.164) were also reported between the IE and the Control groups in the three testing conditions. Figure 4.4 next page illustrates groups' performance in the three testing conditions in the fill in the gap task. The blue line corresponds to the pre-test, the green to the immediate post-test, while the yellow line corresponds to the delayed post-test. The horizontal axis depicts groups, while the vertical axis depicts raw scores.

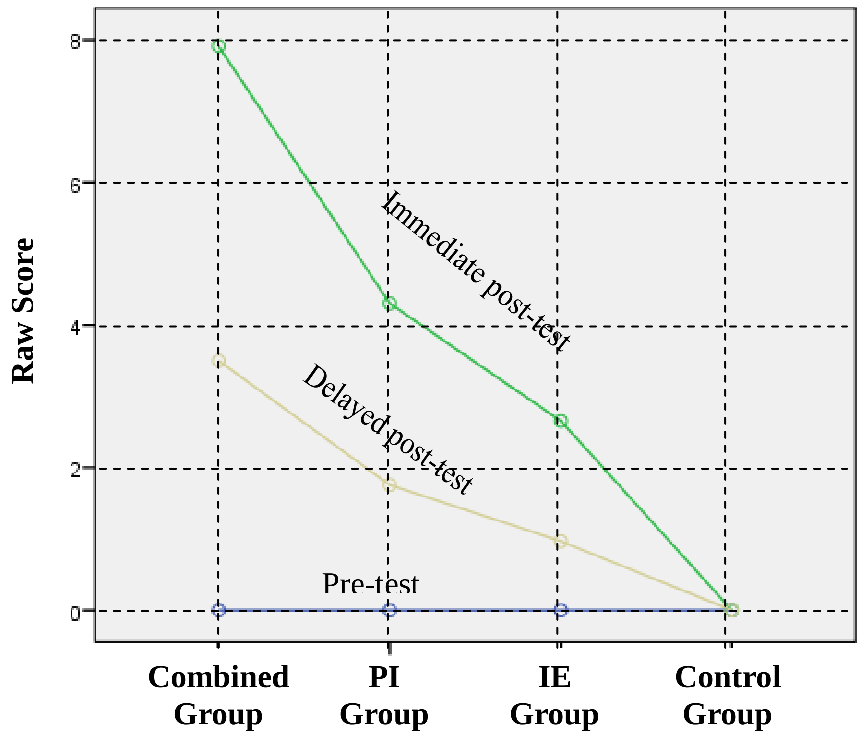


Figure 4.4: Fill in the gap task- Repeated Measures ANOVA

Figure 4.4 clearly depicts that groups had no prior knowledge of the target form prior to instruction based on the raw scores performance of zero. Post-instruction, it is evident that all three experimental groups made gains in the immediate post-test which were maintained, at a lower rate, in the delayed post-test condition. The obtained hierarchy in the immediate post-test is maintained also in the delayed post-test showing that the Combined group outperformed the PI group, which outperformed the IE group, which in turn outperformed the Control group (Combined>PI>IE>Control).

4.5 Interpretation task results

The interpretation task was the fourth task in the battery of testing tasks. The task included six target forms. Participants were required to choose the correct interpretation out of four possible options, one being 'I am not sure' to control for guessing. Each correct response received two points, while all other responses received zero points. The task was biased towards the PI and the Combined groups, as it was available in the two groups' treatment packages. Therefore it was anticipated that the two groups would outperform the IE and Control groups.

Descriptive analysis of raw score performance showed that groups did not significantly vary in performance in the interpretation task prior to instruction. For the Combined and the PI groups, SD scores were below the range of one SD; while for the IE and Control groups SD numbers minimally exceeded the threshold of one SD. This means that there was great variability within the IE and Control groups regarding participants' performance in the pre-test condition. This great variability is also evident from the mean scores.

Table 4.13: Interpretation task pre-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	4	.29	.871
PI	33	0	2	.36	.783
IE	32	0	4	.69	1.203
Control	32	0	4	.50	1.136

The statistical analysis carried out with the Kruskal-Wallis test confirmed that there were no significant differences ($p \leq .408$) between groups prior to instruction. Further paired comparisons between the Combined and the PI, IE and Control groups showed that there were no significant differences ($p \leq .503$). Paired comparisons between the PI and IE and Control groups also reported non-significant differences ($p \leq .861$). Similarly, paired comparisons between the IE and Control groups also presented no differences ($p \leq .414$). Therefore it is valid to assume that the target form was a novel structure for participants in all four groups.

Post-instruction descriptive analysis revealed that instruction did make a difference, as mean score performance significantly improved within groups and varied between groups. Mean score performance reveals a hierarchy showing that the Combined group outperformed the PI group, which outperformed the IE group, which outperformed the Control group (Combined>PI>IE>Control), as Table 4.14 below depicts:

Table 4.14: Interpretation task immediate post-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	12	9.18	3.912
PI	33	0	12	7.76	4.994
IE	32	0	12	3.81	4.497
Control	32	0	4	.88	1.519

The Kruskal-Wallis test showed significant differences (p=.001) between groups post instruction. However, paired group comparisons between the Combined and the PI showed non-significant differences (p=.379) between the two groups. Significant differences (p=.001) were reported between the Combined and the IE and Control groups. Paired group comparisons between the PI and the IE and Control groups revealed significant differences (p≤.005) between groups. Similarly, comparisons between the IE and the Control groups also showed significant differences (p=.004) between groups.

Delayed post-test performance, according to descriptive analysis, showed that groups maintained gains twelve to fourteen weeks post instruction, though at a lower rate than in the immediate post-test. Mean score performance showed that the hierarchy obtained in the immediate post-test was maintained in the delayed post-test condition (Combined>PI>IE>Control). However, SD scores showed that there was great variability within groups, as scores were above the threshold of one standard deviation, as Table 4.15 summarises:

Table 4.15: Interpretation task delayed post-test raw scores

Group	N	Minimum	Maximum	Mean	SD
Combined	34	0	12	5.18	5.396
PI	33	0	12	3.94	4.541
IE	32	0	12	2.81	3.126
Control	32	0	6	.50	1.344

Statistical analysis using the Kruskal-Wallis test showed that groups were significantly different (p=.001) in the delayed post-test. However, paired comparisons showed significant differences (p=.001) between the Combined and the Control groups, the PI and the Control groups, as well as between the IE and the Control groups. Paired comparisons between the three experimental groups showed non significant differences (p≤.604) between the Combined the PI and the IE groups.

Overall, RMA revealed a positive interaction between the interpretation task and groups from the pre- to the immediate post-test and from the immediate to the delayed post-test conditions (p≤.007). Post-hoc Schefe and Bonferonni tests showed that there were non-significant differences (p≤.451) between the Combined and PI groups in the interpretation task. However, the Combined group was overall significantly different (p=.001) from the IE and the Control groups. Overall, significant differences were reported from post-hoc comparisons between the PI and IE and between the PI and the Control groups (p≤.017) and the IE and the Control groups (p≤.005). Figure 4.5 below illustrates groups' performance in the three testing conditions in the interpretation task. The blue line corresponds to the pre-test, the green to the immediate post-test, while the yellow line corresponds to the delayed post-test. The horizontal axis depicts groups, while the vertical axis depicts raw scores.

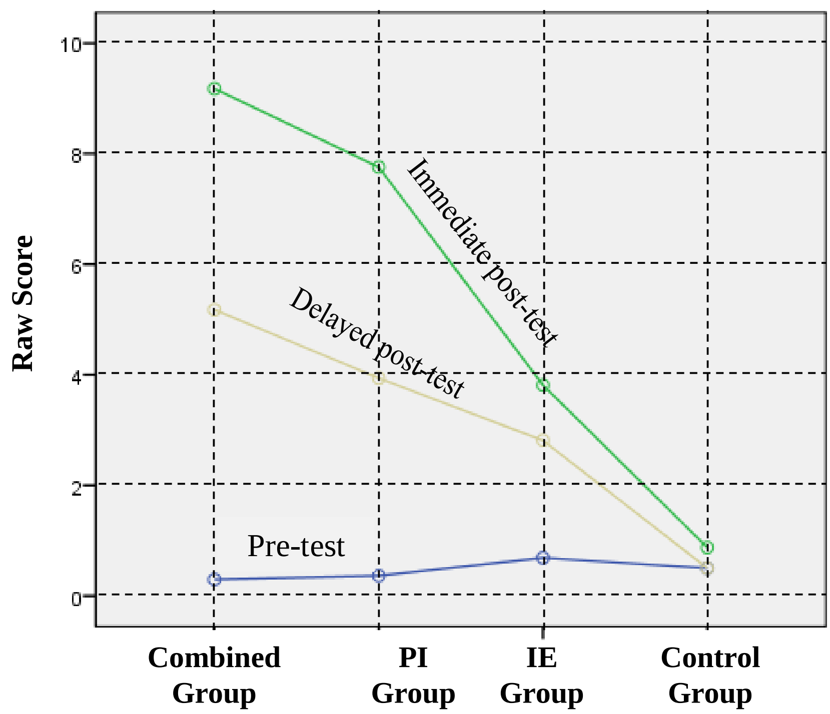


Figure 4.5: Interpretation task- Repeated Measures ANOVA

Figure 4.5 clearly illustrates groups' performance in the three testing conditions. Based on the figure, it is evident that in the pre-test condition, the IE group outperformed the Control group. In turn the Control group outperformed the PI group, and the PI group outperformed the Combined group (IE>Control>PI>Combined). However, in the immediate post-test condition the hierarchy is not maintained, as the Combined group outperformed the PI group, which in turn outperformed the IE group, which outperformed the Control group (Combined>PI>IE>Control). The hierarchy obtained in the immediate post-test is maintained in the delayed post-test condition. However, gains are maintained at a lower rate by comparison.

4.6 Summary of results

131 participants were allocated into four groups, namely a Combined; a PI; an IE; and a Control group. A battery of testing tasks was applied using an error correction; a reading comprehension; a fill in the gap; and an interpretation task. Schools curricula, teachers' feedback, the Goethe proficiency test and pre-test verified that participants had no prior knowledge of the target form. The Goethe proficiency test, pre-test, standard deviation scores and the statistical tests of normality and homogeneity of variance showed that groups were not normally distributed prior to instruction and therefore non-parametric tests were carried out for the analysis and the reporting of results. Results were analysed and reported using descriptive analysis; non-parametric statistical tests exploring overall performance of groups (Kruskal-Wallis) and paired group comparisons (Mann-Whitney); as well as overall performance of groups using parametric Repeated Measures ANOVA (RMA) statistical tests. Findings of the present study are summarised in the next page:

Overall Performance of the four groups
(Descriptive and non-parametric tests)

<table>
<tr><th>Pre-test</th></tr>
<tr><td>All groups were not significantly different (p=.974).

Paired Mann-Whitney tests showed that groups were not statistically significantly different (p>.05) from each other in the four testing tasks and in overall performance.

Findings from this analysis are in line with the placement tests measuring participants' proficiency in the L2.

Therefore it is valid to conclude that the target structure was a novel structure and that any improvements observed post-instruction would be the outcome of instruction; while variations in performance would be the outcome of type of instruction.</td></tr>
</table>

Immediate post-test
Raw score data revealed that the Combined group post-instruction outperformed all three groups. The Kruskal-Wallis revealed highly significant differences ($p=.001$), while paired Mann-Whitney comparisons showed that: The Combined group was statistically significant different ($p \leq .003$) from the PI, IE and the Control groups. The PI was significantly different ($p \leq .007$) from the IE and the Control groups; while the IE was significantly different ($p=.001$) from the Control group. Based on mean score performance, the following hierarchy is obtained for the immediate post-test: Combined>PI>IE>C for overall performance. However, SD deviation shows a great range suggesting that there is great variability within groups.

Delayed post-test
Analysis showed that gains made post-treatment were maintained; though at a lower rate than in the immediate post-test condition. Nonetheless, the hierarchy obtained in the immediate post-test was maintained in the delayed post-test condition. The Kruskal-Wallis analysis showed significant differences ($p=.001$). Paired Mann-Whitney comparisons indicated significant differences between the Combined and the IE; the Combined and the Control; the PI and the Control; and the IE and the Control groups ($p \leq .004$). However, non-significant differences were reported from paired Mann-Whitney comparisons for the Combined and PI groups ($p=0.66$) and between the PI and the IE groups ($p=.305$).

Repeated Measures ANOVA (RMA) for overall performance in the three testing conditions (pre-, immediate- and delayed post-tests)
Significant differences (p=.001) were revealed across groups from the pre- to the immediate post-test and from the immediate to the delayed post-test. Post-hoc Scheffe and Bonferonni tests comparing paired groups performance revealed that groups were significantly different (p<.05).

Error-Correction task		
Pre-test	Immediate post-test	Delayed post-test
Groups did not significantly differed in performance (p>.05).	Groups significantly (p=.001) improved, while median score performance revealed the hierarchy Combined>PI>IE>C. Paired Mann-Whitney group comparisons revealed significant differences (p≤.007) between groups' performance.	Significant differences (p=.001) were maintained, however based on paired Mann-Whitney group comparisons, the hierarchy was not maintained as it revealed that the Combined=PI>IE>Control.
Notes for all three conditions	SD shows that there was great variability within groups in all three conditions. RMA confirmed that groups were significantly different (p=.001) post-instruction in both conditions. Post-hoc Scheffe and Bonferonni tests showed that overall groups were significantly different (p≤.05) in the error correction task.	

Reading Comprehension task		
Pre-test	Immediate post-test	Delayed post-test
Groups did not significantly differed in performance (p>.05). The hierarchy obtained based on median score performance is: PI>Combined>Control >IE	Groups were significantly different (p=.011) only in the immediate post-test. Paired group Mann-Whitney comparisons showed significant differences only in the immediate post-test and only for the Combined and the IE groups (p=.047), as well as for the Combined and the Control groups (p=.001). Based on median score performance, it is evident that all four groups improved post-instruction. The hierarchy obtained based on median score performance differed for the immediate post-test is: Combined>IE>PI>Control	Groups were not significantly different in performance (p>.05). Based on median score performance, it is evident that all four groups improved and maintained gains post-instruction. The hierarchy obtained based on median score performance is: Combined>Control>PI>IE
Notes for all three conditions	SD shows that there was great variability and range in scoring within all four groups. RMA tests showed that there was a positive interaction for comprehension from the pre- to the immediate condition and from the immediate to the delayed post-test condition ($p \leq .05$) though non-significant in relation to groups ($p \geq .05$). Post-hoc Scheffe and Bonferonni tests also demonstrated non-significant differences ($p \geq .05$) between groups in the three testing conditions.	

Fill in the gap task		
Pre-test	Immediate post-test	Delayed post-test
Groups did not significantly differed in performance ($p>.05$).	Groups were significantly different ($p=.001$). Paired Mann-Whitney group comparisons between groups showed significant differences ($p\leq.05$) for the four groups. The hierarchy obtained based on median score performance is: Combined>PI>IE>Control	Groups were significantly different ($p=.001$). Paired Mann-Whitney group comparisons between groups showed significant differences ($p\leq.05$) for the four groups in the two conditions, except from comparisons between the PI and the IE groups ($p=.679$). The hierarchy obtained based on median score performance is: Combined>PI>IE>Control
Notes for all three conditions	SD shows that there is dispersion from the average within groups. RMA confirmed the descriptive and non-parametric analysis of results regarding the fill in the gap task.	

Interpretation task		
Pre-test	Immediate post-test	Delayed post-test
Groups did not significantly differed in performance (p>.05). Based on mean score performance, the hierarchy in the pre-test was: IE>Control>PI>Combined	Significant differences (p=.001) were revealed. Paired Mann-Whitney comparisons in the immediate post test showed that groups were significantly different ($p \leq .005$) from each other. Based on mean score performance, the hierarchy in the immediate post-test was: Combined>PI>IE>Control	Groups were significantly different (p=.001). Paired Mann-Whitney tests were significantly different (p=.001) only for the Combined and the Control groups, the PI and the Control groups, and between the IE and the Control groups. No significant differences (p>.05) were reported from paired comparisons between the experimental groups, i.e. Combined, PI and IE groups. Based on mean score performance, the hierarchy in the delayed post-test was: Combined>PI>IE>Control
Notes for all three conditions	SD scores showed that there was great variability within groups. RMA revealed a positive interaction for the interpretation task and groups from the pre- to the immediate post-test and from the immediate to the delayed post-test conditions ($p \leq .007$). Post-hoc Scheffe and Bonferonni tests showed that groups were significantly different ($p \leq .017$) apart from comparisons between the Combined and PI groups ($p \leq .451$).	

4.7 Conclusion

In conclusion, the present study provides robust significant evidence partially confirming the four hypotheses driving the present study. In terms of the combined teaching intervention, being more elaborate, obtrusive and explicit than PI and IE (Hypothesis 1), mean score performance and paired Mann-Whitney comparisons confirmed that the Combined teaching intervention is indeed more effective than PI, IE and non-instruction. Based on mean score performance and statistical analysis using non-parametric (Kruskal-Wallis) and parametric (RMA) tests, groups are significantly different ($p=.001$) in general from each other post-instruction in the two conditions (immediate and delayed). Thus, Hypothesis 1 is verified both in the short and the long-term. However, a closer examination of the dataset in order to establish where the differences lie showed that groups were significantly different ($p \leq .007$) in the immediate post-test and that the Combined group outperformed the PI, IE and Control groups. However, in the delayed post-test, despite the fact that the Combined group still outperformed the other groups based on mean scores, significant differences ($p \leq .004$) are reported between the Combined and the IE and the Combined and the Control groups, but not for the Combined and PI groups ($p=0.66$). Thus, Hypothesis 1 is not fully verified in terms of the Combined group being more effective than the PI group in the long term.

The second hypothesis of the present study assumes that PI, being more elaborate, obtrusive and explicit than IE, will be a more effective teaching intervention. This second hypothesis is also partially verified, since the PI group is significantly different ($p \leq .007$) from the IE group in the immediate post-test, though non-significantly different ($p=.305$) in the delayed post-test condition. As in the case of the Combined group, the PI group is also more effective than the IE group in terms of performance in the short term but not in the long term based on the statistical analysis. Similarly, as in the case of the Combined group, the PI group is also overall significantly different (p.05) from the IE group based on the Kruskal-Wallis and RMA tests. This fact was also confirmed from the mean score outperformance of the PI group over the IE group.

With regard to the third hypothesis of the present study, the assumption was that IE would be a more effective teaching intervention than non-instruction. This was verified both in the short and the long term, as statistically significantly differences (p=.001) were reported both from the Kruskal-Wallis and the RMA, as well as from paired Mann-Whitney group comparisons in the immediate and the delayed post-test conditions.

Despite the fact that there was a consistency in terms of the Combined group outperforming the PI group, which in turn outperformed the IE group, which outperformed the Control group (Combined>PI>IE>Control) based on mean score performance,; statistical analysis verified the fourth hypothesis of the present study only for the short term. In the long term, groups did maintain their gains made post-instruction and in the immediate post-test condition; however, raw scores were lower in the delayed post test in comparison to the immediate post-test. Despite the fact that both the Kruskal-Wallis and the RMA showed significant differences (p=.001) across groups, a closer exploration of the dataset with paired Mann-Whitney group comparisons showed that the differences were significant (p<.05) in the short term. However, in the long term, comparisons between the Combined and the PI and the PI and the IE groups showed non-significant differences (p>.05).

Overall, it could be argued that the four hypotheses of the present study are confirmed. On the other hand, exploring the dataset in greater detail shows that the hypotheses are mostly validated in terms of short-term effects. In the long-term, results show a positive effect of the teaching interventions, which is also in line with the hypothesised and confirmed hierarchy (Combined>PI>IE>Control). However in depth-analysis of the dataset indicates that there were no significant differences between the Combined and the PI and the PI and the IE groups' effectiveness in the long-term.

In the next chapter, results will be further discussed in more detail looking also at groups' performance in each of the four testing tasks. Results will also be examined in the light of the original theories and models of acquisition that drove the design of the present study. Furthermore, the methodology applied in the present study will be reassessed, focusing on possible factors that might have affected the obtained results.

5 Discussion

Introduction

The present study set to investigate the effectiveness of three different, though interrelated, teaching interventions, namely PI, IE and a combination of PI and IE. The aim was to investigate to what extent the sole application and/or the combination of the two approaches can be effective in helping English learners of German acquire word order and case marking in the short and the long term. Therefore for the purposes of the present study, the following research questions were investigated, as discussed in the literature review section (Chapter 2):

1. Is instruction using colour as Typographical/Textual Input Enhancement (+IE-PI) of the target linguistic form more effective than Processing Instruction (-IE+PI) and/or no instruction (-IE-PI) in the short and long term?

2. Does instruction using Processing Instruction (-IE+PI) prove to be more effective than Colour Input Enhancement (+IE-PI) and / or no instruction (-IE-PI) in the short and long term?

3. Is the combined instructional method (+IE+PI) more effective than the application of each technique separately and/or no instruction in the short and long term?

For ease of reference the hypotheses of the present study, as presented in the literature review chapter are also provided below:

Hypothesis 1: The combined teaching intervention, being more elaborate, obtrusive and explicit than PI and IE will be a more effective teaching intervention.

Hypothesis 2: PI, being elaborate, more obtrusive and more explicit than IE will be a more effective teaching intervention than IE.

Hypothesis 3: IE, will be a more effective teaching intervention than no instruction based on IE studies and meta-analyses conducted to date.

Hypothesis 4a: Effects will be found in the short term.

Hypothesis 4b: Effects will be found in the long term.

This chapter discusses the findings of the present study. In the first section, the three types of instruction will be discussed with reference to the original theoretical frameworks and models of language processing and acquisition on which they are based (section 5.1). Furthermore, an analysis of the results will be provided based on other theoretical models and/or frameworks in order to interpret the groups' performance in greater detail and link theory with practice based on the various theoretical models and frameworks presented in the literature review section (sections 5.2-5.5; see also Chapter 2). Findings of the study will also be discussed based on the theoretical models and frameworks in order to explore if findings can support claims regarding acquisition (section 5.5). The methodology applied in the present study will also be re-assessed in order to evaluate if there have been factors that might have affected the findings of the present study (section 5.6). Conclusions and overall contribution of the present study to IE and PI research will be discussed in section 5.7.

5.1 Overall group performance

The findings of the present study, based on mean score overall performance in the four testing tasks, show that the Combined teaching intervention was more effective both in the short and the long term, than the PI, IE and non-instruction. In turn, the findings show that the PI group is a more effective teaching intervention than IE and non-instruction (based on mean scores). The results from the present study also show that IE is more effective than non-instruction. Mean score performance reveals an apparent hierarchy, namely Combined>PI>IE>Control group, which is valid both in the short term, one week post-instruction, and in the long term, twelve to fourteen weeks post-instruction. Statistical analysis focusing on the overall performance of groups showed that there were significant differences ($p=.001$) both in the short and the long term for all four groups (Kruskal-Wallis test). However, a closer analysis of the dataset with paired group comparisons (Mann-Whitney test) showed that the Experimental groups significantly ($p=.001$) outperformed the Control group in all four tasks in the immediate and delayed post-test conditions. Paired comparisons between the experimental groups showed significant differences ($p \leq .007$) in the short term. However, in the long term significant differences ($p \leq .004$) were only reported for comparisons between the Combined and the IE groups. Therefore results from the present study support the claim that instruction matters, while the hierarchy indicates that type of instruction has a significant role in performance. Hence, findings partially support Hypotheses 1 and 2 while they confirm Hypothesis 3.

PI aims to force learners to process the form in order to decode the meaning of the sentence. IP Principles predict the cognitive steps that learners make when processing the input. In turn, PI aims to force learners to avoid the predicted cognitive steps by altering the processing strategies (VanPatten, 1996, 2004, 2007). Findings from the present study show that the PI treatment package has successfully achieved altering the predicted ‘First Noun Principle’ (FNP) strategy both in the short and the long term, as mean scores improved post-instruction in the immediate and delayed post-test conditions. Descriptive and statistical analysis with non-parametric (Kruskal-Wallis) and parametric (Repeated Measures ANOVA-RMA) tests confirmed Hypothesis 2, in terms of the PI group outperforming the IE group. It seems that the use of ‘full PI’ type of instruction with the inclusion of explicit rule explanation has succeeded in helping learners make initial form-meaning connections. Referential activities seem to have assisted learners in establishing correct form-meaning connections, whereas affective activities reinforced correct processing (see Lee and VanPatten, 1995, 2003; VanPatten, 1996, 2002, 2004, 2007, 2009). In line with previous PI studies, results show that the PI group were able to produce the target form in the fill in the gap task, despite the fact that participants of the group were only engaged in input based activities, as production of the target form was not part of the PI treatment package (Culman et al. 2009; Fernandez, 2008; VanPatten and Cadierno, 1993; VanPatten and Wong, 2004).

On the other hand, the findings of the present study indicate that IE has also been successful in making the target form salient to participants of the IE group. The external input manipulation has had a positive effect in making input salient internally and in getting further processed by the language learning mechanisms (Sharwood Smith, 1991, 1993). Mean score performance and paired group analysis showed that the IE was more effective than non-instruction confirming Hypothesis 3. Findings are in line with most

IE studies investigating the effectiveness of IE over non-instruction (Alanen, 1995; Doughty, 1991; Jourdenais et al. 1995; Leow, 1997, among many).

