Polish Studies in English Language and Literature
Edited by Jacek Fisiak

Anna Turula

Form-Focused Instruction and the Advanced Language Learner

On the Importance of the Semantics of Grammar

33

PETER LANG
Internationaler Verlag der Wissenschaften

Form-Focused Instruction and the Advanced Language Learner looks at the role of FFI at higher levels of foreign language learning. It argues that – contrary to a common belief – there *are* aspects of grammar to be taught to the proficient FL user. While such a learner may be familiar with formal properties of different structures, (s)he still needs to focus on the semantics of target language grammar. Considering this, the book investigates the efficiency of a FFI treatment called *Organic Approach Deductivised* or *3-D language pedagogy*, devised to teach the semantics of the English tense-and-aspect system to the advanced EFL learner. In doing so, the book takes the reader through different aspects of focus on form, looking at the semantics of the English time talk from the cognitive perspective.

Anna Turula holds a PhD in applied linguistics. She has more than 20 years experience in teaching English as a foreign language and has worked in TEFL teacher training for more than 10 years. At present, she is head of the Department of Modern Languages and Literature and a teacher trainer at the College of Foreign Languages, Czestochowa (Poland).

www.peterlang.de

Form-Focused Instruction and the Advanced Language Learner

Polish Studies in English Language and Literature

Edited by Jacek Fisiak

Vol. 33

PETER LANG
Frankfurt am Main · Berlin · Bern · Bruxelles · New York · Oxford · Wien

Anna Turula

Form-Focused Instruction and the Advanced Language Learner

On the Importance of the Semantics of Grammar

PETER LANG
Internationaler Verlag der Wissenschaften

Bibliographic Information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the internet at http://dnb.d-nb.de.

The publication was financially supported
by the College of Foreign Languages Czestochowa.

ISBN 978-3-653-01876-9 (eBook)
ISSN 1436-7513
ISBN 978-3-631-60749-7

www.peterlang.de

mojemu Mężowi, który musiał chodzić na palcach i nie mógł oddychać (ale wytrzymał)
moim Synom, których problemy musiały czekać, aż skończę zdanie
moim Rodzicom, do których dzwoniłam zdecydowanie za rzadko
i moim Przyjaciołom, którzy myśleli, że o nich zapomniałam
@

Contents

ACKNOWLEDGEMENTS

I would like to thank all of those without whom the present research effort could have never been implemented and brought to a successful completion.

First and foremost, I want to express my gratitude to my students from the E1, E2 and C groups for taking part in all of the studies, with a special thank you to the two experimental groups for their willingness to cooperate in a study project that went beyond their curricular requirements.

I would also like to thank all of the native speakers who participated in the interviews and provided grammaticality judgments. They were (in the alphabetical order) Richard Flack, Carl Humphries, Tom Kay, James Lambie, Robert Lawler, Tom Linton, Jonathan Lowe, Lindsay Pinnell, Eva Piotrowska, Russell Preston, Hannah Richardson, Michele Simmons, Yahya Tamezouit, Scott Tattam and Christopher Walker. In addition to this, my most sincere appreciation is offered to Michele Simmons for proofreading my work and making sure it sounds decently English.

Finally, a word of thank you to fellow academics who supported me in my research: to Bogush Bierwiaczonek for his inspiration and guidance; to Ewa Piechurska-Kuciel and Joanna Nijakowska for encouraging my efforts in statistics and for all the patience they showed me; and to Marek Oziewicz and Roman Ociepa for their help concerning various research practicalities. I would also like to express my gratitude to my reviewers, professor Kamila Turewicz and professor Mirosław Pawlak, for all their comments, which helped me improve the book.

Last but not least, a special appreciation to my friend Bożena Pawłowska for all of the warm support, encouragement and the shoulder to cry on she was always ready to offer.