Results from the present study have also confirmed Hypothesis 1, which assumed that the Combined group would outperform both the PI and IE groups, since participants of the Combined group would benefit more from the integration of the two attention drawing interventions. As noted in the literature review section, both PI and IE approaches have failed to provide a detailed theoretical platform that can link empirical data with theory and thus provide a coherent explanation, regarding for instance the processes involved in making one or the other a more effective teaching intervention than other types of instruction and/or no instruction. To some extent, PI can provide an interpretation regarding why the PI group outperformed the IE group. It could be claimed that PI has, compared to IE, successfully predicted through Input Processing the cognitive steps that learners make when processing OVS sentences. Indeed it seems that with the aid of Structured Input (SI) activities, learners were forced to alter the predicted processing strategies and thus make correct form-meaning connections, confirming Hypothesis 2 (see also VanPatten, 1996, 2004, 2007, 2009). On the other hand, it could be argued that coloured IE has made salient the target form and may have induced attention to the target form. However, by comparison, IE has not been shown to be more effective than PI, as probably the external manipulation of the target form was probably not as well perceived by the language learning mechanisms as the one generated by PI, confirming Hypothesis 3. Similarly, it could be argued that IE in the Combined type of instruction has been more effective in inducing saliency than the IE type of instruction (see also Sharwood Smith, 1991, 1993). However, explaining why the Combined group outperformed the PI group (Hypothesis 1) remains problematic. More importantly, neither PI nor IE can provide hard explanations regarding groups' performance based on the cognitive processes involved during on-line processing and their impact, e.g. how and when learners have made form-meaning connections (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Sharwood Smith and Trenkic, 2001).

Many researchers define PI as one type of IE (Benati and Lee, 2007, Polio, 2007, Wong, 2004), while others also distinguish between explicit and implicit type(s) of L2 instruction (Lee and Huang, 2008; Norris and Ortega 2000); and others distinguish between obtrusive and less-obtrusive instructional methods (Doughty, 2003; Doughty and Williams, 1998; Housen et al. 2005; see also references in Sharwood Smith, 1991, 1993 to 'elaborate' vs. 'less elaborate' types of instruction). The aim of PI is to force learners to make (correct) form-meaning connections by paying attention to the form. In the present study, participants in the PI group were forced to pay attention to the case marking and word order in order to decode the meaning of the sentence. In this way, PI can be seen as an explicit and more obtrusive and/or elaborate type of L2 instruction than IE, since it contains explicit rule explanation and its purpose is to force learners to attend to the form. On the other hand, IE has been defined in this study as a less explicit and less obtrusive/elaborate form of instruction, when compared to PI, since its purpose is to make the input salient to the learner by typographically/textually enhancing the target form (Doughty, 2003; Doughty and Williams, 1998; Housen et al. 2005). However, it should be reminded that there is no guarantee how the learner and the learner's internal language mechanisms will perceive the enhancement and if salience will have an impact and/or to what extent (see Sharwood Smith, 1991, 1993).

The Combined instructional method seems to integrate the implicit with explicit types of both PI and IE. The fact that the Combined group received both explicit rule explanation and the combination of the attention drawing techniques used in PI and IE, e.g. forcing learners to pay attention to the target form and at the same time making the target form salient, may have resulted in the enhancement and/or reinforcement of processing of the target structure in a more explicit, obtrusive and elaborate way than the sole application of PI and/or IE.

Arguing that the reason why the Combined group outperformed the PI group, which in turn outperformed the IE group is because of the degree of explicitness, obtrusiveness and elaboration is in line with the hierarchy observed in metanalyses conducted on FonF vs. FonfS studies (see section 3.3.), claiming that explicit types of focus on form and focus on forms types of instruction are more effective than implicit types of instruction, i.e. FonF explicit>FonfS explicit> FonF implicit> FonfS implicit (Doughty, 2003:267; Lee and Huang, 2008; Norris and Ortega, 2000:465). Furthermore, Doughty and Williams (1998) have provided a taxonomy arguing that PI is, by comparison to IE, a more obtrusive type of instruction, resembling Sharwood Smith's (1991) proposal on more elaborated vs. less elaborated types of instruction. Based on the provided taxonomy and metanalyses also discussed previously (see section 3.3. for further details), it could be argued that PI can be a more effective teaching intervention than IE as it is more obtrusive, more elaborate and therefore a more explicit type of FonF than IE (see also Doughty, 2003, 2004; Doughty and Williams, 1998; Lee and Huang, 2008; Norris and Ortega, 2000).

Doughty (2004:263) argues that PI in its original conceptualisation entails both a FonF and FonfS type of processing. FonfS or metalinguistic processing is available through the EI and/or metalinguistic explanation component (e.g. Did you notice that the word order is not the same?). FonF processing is available through SI activities that make meaning available to the learner. In sections 2.2, 2.2.1 and 2.3, the fact that many IE studies have been primarily self-defined as FonF studies has been extensively discussed. In the metanalyses mentioned in the present study, IE studies have been considered and analysed as FonF studies, despite the fact that their classification as FonF studies is problematic (Norris and Ortega, 2000; Sharwood Smith and Trenkic, 2001; Sheen, 2005; Wong, 2003, 2005). Nonetheless, findings of meta-analyses show that explicit types of instruction are more effective than implicit types, in that they draw learner's attention to the target form and the target form being further processed by the learning mechanisms (Schmidt, 1990, 1993, 1995, 2001; Sharwood Smith, 1991; 1993; Tomlin and Villa, 1994; VanPatten, 1996, 2002, 2004, 2007, 2009 among many).

A common point that could partially explain the beneficial effects of PI and IE is the attention drawing factor. Attention has a facilitative role in language learning and acquisition both for PI and IE. However, its role and effect in the learning and acquisition process is not adequately explained, in terms of how exactly it operates in PI and/or IE, and how exactly it can be operationalised within each framework in order for a teaching intervention to become more effective (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Sharwood Smith and Trenkic, 2001). Despite the fact that both IE and PI approaches have been informed by Schmidt (1990, 1993, 1995, 2001) or Tomlin and Villa (1994), neither IE nor PI studies have explained results focusing on the role of attention and possible operationalising through one of the two approaches. Thus, the role of attention in IE and PI remains vague and not adequately explained. Stepping into the debate concerning whether attention and awareness form one entity, or are two different concepts is beyond the scope of the present study. However, as one of the aims of this study is to provide a theoretical explanation for the results, findings will be discussed at a theoretical level.

Raising awareness by intentionally focusing attention on specific elements of the input would be interpreted according to Schmidt (1990, 1993, 1995, 2001) as the only and sufficient condition for acquisition to take place. However, mean score performance and paired group comparisons show a tendency for gains to be maintained at a lower rate in the long term than the short-term.

Truscott (1998) suggests that the 'Noticing Hypothesis' should be limited to claims about the acquisition of metalinguistic knowledge (see also Schwartz, 1993), for which conscious noticing is important: 'the acquisition of metalinguistic knowledge is tied to (conscious) noticing; development of competence is not' (Truscott, 1998:124). Comparisons between the Combined and PI as well as the PI and IE groups show that in the long term, there are no significant differences between the groups ($p>.05$). If 'noticing' is equivalent to acquisition, then effects should be present in the delayed post-test. Results from the present study suggest that 'noticing' has an impact on the learning process, however it might not result in acquisition (Carroll, 2004, 2006; DeKeyser et al. 2002; Harrington, 2004; Sharwood Smith and Trenkic, 2002; Sharwood Smith and Truscott, in prep; Truscott, 1998). Comparing overall performance from the pre- to the post-test conditions it could be argued that participants of the three groups (IE, PI and Combined) have successfully 'noticed' the form. The fact that results show a decline in scores from the immediate to the delayed post-test condition cannot support that participants have acquired the target form. However, findings clearly point out that the target form has been processed. Looking at individual scores there is great variability, as the mean average shows a decline in performance but there are also cases where the scores are maintained at the same level in both conditions, or are higher from the immediate post-test. Therefore findings from the present study can lend support to the argument that participants of the three experimental groups have internalised the target form. However to what extent and/or whether they have acquired the target form

(Hypothesis 4) remains a highly debatable issue, which will be discussed in greater detail in the next section (see section 5.3). Given the results of the present study, it seems more plausible to claim that participants are in a learning stage, i.e. they are currently learning the target form. Perhaps, if this study was a longitudinal study and further evidence were provided claims could be more robust arguing whether participants have or not learned/acquired the target form. Hence, the present study's findings could potentially support Schmidt's weaker proposition on 'noticing', according to which attention has a facilitative effect on learning (Carroll, 2006; Hulstjin, 2005; Schmidt, 2001; VanPatten, 1996, 2009).

Tomlin and Villa (1994), contrary to Schmidt, dissociate attention from awareness and provide a definition of awareness and consciousness. Awareness is defined as 'a particular state of mind in which an individual has undergone a specific subjective experience of some cognitive content or external stimulus' (Tomlin and Villa, 1994:193), whereas consciousness has a richer definition, with multiple associations, such as perception, awareness, understanding, intention or specific knowledge (ibid). For Tomlin and Villa (1994: 194) 'awareness requires attention but attention does not require awareness' (see also Carroll, 2001, 2006, 2007; Sharwood Smith, 1991, 1993; VanPatten, 1996; among many). 'Noticing' in this sense would be for Tomlin and Villa (1994) synonymous with detecting information. Interpreting the present study's findings in the light of Tomlin and Villa's (1994) propositions, it could be argued that external input manipulation either through colour enhancement, or by forcing learners to process the form to obtain meaning, has had a positive impact in orienting attention to the target form with or without conscious awareness, without necessarily leading to acquisition of the target form. However, it could be argued that may be orienting attention has assisted participants in correctly mapping the target form to a meaning (as Doughty 1991, 1995, 2003; Sharwood Smith 1991, 1993; Sharwood Smith and Rutherford, 1987; VanPatten, 1993, 1996, 2002, 2004, 2008; amongst others have argued).

So far, the theoretical approaches discussed are able to provide interpretations in terms of why groups have improved post-instruction, based on factors such as the role of type of instruction, attention, explicitness and/or obtrusiveness and elaboration. We can conclude that both IE and PI have been successful in drawing learners' attention to the target linguistic form; that PI has been more effective than IE due to the fact that the PI teaching intervention was more explicit, obtrusive and/or elaborated than IE; or that PI contrary to IE forces learners to map form to meaning. We can conclude that attention does have a facilitative role in learning; however we cannot know whether conscious awareness and conscious processing of the target form have an impact on acquisition. More importantly, conclusions drawn remain at points vague, lacking a theoretical platform that could link PI and/or IE propositions with psycholinguistic constructs and thus provide a detailed interpretation of the findings of the present study. We cannot provide claims about how type of instruction has had an impact on the developing system; or how the information has been processed by the language learning system (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Doughty, 2004; Harrington, 2004; Sharwood Smith and Trenkic, 2001).

Furthermore, up to this point there has been no adequate explanation about why the Combined group has been more effective than the PI group, given the fact that the two groups varied only in the colour enhancement of the target form and no instruction or why the IE group is so different in performance from the Combined group, since both groups had colour enhancement in common. Responding to the issues raised will ultimately assist in tackling more difficult questions posed in the present study, referring to whether the target form has been acquired and in how far the type of instruction has had an impact in the acquisition of the target form.

5.2 Interpreting findings with the Competition Model (CM)

The Competition Model (Bates and MacWhinney, 1989) can provide a theoretical platform in order to partially explain how PI operates (see also VanPatten, 1996, 2004, 2007, 2009). As Carroll (2004:305) notes, an advantage of the CM is the precise predictions it can offer regarding which cues can “win-out” the competition in comparison to IP sub-principles of the ‘First Noun Principle’ (FNP), which are provided as alternatives to word order.

VanPatten’s (1993, 1996, 2004, 2007, 2009) proposal of the IP model encompasses the idea of cue validity and cue reliability, as proposed in the CM. His formulation of the FNP suggests that learners will process the first noun or pronoun they encounter as the subject/agent of the sentence is directly comparable to cue availability (VanPatten, 1993, 1996, 2004, 2007, 2009). In order for learners to overcome this psycholinguistic principle, VanPatten (1993, 1996, 2004, 2007, 2009) proposes PI, which will force learners to process the form, e.g. case marking and word order, in order to correctly decode the meaning of the sentence and correctly assign the subject/agent roles of the particular sentence. Therefore it is valid to assume that, in this case, PI has successfully enhanced cue reliability. Moreover, conflict validity has also been operationalized through Structured Input (SI) activities, where learners have been presented with two sets of sentences, one containing the target form and another sentence containing contrasting forms of the target form. For instance, for the present study, the SI activities entailed paradigms containing pairs of SVO versus OVS sentences. Therefore findings from PI studies can be attributed to competition and the strength of cues, explaining why PI learners are performing better than learners that received other types of L2 instruction, such as IE, and/or no instruction.

In English, the agent role of a lexical item is indicated by the word order, placing the agent in the preverbal position of an utterance. In German, the word order is not always reliable, since German allows both Subject-Verb-Object (SVO) and Object-Verb-Subject (OVS) sentences and therefore case marking should be used in order to indicate the correct functional relation of lexical items. Going back to the psycholinguistic IP Principles, two sub-principles of the FNP have been considered in the design of the present study. The 'Lexical Preference Principle' (sub-principle 2a), which implies that learners will process redundant lexical items instead of grammatical forms in order to decode the meaning of the sentence; and the 'The Event Probabilities Principle' (sub-principle 2b), which predicts that, where possible, learners may rely on event probabilities instead of word order to interpret sentences (VanPatten, 1996, 2004). Considering the two sub-principles, lexical items and cue animacy (that could be utilized as cues in decoding the meaning instead of processing the target form) were excluded from treatment materials for all three experimental groups.

Findings of the study showed that the PI treatment materials modified the relative weight of cues through SI activities. Cue validity and cue reliability were successfully enhanced, as mean score performance substantially improved post-treatment, while gains were maintained also 12-14 weeks post-instruction. It could therefore be argued that based on propositions of the CM, PI enhanced the validity and reliability cues regarding case marking and diminished the validity and reliability cues for word order, event probabilities and lexical items. This means that participants improved in their ability to rely on case marking instead of the unreliable German preverbal position cue. Thus, they were able to correctly decode the meaning of the sentence and make correct form-meaning connections. In this process, conflict validity has played an important role. The fact that, in the referential activities, participants were always presented with an SVO and an OVS sentence and a picture and were required to decide which sentence corresponded best to the picture and vice versa has further promoted their ability to resolve such conflicts. Thus, with the help of frequency, participants of the PI group learned to assign correct roles and make correct mappings.

On the other hand, both the CM and PI have been criticized for the fact that learners are engaged at sentence-level processing with simple sentences not reflecting actual communicative situations. With this in mind it has, however, been argued that this type of processing/language learning encourages learners to develop a ‘particular’ strategy to obtain meaning (DeKeyser et al. 2002; Gibson, 1992; Harrington, 2004; McLaughlin and Harrington, 1989). To what extent the treatment packages and testing materials promoted task-taking strategies to obtain meaning remains an unresolved issue in the present study. Further research is therefore necessary in order to explore PI/and or Combined treatment effects in a communicative setting. Moreover, spontaneous oral data could provide an insight into whether and to what extent PI promotes only task-taking strategies, as they can tap into implicit knowledge. Nonetheless, in the present study, the fact that in the immediate post-test, scores are substantially high in all three groups (though higher for the Combined and the PI groups) can provide support for such claims. On the other hand, scores are maintained at a lower rate in the delayed post-test suggesting that even if learners are trained to develop a task taking strategy, effects can be seen in the long term. The difficulty is to determine whether these effects can be attributed to language acquisition.

In the field of IE, there is no reference to the CM and/or how it can be operationalized through IE. However, one can assume that enhancing the target form could have an impact in enhancing cue strength. Using color to typographically enhance the target form, salience of the form could result in enhancing the cue availability, in this case the OVS word order in German. The application of different colors to typographically enhance case marking and thus make the difference between an SVO and OVS sentence salient can promote cue reliability and/or conflict validity, depending on the cue strengths of learners.

Despite the idea that the CM could serve as a theoretical platform for both IP and PI, neither VanPatten nor PI studies have attempted to link findings of PI studies with the theoretical underpinnings of the CM as expressed through cues and mathematical computations. VanPatten (1996, 2004, 2007, 2009) does acknowledge the fact that the CM can serve as a theoretical platform for the First Noun Principle in his model; however he argues that the CM fails to provide information regarding the formation of cues at the initial state and more specifically ‘how the cues are initially detected and made available for those mechanisms that compute cue validity for the developing system’ (VanPatten, 1996:52).In contrast to PI and IP, in the original writings of IE, there is no direct or indirect reference to competition and how it could be operationalized within IE (Sharwood Smith, 1991, 1993). The methodology applied in studies investigating the CM involves an agent-identification task. Participants are presented with simple transitive sentences consisting of two nouns and a verb (e.g. the man calls the woman) and are asked to decide which noun refers to the agent of the sentence, while sentences contain converging and competing combinations of cues (e.g., Bates, Devescovi and D’Amico, 1999; Berger, et al. 1996; Devescovi, D’Amico and Gentile, 1999). Given the task design followed in the present study for the IE package, materials mostly included reading comprehension and a few true/false tasks at sentence level. Thus, the design materials of the IE group could not support the methodology applied when investigating groups’ performance through the CM. However, I have attempted here to illustrate how the CM could be operationalized in IE, based on the fact that later writings proposed competition as a process aiding in language processing and acquisition (see Sharwood Smith and Truscott, 2004, 2005 and in prep; Truscott and Sharwood Smith, 2004, 2005).

The CM can therefore provide a limited theoretical explanation for PI and IE instruction, as discussed in the previous paragraph. In both cases, prominent constructs underpinning PI and IE, such as the role of attention, cannot be explained with the CM. On the other hand, it should be acknowledged that the CM is not meant to be a model of attention and/or detection (see also VanPatten, 1996). More importantly, the CM does not explain the functional architecture of the language faculty and therefore cannot provide information about the initial state of language processing and the mechanisms involved in converting input into intake leading to restructuring and acquisition (Carroll, 2004; Sharwood Smith and Truscott, in prep; VanPatten, 1996, 2004, 2007, 2009).

5.3 Can findings of the present study be interpreted with AIT?

Carroll (2004:297) has criticised the fact that the IP model lacks a theoretical platform, which should include ‘a theory of perception and parsing’, in order to be able to interpret and evaluate PI studies’ findings and IP propositions (see also Collentine, 2004; DeKeyser et al. 2002; Doughty, 2004; Harrington, 2004). She argues that the ‘Autonomous Induction Theory’ (AIT), which is based on ‘acquisition being failure driven’ can provide such a theoretical account (Carroll, 2004).

For Carroll (19991, 2004, 2006, 2007), grammatical knowledge is viewed in the Autonomous Induction Theory (AIT) as formal features, which can be combined (unified) in order to form formal categories such as nouns, and in turn can be unified to form hierarchical structures, such as Noun Phrases. These features are available to the Language Acquisition Device (LAD), whose role is to build structures and change them, through UG. In the case of gender, the assortment of features for a specific language is unique. For instance, English has no grammatical gender distinction in its nouns. German nouns on the contrary, have features such as masculine, feminine and neuter gender, i.e. ‘der Tisch’ (the table), where the determiner *der* specifies that the noun has a MASCULINE and NOMINATIVE feature. Thus, German modifiers act as cues, i.e. der schwarze Tisch (the black table) indicating gender class and agreement. Therefore for AIT, an English learner of German must acquire the gender system of German unconsciously. LAD is a psychological mechanism operating implicitly, and is constrained by modular processing. The role of LAD in this case is to discover the properties of German nouns, if there is ‘parsing failure’. At the same time UG simplifies LAD’s discovery process through the available linguistic primitives (Carroll, 1999, 2001, 2007:157).

Restructuring of grammar is therefore argued to be taking place in the LAD and occurring unconsciously and outside of a learner’s control (Carroll, 2007:155). LAD is triggered only when there is a parsing failure, which can only occur during comprehension (Carroll, 2007). For the present study, this means that explicit instruction and comprehension practice can lead to changes in the mental grammar, only if external manipulation of the target form leads to changes in the processing of primary linguistic data, in other words to parsing failure (see Carroll, 2001, 2004, 2007). According to Carroll (2004, 2007:170), PI is a type of instruction that leads to parsing failure. Similarly, it could be argued that the Combined type of instruction, differing only from PI in the coloured IE, could also lead to parsing failure. However, in the case of the IE group, treatment mainly focused participants’ attention towards general understanding of the information provided in the reading comprehension texts, while the target form was manipulated with colours to induce saliency. Participants were not therefore required or forced to process the target form further and therefore one could assume that processing in the IE group was not based on parsing failure.

Key to Carroll’s account is the role of L1 transfer during L2 comprehension (see also VanPatten, 2009). In the present study learners were English learners of German and the target form was case marking and word order. This means that in this case L1 transfer could not assist processing, since English has a strict SVO word order and no gender features as German has. Therefore as Carroll (2001) also acknowledges the information contained in the input cannot be parsed as it cannot be detected in the first place, i.e. no parsing failure can occur simply because there is no L1 transfer alerting the parser that there is a problem (see also VanPatten, 2009). However, she argues that grammatical metalinguistic instruction, feedback and/or instruction may have variable outcomes, leading also to restructuring. According to Selinker et al. (2004:85), ‘such restructuring may not be maintained in the interlanguage, which again brings up the issue of existence vs. persistence as fossilization[54]' (see also Lardiere, 2004). Can we therefore assume that participants in all three experimental groups experienced parsing failure? Based on the information obtained from the questionnaire, interviews and the results of the pre-test there was no prior knowledge of the target form, nor instruction or exposure to the best of my knowledge during the experimental period. Fossilization implies that learners may continue to make progress in certain areas, and yet return again and again to the same error. For Carroll (2001:169) fossilization occurs 'when the organism fails to detect errors'. But what if participants of the present study are not able to detect the errors in the first place, given the limitations of the present study as explained above? Again as Selinker et al. (2004) and Lardiere (2004) note there is certainly an issue of existence vs. persistence that remains unresolved for the present study when findings are interpreted solely based on AIT.

VanPatten (2009) does not explicitly adopt Carroll's (2001) perspectives on parsing and acquisition though he does not dismiss them. He argues that IP Principles can be directly linked to Carroll's (2001) claims that 'acquisition being failure driven'. According to VanPatten (2009:53) there are differences between Carroll and his approach in the factors leading to 'parsing failure'. In the case of the FNP, Carroll assumes that the problem resides in the L1 parser that expects an SVO sentence and through interaction with language the parsing problem surfaces and thus processing proceeds accordingly (see section 2.4.4.1). For VanPatten the problem resides in the universal strategy of expecting the first (pro)noun to be the subject of the sentence. IP Principles assist learners in overcoming the parsing problems they encounter, since learners realise that incoming information does not match their beliefs. Through explicit instruction the parsing principle that needs to be altered is identified, e.g. in English every (pro)noun is the subject of the sentence, whereas in German the first (pro)noun maybe the subject or the object of the sentence. In addition, the reminders of the processing principle, the feedback learners receive, as well as the structured input activities during full PI instruction alert the parser that there is a mismatch and a

[54] Selinker (1972:215) argues that: *'Fossilizable linguistic phenomena are linguistic items, rules and subsystems which speakers of a particular NL will tend to keep in their IL relative to a particular TL, no matter what the age of the learner or amount of explanation and instruction he receives in the TL'* (NL - Native Language; IL - Interlanguage; TL - Target Language).

problem that needs to be resolved (VanPatten, 2009). For Carroll a good deal of this process happens incidentally and therefore conscious processing of grammar (except with PI) has little to offer in the development of mental representations of language. With this in mind it could be argued that both approaches are compatible (see Carroll, 2004; VanPatten, 2009). However, for VanPatten (2009) the difference between PI that forces parsing failure and AIT that induces parsing failure is the reason why PI can help learners establish correct-form meaning connections whereas AIT cannot.

The present study has followed the full PI approach and has also provided reminders for the PI and the Combined groups. However for the latter, coloured IE was also included. In order to control for the input exposure variable, the IE group also received similar reminders (see section 3.4.2.1)**.** Interpreting findings through AIT, one should conclude that the three instructional types were effective because the external manipulation they have used was successful in leading to changes in the processing of primary linguistic data, i.e. to parsing failure during comprehension. However, AIT cannot provide sufficient explanation regarding the obtained hierarchy, namely Combined>PI>IE>Control group. Is it valid to assume that the hierarchy is the outcome of a teaching intervention being more successful in creating more instances of parsing failure processing? One could conclude that the Combined group promotes more instances of parsing failure processing during comprehension and therefore is more effective than PI and IE. In turn, PI outperforms the IE group because it induces more instances of parsing failure processing during comprehension. However, what are the qualitative differences and how can we adequately explain these differences at a theoretical level using AIT in order to claim that one teaching intervention is more effective in promoting parsing failure than another; especially when AIT argues that instruction has limited effects on SLA?

IE and PI (and as a result the Combined intervention) are attention-drawing types of instruction. Especially in the case of the PI and/or Combined groups it could be further argued that attention is oriented and participants are forced to process the form that they were primarily forced to attend to. Therefore the role of ‘noticing’ with or without conscious awareness is an important factor that needs to be addressed when comparing the particular instructional techniques in terms of effectiveness of L2 instruction.

Carroll (2007:164) views noticing as result of processing, where she assumes that noticing is synonymous with conscious awareness and is a 'by-product of the processing of phonological representations' following Jackendoff's (1987, 1997, 2002, 2007) definition.

Despite the fact that Carroll's (1999, 2001, 2002, 2007) AIT is a promising candidate in providing a theoretical platform that could potentially give a more detailed interpretation of language development, it seems that the constructs of AIT can be partially compatible in interpreting the findings of the present study. The main problem that surfaces for the present study is that it seems that acquisition and fossilization to some extent coincide, i.e. they both stem from the fact that the 'organism fails to detect errors'. This does not mean that for Carroll the two terms coincide but given the facts of the present study and as described in this section and in Chapter 3, there can be no clear distinction between the two terms for the present study.

5.4 Can findings of the present study be interpreted with MOGUL?

Modular Online Growth and Use of Language (MOGUL), through Perceptual Output Structures (POpS), provides perhaps a more detailed answer to the following questions. Why and how did the three groups differ in degrees of explicitness and elaboration? What is the role of consciousness? How far can conscious processing affect language processing? What is the role of activation levels and competition in language acquisition when only written input is available?

According to MOGUL, information entering through our sensory system, e.g. through POpS, can serve as the basis for higher level processing due to the rich interconnectivity of POpS, which give rise to conscious experience (Sharwood Smith and Truscott in prep; a detailed presentation and review is provided in section 2.4.4.2. and 2.4.4.2.1). In the present study, all four groups received written input only and therefore information entering the perceptual system was available through the visual system, part of POpS. The three experimental groups did not vary quantitatively in the amount of input received. However, they did vary qualitatively, in terms of the way the input was manipulated in order to draw learners' attention and stimulate processing of the target form in the language learning system.

The IE group received input through POpS, in terms of the activities being available through written input, as well as from the coloured enhancement of the target form in the same way as in the Combined group. The fact that both the Combined and the PI groups outperformed the IE group in the two conditions indicates an effect of type of instruction. Sharwood Smith (1991, 1993) has argued that manipulating the form externally does not guarantee that the form will also be enhanced internally, as we cannot know how the internal language mechanisms will perceive the manipulated input (see also Tomlin and Villa, 1994; Sharwood Smith and Truscott, in prep). In other words, external manipulation of the target form to make it salient does not guarantee internal saliency. The treatment package for the IE group included explicit rule explanation along with coloured enhancement of the target form accessible only at the beginning of each session, as for the PI and the Combined groups. Treatment activities mainly involved reading comprehension and true/false tasks that focused participants' attention towards general understanding of the information provided in the text, while the target form was manipulated with colours to induce saliency. Hence, participants were not required or forced, as under PI, to process the target form further in order to assign meaning (CS), to refer back to the explicit instruction or process the form in the language module (SS) in order to assign a syntactical role/category.

Exploring processing of the target structure in the IE group using MOGUL, it could be assumed that perceptual structures entering through the visual structure stimulated activation in the visual, syntactical and conceptual structures (VS, SS and CS). However, much of the received incoming information was not able to be matched and coindexed to a specific representation. At this point it should be noted that the treatment activities in the IE group required general comprehension of a reading text, and participants were provided with hard copies containing assumed unknown words. In contrast to the other two treatment packages, activities in this package were not at sentence level but involved a ten-line reading comprehension text. In this task, participants were not forced to process the form in order to establish morphosyntactical properties of the enhanced target form. They could easily complete the task by relying on the vocabulary. In addition, all OVS target sentences that were included in the text were placed at the beginning of the sentence and therefore it could be assumed that participants were processing the OVS sentences as SVO, supporting the 'First Noun Principle' prediction.

The few true/false tasks included in the IE treatment package involved a picture and a sentence requiring participants to decide if the picture corresponded to the meaning expressed in the sentence, where the target form was highlighted with colour[55]. It could be hypothesised that in this true/false task participants were pushed to match form with meaning in order to complete the task. The target structure should have generated stimulation and activation at the syntactical and conceptual structures, as in the case of the PI and the Combined groups, in order to match with to meaning (unconscious processing). In addition, it should also have stimulated POpS, as described for the Combined group, i.e. process and match colour with syntactical and conceptual structures (conscious metalinguistic processing). With the true/false tasks, participants of the IE group could have been able to match the colours with the relevant cases at the syntactical and conceptual structures. However, they were not able to establish a strong chain that represents the use and nature of the target form, i.e. a VS-SS-CS structure like the other two groups, because the majority of tasks being reading comprehension tasks did not promote this type of processing of the target form. Figure 5.1 next page illustrates processing of the target structure with MOGUL for the IE group by isolating a sentence that could be incorporated either in the reading comprehension text or the true or false task. The question marks depicted in the syntactical structure (SS) and in the conceptual structure (CS) signal the difficulty of matching and coindexing the target structure with the appropriate colour, according to the coding.

[55] The reader is also reminded that participants in all four groups did not receive any type of feedback regarding their answers and/or scoring.

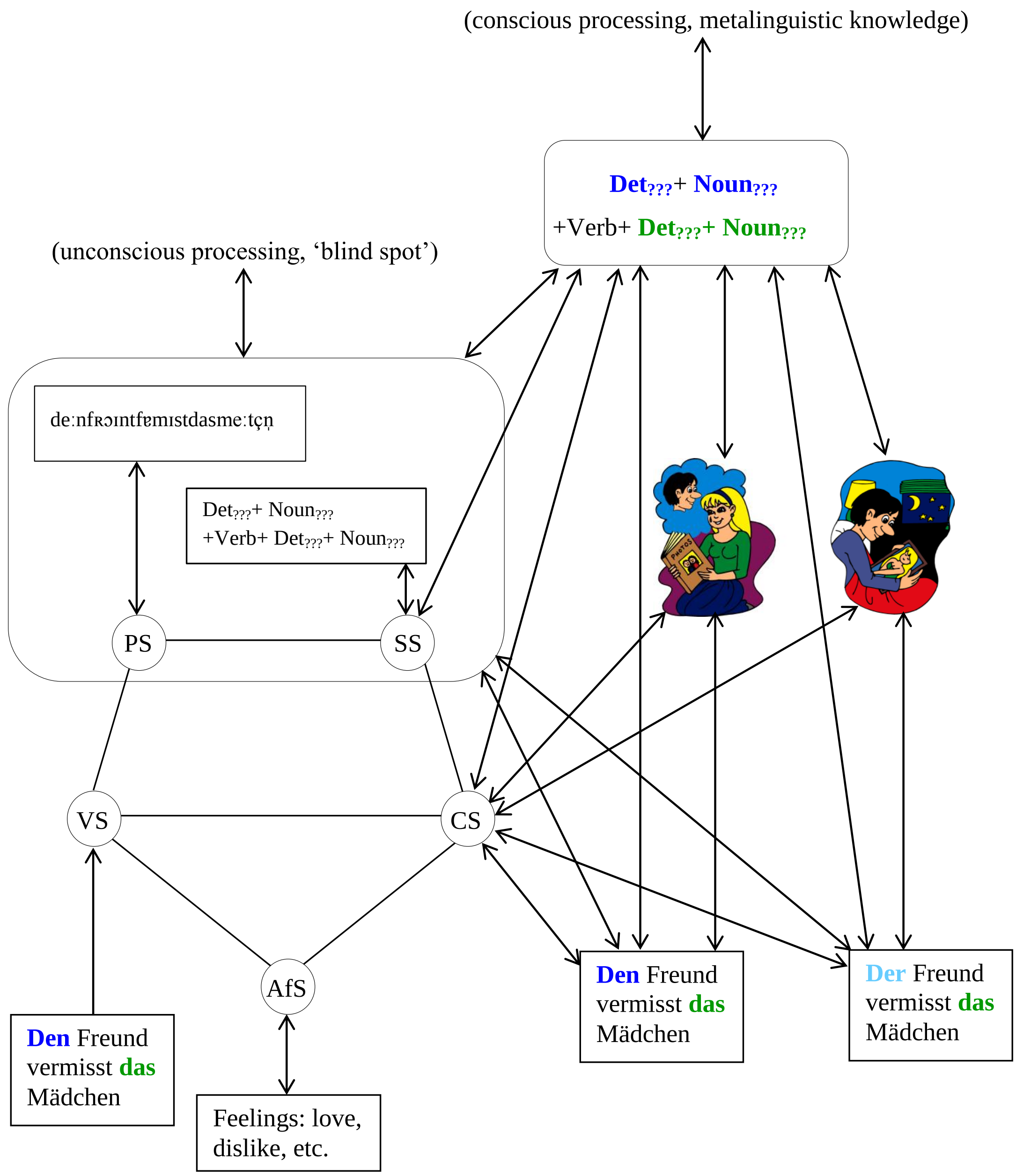


Figure 5.1: Processing of the target structure for the IE group with MOGUL

The fact that coindexing might not have been possible in all of the tasks included in the IE treatment package, as depicted in Figure 5.1, can provide various interpretations regarding processing in the IE group: a) participants were processing the first determiner and noun as agent and the second determiner and noun they encountered as patient (incorrect form meaning connections); b) participants were processing the first determiner and noun in the accusative case and the second determiner and noun they encountered in the nominative case (correct form meaning connections); and c) participants were confused and uncertain and assigned cases purely based on chance. Hence, the two available pictures in the conceptual structure are provided in order to describe the confusion that learners might have experienced during processing of the target structure in the IE group. Clearly, some learners processed an SVO sentence (the friend misses the girl), while others an OVS sentence (the girl misses the friend). There might also have been instances where learners were uncertain which one of the possible options was the correct interpretation. However, it should be noted that for the IE group, this type of processing was minimally possible and only in the few true or false tasks. The majority of the tasks included in the treatment comprised reading comprehension tasks that included the enhanced target forms but required general understanding of the text.

Regarding the Combined group outperforming the PI group, the two groups only differed in the amount of POpS available for processing. Both groups received a) the same amount of explicit instruction; b) the same Structured Input (SI) activities at a sentence level; and c) the same set of pictures. Thus, at this level groups were exactly the same in terms of exposure to input. Both groups received information through the visual structure via written input, i.e. sentences and pictures. Information entering through the visual processor was stimulating and at the same time was trying to match information from the language module, concentrating on the syntactical memory processor (as there was no oral input available) in order to formulate a conceptual structure. In IP and PI terms, this process would be translated into processing the form in order to decode the meaning of the sentence/picture. Successful processing in terms of participants' correct responses can be interpreted as the making of correct form-meaning connections for IP and PI. For MOGUL, successful processing is the build up of a chain of structures consisting of VS-PS-SS-CS[56]. At this point processing is common for both the PI and the Combined groups, focusing mainly on the 'blind spot', i.e. the syntactical structure in order to stimulate and match activation at the conceptual structure to decode meaning. Processing at this stage is considered to take place without conscious awareness (Carroll, 1999, 2001, 2007; Jackendoff, 1987, 1997, 2002, 2007; Sharwood Smith and Truscott, 2004 and in prep). The following figure (Figure 5.2)

[56] Visual Structure (VS) –Phonological Structure (PS) - Syntactical Structure (SS) - Conceptual Structure (CS) (for further details please refer to the literature review chapter).

illustrates the processing of the target form (a German OVS sentence, *Den Freund vermisst das Mädchen* ‘the girl misses the friend’) as described in this paragraph using an example[57] from the present study:

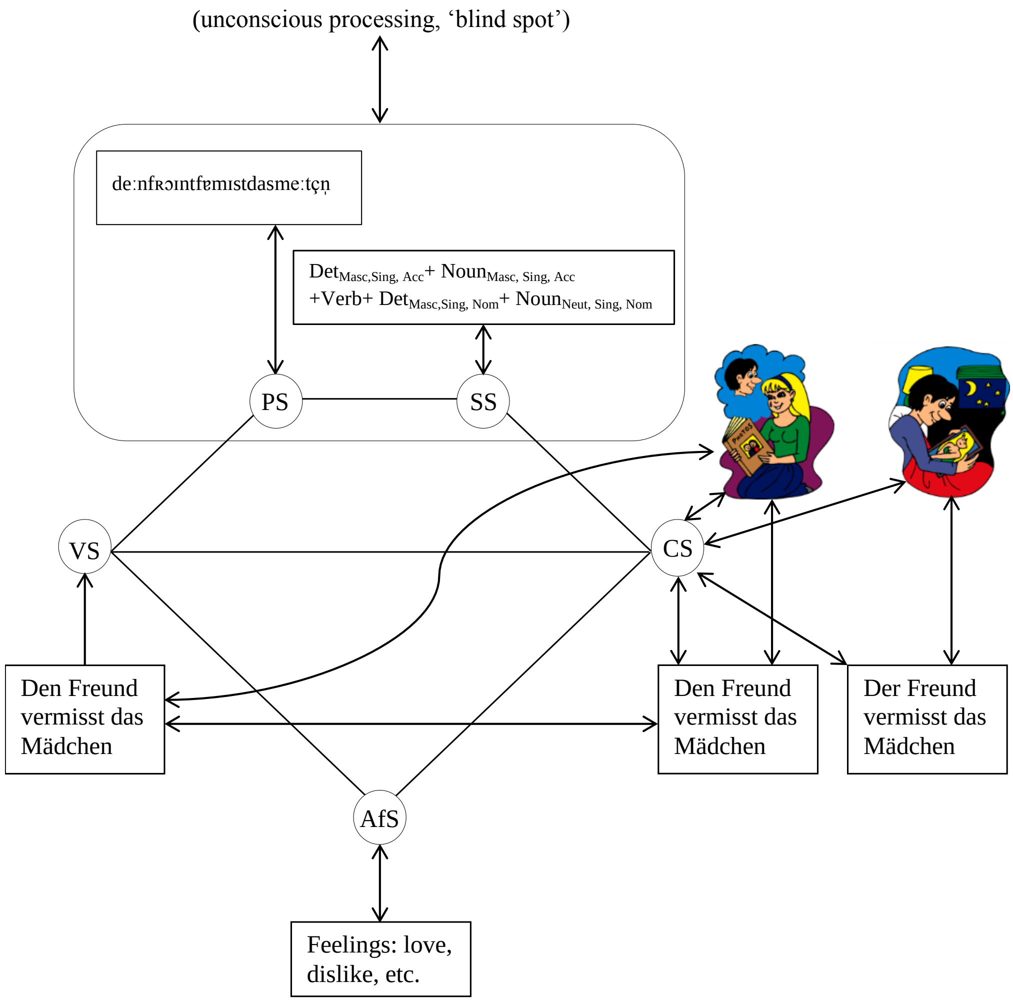


Figure 5.2: Processing of the target structure for the PI group with MOGUL

Processing depicted in Figure 5.2 above presents the way processing in the PI group could be described with MOGUL. The example depicted above partially applies to the Combined group, as the difference between the two groups lies in the coloured IE of the target form for the Combined group. It could be hypothesised that through the coloured IE, the Combined group received a greater stimulation of POpS and therefore an increase in the activation levels of conscious awareness of the processing of the

[57] The Affective structure (AfS) incorporated in the illustration does not directly apply to the aims and purposes of the present study, however it is provided, as according to MOGUL, it can be part of processing of the target structure.

structure. The written input entered the sensory system through the visual structure, while stimulation and activation generated in the language module and the conceptual structure began the process of 'indexing' (i.e. match perceptual structures and/or create new nodes), in the same way as in the PI group. In contrast to the PI group, coloured enhancement of the target form could have generated a parallel activation of another set of POpS processing, which focused on decoding colour from form, as well as matching form to one of the few possibly activated structures in the corresponding language module, conceptual and/or visual structures. In other words, for the second set of POpS, information that entered through the visual structure was stimulating activation in the syntactical memory store in order to match the colour to a syntactical structure (i.e. assign the nominative for light blue, the accusative case for the masculine nouns for dark blue, while pink was used for feminine and green for neuter nouns) as well as activation to the conceptual structure in order to assign meaning. The type of processing described for the Combined group, generates greater POpS interconnectivity than the POpS interconnectivity that was provided for the PI group, as Figure 5.3 next page depicts:

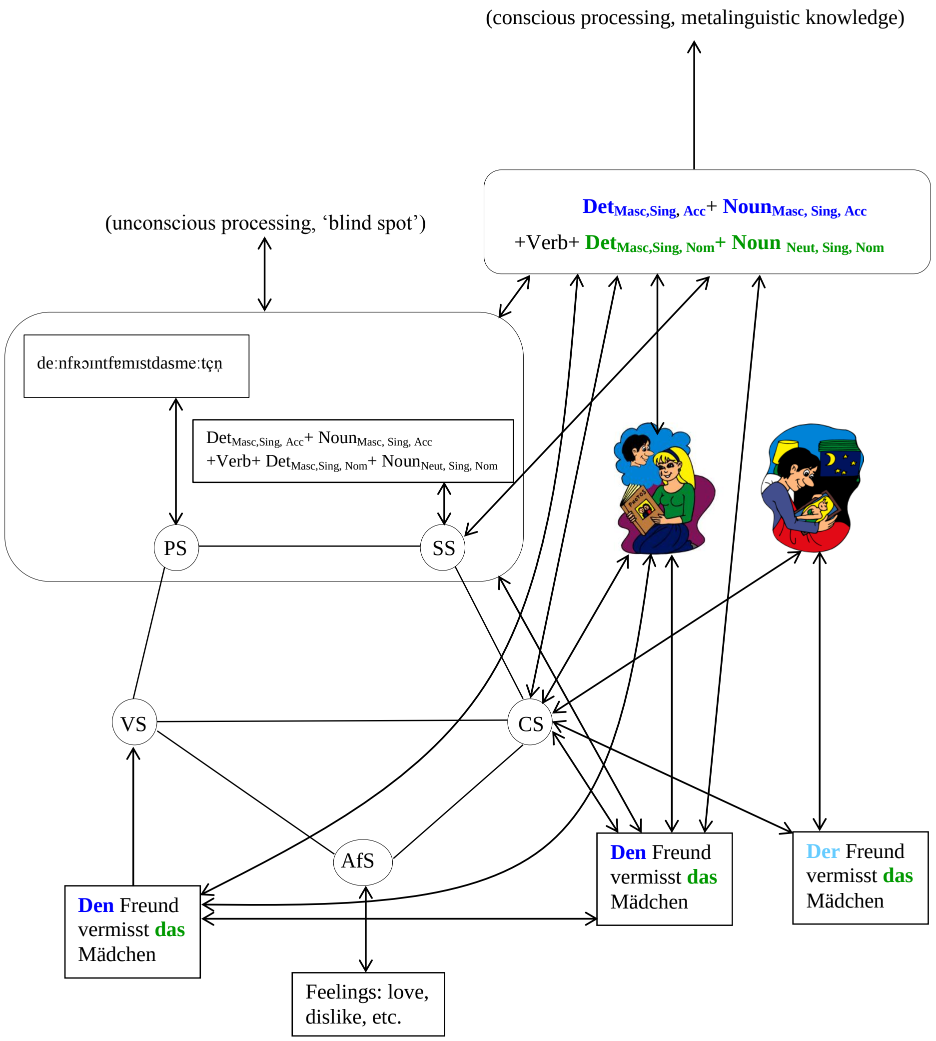


Figure 5.3: Processing of the target structure for the Combined group with MOGUL

According to MOGUL and as depicted by Figure 5.2 and Figure 5.3, the PI group participants had higher levels of activation in their syntactical and conceptual structures, since the focus of this type of instruction was both to force learners attend to the target form and process the form to make correct form-meaning connections. Through the visual structure, POpS enabled the incoming linguistic information for further processing. SI activities included POpS, in terms of written input containing picture matching with linguistic information, e.g. sentences containing the target form requesting participants to decide if the meaning expressed in the picture corresponded to the target structure and vice versa. Processing in these types of activities could therefore be argued to also occur unconsciously in the language module despite incoming information from the visual structure, i.e. via pictures and/or written sentences.

MOGUL argues that the rich interconnectivity of POpS generates metalinguistic knowledge and therefore conscious processing (Sharwood Smith and Truscott, in prep: 18; see also Baars, 1987, 2001; Carroll, 1999, 2001, 2004, 2007; Schmidt, 2001). In the case of the PI group the benefits of POpS are not as important, when compared to the benefits that available POpS structures provide for the Combined group; whereas at the same time POpS are more robust when compared to the IE group (supporting Hypotheses, 1, 2 and 3).

The obtained hierarchy Combined>PI>IE>Control group provided from mean score performance and statistical paired group comparisons can be explained with MOGUL through successful indexing, matching (coindexing) and high activation levels of the target structure. The differences in performance between groups indicate that the Combined instructional technique has been more successful first of all in creating an index for the target structure in the lexicon. For example, an index is created in the syntactical memory store for the perceptual structure *Den Freund vermisst das Mädchen* ‘the girl misses the friend’ the index DETERMINER, MASCULINE, ACCUSATIVE, SINGULAR + NOUN, MASCULINE, ACCUSATIVE, SINGULAR + VERB, SINGULAR + DETERMINER, NEUTER, NOMINATIVE, SINGULAR [58]. As Figure 5.4 next page depicts:

[58] At the initial stage of processing an index will be created for each lexical item while putting together the information to formulate the sentence can be considered as coindexing. For ease of explanation I have summarised the process.

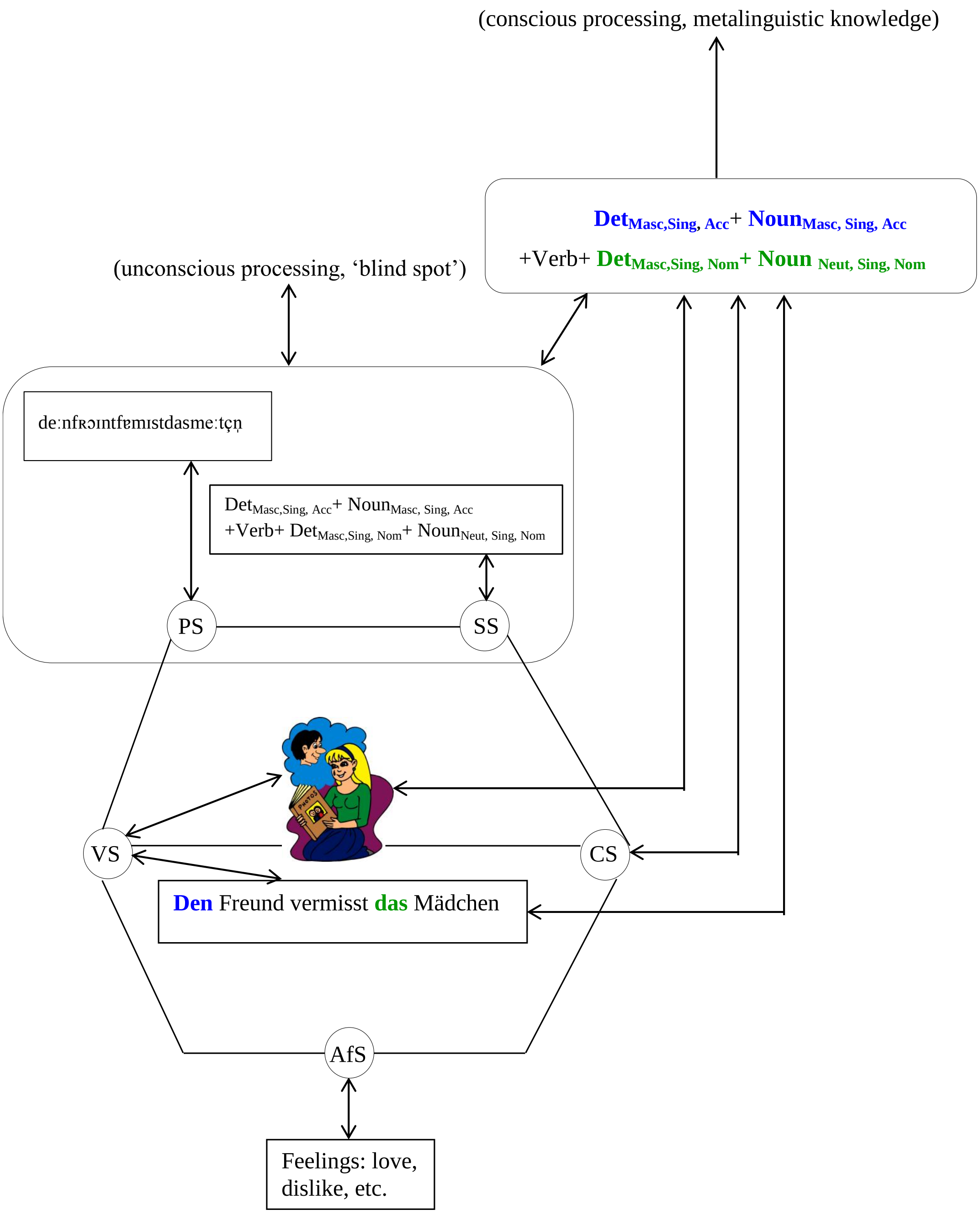


Figure 5.4: Coindexing perceptual structures in MOGUL

In turn, the Combined teaching intervention has been effective in stimulating processing of the structure, raising its current activation levels, as well as in increasing the structure's resting activation levels more frequently. In other words, the combined type of intervention had increased processing of the target structure, in terms of triggering processing, winning the competition over other competing structures as well as in maintaining the high levels of activation, which in turn increased the chances of the particular structure to be processed more frequently in the future. It could also be argued that the Combined teaching intervention with the integration of PI and coloured IE has achieved greater POpS interconnectivity and has brought the target structure to higher levels of conscious awareness during on-line processing. In this way matching of perceptual structures with syntactical and conceptual memory structures (coindexing) achieved through stimulation has been more effective due to frequency of activation, resulting in the making of correct form-meaning connections.

The PI group differed from the Combined group only in the variable coloured IE. The fact that, through SI activities, participants were forced to process the form in order to decode the meaning of the sentence shows that POpS stimulated activation in the syntactical and conceptual structures. Processing of the incoming information did not involve the same type of processing as for the Combined group. Indeed, the Combined group had to process a second set of POpS in order to assign colours to syntactical and conceptual properties. Compared to the Combined group, POpS in the PI group did not seem to achieve as rich interconnectivity as in the Combined group and therefore may not have reached as high levels of conscious processing as the Combined group. However, current and resting activation levels of the target structure seemed to be substantially higher than those of the IE group's, as mean scores and statistical analysis suggest.

Compared to the IE group, the PI group was shown to be more effective indicating that at early stages of language learning, form-focused processing at a sentence level (as proposed in PI) is more efficient than reading comprehension tasks involving general understanding of a written text, as applied in IE. Evidently, PI has been more effective in stimulating but more importantly in this case in matching (co-indexing) incoming information with syntactical and conceptual structures. At beginner levels of language learning, it seems that focusing learners' attention on specific aspects of the input while eliminating other factors that can potentially distract learners from the linguistic properties of the target form, and/or deprive learners of essential information that can further assist the process of co-indexing and seems to be more effective than less focused teaching interventions requiring general understanding of the input, which can be assisted either through context and/or vocabulary. On the other hand, the latter is more effective than non-instruction based on the findings of the present study. These results advocate the importance of some sort of instruction in second language learning and acquisition.