INTRODUCTION

"There is no grammar to teach to the advanced learner," is a frequently made assertion based on the understanding that at higher levels of language proficiency formal instruction is reduced to mere recycling of what has already been successfully acquired. However, while it is true that the learned structures actually get automatised with exposure and practice, their retrieval and application becoming increasingly fluent, many non-native (NNS) advanced target language users have the feeling that their choices as regards form/context connection are very much unlike those of a native speaker (NS). This feeling is reinforced by exposure to authentic target language discourse in which numerous examples of use seem to contradict the automatised form-function mappings learned from respectable tutors and grammar books over the years of target language education. Such a perceivable NS/NNS gap is highly frustrating for the proficient language user, who expects to be near-native-like in every language area, grammar included. This feeling of dissatisfaction is further strengthened by the fact that an effective solution to the problem does not seem to be offered by the traditional grammar instruction.

The research described in the present work, rooted in the above-described frustration, was undertaken in search of appropriate grammar instruction, instruction that could help to level the said NS/NNS gap in the advanced language learner. Considering the fact that in many examples of use the difference amounts to native speakers going beyond the rules to which non-natives have been taught to adhere, it has to be hypothesised that, to bring desirable results, the pedagogy at issue has to exceed purely formal tuition and enter into the sphere of bending the forms towards the meaning intended by the speaker. Such a domain, in which grammatical forms subserve the speaker's intentions, is the semantics of grammar and it – contrary to the popular belief quoted above – constitutes an aspect of grammar to be taught to the advanced language learner. The three studies presented in the present work were carried out with the intention of confirming this claim; they also sought the answer to the question of how teaching grammar semantics can be implemented effectively. To be precise, it was investigated to what extent form-focused instruction can be applied to the successful teaching of form-meaning mappings at higher levels of language proficiency. In addition to this, the studies looked at the quality and the dynamics of such change.

Paving the way for the research chapters of the book are two theoretical parts, in which focus on form is analysed from two perspectives: psychological and psycholinguistic (Part 1) as well as linguistic (Part 2).

Part 1, entitled *FOCUS on form*, deals with two aspects of the uppercased focus: cognitive processing with special regard to attention (Chapter 1) as well as pedagogic procedures (Chapter 2) known as focus on form or form-focused

instruction. Chapter 1 looks at attention through its two main characteristics, selectivity and conscious awareness. While the pedagogic importance of both is fully acknowledged based on relevant research in both psychology and language acquisition, the chapter argues for a fuller recognition of the importance of global attention and the role of f-conscious processes in learning. In doing so, it examines models of attention in relation to working memory, the latter seen as spreading activation patterns in the LTM store. Some constraints on attention, with special regard to frequency and reliability are discussed as well. Hinged on the global/focal, f-conscious/conscious distinction – discussed in Chapter 1 and treated as extremes on a certain continuum rather than oppositions – Chapter 2 takes the reader through language acquisition/learning approaches, starting, respectively, from the more non-interventionist and moving towards the more teacher-fronted and explicit ones. Alongside the different instructional modes, the two types of learning – implicit and explicit – together with the properties of the resulting knowledge are discussed, with special regard to the problem of interface between the two and their mutual, manifold interplay, which, it seems, have to be taken into account in any attempt at effective language pedagogy.

Part 2 – under the title of *Focus on FORM* – is an attempt to linguistically delineate the object of the instruction whose pedagogic bases were discussed in Part 1. Since the focus is actually on the *meaning* of form, the area of investigation is grammar semantics. That is why, first of all, Part 2 looks at the broad meaning of meaning, describing it as encyclopeadic rather than dictionary. Such meaning is best interpreted within a conceptual architecture whose structure is built by various linguistic means and is filled in by holistic semantic models, such as frames or scripts, and in relation to the processes these models are subject to, including frame shifting and conceptual blending. Chapter 3, which describes this, creates the background for Chapter 4, in which all of the said theoretical prerequisites are applied to the area of language learning that is of special interest to the present work: the English tense and aspect or the so-called time talk.