It should be noted that paired group comparisons showed that the Combined and the PI as well as the PI and IE groups were not significantly different in the long term. In the next section (section 5.5) there will be extensive analysis explaining the reasons why as well as their impact on the hypotheses of the present study.

Despite the fact that MOGUL offers a theoretical platform that can explain the role of IE, PI and the combination of the two approaches, the processes involved when applying these types of instruction and their impact in language processing and learning, as well as explaining the obtained hierarchy based on the mean score performance in the present study, it should be acknowledged that it is a new account of language processing, as is the present study, and therefore, further empirical evidence is necessary. However, compared to the other models and frameworks that have been mentioned in this section, it should be noted that MOGUL seems to offer a clear interpretation and distinction of conscious and subconscious processing, whereas in the other models/frameworks the explanations about the two processes remain vague and/or not as explicit.

In this section the overall performance of the experimental groups has been discussed. Findings of the present study confirmed the original Hypotheses 1, 2 and 3 and hierarchy predicted for types of instruction. Findings were primarily discussed based on the theoretical models originally informing the three types of instruction applied in the study, as well as the psycholinguistic constructs underpinning namely IP, PI and IE. The Competition Model (CM), the Autonomous Induction Theory (AIT) and Modular Online Growth and Use of Language (MOGUL) have provided theoretical platforms for interpreting and evaluating the effectiveness of the teaching interventions, as well as explaining why one type of instruction has been more effective than the other. MOGUL seems to provide the most detailed explanation for the present study’s findings thanks to its interdisciplinary and modular approach to language development.

In the next section, the findings of the present study will be discussed in order to assess whether the target form has been acquired or simply learned.

5.5 Can we claim acquisition of the target form?

Findings of the present study provide robust evidence confirming that instruction does play a role in changing behaviour, whereas types of instruction differ in degree of effectiveness; however, it remains an open question whether it results in learned knowledge or linguistic competence/acquisition (see Schwartz, 1993). All four groups were significantly different ($p=.001$) in the immediate and delayed post- test. A more detailed statistical analysis of results comparing two sets of groups showed that groups were significantly different ($p\leq.007$) in the immediate post-test. However, in the delayed post-test significant differences ($p\leq.05$) are reported only for comparisons between the Combined and the IE, the Combined and the Control, the PI and the Control, and the IE and the Control groups. No significant differences have been found from paired comparisons between the Combined and the PI groups and between the PI and the IE groups.

According to the study design, the immediate post-test was administered immediately after the last treatment, measuring short-term effects. The delayed post-test was administered 12-14 weeks post-instruction measuring long-term effects. Based on the present study's findings, Hypothesis 4 regarding short and long term effectiveness of the three teaching interventions is partially confirmed.

The majority of PI and IE studies have adopted a different study design, measuring short-term effects one week post-instruction; while for long term effects one month (in some cases three weeks) post-instruction (Alanen, 1995; Lee and Benati, 2007; Leeman et al. 1995; VanPatten and Cadierno, 1993; VanPatten and Oikennon, 1996; among the majority of PI and IE studies; see also sections 2.3.3; 2.4.3; 3.3). Apart from VanPatten and Fernandez (2008) and White (1998), who have administered delayed post-tests eight months post-instruction, the remaining studies' claims regarding acquisition of the target form are based in too short elapsed time (considering treatment and delayed post-test administration being maximum up to one month post instruction). Thus, minimizing validity and reliability of claims regarding acquisition (see Mitchell and Myles, 2004).

VanPatten (2009:59) defines acquisition under IP as 'the by-product of comprehension'. In the original writings on IE (Sharwood Smith, 1991, 1993), the discussion focuses on how IE can have an impact on input being internalised and further processed in the language learning mechanisms. It is therefore clear that in the field of IE a processing perspective of language could be supported (see Sharwood Smith, 1991, 1993). However, IE studies have claimed acquisition of the target form (see sections 2.3.3), by adopting Schmidt's (1990, 1993, 1995, 2001) strong view of the Noticing Hypothesis where conscious awareness, i.e. paying attention to the target form, is the only and sufficient condition for acquisition to take place. However, the fact that neither PI nor IE provides an adequate theoretical explanation of what constitutes acquisition is problematic for the purposes of the present study. According to the present study, neither noticing', as previously defined (strong view), nor measuring performance three weeks post-instruction can validate claims regarding type of knowledge of the target form (see also critique from Collentine, 2004; DeKeyser et. al, 2002; Doughty, 2003, 2004; Lee, 2004; Lee and Huang, 2008; Mitchell and Myles, 2004; Norris and Ortega, 2000; Sharwood Smith and Trenkic, 2001; Sheen, 2005).

According to Pallotti (2007:361), establishing specific criteria in order to assess if a form has been acquired is 'arbitrary'. Researchers tend to use thresholds to decide on the acquisition of a certain target form. These thresholds are often expressed as accuracy percentages, claiming that a structure has been acquired if there is 60 per cent of correct use (Vainikka and Young-Scholten 1994); 75 per cent (Ellis 1988); or 80 per cent (Andersen 1978); while others consider the target form to be acquired above 90 per cent of accuracy (Dulay and Burt 1974; Bahns 1983). Clearly emergence has for many researchers different criteria in order to be established. In more general terms, emergence has been defined as a threshold used to indicate the appearance of a previously not available (target) form (see Hatch and Faraday, 1982; Pallotti, 2007; Pienemann 1998; among many). The lack of convincing theoretical explanation regarding the justification that a certain threshold is more valid than another raises an issue in the choice of a valid and reliable criterion in making claims about acquisition (Pallotti, 2007:362). Hatch and Faraday (1982: 182ff) have shown the different conclusions that one can draw about acquisition orders of the same target structure with the same dataset, when applying two different acquisition criteria. The very fact that there is great range between the given thresholds (60-90 per cent accuracy) indicates that researchers seem to equate 'acquisition' with 'mastery' (Pallotti, 2007:362).

Instead, the ‘emergence criterion’, i.e. measuring acquisition based on emergence of the target structure, is proposed as a less arbitrary criterion representing a qualitative restructuring of the interlanguage (e.g. Bahns 1983; Bardovi-Harlig 2000; Hammarberg 1996; Meisel et al. 1981; Pallotti, 2007; Pienemann 1998).

Findings from the present study show that score performance has significantly improved post-instruction. In the error-correction and the fill in the gap tasks, where participants were required to correct and/or supply the target form, score comparisons showed that in the pre-test participants of the four groups were not able to supply the target form (zero points were given to all four groups in these two tasks). However, in the immediate post-test the three experimental groups significantly improved as the mean score performance ranged between 3 to 10 out of the 15 points for the error correction task and 2-7 points for the fill in gap task out of the 16 points they could maximally score. The Combined group achieved the maximum scores, followed by the PI and the IE groups. Comparing groups’ pre-test performance to the delayed post-test, scores are higher, ranging from 2-5 for the error-correction 1-4 points for the fill in the gap tasks; however scores are lower when comparisons are drawn between the immediate and the delayed post-test. Even if the data of the present study had been analysed for each participant’s performance and not just for groups, claiming acquisition of the target form based on the ‘emergence criterion’ would be questionable. On the other hand, it could be argued that results suggest that there are evidence of syntactic restructuring and production of inflectional morphology (see Schwartz and Gubbala-Ryzack, 1991).

McLaughlin (1990) has argued that interlanguage development is non-linear, following a ‘U-shape’ pattern, as representations change over time through ‘restructuring’. The fact that groups’ performance has ‘backslided’ in the delayed post-test would be, according to McLaughlin (1987, 1990),the result of restructuring; a cognitive process ‘in which the components of a task are coordinated, integrated or reorganised into new units’, thereby allowing old components to be replaced with new ones (McLaughlin, 1990:118). McLaughlin (1990) and VanPatten (1996, 2004, 2007, 2009) conceptualise second language acquisition as the result of internal mechanisms that consist of a set of processes, i.e. input processing (Phase 1); accommodation (Phase 2); restructuring (Phase 3); while VanPatten (1996, 2004) also adds a fourth phase, ‘access’, in which linguistic data have been integrated into the developing system and can be accessed through output and/or production. McLaughlin (1990:121),referring to the third possible state, argues that ‘learners restructure their knowledge until they sort out form-function relationships’ Carroll (2001, 2006, 2007) also supports the notion of restructuring in the AIT through ‘i-learning’, an on-line process that can change perceptual and parsing procedures in order to build a representation, when parsing fails. She does also distinguish between restructuring and acquisition, arguing that acquisition occurs incidentally, as it is a byproduct of learners’ interaction with language in different settings (Carroll, 2007). On the other hand VanPatten (1996, 2004, 2007, 2009) states, as noted in Chapter 2, that IP is not a model of acquisition, nor does it attempt to describe L2 parsing and processing. Instead, IP is concerned with the initial stage of input processing. Findings of the present study exhibit significant improvement in the short-term, thus advocating that the target form has been internalised, at some level, as some sort of knowledge. Accommodation and restructuring seem to have taken place. Due to the gap between instruction and delayed post-test, in which participants to the best of my knowledge had no exposure of the target form, performance scores declined. Therefore assumptions regarding acquisition cannot be valid. However, it could be hypothesised that participants of the three groups are in a state where they are restructuring their knowledge until they sort out form-function relationships; and/or in a stage where ‘i-learning’ takes place (similar to Carroll, 1999, 2001, 2006, 2007; McLaughlin, 1990; VanPatten, 1996, 2004, 207, 2009).

A further interpretation regarding the decline in long-term performance is provided from VanPatten (1996, 2004) through the concept of ‘Communicative Value’ (CV), which refers to the meaningfulness of the form in contributing to the overall sentence meaning (see section 2.4.1). Case marking has a low communicative value for English learners. The aim of PI is to force learners to pay attention and process this low CV structure in order to decode the meaning by excluding redundant features. Findings of the present study support the idea that PI has been more successful in the short than in the long term in making learners process this low CV structure, as scores were lower in the delayed post-test compared to the immediate post-test. However, PI has been more successful than IE in dealing with the issue of CV both in the short and the long term. Based on the findings, the Combined group, which differed from PI only in the variable coloured IE, has been more successful than IE both in the short and the long term. However statistical analysis from comparisons between the Combined and the PI groups showed significant differences only in the short term. The findings of the present study are in line with VanPatten and Fernandez (2004), who also explored long term effects of PI instruction, as they too observed a decline in performance in the long-term. However, important questions remain open: How can we explain the significant differences in performance between the Combined and the PI groups in the short term? Can we claim acquisition, i.e. restructuring of ix competence? And if yes, how can we theoretically link CV with acquisition in order to explain the obtained results?

Based on the Competition Model (Bates and MacWhinney, 1989), short and long term effects are the outcome of successful 'win-out' of competing "parent" representations (see also Carroll, 2001, 2004, 2007; Sharwood Smith and Truscott, 2004, 2005 and in prep; VanPatten, 1996, 2004, 2007, 2009). Fluctuation in score performance, in terms of increase in the short-term and decrease in the long-term, result from the representation's 'cue strength'. These arguments can support the short-term findings in the present study, i.e. explain that the Combined group outperformed the other two groups because type of instruction has been more successful in enhancing 'cue strength' therefore reinforcing the target structure in winning the competition from 'parent' structures. However, questions still remain unanswered: Why is the Combined group more effective than the PI group in the short term and not also in the long term? Why is the PI group more effective than the IE group in the short term and but not also in the long term? How can we explain delayed post-test performance? And, more importantly, can we claim acquisition of the target form based on the given findings?

Through the interdisciplinary approach it adopts, MOGUL can offer a better explanation to the issues raised regarding the validity of claiming acquisition of the target form based on the findings of the present study At the same time MOGUL can offer a theoretical platform that can explain in more detail the processes involved during on-line processing and their impact on language learning in the short and the long term. Sharwood Smith and Truscott (2004; in prep: 59) describe acquisition as 'the lingering effect of processing' and propose the use of the term 'growth' instead of 'acquisition' to better describe this process. Growth in MOGUL occurs both within and between the modules and is influenced by the changes of the activation levels of the structural combinations, while it is constrained by the principles of each module (Sharwood Smith and Truscott, 2004; in prep). With this in mind, it can be argued that the Combined teaching intervention has generated more ‘growth’ in the language module and the conceptual structure through the greater activation levels and the richer interconnectivity of POpS than PI or IE alone. In turn, the PI group has been more effective in achieving ‘growth’ as described above than the IE group; whereas the IE group was also more efficient in developing ‘growth’ of the ‘perceptual representation’, i.e. the target form, in the relevant sub-modules than non-instruction. In addition, it can be assumed based on MOGUL’s propositions that richer POpS interconnectivity has promoted processing at a conscious level (see SST in prep; see also Baars, 1988, 1997, 2007), which based on findings from the present study, is shown to have an impact in groups’ performance, i.e. the greater POpS interconnectivity the highest mean scores; and thus the more effective the type of instruction.

As a result of the differences in processing due to POpS in each of the three types of instruction (please refer to section 5.3 for further details) the activation levels, i.e. current and resting levels, of the target form are the highest for participants of the Combined group, followed by participants of the PI group and the IE group respectively. The immediate post-test performance exhibits higher mean scores than the delayed post-test because testing immediately followed instruction, whereas instruction and immediate testing conditions were completed in the same week. Consequently, the resting levels of the target structure in each of the three experimental groups were high, however they were higher for the Combined, lower for the PI, and even lower for the IE, when comparing the three groups, and higher for all three groups compared to the non-instruction group. This means that the target structure in each of the three experimental groups had more chances to win the competition when other ‘parent’ structures were triggered during on line processing. The fact that activation levels were stimulated and, because of strength of activation, won the competition further increased the resting levels of the target structure each time. In other words, within one week the resting and current levels of the target structure in the three groups were constantly increasing, assuming that there was stimulation of activation and ‘winning out’ of the target structure. The fluctuations in the mean scores for each of the groups therefore mirror the activation level fluctuations of the target structure between groups.

During the time span of 12-14 weeks from the immediate to the delayed post-test condition, it can be supported based on MOGUL (SST 2005, in prep; TSS, 2004) that the resting levels of the target structure were constantly decreasing, since there was no stimulation based on instruction (see also Bates and MacWhinney,1987; and other connectionist approaches). According to the design of the study, during the 12-14 weeks there was no instruction of the target form. To the best of my knowledge, teachers did not give participants any information regarding the target structure of the study. It can be assumed that the target structure might have been activated by chance during processing of a competing structure; however if the target structure in focus had been activated, it would have been highly unlikely that the activation levels were increasing and/or winning the competition. Therefore as activation levels decreased, it became difficult for the target structure to win the competition in the delayed post-test condition from other competing structures that could have been triggered. The situation described for the delayed post-test findings can be better explained with the *'use it or lose it notion'* in MOGUL terms, which suggest that structures that are not highly activated and/ or frequently used are less likely to win the competition when entering one of the processors (SST, in prep).

The decrease in score performance from the immediate to the delayed post-test condition is a recurring phenomenon in the FonF/IE and PI literature. Metanalyses attribute the obtained findings in the short-term nature of most studies, in terms of exposure to and testing of the target form (e.g. mean hour treatment exposure of 1.5- 2 hours, followed by an immediate, 1 week, and delayed post-test, 4 weeks post instruction; see Doughty 2003; Lee and Huang, 2008; Norris and Ortega, 2000; see also Sharwood Smith and Trenkic, 2000; VanPatten and Fernandez, 2004). Despite the fact that effects of instruction seem to be durable there are observable decreases in the effect sizes (Lee and Huang, 2008; Norris and Ortega, 2000). Findings from the Spada and Lightbown (1993) study reported in White et al. (1991) seem to suggest that gains made during the intervention can be maintained in the long term, if treatment is not limited to a 'one off' type of intervention. Unlike common experimental studies (in the field of

FonF/IE/PI) in the Spada and Lightbown study (1993) participants continued to receive interaction and feedback for the targeted structure during the five month post-instructional intervention. Thus, as VanPatten and Fernandez (2004) argue long-term effects should be attributed to the effects of intervention plus continued feedback. Trahey and White (1993) and Trahey (1996) studies have also shown that positive evidence including frequent exposure to the target form can lead to acquisition of the target form. Perhaps if longitudinal studies were carried out and if studies' design included more exposure to the target form during the treatment period effects could be maintained from the immediate to the delayed post-test condition.

In the field of second language research the role and influence of external factors in language development is widely debated. It is highly questionable whether external efforts to 'teach' L2 knowledge can truly influence learners' developing L2 competence. Within this debate, theorists (such as Gubala-Ryzak, 1992; Schwartz, 1993) argue that true linguistic competence is not affected by instruction (see also Carroll, 1999, 2001, 2004). On the contrary, the potential role and contribution of negative evidence to the learners' interlanguage development is not fully dismissed (White, 1992). Chomsky (1981) supports the idea that direct negative evidence is not necessary for language acquisition, but indirect negative evidence may be relevant. Furthermore, metanalyses of studies show a positive effect of instruction in the acquisition of a target form regardless of instructional variables (Norris and Ortega, 2000; Lee and Huang, 2008).

On the other hand, it could also be argued that 'learners do not jump neatly from one discrete stage to another but undergo periods of optionality, sometimes quite long ones, where both new and old forms occur in learner performance' (Sharwood Smith and Truscott, 2005: 205).

5.6 Issues arising from the methodology of the present study

Doughty (2004:265-266) argues that PI studies' assessment tasks cannot provide valid measures assessing whether the learner's developing system has indeed changed due to type of instruction. PI assessment tasks, like other L2 studies investigating effectiveness of different types of instruction, attest metalinguistic knowledge, which is declarative knowledge about language, (see also Doughty, 2003, 2004; Norris and Ortega, 2000;

Paradis, 2004). In the present study, a battery of testing tasks was used instead. For comparability reasons, a reading comprehension task and an interpretation task, each biased towards IE and PI respectively, were included. Furthermore, a fill in the gap was included for comparability reasons as it is used as an 'independent' task to measure 'production' in PI studies (the task is also often used in IE studies for the same reason). The error-correction task was also used as an independent task to measure participants' ability to recognise, correct and produce the target form. With hindsight, the battery of testing tasks could have been modified in order to include either timed constrained tasks and/or response time measures. In this way, information about real time processing of the target structure and therefore a better understanding about any changes in learners' developing system through types of L2 instruction could have been provided, since the particular target form chosen did not allow for the use of oral tasks (which are traditionally seen as a good way of investigating implicit knowledge). However, further research is necessary in order to be able to include oral data that can provide a greater insight into the processes involved in PI and their impact in SLA.

In the present study, feedback was minimally included in the treatment packages, whereas no feedback was available in the testing conditions. To date the role of feedback in PI has been vaguely described. In IE studies, the role of feedback has not been considered. Recent PI studies indicate that there is a link between the presence of feedback in the treatment materials and speedier and more accurate processing (see Culman et al. 2009; Fernandez, 2008). Therefore it could be hypothesised that including more feedback might have resulted in different results. Further research is necessary in order to investigate and determine the role and the amount of feedback necessary in order to have an impact on the language learning or acquisition mechanisms and linguistic development.

Based on the participants' comments, the battery of testing tasks was too long. In retrospect, fewer items could have been included in the error-correction and fill in the gap tasks. The fact that the error-correction task was too long (it included 15 target forms and 15 distracters) may have had an impact on participants' performance. The PI and Combined groups had substantially higher scores in this task than the IE group. Taking into consideration that the first step of the error-correction task resembled referential activities, one could argue that in the immediate post-test it could have reinforced task taking strategies, regardless of distracters and based on frequency of recurrence of similar tasks (Ellis, 2002; Gass and Mackey, 2002; Schmidt, 2001).

The fill in the gap task, although third in the series of testing tasks, was also long (16 target items and 14 distracters). Participants could have filled in the task based on the exercises they had previously completed, or because ‘it felt right’, and/or because they might have assumed that all sentences should start with ‘den’. Redesigning the task using a timed constraint and/or measuring response time could have provided different findings and a better insight regarding participants automatised vs. controlled and/or explicit vs. implicit knowledge (Anderson, 1983; Bialystok, 1981b; Gass and Mackey, 2005; McLaughlin, 1983, 1990).

The use of simpler vocabulary in the battery of testing tasks might also have provided different findings. The vocabulary hard copy might have had a positive impact for some participants, acting as an additional cue, while problems with vocabulary might have hindered processing of the target form, as might have been the case for some participants in the reading comprehension testing task. Perhaps the hard copy was one of the reasons why there were no significant differences between groups, as participants in all four groups (including the Control group) showed improvement. On the other hand, the hard copy might have also had a negative impact on language processing, as participants may have relied to much on the copy rather than focusing on the task applying the cues taught during instruction to complete the task. The latter can be supported based on delayed post-test performance, where scores in the reading comprehension task were lower than in the immediate post-test, while within and between groups’ performance exhibited great range in variability.

Finally, direct comparisons could have been drawn with existing PI and IE studies, if the present study design had also included a post-test three to four weeks post instruction.

5.7 Conclusions and overall contribution of the present study to IE and PI research

Lee and Huang (2008) conducted a meta-analysis of studies that have applied visual input enhancement, namely IE. A total of sixteen studies were analysed, following Norris and Ortega's (2000) methodology measuring of effect sizes. The meta-analysis investigated the magnitude of the impact of IE on grammar learning. Analysis showed a small effect for learners exposed to enhanced texts over learners exposed to unenhanced texts, while a small but negative effect was found on learners' processing meaning in the reading comprehension measures. Thus, Lee and Huang's (2008) meta-analysis of IE studies, similarly to Norris and Ortega (2000), shows an outperformance of explicit grammar instruction over implicit IE[59] (see also Alanen, 1995, Doughty, 1991; Leeman et al. 1995). Furthermore, no significant differences are reported in comprehension, when IE is compared with other types of instruction (Jourdenais, 1995; Lee, 2007; Leow, 1997, 2001; Leow et al. 2003; Overstreet, 1998, 2002; Wong, 2003). Hence, it is valid to conclude that meta-analyses suggest an overall outperformance of elaborate, obtrusive and/or explicit IE over less elaborate, less obtrusive and/or implicit types of IE L2 instruction (see also Doughty, 2003:267; Lee and Huang, 2008; Norris and Ortega, 2000; Sharwood Smith, 1991).

The design followed in the present study has balanced the experimental groups in terms of amount of types and tokens as well as overall exposure to input. One of the aims and objectives of the present study was to explore the effectiveness of explicit and/or more elaborate and obtrusive vs. implicit and/or less elaborate and less obtrusive types of instruction. Findings from the present study have shown similar effects as the ones reported in the previously discussed meta-analyses, in terms of type of instruction. The present study's findings therefore support that explicit and more elaborate/obtrusive type of instruction, namely the combination of PI and IE, is more effective than less explicit and less elaborate/obtrusive type of instruction, namely PI. In turn, it has been found that a more implicit type of instruction such as IE, compared to the Combined and PI groups is less effective than explicit types, but more effective than non-instruction (Combined>PI>IE>Control).

[59] However, it should be noted that in the field of SLA it remains a vexed issue, whether performance indicates acquisition.

Despite the fruitful area of research and the promising results, methodological issues have undermined the possible contribution of PI and IE studies to SLA. In most studies, the sample sizes are too small and the lack of delayed post-tests exceeding a period of four weeks post-instruction do not allow for valid, reliable and generalizable conclusions arguing for the effectiveness of the proposed teaching interventions. All studies to date, apart from VanPatten and Fernandez (2008) and White (1998) have provided data for short-term effects of PI, since delayed post-tests range from one to three weeks post-instruction. Therefore claims regarding acquisition of the target form are invalid (see critique from Collentine, 2004; DeKeyser et. al, 2002; Doughty, 2003, 2004; Lee, 2004; Lee and Huang, 2008; Mitchell and Myles, 2004; Norris and Ortega, 2000; Sharwood Smith and Trenkic, 2001; Sheen, 2005). The present study aimed to contribute to the investigation of possible long-term effects of IE, PI and or the combination of the two approaches by administrating the delayed post-test 12 weeks post-instruction. Findings from the present study have shown that three teaching interventions have been more effective than non-instruction in the long-term. Similarly, the Combined teaching intervention has been shown to be more effective than the sole application of IE in the long term. However, no differences were reported between the Combined and the PI and the PI and IE teaching interventions in the long term. Further research is therefore necessary to establish if this is due to type of instruction, testing methods and materials and/or other factors that the present study has not accounted for.

The majority of the research studies both in the field of IE and PI have investigated adult SLA. In the present study, participants were beginning secondary school students. In this way, the present study aimed to contribute to IE and PI research by providing data regarding earlier developmental stages of learners in SLA.

The present study has also aimed to expand research in the field of PI and IE to another target language and target form by investigating the acquisition of German word order and case marking. For IE, this is a novel structure and target language under investigation, while for PI it will be the second study to the best of my knowledge investigating the particular target language and form (see Culman et al. 2009).

Concluding, the present study has attempted to analyse and discuss its findings based on theoretical frameworks and models of language processing and acquisition. In this way, it has aimed to address previous criticism of PI and IE studies regarding lack of theoretical support in evaluating obtained results and making claims about second language development (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Harrington, 2004; Sharwood Smith and Trenkic, 2001).

6 Conclusion

Instruction and types of instruction play a role in language development. Comparisons between the three experimental types of instruction confirm that the combination of PI with coloured TTIE is more effective than the sole application of PI and/or coloured TTIE types of instruction in the acquisition of case marking and word order in German. In turn, it has been shown that the sole application of PI is more effective than coloured TTIE. Although the findings of the present study cannot argue for ultimate attainment of the target structure, they seem to suggest that participants have reached higher levels of language development, especially in the short term. Gains were also maintained in the long term, though at a lower rate. Despite the fact that both descriptive and statistical analysis have shown that there is an effect of the proposed types of instruction closer analysis shows that there are no differences in effectiveness between the Combined and the PI and the PI and the IE paired instruction comparisons (Ellis, 1994; Housen and Pierrard, 2005; Klein, 1986; Mitchell and Myles, 2004). Given the findings of the present study, the term ‘language growth’, as proposed in Modular Online Growth and Use of Language, MOGUL, (Sharwood Smith and Truscott, 2004 and in prep; Truscott and Sharwood Smith, 2004, 2005) provides a better description than ‘language acquisition’ of participants ‘current state’ regarding the target structure. However, for ease of reference the terms language ‘acquisition’, ‘development’ and ‘growth’ have been used interchangeably in the present study.

The findings of the present study show that English learners of German can benefit from the combination of Processing Instruction (PI) and coloured Textual/Typographical Input Enhancement (TTIE) in ‘acquiring’ German case marking and word order. The findings suggest that the particular teaching intervention assist learners in restructuring existing linguistic representations available from L1 (Carroll, 2001; McLaughlin, 1987, 1990; VanPatten, 1996, 2004, 2007, 2009). PI can assist learners in the cognitive steps they take when processing input and it seems that enhanced PI with colour can be even more effective than PI alone, at least in the short term. Similarly, PI through Structured Input (SI) activities can enhance cue strengths in terms of cue validity and cue reliability therefore promoting the establishment of correct form-meaning connections (Bates and MacWhinney, 1987; MacWhinney and Bates, 1989; VanPatten, 1996, 2004, 2007, 2009). However, further research is necessary in order to investigate whether enhanced PI can be more effective than unenhanced PI in the long term.

Psycholinguistic concepts, such as attention, noticing, consciousness and awareness, play a facilitative role in language learning (Schmidt, 1990, 1993, 1995, 2001; Sharwood Smith, 1991, 1993; Tomlin and Villa, 1994; VanPatten, 1996, 2004, 2007, 2009). The teaching interventions applied in the present study are highly interconnected with these psycholinguistic concepts. Findings from the present study suggest that the level of interaction and access to these psycholinguistic concepts in each of the three different types of instruction results in different levels of effectiveness. In other words, the combined teaching intervention has been shown to be more effective than the sole application of PI and IE because of the higher levels of conscious awareness achieved when processing the target structure. MOGUL's inter-disciplinary approach possibly offers a good explanation regarding the effect of psycholinguistic processes in conjunction with the specific types of instruction on the overall processing of the target form (Sharwood Smith and Truscott, 2004 and in prep; Truscott and Sharwood Smith, 2004, 2005). However, as this is a novel proposal in the field of SLA and this is the only study to the best of my knowledge that has attempted to explain empirical results with MOGUL, further empirical research is needed in order to provide support to these claims. Nonetheless, it seems that MOGUL can offer a theoretical platform for both PI and IE research therefore overcoming criticism regarding interpreting the processes involved when processing input with particular types of instruction, as well as in explaining results obtained from PI and IE studies (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Harrington, 2004;Sharwood Smith and Trenkic, 2001).

Overall, foreign language teaching could benefit from the proposed teaching interventions and especially the combined type of instruction. Foreign language teaching curricula could incorporate the particular types of instruction for the teaching of problematic grammatical forms. In the case of beginner learners, it is evident that processing of new and difficult grammatical forms, processing of new structures at a sentence level, as is the case in the combined and the PI types of instruction can be more effective than non-instruction. Providing a very brief and focused presentation of explicit instruction followed by SI activities with the target form being coloured seems to be promoting more effective processing of the target form and more possibilities for learners to establish correct form-meaning connections. However, further research is necessary in order to explore long term effects of the proposed types of instruction. Moreover, performance variability needs to be further investigated.

The instruction of grammar through brief explicit rule presentation and SI activities with coloured enhancement of the target form in the form of computerised materials can become an additional way of attracting and motivating students to learn a foreign language, especially for languages that are thought to be difficult to learn, such as German.

In conclusion, despite the promising effects that seem to arise from the combined and the PI type of instruction, further research is necessary in order to expand the existing guidelines of PI and possible application to specific linguistic forms. Input Processing (IP) Principles should be further developed in order to include other possible predictions and strategies in the processing and acquisition of linguistic forms which have not been taken into account yet in the IP model. So far, existing IP Principles and guidelines in developing materials substantially limit research in terms of the choice of grammatical form and target language. Further research is therefore necessary in order to develop the IP model. Moreover, the IP model needs to provide more detailed explanations about the processes involved when processing input with PI. The same can be proposed for IE. In this way replication studies and/or studies wishing to investigate further the two types of instruction can maintain ‘treatment fidelity’ or remain true to the theoretical underpinnings proposed by either PI and/or IE. Adopting and/or proposing a theoretical framework that includes a detailed explanation of the language faculty including mechanisms concerning perception is therefore necessary in order to better understand and benefit from the propositions of the IP and IE models (Carroll, 2004; Collentine, 2004; DeKeyser et al. 2002; Harrington, 2004; Sharwood Smith and Trenkic, 2001).

7 References

Alanen, R. (1995). Input Enhancement and rule presentation in second language acquisition. In Schmidt, R. (Eds.) (1995). *Attention and awareness in foreign language teaching*, pp. 259-301. Honolulu: University of Hawai'i Press.

Alderson, C. (1983). Who Needs Jam? In Hughes, A. and Porter, D. (Eds.) *Current Developments in Language Testing*, pp. 87-92. London: Academic Press.

Alderson, C. (1992). Guidelines for the Evaluation of Language Education. In Alderson, C. and Beretta, A. (1992). *Evaluating second language education*, pp. 274-304. Cambridge, NY: Cambridge University Press.

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

Anderson, J. R. (1985). *Cognitive Psychology and its Implications (2nd Ed.)*. New York: Freeman.

Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Erlbaum.

Anderson, R. T. (2004). Phonological acquisition in preschoolers learning a second language via immersion: a longitudinal study. *Clinical Linguistics and Phonetics,* 18(3): 183-210.

Archibald, J. (Ed.) (2000). *Second language acquisition and linguistic theory*. Malden, Mass.: Blackwell

Baars, B. (1988). *A cognitive theory of consciousness*. NY: Cambridge University Press.

Baars, B. (1997). *In the theater of consciousness: the workspace of the mind*. New York: Oxford University Press.

Baars, B. and Gage, N. (2007). *Cognition, Brain and Consciousness: An Introduction to Cognitive Neuroscience*. London: Academic Press.

Bachman, L. (2006). What is the construct? The dialectic of abilities and context in defining constructs in language assessment. In Fox, J. and Wesche, M. (Eds.) *Language testing reconsidered.* University of Ottawa Press.

Bachman, L. and Palmer, A. (1985). *Basic concerns in language test validation.* Reading, MA: Addison-Wesley.

Bachman, L. and Palmer, A. (1996). *Language testing in practice: designing and developing useful language tests.* Oxford: Oxford University Press.

Bahns, J. (1983). On acquisitional criteria. *International Review of Applied Linguistics,* 21: 57-68.

Bailey, N., Madden, C. and Krashen, S. (1974). Is there a 'natural sequence' in adult second language learning? *Language Learning*, 21(2): 235-243.

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use.* Oxford: Blackwell.

Bates, E., Devescovi, A., and D'Amico, S. (1999). Processing complex sentences: A cross-linguistic study. *Language and Cognitive Processes*, 14(1): 69-123.

Bates, E. McNew, S. MacWhinney, B., Devescovi, A. and Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, 11:245-299.

Bates, E. and MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner and L. Gleitman (Eds.). *Language Acquisition: The State of the Art*, Cambridge University Press.

Bates, E. and MacWhinney, B. (1987). Competition, variation and language learning.

In MacWhinney, B. (Ed.) *Mechanisms of Language Acquisition.* Hillsdale, N.J: Lawrence Erlbaum.

Bates, E. and MacWhinney. B. (1989). Functionalism and the competition model. In E. Bates E. and MacWhinney, B. (Eds.) *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.

Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense. *Language Teaching Research,* 5(2): 95–127.

Benati, A. (2004). The Effects of Structured Input Activities and Explicit Information on the Acquisition of the Italian Future Tense. In VanPatten, B. (Eds.), *Processing Instruction: Theory, research, and commentary,* pp. 207-226. Mahwah, NJ: Lawrence Erlbaum.

Benati, A. (2008). From Processing Instruction on the Acquisition of Italian Noun–Adjective Agreement to Secondary Transfer-of-Training Effects on Italian Future Tense Verb Morphology. In Benati, A. and Lee, J. (2008). *Grammar Acquisition and Processing Instruction: Secondary and Cumulative Effects,* pp. 54-87. Clevedon: Multilingual Matters.

Benati, A. (2008). A Theory of Input Processing: How Learners Work with Input. In Benati, A. and Lee, J. (2008). *Grammar Acquisition and Processing Instruction: Secondary and Cumulative Effects,* pp. 1-23. Clevedon: Multilingual Matters.

Benati, A. (2008). Processing Instruction: Research and Practice in Assessing Primary Effects. In Benati, A. and Lee, J. (2008). *Grammar Acquisition and Processing Instruction: Secondary and Cumulative Effects,* pp. 24-53. Clevedon: Multilingual Matters.

Benati, A. and Laval, C. (2008). From Processing Instruction on the Acquisition of French Imparfait to Secondary Transfer-of-Training Effects on French Subjunctive and to Cumulative Transfer-of-Training Effects with French Causative Constructions. In Benati, A. and Lee, J. (Eds.) *Grammar acquisition and processing instruction:*

secondary and cumulative effects, pp. 121-158. Clevedon: Multilingual Matters.

Benati, A. and Lee, J. (2008). *Grammar Acquisition and Processing Instruction: Secondary and Cumulative Effects.* Clevedon: Multilingual Matters.

Berger, A., Della Pietra, A. and Della Pietra, V. J. (1996). A Maximum Entropy Approach to Natural Language Processing. *Computational Linguistics,* 22(1):39-71.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition,* 16: 157-168.

Bialystok, E. (1994). Representation and ways of knowing: Three issues in second language acquisition. In Ellis, N. (Ed.). *Implicit and explicit learning of languages,* pp. 549-569. London: Academic Press.

Bialystok, E. (2002). Cognitive processes of L2 users. In Cook, V. (Ed.). *Portraits of the L2 User,* pp. 145-166. Clevedon: Multilingual Matters.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In Gass, S. And Schachter, J. (Eds.). *Linguistic perspectives on second language acquisition,* pp. 41-68. Cambridge University Press.

Block, D. (2005). Convergence and resistance in the construction of personal and professional identities: Four French modern language teachers in London. In Suresh Canagarajah, A. (Eds). *Reclaiming the Local in Language Policy and Practice,* pp. 167-196. Mahwah, NJ: Lawrence Erlbaum.

Bowerman, M. (1987). Commentary: Mechanisms of language acquisition. In MacWhinney, B. (Eds). *Mechanisms of language acquisition,* pp. 443-466. Hillsdale, N.J.: Lawrence Erlbaum.

Bowles, M. A. (2008). Task type and reactivity of verbal reports in SLA: A first look at an L2 task other than reading. *Studies in Second Language Acquisition,* 30 (3):359-

387.

Bowles, M. A. and Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology. *Studies in Second Language Acquisition,* 27:415-440.

Bresnan, J. (Ed.) (1982). *The mental representation of grammatical relations.* Cambridge, Mass: MIT Press.

Bresnan, J. (2001). *Lexical Functional Syntax*. Oxford: Blackwell.

Broadbent, D.E. (1958). *Perception and Communication.* London: Pergamon.

Bryman, A. and Cramer, D. (1997). *Quantitative data analysis with SPSS for Windows: A guide for social scientists*. London: Routledge.

Carlson, R.A. and Dulany, D. E. (1985). Conscious attention and abstraction in concept learning. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 11: 45-58.

Carr, T. and Curran, T. (1994). Cognitive factors in learning about structured sequences: applications to syntax. *Studies in Second Language Acquisition,* 16:205-230.

Carroll, S. E. (1999). Putting 'input' in its proper place. *Second Language Research,* 15: 337-388.

Carroll, S. E. (2001). *Input and evidence*. Amsterdam: John Benjamins.

Carroll, S. E. (2002). Induction in a modular learner. *Second Language Research,* 18(3): 224–249.

Carroll, S. E. (2004). Some Comments on Input Processing and Processing Instruction. In VanPatten, B. (Eds.) *Processing Instruction: Theory, research, and* *commentary*, pp. 293-309. Mahwah, NJ : Lawrence Erlbaum.

Carroll, S. E. (2006). Salience, Awareness and SLA. In O'Brien, M. G., Shea, C. and Archibald, J. (Eds). *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, pp. 17-24. Somerville, MA: Cascadilla.

Carroll, S. E. (2007). Autonomous Induction Theory. In VanPatten B. and Williams J. (Eds.) *Theories in Second Language Acquisition*, pp. 155-173. New Jersey: Lawrence Erlbaum and Associates.

Chaudron, C. (1985a). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition*, 7:1-14.

Chaudron, C. (1985b). A method for examining the input/intake distinction. In Gass, S. and Madden C. (Eds.) *Input in second language acquisition*, pp. 285-300. Rowley, MA: Newbury House.

Cheng, P. W. (1985). Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review*, 92: 214-23.

Chomsky, N. (1965). Formal discussion: the development of grammar in child language. In Bellugi U. and Brown R. (Eds.) *The Acquisition of Language*. Indiana: Purdue University.

Chomsky, N. (1972) *Language and Mind*. NY: Harcourt Brace Jovanovitch.

Chomsky, N. (1980). *Rules and Representations.* Oxford: Blackwell.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass: MIT Press.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36 (1): 1-

22.

Chun, M. M., and Wolfe, J. M. (2000). Visual attention. In Goldstein E. B. (Ed.) *Blackwell handbook of perception,* pp. 272–310. Oxford: Blackwell.

Clahsen, H., Meisel, J.M. and Pienemann, M. (1983). *Deutsch als Zweitsprache: Der Zweitspracherwerb ausländischer Arbeiter*. Tübingen: Günter Narr.

Clahsen, H. and Muysken, P. (1986). The availability of universal grammar to adult and child learners - a study of the acquisition of German word order. *Second Language Research,* 2 (2): 93-119.

Collentine, J. (2004): Commentary: Where PI Research Has Been and Where It Should Be Going. In VanPatten, B. (Ed.). *Processing Instruction: Theory, research, and commentary,* pp. 144-169. Mahwah, NJ: Lawrence Erlbaum.

Corbetta, M. and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience,* 3:215-229.

Culman, H., Henry, N. and VanPatten, B. (2009). The Role of Explicit Information in Instructed SLA: An On-Line Study with Processing Instruction and German Accusative Case Inflections. *Die Unterrichtspraxis Teaching German,* 42(1): 19-31.

Cummins, J. (1983). *Heritage Language Education. A Literature Review.* Toronto: Ontario Institute for Studies in Education.

Curran, T. and Keele, S.W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 19:189-202.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J. and Sergant, C. (2006). Conscious, Preconscious, and Subliminal Processing: A Testable Taxonomy. *Trends in Cognitive Science,* 10 (5):204–211.

DeKeyser, R. (1994). How implicit can adult second language learning be? *AILA*, 11:84-96.

DeKeyser, R. (1995). Learning second language grammar rules. *Studies in Second Language Acquisition*, 17: 379-410.

DeKeyser, R. (1997). Beyond explicit rule learning: automatising second language morphosyntax. *Studies in Second Language Acquisition*, 19: 195-221.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In Doughty C. and Williams J. (Eds.), *Focus on form in classroom language acquisition*, pp. 42-63. New York: Cambridge University Press.

DeKeyser, R. (2001). Automaticity and automatisation. In P. Robinson (Ed.) *Cognition and second language instruction*, pp. 125-151. Cambridge: Cambridge University Press.

DeKeyser, R. (2003). Implicit and explicit learning. In Doughty C. and Long M. (Eds.) *Handbook of second language acquisition*, pp. 313-348. Oxford: Blackwell.

DeKeyser, R. (2005). What makes learning second language grammar difficult? *Language Learning*, 55:1-25.

DeKeyser, R. (2007). Skill Acquisition Theory. In VanPatten, B. and Williams, J. (Eds.)*Theories in Second Language Acquisition*, pp. 97-113. Lawrence Erlbaum, Mahwah: New Jersey.

DeKeyser, R., Salaberry, R., Robinson, P. and Harrington, M. (2002). What gets processed in processing instruction? A commentary on Bill VanPatten's processing instruction: an update. *Language Learning*, 52(4):805–823.

DeKeyser, R. and Sokalski, K. (1996). The differential role of comprehension and production practice. *Language Learning,* 46(4):613-642.

Deutsch, J. and Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review,* 70: 80-90.

Devescovi, A., D'Amico, S., and Gentili, S. (1999). The development of sentence comprehension in Italian: A reaction time study. *First Language,* 19:129-163.

DiBiase, B. (2008). Focus-on-form and development in L2 learning. In Keβler, J. U. (Ed). *Processability Approaches to Second Language Development and Second Language Learning*, pp. 197-220. Cambridge Scholars.

Dienes, Z., Broadbent, D.E. and Berry, D. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17:875-878.

Dörnyei, Z. (2003). *Questionnaires in second language research*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. (2005). *The psychology of the language learner.* Mahwah, NJ: Lawrence Erlbaum.

Doughty, C. (1991). Second language instruction does make a difference, *Studies in Second Language Acquisition,* 13: 431-469.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In Robinson, P. (Ed.). *Cognition and second language instruction,* pp. 206-257. Cambridge: Cambridge University Press.

Doughty, C. J. (2003). Instructed SLA: Constraints, Compensation, and Enhancement. In Doughty, C. J. and Long, M. H. (Eds.). *The Handbook of Second Language Acquisition*, pp. 256-310. Oxford: Blackwell.

Doughty, C. and Long, M. (2003). *Handbook of Second Language Acquisition,* Oxford: Blackwell.

Doughty, C. and Williams, J. (Eds.) (1998). *Focus on Form in Classroom Second Language Acquisition,* Cambridge: Cambridge University Press.

Dulay, H. and Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning,* 24: 37-53.

Ellis, N. C. (1994). *Implicit and explicit learning of languages*. London: Academic Press.

Ellis, N. C. (1995). Consciousness in second language acquisition: A review of field studies and laboratory experiments. *Language Awareness*, 4(3):123-146.

Ellis, N. C. (1996). Analysing language sequence in the sequence of language acquisition: Some comments on Major and Ioup. *Studies in Second Language Acquisition,* 18:361-36.

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language Learning,* 48:631-664.

Ellis, N. C. (2001). Memory for language. In: Robinson, P. (Ed.). *Cognition and second language instruction,* pp. 33-68. Cambridge: Cambridge University Press.

Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition,* 24:143-188.

Ellis, N. C. (2003). Incidentals that build fluency: Optimal word processing and its implications for vocabulary acquisition. *Studies in Applied Linguistics*, 9:1-18.

Ellis, N. C. (2005). At the interface: dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27: 305-352.

Ellis, N. C. and Larsen-Freeman, D. (2006). Language Emergence: Implications for Applied Linguistics. Introduction to the Special Issue. *Applied Linguistics,* 27(4): 558-589.

Ellis, R. (1985), *Understanding second language acquisition.* Oxford: Oxford University Press.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition,* 27 (2): 141-172.

Ellis, R. (2003). *Task-based language learning and teaching.* Oxford: Oxford University Press.

Ericsson, K. A., and Simon, H. A. (1993). *Protocol analysis; Verbal reports as data* (revised edition). Cambridge, MA: Bradford books/MIT Press.

Ervin-Tripp, S. (1974). Is second language learning like the first? *TESOL Quarterly,* 8:111-27.

Farley A. P. (2004). The Relative Effects of Processing Instruction and Meaning-Based Output Instruction. In VanPatten, B. (Eds.). *Processing Instruction: Theory, research, and commentary,* pp. 143-168. Mahwah, NJ: Lawrence Erlbaum.

Farley A. P. (2004). Processing Instruction and the Spanish Subjunctive: Is Explicit Information Needed? In VanPatten, B. (Eds.), *Processing Instruction: Theory, research, and commentary,* pp. 227-239. Mahwah, NJ: Lawrence Erlbaum.

Felix, S. (1985). More evidence on competing cognitive systems. *Second Language Research,* 1: 47-72.

Fernández, C. (2008). Re-examining the Role of Explicit Information in Processing

Instruction. *Studies in Second Language Acquisition,* 30 (3):277-305.

Field, A. (2009).*Discovering Statistics Using SPSS. London*. Thousand Oaks, New Delhi: Sage Publications.

Fodor, J. (1981).The mind-body problem. *Scientific American*, 244:114-25.

Fodor, J. (1983). *The modularity of mind*. Cambridge, Massachusetts: MIT Press.

Fotos, S.S. (1993). Consciousness-raising through focus on form: Grammar task performance versus formal instruction. *Applied Linguistics*, 14: 385-407.

Gall, M. D., Borg, W. R., and Gall, J. P. (1996). *Educational research: An introduction.* White Plains, NY: Longman.

Gardner, R. (1985). *Social psychology and second language learning: the role of attitudes and motivation*. London: Arnold.

Garret, M.F. (1976). Syntactic processes in sentence production. In Wales, R. and Walker E. (Eds.). *New Approaches to language mechanisms,* pp. 231-256. Amsterdam: North Holland.

Garret, M.F. (1980). Levels of processing in language production. In Butterworth (Ed.), *Language production, Vol. 1. Speech and Talk,* pp. 170-220. London: Academic Press.

Garret, M.F. (1982). Production of Speech: Observations from normal and pathological language use. In Ellis, A.W. (Ed.), *Normality and pathology in cognitive functions*. London Academic Press.

Gass, S. 1989. Language universals and second language acquisition. *Language Learning,* 39:497-534.

Gass, S. (1998). Apples and oranges; or why apples are not oranges and don't need to be, a response to Firth and Wagner, *The Modern Language Journal*, 82(1): 83-90.

Gass S. M. and Mackey, A. (2005): *Second Language Research: Methodology and Design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gathercole, S. and Baddeley, A. (1993). *Working memory and language.* Hillsdale, NJ: Lawrence Erlbaum.

Givón, T. (1979). From discourse to syntax: grammar as a processing strategy. In Givón, T. (Ed.) *Syntax and semantics, Vol. 12: Discourse and syntax.* Academic Press, New York.

Godfroid, A., Housen, A. and Boers, F. (2010). A procedure for testing the Noticing Hypothesis in the context of vocabulary acquisition. In Pütz, M. and Sicola L. (Eds.) *Cognitive Processing and Second Language Acquisition: Inside the Learner's Mind*, pp. 169-197. Amsterdam and Philadelphia: John Benjamins.

Gregg, K. R. (1993). Taking Explanation seriously; or, Let a couple of flowers bloom. *Applied Linguistics*, 15:276-294.

Gregg, K. R. (1996). The logical and developmental problems of second language acquisition. In Ritchie, W. and Bhatia, T. (Eds.) *Handbook of second language acquisition*, pp. 50-84. San Diego: Academic Press.

Gregg, K. R. (2001). Learnability and second language acquisition theory. In Robinson, P. (Ed.) *Cognition and second language instruction*, pp. 152-182. Cambridge: Cambridge University Press.

Gregg, K. R. (2003). SLA Theory: Construction and Assessment. In Doughty, C. J. and Long, M. H. (Eds.) *The Handbook of Second Language Acquisition*, pp. 831-865. Oxford: Blackwell.

Håkansson, G., Pienemann, M., and Sayehli, S. (2002). Transfer and typological proximity in the context of L2 processing. *Second Language Research*, 18(3):250-273.

Hakuta, K. (1976). Becoming bilingual: A case study of a Japanese child learning English. *Language Learning,* 26: 321-51.

Hammarberg, B. (1996). Examining the processability theory: the case of adjective agreement in Swedish. In Kellermann, E., Weltens, B. and Bongaerts, T. (Eds.). EUROSLA 6. A selection of papers. *Toegepaste talwetenschap in artikelen,* 55(2):75-88.

Harley, B. and Swain, M. (1984). The interlanguage of immersion students and its implication for second language teaching. In Davies, A., Criper, C. and Howatt, A. (Eds.) *Interlanguage,* pp. 291-311. Edinburgh: Edinburgh University Press.

Harrington, M. (2001). Sentence processing. In Robinson, P. (Ed.) *Cognition and second language instruction,* pp. 91-124. Cambridge: Cambridge University Press.

Harrington, M. (2004). Commentary: Input Processing as Theory of Processing Input. In VanPatten, B. (Eds.). *Processing Instruction: Theory, research, and commentary,* pp. 79-93. Mahwah, NJ: Lawrence Erlbaum.

Hatch, E. and H. Faraday (1982). *Research Design and Statistics for Applied Linguistics*. Rowley: Newbury House.

Hawkins, R. (2001). *Second language syntax*. Oxford: Blackwell.

Heilenman, L. K. and McDonald, J. L. (1993). The processing of dislocated sequences by native speakers of French. *Journal of French Language Studies,* 3:165-190.

Heredia, R. and McLaughlin, B. (1992). Bilingual memory revisited. In Harris, R. (Ed.) *Cognitive processing in bilinguals,* pp. 91-104. Amsterdam: Elsevier.

Herschensohn, J. (2000). *The Second Time Around: Minimalism and Second Language Acquisition.* Amsterdam: John Benjamins.

Holland, J., Holyoak, K., Nisbett, R. and Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.

Housen, A. and Pierrard, M. (Eds.) (2005). *Investigating Instructed Second Language*

Acquisition. Berlin, New York: Mouton de Gruyter.

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research,* 18 (3): 193–223.

Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition,* 27: 129-140.

Ioup, G., Boustagui, E., El Tigi, M., and Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition*, 16 (1): 73-98.

Iwasaki, S. (1993). The Structure of the Intonation Unit in Japanese. In Choi, S. (Ed.), *Japanese/Korean Linguistics*: 39-53.

Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge, MA: MIT Press.

Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.

Jackendoff, R. (2002). *Foundations of language: brain, meaning, grammar, evolution.* Oxford: Oxford University Press.

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Research,* 1146: 2-22.

Jourdenais, R, Ota, M., Stauffer, S., Boyson, B. and Doughty, C. (1995). Does textual enhancement promote noticing? In Schmidt, R. (Eds.) (1995). *Attention and awareness in foreign language teaching,* pp. 183-216. Honolulu: University of Hawai'i Press.

Just, M.A., and Carpenter, P.A. (1993). A capacity theory of comprehension:

Individual differences in working memory. *Psychological Review*, 99: 122-149.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, New Jersey: Prentice-Hall.

Kail, M. (1989). Cue validity, cue cost, and processing types in French sentence comprehension. In: MacWhinney, B. and Bates E. (Eds.) *The crosslinguistic study of language processing*, pp. 77-117. New York: Cambridge University Press.

Kail, R. and Hall, L. (2001). Distinguishing short-term memory from working memory. *Memory and Cognition*, 29 (1): 1-9.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: evidence from metalinguistic and repair data. *Cognition,* 23: 95-147.

Kawaguchi, S. (2005). Argument structure and syntactic development in Japanese as a second language. In Pienemann, M. (2005). *Cross-linguistic aspects of Processability Theory,* pp. 253-298. Amsterdam: John Benjamins.

Keil, F. C. (1983). *Semantic Influences and the Acquisition of Word Meaning.* Berlin: Springer-Verlag.

Keil, F. C. and Carroll, J. J. (1980). The child's conception of "tall": Implications for an alternative view of semantic development. *Papers and Reports on Child Language Development,* 19:21-8.

Kempen, G. and Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science,* 11:201-258.

Kentridge R.W., Heywood, C.A., Weiskrantz, L. (2004). Spatial attention speeds discrimination without awareness in blind sight. *Neuropsychologia,* 42:831-835.

Kihlstrom, J. (1984). *Conscious, subconscious, unconscious: A cognitive perspective*, In Bowers, K. and Meichenbaum, D. (Eds.) *The unconscious reconsidered*. New York: Wiley.

Klein, W. (1986). *Second Language Acquisition.* Cambridge: Cambridge University Press.

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Denver: Roberts and Co.

Koch, C. (2007) Neural correlates of consciousness. *Scholarpedia*, 2(12):1740.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

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

LaBerge, D. (1995). *Attentional Processing: The Brain's Art of Mindfulness*. Harvard University Press.

Lachman, R., Lachman, J.L., and Butterfield, E.C. (1979). *Cognitive psychology and information processing.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago.

Lamme, V.A.F (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences,* 7(1):12-18.

Lardiere, D. (2004). Representation and processing really are stuck with each other: Review article of Input and Evidence by Susanne E. Carroll. *Studies in Second Language Acquisition,* 26: 461–467.

Larsen-Freeman, D. (1991). Second language acquisition research: staking out the territory. *TESOL Quarterly,* 25: 315-50.

Larsen-Freeman, D. (2003). *Teaching Language: From grammar to grammaring*. Boston, MA: Heinle.

Larsen-Freeman, D. and Long, M. (1991). *An Introduction to Second Language Acquisition Research.* London and New York: Longman.

Lee, J.F. (2004). On the Generalizability, Limits, and Potential Future Directions of Processing Instruction Research. In VanPatten, B. (Eds.), *Processing Instruction: Theory, research, and commentary*, pp. 311-323. Mahwah, NJ: Lawrence Erlbaum.

Lee, J. and. Benati, A. (2007) *Second Language Processing: An analysis of Theory, Problems and Possible Solutions.* London: Continuum.

Lee, J. F. and. Benati, A. (2007). *Delivering Processing Instruction in Classroom and Virtual Contexts: Research and Practice*. London: Equinox.

Leeman, J., Arteagoitia, I., Fridman, B. and Doughty, C. (1995). Integrating attention to form and meaning: Focus on form in content- based Spanish instruction. In Schmidt, R. (Eds.) (1995). *Attention and awareness in foreign language teaching*, pp. 217-258. Honolulu: University of Hawai'i Press.

Leow, R. (1997a). The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition. *Applied Language Learning*, 8: 151-182.

Leow, R. (1997b) Attention, awareness, and foreign language behavior. *Language Learning*, 47: 467-506.

Leow, R. (1998a). Toward operationalizing the process of attention in second language acquisition: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics*, 19: 133-159.

Leow, R. (1998b). The effects of amount and type of exposure on adult learners' L2 development in SLA. *Modern Language Journal*, 82:49-68.

Leow, R. (2007). Input Enhancement in Classroom-Based SLA Research: An Attentional Perspective. In Gascoigne, C. (2007) (Ed.). *Assessing the Impact of Input*

Enhancement in Second Language Education, Stillwater: New Forums Press, Inc.

Leow, R., and Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition,* 26(1): 35-57.

Levelt, W. (1989). *Speaking: from intention to articulation*. Cambridge, Mass: Bradford Books.

Levelt, W. (1999). Producing spoken language: a blueprint of the speaker. In Brown, C. and Hagoort, P. (Eds.) *The neurocognition of language,* pp. 83-154. Oxford: Oxford University Press.

Levelt, W.J.M., Roelofs, A.P.A., and Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences,* 22 (1): 1-37.

Lightbown, P. (1985). Great expectations: second language acquisition research and classroom teaching. *Applied Linguistics,* 6:173-89.

Lightbown, P. (1985). Can language acquisition be altered by instruction? In Hyltenstam, K. and Pienemann, M. (Eds.). *Modelling and assessing second language acquisition,* pp. 101-112. Clevedon, Avon: Multilingual Matters.

Lightbown, P. (1998). The importance of timing in focus on form. In Doughty, C. and Williams, J. (Eds.) (1998). *Focus on Form in Classroom Second Language Acquisition,* pp. 177-196. Cambridge: Cambridge University Press.

Lightbown, P. and Spada, N. (1993). *How languages are learned*. Oxford: Oxford University Press.

LoCoco, V. (1987). Learner comprehension of oral and written importance of word order. In: VanPatten, B., Dvorak, T. and Lee, J. F. (Eds.). *Foreign language learning: A research perspective,* pp. 119-129. Rowley, MA: Newbury House.

Logan, G. D. (1988). Automaticity, resources and memory: Theoretical controversies and practical implications. *Human Factors,* 30: 583-598.

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

Long, M.H. (1981). Input, interaction and second language acquisition. In Winitz, H. (Ed.) *Native language and foreign language acquisition*. *Annals of the New York Academy of Sciences,* 379:259-278.

Long, M. H. (1983). Does second language instruction make a difference? A review of research'. *TESOL Quarterly,* 17 (3): 359-82.

Long, M. H. (1983). Native speaker/non-native-speaker conversation and the negotiation of comprehensible input. *Applied Linguistics,* 4(2):126-141.

Long, M.H. (1983). Training the second language teacher as classroom researcher. In Alatis, J, Stern, H, and Strevens, P. (Eds.) *Applied linguistics and the preparation of second language teachers: toward a rationale. GURT '83,* pp. 281-297. Washington, D.C.: Georgetown University Press.

Long, M. H. (1985). Input and second language acquisition theory. In Gass, S. and Madden, C. (Eds.) *Input in second language acquisition,* pp. 377-393. Newbury House, Rowley, Mass.

Long, M.H. (1985). A role for instruction in second language acquisition: task-based language teaching. In Hyltenstam, K. and Pienemann, M. (Eds.). *Modelling and assessing second language development,* pp. 77-99. Clevedon, Avon: Multilingual Matters.

Long, M.H. (1991). Focus on form: A design feature in language teaching methodology. In de Bot, K., Ginsberg, R. and Kramsch, C. (Eds.). *Foreign language research in cross-cultural perspective,* pp. 39-52. Amsterdam: John Benjamins.

Long, M.H. (1996). The Role of the Linguistic Environment in Second Language Acquisition. In Ritchie, W. and Bhatia, T. (Eds.). *Handbook of Second Language Acquisition*, pp. 413-468. San Diego: Academic Press.

Mackey, A. and Gass, S. M. (2005). *Second Language Research: Methodology and Design*, New Jersey: Lawrence Erlbaum Associates, Inc.

MacWhinney, B. (1987). Applying the Competition Model to bilingualism. *Applied Psycholinguistics*, 8: 315-327.

MacWhinney, B. (Ed.) (1987). *Mechanisms of Language Acquisition.* Hillsdale, N.J.: Lawrence Erlbaum.

MacWhinney, B. (1989). Competition and connectionism. In MacWhinney, B. and Bates, E. (Eds.). *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

MacWhinney, B. (2005b). A unified model of language acquisition. In Kroll, J. F. and de Groot, A. M. B. (Eds.). *Handbook of bilingualism: Psycholinguistic approaches*, pp. 49-67. New York: Oxford University Press.

MacWhinney, B., Bates, E. and Kliegl, R. (1984). Cue validity and sentence interpretation in English, German and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23:127-150.

MacWhinney, B. and Bates, E. (Eds.) (1989). *The Crosslinguistic Study of Sentence Processing.* Cambridge: Cambridge University Press.

MacWhinney, B., Pléh, C., and Bates, E. (1985). The development of sentence interpretation in Hungarian. *Cognitive Psychology*, 17: 178-209.

MacWhinney, B., and Pléh, C. (1988). The processing of restricted relative clauses in Hungarian, *Cognition*, 29: 95-141.

Marcel, A. (1983). Conscious and unconscious perception: an approach to the relations between phenomenal experience and cognitive processes. *Cognitive Psychology*, 15: 238-300.

Marsden, E. (2004). *Teaching and learning of French verb inflections: a classroom experiment using processing instruction*. Unpublished manuscript. University of Southampton.

Marsden, E. (2006) Exploring Input Processing in the classroom: An experimental comparison of Processing Instruction and Enriched Input. *Language Learning*, 56(3): 507-566.

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

McDonald, J. (1986). The development of sentence comprehension strategies in English and Dutch. *Journal of Experimental Child Psychology,* 41: 317-35.

McDonald, J. (1987). Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics,* 8: 379-413.

McDonald, J. L and Heilenman, L. K. (1992). Changes in sentence processing as second language proficiency increases. In Harris, R. J. (Ed.). *Cognitive processing in bilinguals,* pp. 325-336. Amsterdam: Elsevier.

McLaughlin, B. (1987). *Theories of Second-Language Learning.* Edward Arnold, London.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics,* 11(2): 113-28.

McLaughlin, B. (1990). ’Conscious’ vs. 'unconscious" learning. *TESOL Quarterly,* 24: 617-634.

McLaughlin, B. and Harrington, M. (1989). Second-language acquisition. *Annual Review of Applied Linguistics,* 10:122-134.

McLaughlin, B. and Heredia, J.L.C. (1996). Information-processing approaches to research on second language acquisition and use. In Ritchie, W C. and Bhatia, T. K. (Eds.). *Handbook of second language acquisition,* pp. 213-228. San Diego: Academic Press.

McLaughlin, B. and Nayak, N. (1989). Processing a New Language: Does Knowing Other Languages make a difference? In Dechert, H. and Raupach, M. (Eds.), *Interlingual Processes,* pp. 5-16. Tübingen, Gunter Narr.

McLaughlin, B., Rossman, T. and McLeod. B. (1983). Second-language learning: An information-processing perspective. *Language Learning,* 33: 135-158.

Meisel, J., Clahsen, H. and Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition*, 3(2): 109-135.

Merikle, P. M. and Joordens, S. (1997). Parallels between perception without attention and perception without awareness. *Consciousness and Cognition*, 6: 219-236.

Mitchell, R. and Myles, F. (2004). *Second language learning theories*. London: Oxford University Press.

Miyake, A. and Shah, P. (1999).*Models of working memory.* Cambridge University Press, New York.

Morgan, J.L. and Newport, E.L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior,* 20: 67-85.

Muijs, D. (2004). *Doing Quantitative Research in Education: with SPSS*. London: Sage.

Muranoi, H. (1996). *Effects of interaction enhancement on restructuring of interlanguage grammar: A cognitive approach to foreign language instruction*. Ph.D.

dissertation, Georgetown University, Washington, DC.

Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning,* 50: 617-673.

Myles, F. (1995). Interaction between linguistic theory and language processing in SLA. *Second Language Research,* 11 (3): 235-266.

Myles, F. (2004). From data to theory: the over-representation of linguistic knowledge in SLA. *Transactions of the Philological Society*, 102 (2): 139-168.

Myles, F., Hooper, J. and Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, 48(3): 323-363.

Myles, F., Mitchell, R. and Hooper, J. (1999). Interrogative chunks in French L2: A basis for creative construction? *Studies in Second Language Acquisition*, 21: 49-80.

Naccache, L., Blandin, E, Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychol Sci,* 13:416–424.

Nation, R. and McLaughlin, B. (1986). Experts and novices: an information-processing approach to the "good language learner" problem. *Applied Psycholinguistics,* 7: 41-56.

Neumann, O. (1996).Theories of attention. In Neumann, O. and Sanders, A.F. (Eds). *Handbook of perception and action,* pp. 389–446. Academic Press, San Diego.

Norris, J. and Ortega, L. (2000). Effectiveness of L2 instruction. *Language Learning,* 50 (3): 417-528.

O'Grady, W. (2005). *Syntactic carpentry: an emergentist approach to syntax.*

Mahwah, NJ: Lawrence Erlbaum.

O'Grady, W. (2008). The emergentist program. *Lingua,* 118 (4): 447-464.

O'Regan, J.K. and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences,* 24(5): 939-1011.

Overstreet, M. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics,* 2: 229–258.

Pallotti, G. (2007). An Operational Definition of the Emergence Criterion. *Applied Linguistics,* 28(3): 361–382.

Paradis, M. (1997). The cognitive neuropsychology of bilingualism. In de Groot, A. and Kroll, J. (Eds). *Tutorials in bilingualism: psycholinguistic perspectives,* pp. 331-354. Mahwah, NJ: Lawrence Erlbaum.

Paradis M. (2004). *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.

Paradis, M. (2009) *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins.

Pica, T. (1994). Questions from the classroom: research perspectives. *TESOL Quarterly,* 29: 49-79.

Pica, T. (1994). Research on negotiation: what does it reveal about second-language learning conditions, processes and outcomes? *Language Learning,* 44:493-527.

Pienemann, M. (1981). *Der Zweitsprachenerwerb ausländischer Arbeiterkinder.* Bonn: Bouvier.

Pienemann, M. (1984). Psychological constraints on the teachability of languages.

Studies in Second Language Acquisition, 6(2): 186-214.

Pienemann, M. (1987). Determining the Influence of Instruction on L2 Speech Processing, *Australian Review of Applied Linguistics*, 10 (2): 83-113.

Pienemann, M. (1987). Psychological constraints on the teachability of languages. In Pfaff, C. (Ed.). *First and second language acquisition processes*, pp. 143-168. Cambridge, MA: Newbury House.

Pienemann, M. (1998). *Language processing and second-language development: Processability Theory*. Amsterdam: John Benjamins.

Pienemann, M. (2005). *Cross-linguistic aspects of Processability Theory*. Amsterdam: John Benjamins.

Pienemann, M. (2006). Was in den Köpfen von Kindern vorgeht. In Pienemann, M., Keßler, J.-U. and Roos, E. (Hrsg.). *Englischerwerb in der Grundschule*. UTB: Schöningh

Pienemann, M. (2007). *Processability Theory*. In VanPatten B. and Williams J. (Eds.)*Theories in Second Language Acquisition*, pp. 137-154. New Jersey: Lawrence Erlbaum and Associates.

Pienemann, M. and Håkansson, G. (1999). A unified approach towards the development of Swedish as L2: a processability account. *Studies in Second Language Acquisition*, 21:383-420.

Pienemann, M., Di Biase, B., Kawaguchi, S. (2005). Extending Processability Theory. In Pienemann, M. (2005). *Cross-linguistic aspects of Processability Theory*, pp. 199-253. Amsterdam: John Benjamins.

Pienemann, M., Di Biase, B., Kawaguchi, S., and Håkanson, G. (2005). Processing constraints on L1 transfer. In Kroll J. F. and de Groot, A. M. B. (Eds.). *Handbook of bilingualism: Psycholinguistic approaches*, pp. 128-153. New York: Oxford

University Press.

Pinker, S. (1998) Words and rules. *Lingua*, 106:219-242.

Pinker, S. and Ullman, M. (2002) The past and future of the past tense. *Trends in Cognitive Science*, 6*:456-463.*

Piske, T. and Young-Scholten, M. (Eds.) (2009). *Input Matters in SLA*. Clevedon: Multilingual Matters.

Polio, C. (2007). A History of Input Enhancement: Defining an Evolving Concept. In Gascoigne C. (Eds.) *Assessing the Impact of Input Enhancement in Second Language Education*, pp. 1-18. USA: New Forums Press.

Polio, C., and Gass, S. M. (1997). Replication and reporting: A commentary. *Studies in Second Language Acquisition*, 19: 499-508.

Posner, M.I. (1992). Attention as a cognitive and neural system. *Current directions in Psychological Science*, 1: 11-14.

Posner, M.I. (1994). Attention: the mechanism of consciousness. *Proceedings of the National Academy of Sciences, U.S.A.*, 91(16): 7398-7402.

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

Posner, M.I. and Rothbart, M.K. (1992). Attention and conscious experience. In Milner, A.D. and Rugg, M.D. (Eds) *The Neuropsychology of Consciousness*, pp. 91-112. London: Academic Press.

Posner, M. and Snyder, C. (1975). Attention and cognitive control. In Solso, R. (Ed.) *Information Processing and Cognition: The Loyola Symposium.* Hillsdale, NJ: Erlbaum.

Prinz, J. (2004). *Gut reactions*. Oxford, UK: Oxford University Press.

Ramsey, R. M. G. (1980). Language-learning approach styles of adult multilinguals and successful language learners. *Annals of the New York Academy of Sciences,* 345: 73-96.

Rast, R. (2008). *Foreign Language Input: Initial Processing*. Multilingual Matters.

Reber, A. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior,* 6: 855-863.

Roberts, R. and Gibson, E. (2002). Individual differences in sentence memory. *Journal of Psycholinguistic Research,* 31 (6): 573-598.

Robinson, P. (1995). Attention, memory, and the noticing hypothesis. *Language Learning,* 45(2): 283-331.

Robinson, P. (1996). *Consciousness, rules and instructed second language acquisition.* New York: Peter Lang.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit language learning. *Language Learning,* 47: 45-99.

Robinson, P. (2001). *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Robinson, P. (Ed.) (2002). *Individual differences and instructed language learning*. Amsterdam, John Benjamins.

Robinson, P. (2003). Attention and memory during SLA. In Doughty, C. and Long, M. (Eds.) *Handbook of second language acquisition,* pp. 631-678. Oxford: Blackwell.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics,* 25: 46-76.

Rosch, E. (1977). Human categorization. In Warren, N. (Ed.). *Advances in cross-cultural psychology* (Vol. 1). London: Academic Press.

Ruchkin, D.S, Grafman, J., Cameron, K., and Berndt, R.S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences*, 26: 709–728.

Rutherford, W. (1982). Markedness in second language acquisition, *Language Learning*, 32: 85-108.

Rutherford, W. (1983). Language typology and language transfer. In Gass, S. and Selinker, L. (Eds.) *Language transfer in language learning*, pp. 358-370. Newbury House, Rowley, Mass.

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

Rutherford, W. (1988). Grammatical theory and L2 acquisition: a brief overview. In Flynn, S. and O'Neal, W. (Eds.). *Linguistic Theory in Second Language Acquisition*, pp. 404-416. The Netherlands: Kluwer.

Sachs, R., and Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, 29: 67-100.

Salaberry, R. (1997). The role of input and output practice in second language acquisition. *Canadian Modern Language Review*, 53:422-453.

Sanz, C. (1997). Experimental tasks in SLA research: Amount of production, modality, memory, and production processes. In Glass, W.R. and Pérez-Lerroux, A.T. (Eds.)*Contemporary perspectives on the acquisition of Spanish*, vol. 2, pp. 41-56. Somerville: Cascadilla Press.

Sanz, C. (2004). Computer Delivered Implicit Versus Explicit Feedback in Processing Instruction. In VanPatten, B. (Ed.). *Processing Instruction: Theory, research, and commentary*, pp. 241-255. Mahwah, NJ: Lawrence Erlbaum.

Sanz, C. (2005). Adult SLA: The interaction between internal and external factors. In Sanz, C. (Ed.) *Mind and Context in Adult Second Language Acquisition*, pp. 3-20.

Washington, DC: Georgetown University Press.

Sanz, C., Lin, H-J., Lado, B., Bowden, H.W., and Stafford, C.A. (2009). Concurrent verbalizations, pedagogical conditions, and reactivity: two CALL studies. *Language Learning, 59* (1): 33-71.

Sanz, C. and Kara Morgan-Short (2004). Positive Evidence versus Explicit Rule Presentation and Explicit Negative Feedback: A Computer-Assisted Study. *Language Learning*, 54 (3): 5–79.

Sanz, C. and Morgan-Short, K. (2005). Explicitness in pedagogical interventions: input, practice and feedback. In Sanz, C. (Ed.) *Mind and context in adult second language acquisition*, pp. 234-266. Washington, DC: Georgetown University Press.

Schelstraete, M.A. and Degand, L. (1998). Assignment of grammatical functions in French relative clauses. *Language Sciences* 20 (2):163-188.

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

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13: 206-226.

Schmidt, R. (Eds.) (1995). *Attention and awareness in foreign language teaching*. Honolulu: University of Hawai'i Press.

Schmidt, R. (2001). Attention. In Robinson P. (Ed.) *Cognition and second language instruction*, pp. 3-32. Cambridge: Cambridge University Press.

Schmidt, R., and Frota, S. (1986). *Developing basic conversational ability in a second language. A case study of an adult learner of Portuguese*. In Day R. (Ed.). *Talking to learn: Conversation in second language acquisition*, pp. 237-326. Rowley, MA: Newbury House.

Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behaviour. *Studies in Second Language Acquisition,* 15: 147-163.

Schwartz, B. (1998). On two hypotheses of 'transfer' in L2A: minimal trees and absolute L1 influence. In Flynn, S., Martohardjono, G. and O'Neil, W. (Eds). *The Generative Study of Second Language Acquisition*, pp. 35-59. Mahwah, NJ: Lawrence Erlbaum.

Schwartz, B. D. and Gubala-Ryzak, M. (1992). Learnability and grammar reorganization in L2A: against negative evidence causing the unlearning of verb movement. *Second Language Research,* 8:1-38.

Selinker, L. (1972). Interlanguage. *IRAL,* 10, (3): 209-231.

Selinker, L., Kim, D.E., Bandi-Rao, S. (2004). Linguistic structure with processing in second language research: is a 'unified theory' possible? *Second Language Research,* 20 (1):77-94.

Shanks, D. R. and St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17: 367-447.

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 (2): 118-132.

Sharwood Smith, M. (1993): Input enhancement in instructed SLA. Theoretical Bases. *Studies in Second Language Acquisition,* 15: 165-179.

Sharwood Smith M. (2006). Enhancing Input and Focusing on Form. A MOGUL perspective, paper presented at Herriot-Watt University, Edinburgh.

Sharwood Smith, M. (2008). *Revisiting the role of consciousness with MOGUL*. In Han, Z., Park, E., Revesz, A., Combs C, .and Kim, J. (Eds.). U*nderstanding second language process,* pp. 1-15. Clevedon: Multilingual Matters.

Sharwood Smith M. (2009). Metalinguistic processing and acquisition within the

MOGUL framework, paper presented for CRiLLS, Newcastle.

Sharwood Smith, M. (2009). A MOGUL perspective on consciousness and input processing. Plenary presentation at the Second Language Processing and Parsing: State of the Science conference, Texas Tech University, Lubbock.

Sharwood Smith, M. and Rutherford, W. (1987). Consciousness raising and universal grammar (reprinted). In Rutherford, W. and Sharwood-Smith, M. (Eds.) (1988). Grammar and Second Language Teaching. New York: Newbury House.

Sharwood Smith, M. and Trenkic, D. (2001). Re-evaluating Theoretical and Methodological Aspects of Focus on Form Research, paper presented at EUROSLA 11, Paderborn, 26-29, September 2001.

Sharwood Smith M. and Truscott, J. (2005): Stages or Continua in Second Language Acquisition: A Mogul Solution. *Applied Linguistics,* 22 (2): 219-240.

Sharwood Smith M. and Truscott, J. (2009). Unpublished draft manuscript. *Modular growth and use of language: a processing perspective* (cited with permission).

Sheen, R. (2005). Focus on forms as a means of improving accurate oral production. In Housen, A. and Pierrard, M. (Eds). *Investigating Instructed Second Language Acquisition,* pp. 271-310. Berlin, New York: Mouton de Gruyter.

Shiffrin, R. and Schneider, W. (1977). Controlled and automatic human information processing. *Psychological Review,* 84: 127-90.

Shook, D. (1994). FL/L2 reading, grammatical information, and the input to intake phenomenon. *Applied Language Learning,* 5(2): 57-93.

Simard, D. and Wong, W. (2001). Alertness, orientation and detection. *Studies in Second Language Acquisition*, 23:103-124.

Skehan, P. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford

University Press.

Slobin, D. (Ed.) (1985). *The crosslinguistic study of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Slobin, D. I. (1987). Thinking for speaking. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society,* pp. 435-444.

Sorace, A. (2003). Near-nativeness. In Doughty, C.J. and Long, M.H. (Eds.) *Handbook of second language acquisition,* pp. 130-152. Oxford: Blackwell.

Stratman, James, F. and Hamp-Lyons L. (1994). Reactivity in concurrent think-aloud protocols. In Samagorinsky, P. (Ed.). *Speaking about writing: Reflections on research methodology,* pp. 89-112. London: Sage.

Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In Gass, S. and Madden, C. (Eds). *Input in second language acquisition,* pp. 235-253. Newbury House, Rowley, Mass.

Swain, M. (1995) Three functions of output in second language learning. In Cook, G. and Seidelhofer, B. (Eds.). *Principle and Practice in Applied Linguistics: Studies in Honor of H.G. Widdowson*, pp. 125-144. Oxford: Oxford University Press.

Swain, M. (2005). The Output Hypothesis: Theory and Research. In Hinkel, E. (Ed.). *The Handbook of Research in Second Language Teaching and Learning,* pp. 471-483. Hillsdale, NJ: Lawrence Erlbaum Associates.

Swain, M. and Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics,* 16: 371-391.

Tomasello, M. (Ed.) (1998). *The new psychology of language: Cognitive and functional approaches to language structure.* Mahwah, NJ: Erlbaum.

Tomasello, M. (2000). *The Cultural Origins of Human Cognition,* Harvard University Press.

Tomasello, M. (2003). *Constructing a Language*. Harvard University Press.

Tomlin, R., and Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16: 183-204.

Towell, R. and Hawkins, R. (1994). *Approaches to Second Language Acquisition.* Clevedon: Multilingual Matters.

Towell, R. and Dewaele, J.M. (2002). The role of psycholinguistic factors in the development of fluency amongst advanced learners of French. In Dewaele, J.-M. (Ed.). *Focus on French a Foreign Language: Multidisciplinary Approaches.* Clevedon: Multilingual Matters.

Truscott, J. (1998). Noticing in second language acquisition: A critical review. *SLA Research* 14: 103-135.

Truscott, J. (2007). Optionality in Second Language Acquisition: A Generative Processing-Orientated Account. *International Review of Applied Linguistics 44* (4): 311-330.

Truscott, J. and Sharwood Smith, M. (2004). How APT is your theory: present status and future prospects. *Bilingualism: Language and Cognition*, 7 (2): 43-47.

Truscott, J. and Sharwood Smith, M. (2004). Acquisition by processing: a modular perspective on language development. *Bilingualism: Language and Cognition* 7 (2): 1-20.

Truscott, J. and Sharwood Smith, M. (2009). *Noticing in MOGUL*, paper presented for the International Symposium of Bilingualism, University of Utrecht.

Vainikka, A. and Young-Scholten, M. (1995). *Tree Growth and Morphosyntactic Triggers in Adult SLA*. IRCS Technical Report Series available at:http://repository.upenn.edu/cgi/viewcontent.cgi?article=1136&context=ircs_reports.

Vainikka, A. and Young-Scholten, M. (1998). *Morphosyntactic triggers in adult SLA*. In Beck, M.-L. (Ed.) *Morphology and its interfaces in second language knowledge,* pp. 89-113. Amsterdam/Philadelphia: John Benjamins.

Vainikka, A., and Young-Scholten, M. (2002). Restructuring the CP in L2 German. In Skrabela, B, Fish, S., and Do, A.H.J. (Eds.). *Proceedings of the XXVI Boston University Conference on Language Development*, pp. 712-722. Sommerville, MA: Cascadilla Press.

Vainikka, A., and Young-Scholten, M. (2007b). The role of literacy in the development of L2 morphosyntax from an organic grammar perspective. In Faux, N. (Ed.). *Low educated second language and literacy acquisition. Proceedings of the Second Annual LESLLA Conference*. Richmond, VA: Virginia Commonwealth University, Literacy Institute.

VanPatten, B. (1996). *Input processing and grammar instruction.* Norwood, NJ: Ablex.

VanPatten, B. (2000). Thirty years of input (or intake, the neglected sibling). In Swierzbin, B., Morris, F., Anderson, M., Klee, C., and Tarone, E. (Eds). *Social and cognitive factors in second language acquisition*, pp. 287-311. Somerville, MA: Cascadilla Press.

VanPatten, B. (2003). *From Input to Output: A teacher's guide to second language acquisition*. USA: McGraw-Hill.

VanPatten, B. (2004). *Processing instruction: theory, research and commentary.* Mahwah, NJ: Lawrence Erlbaum.

VanPatten, B. (2004). Input Processing in SLA. In VanPatten, B. (Eds). *Processing Instruction: Theory, research, and commentary,* pp. 5-33. Mahwah, NJ: Lawrence Erlbaum.

VanPatten, B. (2004). Several Reflections on Why There Is Good Reason to Continue Researching the Effects of Processing Instruction. In VanPatten, B. (Eds.). *Processing Instruction: Theory, research, and commentary,* pp. 325-335. Mahwah, NJ: Lawrence

Erlbaum.

VanPatten, B. (2005). Processing instruction. In Sanz, C. (Ed.) *Mind and context in adult second language acquisition*, pp. 267-282. Washington, DC: Georgetown University Press.

VanPatten, B. (2007). Input Processing in Adult Second Language Acquisition. In VanPatten, B. and Williams, J. (Eds). *Theories in Second Language Acquisition*, pp. 115-135. Mahwah, NJ: Lawrence Erlbaum.

VanPatten, B. (2009). Processing Matters in Input Enhancement. In Piske, T. and Young-Scholten, M. (Eds.). *Input Matters in SLA*, pp. 47–61. Clevedon, UK: Multilingual Matters.

VanPatten, B., and Cadierno, T. (1993a). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15: 225–244.

VanPatten, B., and Cadierno, T. (1993b). Input processing and second language acquisition: A role for instruction. *The Modern Language Journal*, 77: 45–57.

VanPatten, B., Fernandez, C. (2004). The Long-Term Effects of Processing Instruction. In VanPatten, B. (Eds). *Processing Instruction: Theory, research, and commentary*, pp. 273-289. Mahwah, NJ: Lawrence Erlbaum.

VanPatten, B, and Oikkenon, S. (1996). Explanation vs. Structured Input in Processing Instruction. *Studies in Second Language Acquisition*, 18: 495–510.

VanPatten, B. and Sanz, C. (1995). From Input to Output: Processing Instruction in Communicative Tasks. In Eckman, F. R., Highland, D., Lee, P. W., Mileham, J., and Weber, R. R. (Eds). *Second Language Acquisition Theory and Pedagogy*, pp. 169-186. Hillsdale, NY: Lawrence Erlbaum Associates.

VanPatten B. and Williams J. (Eds.) (2007). *Theories in Second Language*

Acquisition. New Jersey: Lawrence Erlbaum and Associates.

VanPatten, B., and Wong, W. (2004). Processing Instruction and the French Causative: Another Replication. In VanPatten, B. (Ed). *Processing Instruction: Theory, Research, and Commentary*, pp. 97–118. Mahwah, NJ: Lawrence Erlbaum.

Velmans, M. (1996). What and where are conscious experiences? In Velmans, M. (Ed.). *The Science of Consciousness: Psychological, Neuropsychological and Clinical Reviews.* London: Routledge.

Verspoor, M., Lowie, W. and De Bot, K. (2009). Input and Second Language Development from a Dynamic Perspective. In Piske, T. and Young-Scholten, M. (Eds.) (2009). *Input Matters in SLA,* pp. 62-80. Clevedon: Multilingual Matters.

Vygotsky, L. S. (1986). *Thought and Language.* Cambridge, MA: MIT Press.

White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In Doughty, C., and Williams, J. (Eds.). *Focus on form in second language classroom acquisition,* pp. 85-113. Cambridge: Cambridge University Press.

White, L. (1989). *Universal Grammar and Second Language Acquisition*. Amsterdam: Benjamins.

White, L. (1996). Universal Grammar and Second Language Acquisition: Current Trends and New Directions. In Ritchie, W. and Bhatia, T. (Eds). *Handbook of Second Language Acquisition*, pp. 85-120. San Diego: Academic Press.

White, L. (2000). Second Language Acquisition: From Initial to Final State. In Archibald J. *Second Language Acquisition and Linguistic Theory*, pp. 130-155. Oxford: Blackwell

White, L. (2003). *Second Language Acquisition and Universal Grammar*, Cambridge University Press.

Whorf, B. L. (1956). In Carroll, J. B. (Eds). *Thought and Reality: Selected Writings of*

Benjamin Lee Whorf. Cambridge, Massachusetts: MIT Press.

Wickens, C.D. (1984). Processing resources in attention. In Parasuraman, R. and Davies, R. (Eds.), *Varieties of attention*. New York: Academy Press.

Williams, J. and Evans, J. (1998): What kind of focus and on which forms. In Doughty, C. and Williams, J. (Eds.) (1998). *Focus on Form in Classroom Second Language Acquisition,* pp. 139-155. Cambridge: Cambridge University Press.

Winitz, H. (Ed.) (1981). *The Comprehension Approach to Foreign Language Instruction,* Rowley MA: Newbury House.

Winter, B. and Reber, A. S. (1994). Implicit learning and the acquisition of natural languages. In Ellis, N.C. (Eds). *Implicit and explicit learning of languages,* pp. 115-147. London: Academic Press.

Wode, H., Bahns, J., Bedey, H. and Frank, W. (1978). Developmental sequences: An alternative approach to morpheme orders. *Language Learning,* 28:175-185.

Wong Fillmore, L. (1979). Individual differences in second language acquisition. In Fillmore, C., Kempler, D. and Wang, W. (Eds). *Individual Differences in Language Ability and Language Behaviour.* New York: Academic Press.

Wong, W. (2003). The effects of textual enhancement and simplified input on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning,* 14: 109-132.

Wong, W. (2004). The Nature of Processing Instruction. In VanPatten, B. (Ed). *Processing Instruction: Theory, research, and commentary,* pp. 3-63. Mahwah, NJ: Lawrence Erlbaum.

Wong, W. (2005). *Input Enhancement: From Theory and Research to Classroom*. New York: McGraw-Hill.

Woodman, G. F., and Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 29:121-138.

Wray, A. (2002). Dual processing in protolanguage: competence without performance. In Wray, A. (Ed). *The Transition to Language.* Oxford: Oxford University Press.

Wundt, W. (1973). *The language of gestures*. The Hague: Mouton

Young-Scholten, M. and Strom, N. (2006). First-time adult L2 readers: Is there a critical period? In Van de Craats, I., Kurvers, J. and Young-Scholten, M. (Eds). *Low-educated adult second language and literacy acquisition*, pp. 45-68.Utrecht: LOT.

Xiao, Y. (2002). The syntactic development of school-age Chinese-speaking children learning English. *IRAL*, (40): 235-271.

Websites:

http://www.msharwood.pwp.blueyonder.co.uk/mogul

http://www.hw.ac.uk/langWWW/mogul/index.htm

www.theliteracyinstitute.org/projects/pdf/LESLLA%20Research7.pdf

Appendix A

IP Principles-Complete and revised list of Principles

Principle 1. The Primacy of Meaning Principle. *Learners process input for meaning before they process it for form.*
Principle 1a. The Primacy of Content Words Principle. *Learners process content words in the input before anything else.*
Principle 1b. The Lexical Preference Principle. *Learners will tend to rely on lexical items as opposed to grammatical form to get meaning when both encode the same semantic information.*
Principle 1c. The Preferences for Nonredundancy Principle*. Learners are more likely to process nonredundant meaningful grammatical form before they process redundant meaningful forms.l*
Principle 1d. The Meaning-Before-Nonmeaning Principle*. Learner are more likely to process meaningful grammatical forms before nonmenanigful forms irrespective of redundancy.*
Principle 1e. The availability of Resources Principle. *For learners to process either redundant meaningful grammatical forms or nonmeaningful forms, the processing of overall entential meaning must not drain available processing resources*
Priciple 1f. The Sentence Location Principle. *Learners tend to process items in sentence initial position before those in final position and those in medial position.*
Principle 2. The First Noun Principle*. Learners tend to process the first noun or pronoun they encounter in a sentence as the subject/agent.*
Principle 2a. The Lexical Semnatics Principle. *Learners may relay o lexical semantics, where possible, instead of word oder to interpret sentences.*
Principle 2b. The Event Probabilities Principle. *Learners may rely on event probabilities, where possible, instead of word order to interpret sentences.*
Principle 2c. The Contetual Constraint Principle. *Learners may rely less on the First Noun Princiople if preceeding context constrains the possible interpretation of a clause or sentence.*

(VanPatten, 2004:14, 18)

PI Guidelines

Lee and VanPatten (1995) and VanPatten (1996) provide a set of guidelines in order for instructors to develop treatment materials for PI activities. Despite the fact that VanPatten (1996) proposes that the guidelines should be used 'flexibly' and always according to the learning group (ibid:67) the particular guidelines that follow are very rigid since there is no reference as to where and how one can be more flexible. Another issue remains the fact that the guidelines lack theoretical support. Moreover, they refer to American teaching styles, where traditional instruction in the form of providing lists of paradigms and rules of the target form in one large paradigmatic chart are part of the teaching curriculum (see VanPatten, 2002).

1. Teach only one thing at a time.

It is suggested that presenting one paradigm or rule at a time instead of a list of paradigms and rules as in traditional instruction will be more effective, i.e. teaching the past tense according to traditional instruction would mean that all the forms are presented to the learners at once with the verbs conjugated in one large paradigmatic chart. PI recommends to break up paradigms and present them gradually during the course of the lesson[60], i.e. presenting the third singular person for the past tense, and working with structured input activities before proceeding to another verb form. (VanPatten, 1996; see also PI studies such as VanPatten and Cadierno 1993; VanPatten and Oikennon, 1996, Wong, 2004).Although there are no set guidelines referring to how many paradigms should be used per lesson, the norm in PI is that learners are presented briefly with explicit instruction of the target form, e.g. word order in German and the role of the nominative and accusative case, followed by two examples. In the case of OVS versus SVO sentences in German the examples available will be one sentence with the target form (OVS) and another of the contrasting form (SVO) followed by reminders and suitable explanations (following propositions of the Competition Model, Bates and MacWhinney, 1989).

---

[60] An issue that remains unclear is how many lessons are required for each target structure. Although not explicitly defined the majority of treatments in PI studies vary between an hour and two with an hour and a half being the norm.

2. Keep meaning in focus.

The main purpose of PI and structured input activities is to force learners to make form-meaning connections. In order to achieve this aim, learners have to pay attention to meaning in order to process the form and complete the activity. This means that the "input strings must encode some meaning that the learner is required to attend to and respond to in some way" (VanPatten, 1996:67). In the above example 'Den Mann vermisst das Mädchen' the learner has to process the form of the determiner in order to decode whether it is the man or the girl that misses the other.

3. Learners must do something with the input.

With this guideline VanPatten wants to remind teachers that learners must be actively engaged in processing the input sentences and must also demonstrate that they can respond by stating agreement or disagreement, indicate if that applies to them or not, select alternatives, complete surveys, make an association, complete a name, etc. With the learners' active involvement it should be established whether or not they have attended to the meaning which is contained in the input sentence(s) (VanPatten, 1996:68). Despite the fact that the logic behind this guideline is the importance of learners' active engagement with the input, this guideline is paradoxical, when considering that during structured input activities learners must not produce the target form. This constitutes another weakness of the theoretical underpinnings of PI.

4. Use both oral and written input.

This guideline refers to learners' individual differences. Taking into account that some learners benefit more from written and others from oral input (Skehan, 1998), a combination of the two is, according to VanPatten(1996), more beneficial to input processing.

5. Move from sentences to connected discourse.

According to this guideline, early activities should involve sentence level input, whereas connected discourse, such as listening to conversations and monologues, should be introduced later in the lesson. The reason for this is to provide the learners with the opportunity to focus on and detect the target form while engaging with the structured input activities. According to VanPatten (1993, 1996:69) and Terrell (1991) this is crucial for the initial processing of the target form because of the limited capacity to process incoming data. Moreover, connected discourse may not grant learners adequate 'processing time' as the sentences occur one after another in an extended form (VanPatten, 1996:69).

6. Keep the psycholinguistic processing mechanisms in mind.

This guideline is considered as the most important one, as it "intends to ensure that learners' attention is appropriately guided during structured input activities" (VanPatten, 1996:69). Structured Input activities, according to VanPatten (1996), take into consideration the learners' psycholinguistic processing mechanisms and attempt to alter the reliance strategies that lead them into processing the target features incorrectly or not at all.

Appendix B

Questionnaire

Please fill in the following questionnaire. Where necessary please circle as appropriate.

Name:____

Age: ___________________ Gender: a. Male b. Female

Years studying German:________ Year at School:______________

Hours of German per week (in school): _______________________

Do you or have you received outside tutoring of German? If yes, please provide how many years and hours per week: ____

Name of Course book:_____________________________________

Other languages that you speak/study:_________________________

Native speaker of: ____

Do you or your parents speak a language other than English at home? If so, which language(s): _______________________________________ _____

Group you were allocated: A B C D

How did you like the way the grammar information was presented?

1. Poor
2. Not helpful
3. Not easy to understand/follow
4. Complicated
5. Disliked it (includes 1-5)
6. Good
7. Easy to understand/ follow
8. Very Comprehensive
9. Extremely helpful
10. Excellent (includes 6-10)

Other comments:

Was the grammar information presented?

1. In an unclear way
2. In a non helpful way
3. Non comprehensive way
4. In an enjoyable/fun way
5. In an interesting way

Other comments:

Did it provide any helpful tricks that will help you with grammar?

a. Yes b. No

If yes, please describe what the trick(s) was/were:

Did you notice how definite articles are used in German? If yes, please explain briefly:

How would you rate the activities?

1. Poor
2. Not helpful
3. Not easy to understand/follow
4. Complicated
5. Good
6. Easy to understand/ follow
7. Not complicated
8. Enjoyable/Funny
9. Extremely helpful
10. Excellent

Other comments:

____
____
____
____
____

Were the activities presented?

1. In an unclear way
2. In a non helpful way
3. Non comprehensive way
4. In an enjoyable/fun way
5. In an interesting way

Other comments:

____
____

Vocabulary hard copies for the testing materials

Vocabulary

Affe: monkey (masculine)
an der Kasse: at the till
andere: other
anzufassen: to touch
Apfel: apple (masculine)
ärgert: annoys
Arzt: doctor (masculine)
auch: also
Auto : car (neutral)
Ballon: balloon (masculine)
bezahlt: pays
böse: bad
brauchen: need
Brüder: brother (masculine)
Chef : boss (masculine)
Delfin: dolphin (masculine)
Enkelsohn: grandson (masculine)
fällt: falls
fängt: catches
Feurwehrmann: fire fighter (masculine)
findet: finds
fragt: asks
Frau: lady (feminine)
frisiert: to dress the hair
Frosh: frog (masculine)
Fuchs: fox (masculine)
füttert: feeds
grüßt: greets
hält : holds
Hasen: Hare, bunny (masculine)
heiraten: to get married
Hexe : witch (feminine)
Hirsch: deer (masculine)
hört: hears
immitiert: imitates
isst: eats
Jungen: boy (masculine)
Kaninchen: rabbit (neutral)
Katze: cat (feminine)
Kollege: colleague (masculine)
Kollegin: female colleague
König: king (masculine)
König: king (masculine)
Königssohn: son of the king (masculine)
korrigiert: corrects
küsst: kisses
lacht…aus: makes fun of
langweiligt: bores
Lehrerin: female teacher
lieben: loves
Löwe: lion (masculine)
Mädchen: girl (neutral)
malen:paint
Maus: mouse (feminine)
Mitschüler: classmate (masculine)
näht: sews
Oma: grandmother (feminine)
Onkel: uncle (masculine)
Opa: granddad (masculine)
Pelikan: pelican (masculine)
pflegt: takes care of
Prinzessin:princess (female)
pustet: blows
Rad(fahren): ride the bike
reitet: rides
retten: to save
Sänger : singer (masculine)
schickt: sends
schlägt: hits
Schlange: snake (feminine)
Schmetterling: butterfly (masculine)
schön: beautiful
schreibt: writes
schreit…an: shouts at
Schüler: pupil (masculine)
Sekretärin: secretary
sieht: sees
spricht: talks
Stern: star (masculine)
stoppen: to stop
stört:bothers
strahlend: shiny
sucht: searches
surfen: to surf
Tante: aunt (feminine)
Tasche: bag (feminine)
Tiger: tiger (masculine)
trägt: carrys
unterhält: amuses
untersucht: examines
verfolgt: follows
verliebt sich in + Akk: to fall in love in
vermisst: misses
versucht: attempts, tries
verwandelt [sich verwandeln in + Akk): transforms in
Vogel: bird (masculine)
Warum: why
weckt: wakes up
wegzugehen: to leave
werden: to become
Wolf: wolf (masculine)
Zauberland: magic place (neutral)
zuerst: first
zur Unterwelt: to the underworld
Zwerg: dwarf (masculine)
zwingt: compels, forces, bullies

Appendix C

Overall Performance

Non-Parametric Analysis

Kruskal-Wallis Test

Ranks

	Group	N	Mean Rank
PRE - TEST	1	34	65.12
	2	33	68.41
	3	32	64.34
	4	32	66.11
	Total	131	
IMMEDIATE	1	34	98.03
	2	33	78.42
	3	32	59.53
	4	32	25.63
	Total	131	
DELAYED	1	34	88.88
	2	33	73.56
	3	32	63.98
	4	32	35.91
	Total	131	

Test Statistics[a,b]

	PRE - TEST	IMMEDIATE	DELAYED
Chi-Square	.223	64.949	33.988
df	3	3	3
Asymp. Sig.	.974	.000	.000

a. Kruskal Wallis Test

b. Grouping Variable: Group

Paired Group Analysis

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.07	1124.50
	2	33	34.95	1153.50
	Total	67		
IMMEDIATE	1	34	40.96	1392.50
	2	33	26.83	885.50
	Total	67		
DELAYED	1	34	38.31	1302.50
	2	33	29.56	975.50
	Total	67		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	529.500	324.500	414.500
Wilcoxon W	1124.500	885.500	975.500
Z	-.408	-2.968	-1.840
Asymp. Sig. (2-tailed)	.683	.003	.066

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.60	1142.50
	3	32	33.39	1068.50
	Total	66		
IMMEDIATE	1	34	44.57	1515.50
	3	32	21.73	695.50
	Total	66		
DELAYED	1	34	40.09	1363.00
	3	32	26.50	848.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	540.500	167.500	320.000
Wilcoxon W	1068.500	695.500	848.000
Z	-.046	-4.834	-2.879
Asymp. Sig. (2-tailed)	.963	.000	.004

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.44	1137.00
	4	32	33.56	1074.00
	Total	66		
IMMEDIATE	1	34	47.50	1615.00
	4	32	18.63	596.00
	Total	66		
DELAYED	1	34	45.49	1546.50
	4	32	20.77	664.50
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	542.000	68.000	136.500
Wilcoxon W	1137.000	596.000	664.500
Z	-.026	-6.115	-5.245
Asymp. Sig. (2-tailed)	.979	.000	.000

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	34.02	1122.50
	3	32	31.95	1022.50
	Total	65		
IMMEDIATE	2	33	39.21	1294.00
	3	32	26.59	851.00
	Total	65		
DELAYED	2	33	35.36	1167.00
	3	32	30.56	978.00
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	494.500	323.000	450.000
Wilcoxon W	1022.500	851.000	978.000
Z	-.452	-2.692	-1.025
Asymp. Sig. (2-tailed)	.651	.007	.305

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	33.44	1103.50
	4	32	32.55	1041.50
	Total	65		
IMMEDIATE	2	33	46.38	1530.50
	4	32	19.20	614.50
	Total	65		
DELAYED	2	33	42.64	1407.00
	4	32	23.06	738.00
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	513.500	86.500	210.000
Wilcoxon W	1041.500	614.500	738.000
Z	-.195	-5.802	-4.192
Asymp. Sig. (2-tailed)	.846	.000	.000

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	3	32	32.00	1024.00
	4	32	33.00	1056.00
	Total	64		
IMMEDIATE	3	32	44.20	1414.50
	4	32	20.80	665.50
	Total	64		
DELAYED	3	32	39.92	1277.50
	4	32	25.08	802.50
	Total	64		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	496.000	137.500	274.500
Wilcoxon W	1024.000	665.500	802.500
Z	-.219	-5.043	-3.203
Asymp. Sig. (2-tailed)	.826	.000	.001

a. Grouping Variable: Group

RMA Analysis

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	performance	Type III Sum of Squares	df	Mean Square	F	Sig.
performance	Level 1 vs. Level 2	33387.294	1	33387.294	270.479	.000
	Level 2 vs. Level 3	8712.154	1	8712.154	71.355	.000
performance * Group	Level 1 vs. Level 2	14759.988	3	4919.996	39.858	.000
	Level 2 vs. Level 3	4118.716	3	1372.905	11.244	.000
Error(performance)	Level 1 vs. Level 2	15676.592	127	123.438		
	Level 2 vs. Level 3	15506.261	127	122.097		

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	18862.674	1	18862.674	566.873	.000
Group	3540.852	3	1180.284	35.471	.000
Error	4225.917	127	33.275		

Post Hoc Tests

Group

Multiple Comparisons

Measure:MEASURE_1

	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Scheffe	1	2	4.68*	1.410	.014	.68	8.67
		3	9.26*	1.421	.000	5.24	13.29
		4	13.93*	1.421	.000	9.90	17.95
	2	1	-4.68*	1.410	.014	-8.67	-.68
		3	4.58*	1.431	.019	.53	8.64
		4	9.25*	1.431	.000	5.20	13.31
	3	1	-9.26*	1.421	.000	-13.29	-5.24
		2	-4.58*	1.431	.019	-8.64	-.53
		4	4.67*	1.442	.018	.58	8.75
	4	1	-13.93*	1.421	.000	-17.95	-9.90
		2	-9.25*	1.431	.000	-13.31	-5.20
		3	-4.67*	1.442	.018	-8.75	-.58
Bonferroni	1	2	4.68*	1.410	.007	.90	8.46
		3	9.26*	1.421	.000	5.45	13.07
		4	13.93*	1.421	.000	10.12	17.74
	2	1	-4.68*	1.410	.007	-8.46	-.90
		3	4.58*	1.431	.010	.75	8.42
		4	9.25*	1.431	.000	5.42	13.09
	3	1	-9.26*	1.421	.000	-13.07	-5.45
		2	-4.58*	1.431	.010	-8.42	-.75
		4	4.67*	1.442	.009	.80	8.53
	4	1	-13.93*	1.421	.000	-17.74	-10.12
		2	-9.25*	1.431	.000	-13.09	-5.42
		3	-4.67*	1.442	.009	-8.53	-.80

Based on observed means.

The error term is Mean Square(Error) = 33.275.

*. The mean difference is significant at the .05 level.

Error Correction task
Non-Parametric Analysis
Kruskal-Wallis Test

Ranks

	Group	N	Mean Rank
PRE - TEST	1	34	66.00
	2	33	66.00
	3	32	66.00
	4	32	66.00
	Total	131	
IMMEDIATE	1	34	99.07
	2	33	83.82
	3	32	55.69
	4	32	22.80
	Total	131	
DELAYED	1	34	82.46
	2	33	77.95
	3	32	49.86
	4	32	52.33
	Total	131	

Test Statistics[a,b]

	PRE - TEST	IMMEDIATE	DELAYED
Chi-Square	.000	77.967	19.690
df	3	3	3
Asymp. Sig.	1.000	.000	.000

a. Kruskal Wallis Test

b. Grouping Variable: Group

Paired Group Analysis
Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	34.00	1156.00
	2	33	34.00	1122.00
	Total	67		
IMMEDIATE	1	34	40.31	1370.50
	2	33	27.50	907.50
	Total	67		
DELAYED	1	34	35.26	1199.00
	2	33	32.70	1079.00
	Total	67		

Test Statistics[a]

	New_Error_PRET	New_Error_IPT	DPT_ERROR_NEW
Mann-Whitney U	561.000	346.500	518.000
Wilcoxon W	1122.000	907.500	1079.000
Z	.000	-2.702	-.543
Asymp. Sig. (2-tailed)	1.000	.007	.587

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.50	1139.00
	3	32	33.50	1072.00
	Total	66		
IMMEDIATE	1	34	45.74	1555.00
	3	32	20.50	656.00
	Total	66		
DELAYED	1	34	41.29	1404.00
	3	32	25.22	807.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	544.000	128.000	279.000
Wilcoxon W	1072.000	656.000	807.000
Z	.000	-5.356	-3.444
Asymp. Sig. (2-tailed)	1.000	.000	.001

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.50	1139.00
	4	32	33.50	1072.00
	Total	66		
IMMEDIATE	1	34	48.03	1633.00
	4	32	18.06	578.00
	Total	66		
DELAYED	1	34	48.09	1635.00
	4	32	18.00	576.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	544.000	50.000	48.000
Wilcoxon W	1072.000	578.000	576.000
Z	.000	-6.492	-6.906
Asymp. Sig. (2-tailed)	1.000	.000	.000

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	33.00	1089.00
	3	32	33.00	1056.00
	Total	65		
IMMEDIATE	2	33	42.06	1388.00
	3	32	23.66	757.00
	Total	65		
DELAYED	2	33	40.08	1322.50
	3	32	25.70	822.50
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	528.000	229.000	294.500
Wilcoxon W	1056.000	757.000	822.500
Z	.000	-3.946	-3.110
Asymp. Sig. (2-tailed)	1.000	.000	.002

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	33.00	1089.00
	4	32	33.00	1056.00
	Total	65		
IMMEDIATE	2	33	48.26	1592.50
	4	32	17.27	552.50
	Total	65		
DELAYED	2	33	47.55	1569.00
	4	32	18.00	576.00
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	528.000	24.500	48.000
Wilcoxon W	1056.000	552.500	576.000
Z	.000	-6.752	-6.865
Asymp. Sig. (2-tailed)	1.000	.000	.000

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	3	32	32.50	1040.00
	4	32	32.50	1040.00
	Total	64		
IMMEDIATE	3	32	44.53	1425.00
	4	32	20.47	655.00
	Total	64		
DELAYED	3	32	43.50	1392.00
	4	32	21.50	688.00
	Total	64		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	512.000	127.000	160.000
Wilcoxon W	1040.000	655.000	688.000
Z	.000	-5.363	-5.597
Asymp. Sig. (2-tailed)	1.000	.000	.000

a. Grouping Variable: Group

RMA Analysis

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	error	Type III Sum of Squares	df	Mean Square	F	Sig.
error	Linear	424.536	1	424.536	109.755	.000
	Quadratic	1421.452	1	1421.452	251.065	.000
error * Group	Linear	212.631	3	70.877	18.324	.000
	Quadratic	680.205	3	226.735	40.047	.000
Error(error)	Linear	491.239	127	3.868		
	Quadratic	719.034	127	5.662		

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2693.032	1	2693.032	307.226	.000
Group	1306.754	3	435.585	49.692	.000
Error	1113.235	127	8.766		

Post Hoc Tests

Group

Multiple Comparisons

Measure:MEASURE_1

	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Scheffe	1	2	1.19*	.418	.050	.00	2.37
		3	3.16*	.421	.000	1.97	4.36
		4	4.75*	.421	.000	3.55	5.94
	2	1	-1.19*	.418	.050	-2.37	.00
		3	1.98*	.424	.000	.78	3.18
		4	3.56*	.424	.000	2.36	4.76
	3	1	-3.16*	.421	.000	-4.36	-1.97
		2	-1.98*	.424	.000	-3.18	-.78
		4	1.58*	.427	.004	.37	2.79
	4	1	-4.75*	.421	.000	-5.94	-3.55
		2	-3.56*	.424	.000	-4.76	-2.36
		3	-1.58*	.427	.004	-2.79	-.37
Bonferroni	1	2	1.19*	.418	.032	.07	2.30
		3	3.16*	.421	.000	2.03	4.29
		4	4.75*	.421	.000	3.62	5.87
	2	1	-1.19*	.418	.032	-2.30	-.07
		3	1.98*	.424	.000	.84	3.11
		4	3.56*	.424	.000	2.42	4.70
	3	1	-3.16*	.421	.000	-4.29	-2.03
		2	-1.98*	.424	.000	-3.11	-.84
		4	1.58*	.427	.002	.44	2.73
	4	1	-4.75*	.421	.000	-5.87	-3.62
		2	-3.56*	.424	.000	-4.70	-2.42
		3	-1.58*	.427	.002	-2.73	-.44

Based on observed means.

The error term is Mean Square(Error) = 2.922.

*. The mean difference is significant at the .05 level.

Reading Comprehension Task
Non-Parametric Analysis
Kruskal-Wallis Test

Ranks

	Group	N	Mean Rank
PRE - TEST	1	34	67.19
	2	33	69.62
	3	32	62.06
	4	32	64.94
	Total	131	
IMMEDIATE	1	34	82.12
	2	33	64.03
	3	32	65.06
	4	32	51.84
	Total	131	
DELAYED	1	34	73.31
	2	33	61.83
	3	32	61.25
	4	32	67.28
	Total	131	

Test Statistics[a,b]

	PRE - TEST	IMMEDIATE	DELAYED
Chi-Square	.737	11.046	2.262
df	3	3	3
Asymp. Sig.	.864	.011	.520

a. Kruskal Wallis Test

b. Grouping Variable: Group

Paired Group Analysis
Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.28	1131.50
	2	33	34.74	1146.50
	Total	67		
IMMEDIATE	1	34	38.18	1298.00
	2	33	29.70	980.00
	Total	67		
DELAYED	1	34	36.81	1251.50
	2	33	31.11	1026.50
	Total	67		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	536.500	419.000	465.500
Wilcoxon W	1131.500	980.000	1026.500
Z	-.316	-1.811	-1.214
Asymp. Sig. (2-tailed)	.752	.070	.225

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	34.74	1181.00
	3	32	32.19	1030.00
	Total	66		
IMMEDIATE	1	34	37.96	1290.50
	3	32	28.77	920.50
	Total	66		
DELAYED	1	34	36.62	1245.00
	3	32	30.19	966.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	502.000	392.500	438.000
Wilcoxon W	1030.000	920.500	966.000
Z	-.552	-1.990	-1.384
Asymp. Sig. (2-tailed)	.581	.047	.166

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	34.18	1162.00
	4	32	32.78	1049.00
	Total	66		
IMMEDIATE	1	34	40.99	1393.50
	4	32	25.55	817.50
	Total	66		
DELAYED	1	34	34.88	1186.00
	4	32	32.03	1025.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	521.000	289.500	497.000
Wilcoxon W	1049.000	817.500	1025.000
Z	-.304	-3.320	-.612
Asymp. Sig. (2-tailed)	.761	.001	.541

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	34.73	1146.00
	3	32	31.22	999.00
	Total	65		
IMMEDIATE	2	33	32.73	1080.00
	3	32	33.28	1065.00
	Total	65		
DELAYED	2	33	32.92	1086.50
	3	32	33.08	1058.50
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	471.000	519.000	525.500
Wilcoxon W	999.000	1080.000	1086.500
Z	-.767	-.120	-.033
Asymp. Sig. (2-tailed)	.443	.904	.973

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	34.15	1127.00
	4	32	31.81	1018.00
	Total	65		
IMMEDIATE	2	33	35.61	1175.00
	4	32	30.31	970.00
	Total	65		
DELAYED	2	33	31.80	1049.50
	4	32	34.23	1095.50
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	490.000	442.000	488.500
Wilcoxon W	1018.000	970.000	1049.500
Z	-.511	-1.147	-.526
Asymp. Sig. (2-tailed)	.609	.251	.599

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	3	32	31.66	1013.00
	4	32	33.34	1067.00
	Total	64		
IMMEDIATE	3	32	36.02	1152.50
	4	32	28.98	927.50
	Total	64		
DELAYED	3	32	30.98	991.50
	4	32	34.02	1088.50
	Total	64		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	485.000	399.500	463.500
Wilcoxon W	1013.000	927.500	991.500
Z	-.372	-1.543	-.662
Asymp. Sig. (2-tailed)	.710	.123	.508

a. Grouping Variable: Group

RMA Analysis

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	compehension	Type III Sum of Squares	df	Mean Square	F	Sig.
compehension	Level 1 vs. Level 2	506.771	1	506.771	34.401	.000
	Level 2 vs. Level 3	114.661	1	114.661	6.440	.012
compehension * Group	Level 1 vs. Level 2	109.100	3	36.367	2.469	.065
	Level 2 vs. Level 3	69.538	3	23.179	1.302	.277
Error(compehension)	Level 1 vs. Level 2	1870.869	127	14.731		
	Level 2 vs. Level 3	2261.088	127	17.804		

Tests of Between-Subjects Effects

Measure:MEASURE_1
Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2825.717	1	2825.717	648.429	.000
Group	26.103	3	8.701	1.997	.118
Error	553.439	127	4.358		

Post Hoc Tests

Group

Multiple Comparisons

Measure:MEASURE_1

	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Scheffe	1	2	.85	.510	.434	-.60	2.29
		3	1.02	.514	.276	-.44	2.47
		4	1.12	.514	.196	-.34	2.58
	2	1	-.85	.510	.434	-2.29	.60
		3	.17	.518	.991	-1.30	1.64
		4	.27	.518	.963	-1.19	1.74
	3	1	-1.02	.514	.276	-2.47	.44
		2	-.17	.518	.991	-1.64	1.30
		4	.10	.522	.998	-1.37	1.58
	4	1	-1.12	.514	.196	-2.58	.34
		2	-.27	.518	.963	-1.74	1.19
		3	-.10	.522	.998	-1.58	1.37
Bonferroni	1	2	.85	.510	.597	-.52	2.21
		3	1.02	.514	.300	-.36	2.40
		4	1.12	.514	.186	-.26	2.50
	2	1	-.85	.510	.597	-2.21	.52
		3	.17	.518	1.000	-1.22	1.56
		4	.27	.518	1.000	-1.11	1.66
	3	1	-1.02	.514	.300	-2.40	.36
		2	-.17	.518	1.000	-1.56	1.22
		4	.10	.522	1.000	-1.29	1.50
	4	1	-1.12	.514	.186	-2.50	.26
		2	-.27	.518	1.000	-1.66	1.11
		3	-.10	.522	1.000	-1.50	1.29

Based on observed means.
The error term is Mean Square(Error) = 4.358.

Fill in the gap task

Non-Parametric Analysis

Kruskal-Wallis Test

Ranks

	Group	N	Mean Rank
PRE - TEST	1	34	66.00
	2	33	66.00
	3	32	66.00
	4	32	66.00
	Total	131	
IMMEDIATE	1	34	96.03
	2	33	74.53
	3	32	56.80
	4	32	34.50
	Total	131	
DELAYED	1	34	84.96
	2	33	65.05
	3	32	62.34
	4	32	50.50
	Total	131	

Test Statistics[a,b]

	PRE - TEST	IMMEDIATE	DELAYED
Chi-Square	.000	54.521	25.462
df	3	3	3
Asymp. Sig.	1.000	.000	.000

a. Kruskal Wallis Test

b. Grouping Variable: Group

Paired Group Analysis

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	34.00	1156.00
	2	33	34.00	1122.00
	Total	67		
IMMEDIATE	1	34	40.69	1383.50
	2	33	27.11	894.50
	Total	67		
DELAYED	1	34	38.87	1321.50
	2	33	28.98	956.50
	Total	67		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	561.000	333.500	395.500
Wilcoxon W	1122.000	894.500	956.500
Z	.000	-2.874	-2.392
Asymp. Sig. (2-tailed)	1.000	.004	.017

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.50	1139.00
	3	32	33.50	1072.00
	Total	66		
IMMEDIATE	1	34	42.72	1452.50
	3	32	23.70	758.50
	Total	66		
DELAYED	1	34	39.12	1330.00
	3	32	27.53	881.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	544.000	230.500	353.000
Wilcoxon W	1072.000	758.500	881.000
Z	.000	-4.140	-2.846
Asymp. Sig. (2-tailed)	1.000	.000	.004

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.50	1139.00
	4	32	33.50	1072.00
	Total	66		
IMMEDIATE	1	34	47.62	1619.00
	4	32	18.50	592.00
	Total	66		
DELAYED	1	34	41.97	1427.00
	4	32	24.50	784.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	544.000	64.000	256.000
Wilcoxon W	1072.000	592.000	784.000
Z	.000	-6.732	-4.712
Asymp. Sig. (2-tailed)	1.000	.000	.000

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	33.00	1089.00
	3	32	33.00	1056.00
	Total	65		
IMMEDIATE	2	33	37.76	1246.00
	3	32	28.09	899.00
	Total	65		
DELAYED	2	33	33.67	1111.00
	3	32	32.31	1034.00
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	528.000	371.000	506.000
Wilcoxon W	1056.000	899.000	1034.000
Z	.000	-2.197	-.413
Asymp. Sig. (2-tailed)	1.000	.028	.679

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	33.00	1089.00
	4	32	33.00	1056.00
	Total	65		
IMMEDIATE	2	33	43.67	1441.00
	4	32	22.00	704.00
	Total	65		
DELAYED	2	33	36.39	1201.00
	4	32	29.50	944.00
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	528.000	176.000	416.000
Wilcoxon W	1056.000	704.000	944.000
Z	.000	-5.481	-2.731
Asymp. Sig. (2-tailed)	1.000	.000	.006

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	3	32	32.50	1040.00
	4	32	32.50	1040.00
	Total	64		
IMMEDIATE	3	32	38.00	1216.00
	4	32	27.00	864.00
	Total	64		
DELAYED	3	32	35.50	1136.00
	4	32	29.50	944.00
	Total	64		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	512.000	336.000	416.000
Wilcoxon W	1040.000	864.000	944.000
Z	.000	-3.595	-2.549
Asymp. Sig. (2-tailed)	1.000	.000	.011

a. Grouping Variable: Group

RMA Analysis

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	production	Type III Sum of Squares	df	Mean Square	F	Sig.
production	Level 1 vs. Level 2	1809.503	1	1809.503	106.150	.000
	Level 2 vs. Level 3	611.474	1	611.474	34.041	.000
production * Group	Level 1 vs. Level 2	1087.092	3	362.364	21.257	.000
	Level 2 vs. Level 3	333.548	3	111.183	6.190	.001
Error(production)	Level 1 vs. Level 2	2164.924	127	17.047		
	Level 2 vs. Level 3	2281.292	127	17.963		

Tests of Between-Subjects Effects

Measure:MEASURE_1
Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	404.660	1	404.660	89.992	.000
Group	253.006	3	84.335	18.755	.000
Error	571.068	127	4.497		

Post Hoc Tests

Group

Multiple Comparisons

Measure:MEASURE_1

	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Scheffe	1	2	1.78*	.518	.010	.32	3.25
		3	2.60*	.522	.000	1.12	4.08
		4	3.80*	.522	.000	2.32	5.28
	2	1	-1.78*	.518	.010	-3.25	-.32
		3	.81	.526	.500	-.68	2.30
		4	2.02*	.526	.003	.53	3.51
	3	1	-2.60*	.522	.000	-4.08	-1.12
		2	-.81	.526	.500	-2.30	.68
		4	1.21	.530	.164	-.29	2.71
	4	1	-3.80*	.522	.000	-5.28	-2.32
		2	-2.02*	.526	.003	-3.51	-.53
		3	-1.21	.530	.164	-2.71	.29
Bonferroni	1	2	1.78*	.518	.005	.39	3.17
		3	2.60*	.522	.000	1.20	4.00
		4	3.80*	.522	.000	2.40	5.20
	2	1	-1.78*	.518	.005	-3.17	-.39
		3	.81	.526	.752	-.60	2.22
		4	2.02*	.526	.001	.61	3.43
	3	1	-2.60*	.522	.000	-4.00	-1.20
		2	-.81	.526	.752	-2.22	.60
		4	1.21	.530	.146	-.21	2.63
	4	1	-3.80*	.522	.000	-5.20	-2.40
		2	-2.02*	.526	.001	-3.43	-.61
		3	-1.21	.530	.146	-2.63	.21

Based on observed means.

The error term is Mean Square(Error) = 4.497.

*. The mean difference is significant at the .05 level.

Interpretation Task
Non-Parametric Analysis
Kruskal-Wallis Test

Ranks

	Group	N	Mean Rank
PRE - TEST	1	34	61.28
	2	33	64.95
	3	32	72.00
	4	32	66.09
	Total	131	
IMMEDIATE	1	34	90.62
	2	33	81.06
	3	32	56.03
	4	32	34.28
	Total	131	
DELAYED	1	34	78.22
	2	33	71.79
	3	32	69.52
	4	32	43.53
	Total	131	

Test Statistics[a,b]

	PRE - TEST	IMMEDIATE	DELAYED
Chi-Square	2.894	47.382	18.855
df	3	3	3
Asymp. Sig.	.408	.000	.000

a. Kruskal Wallis Test

b. Grouping Variable: Group

Paired Group Analysis

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	33.03	1123.00
	2	33	35.00	1155.00
	Total	67		
IMMEDIATE	1	34	35.96	1222.50
	2	33	31.98	1055.50
	Total	67		
DELAYED	1	34	36.01	1224.50
	2	33	31.92	1053.50
	Total	67		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	528.000	494.500	492.500
Wilcoxon W	1123.000	1055.500	1053.500
Z	-.670	-.880	-.910
Asymp. Sig. (2-tailed)	.503	.379	.363

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	30.90	1050.50
	3	32	36.27	1160.50
	Total	66		
IMMEDIATE	1	34	42.57	1447.50
	3	32	23.86	763.50
	Total	66		
DELAYED	1	34	36.18	1230.00
	3	32	30.66	981.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	455.500	235.500	453.000
Wilcoxon W	1050.500	763.500	981.000
Z	-1.641	-4.067	-1.220
Asymp. Sig. (2-tailed)	.101	.000	.222

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	1	34	32.35	1100.00
	4	32	34.72	1111.00
	Total	66		
IMMEDIATE	1	34	47.09	1601.00
	4	32	19.06	610.00
	Total	66		
DELAYED	1	34	41.03	1395.00
	4	32	25.50	816.00
	Total	66		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	505.000	82.000	288.000
Wilcoxon W	1100.000	610.000	816.000
Z	-.803	-6.196	-3.821
Asymp. Sig. (2-tailed)	.422	.000	.000

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	31.23	1030.50
	3	32	34.83	1114.50
	Total	65		
IMMEDIATE	2	33	39.26	1295.50
	3	32	26.55	849.50
	Total	65		
DELAYED	2	33	34.14	1126.50
	3	32	31.83	1018.50
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	469.500	321.500	490.500
Wilcoxon W	1030.500	849.500	1018.500
Z	-1.048	-2.797	-.518
Asymp. Sig. (2-tailed)	.295	.005	.604

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	2	33	32.73	1080.00
	4	32	33.28	1065.00
	Total	65		
IMMEDIATE	2	33	43.82	1446.00
	4	32	21.84	699.00
	Total	65		
DELAYED	2	33	39.73	1311.00
	4	32	26.06	834.00
	Total	65		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	519.000	171.000	306.000
Wilcoxon W	1080.000	699.000	834.000
Z	-.175	-4.993	-3.460
Asymp. Sig. (2-tailed)	.861	.000	.001

a. Grouping Variable: Group

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
PRE - TEST	3	32	33.91	1085.00
	4	32	31.09	995.00
	Total	64		
IMMEDIATE	3	32	38.63	1236.00
	4	32	26.38	844.00
	Total	64		
DELAYED	3	32	40.03	1281.00
	4	32	24.97	799.00
	Total	64		

Test Statistics[a]

	PRE - TEST	IMMEDIATE	DELAYED
Mann-Whitney U	467.000	316.000	271.000
Wilcoxon W	995.000	844.000	799.000
Z	-.818	-2.914	-3.733
Asymp. Sig. (2-tailed)	.414	.004	.000

a. Grouping Variable: Group

RMA Analysis

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	interpretation	Type III Sum of Squares	df	Mean Square	F	Sig.
interpretation	Level 1 vs. Level 2	3200.118	1	3200.118	180.116	.000
	Level 2 vs. Level 3	691.525	1	691.525	25.248	.000
interpretation * Group	Level 1 vs. Level 2	1498.523	3	499.508	28.114	.000
	Level 2 vs. Level 3	346.812	3	115.604	4.221	.007
Error(interpretation)	Level 1 vs. Level 2	2256.408	127	17.767		
	Level 2 vs. Level 3	3478.409	127	27.389		

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1171.401	1	1171.401	302.783	.000
Group	345.340	3	115.113	29.754	.000
Error	491.335	127	3.869		

Post Hoc Tests

Group

Multiple Comparisons

Measure:MEASURE_1

	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Scheffe	1	2	.86	.481	.363	-.50	2.22
		3	2.44*	.484	.000	1.07	3.82
		4	4.26*	.484	.000	2.88	5.63
	2	1	-.86	.481	.363	-2.22	.50
		3	1.58*	.488	.017	.20	2.97
		4	3.40*	.488	.000	2.01	4.78
	3	1	-2.44*	.484	.000	-3.82	-1.07
		2	-1.58*	.488	.017	-2.97	-.20
		4	1.81*	.492	.005	.42	3.21
	4	1	-4.26*	.484	.000	-5.63	-2.88
		2	-3.40*	.488	.000	-4.78	-2.01
		3	-1.81*	.492	.005	-3.21	-.42
Bonferroni	1	2	.86	.481	.451	-.43	2.15
		3	2.44*	.484	.000	1.15	3.74
		4	4.26*	.484	.000	2.96	5.56
	2	1	-.86	.481	.451	-2.15	.43
		3	1.58*	.488	.009	.27	2.89
		4	3.40*	.488	.000	2.09	4.70
	3	1	-2.44*	.484	.000	-3.74	-1.15
		2	-1.58*	.488	.009	-2.89	-.27
		4	1.81*	.492	.002	.49	3.13
	4	1	-4.26*	.484	.000	-5.56	-2.96
		2	-3.40*	.488	.000	-4.70	-2.09
		3	-1.81*	.492	.002	-3.13	-.49

Based on observed means.

The error term is Mean Square(Error) = 3.869.

Multiple Comparisons

Measure:MEASURE_1

	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Scheffe	1	2	.86	.481	.363	-.50	2.22
		3	2.44*	.484	.000	1.07	3.82
		4	4.26*	.484	.000	2.88	5.63
	2	1	-.86	.481	.363	-2.22	.50
		3	1.58*	.488	.017	.20	2.97
		4	3.40*	.488	.000	2.01	4.78
	3	1	-2.44*	.484	.000	-3.82	-1.07
		2	-1.58*	.488	.017	-2.97	-.20
		4	1.81*	.492	.005	.42	3.21
	4	1	-4.26*	.484	.000	-5.63	-2.88
		2	-3.40*	.488	.000	-4.78	-2.01
		3	-1.81*	.492	.005	-3.21	-.42
Bonferroni	1	2	.86	.481	.451	-.43	2.15
		3	2.44*	.484	.000	1.15	3.74
		4	4.26*	.484	.000	2.96	5.56
	2	1	-.86	.481	.451	-2.15	.43
		3	1.58*	.488	.009	.27	2.89
		4	3.40*	.488	.000	2.09	4.70
	3	1	-2.44*	.484	.000	-3.74	-1.15
		2	-1.58*	.488	.009	-2.89	-.27
		4	1.81*	.492	.002	.49	3.13
	4	1	-4.26*	.484	.000	-5.56	-2.96
		2	-3.40*	.488	.000	-4.70	-2.09
		3	-1.81*	.492	.002	-3.13	-.49

Based on observed means.

The error term is Mean Square(Error) = 3.869.

*. The mean difference is significant at the .05 level.