It has to be admitted that in both these parts the book, whose main focus is foreign language pedagogy, trespasses on foreign territory: psychology and theoretical linguistics. This, however, seemed to be a necessary step to take as one cannot address the problem of teaching grammar semantics without analysing the cognitive process of attention which is responsible for learning as well as delineating the object of such instruction by means of defining semantics per se. In both fields – psychology and linguistics – the defining word is *cognitive*. This is why, even in the absence of overt reference, the reflections presented in parts 1 and 2 are to a considerable degree motivated by the findings of Bruner (1956 and later works), Damasio (1994, 1999), Pinker (1994 and later works) as well as (Wierzbicka 1988, to whose publication I refer in the subtitle of my book), Langacker (1987 and later works), Jackendoff (1975 and later works), Radden and Dirven (2007) and many others.

The pedagogical solutions as regards form-focus on the semantics of the English temporal expressions are presented in Part 3. Its first chapter, Chapter 5, which deals with the research design, starts with a description of research to date into the acquisition of the English tense and aspect. The historical perspective, in addition to offering the diachronic view on the area studied as well as relevant research contexts, helps to pinpoint the issues which need to be addressed in any attempt to design an experimental treatment: the linguistic framework within which the pedagogical intervention can be implemented; and the learner specificity of the treatment, with special regard to ID-determined needs and potential areas of difficulty. Since it is proposed that time talk constitutes a full-structure system rather than a random collection of form-meaning connections, construction grammar with its form-meaning pairings and inheritance hierarchies is proposed as a linguistic point of reference. When it comes to the learner, the proficient student, a highly neglected species, and their needs are in focus. It is speculated that since the gap between such near-native users and native speakers of the target language is caused by conceptual transfer as a result of which certain form-meaning pairings are subject to native-language bias, the necessary pedagogic intervention needs to provide opportunities for conceptual refocusing. Chapter 5 concludes with a presentation of a relevant treatment called the *Organic Approach Deductivised* or *3-D grammar* pedagogy, which hopefully fulfils the above requirement.

The said treatment was put to the test in three experimental studies whose implementation, results, data analysis, conclusions and implications are presented in Chapter 6. However, what needs to be pointed out is that these studies will be accommodated within the research tradition delineated in Chapter 5 in two more ways. First of all, in spite of the fact that the final results of the *Organic Approach Deductivised* will be given and discussed, the main focus of the present work, as in the case of other time talk studies (cf. Bardovi-Harlig 2000), is going be on the cognitive processes accompanying the implementation of the treatment rather than the product of this treatment; on emergence rather than acquisition. Secondly – and very much in relation to the assertion above – the three studies presented in Chapter 6, like the majority of the said research to date, will be small-scale, qualitative rather than quantitative, and based on data collected from a small number of testees through tests and verbal reports. The decisions underlying the research design have been made in the belief that while large-format research efforts provide us with invaluable statistically significant data, small experimental studies are a worthy supplement offering insights into phenomena which would otherwise remain uncovered.

The final part of the book offers a number of conclusions and teaching implications based on the research presented in Part III.

PART I: *FOCUS* ON FORM

This part of the book deals with focus on form emphasising the *focus* part of this pedagogical phenomenon. This is why it seems important to clarify what *focus* may mean. There are two educationally sound interpretational possibilities: focusing on the part of the learner, which amounts to this learner paying attention to what is supposed to be learned; or focusing understood as instruction offered to this learner.

The above-mentioned dual interpretation of *focus* is represented by Chapters 1 and 2: the former looks at attention, its two main aspects – selectivity and conscious awareness – and tries to relate the two to learning in general and foreign language education in particular; the second chapter, in turn, examines instruction in foreign language learning. All these considerations are aimed at answering the question of whether to learn a foreign language the learner needs to *consciously* pay *focal* attention to the object of study; a question that cannot be ignored when reflecting upon one of the most important postulates related to the role of attention in language learning: Schmidt's Noticing Hypothesis (Schmidt 1990 and later works) in its strong version in the light of which allocating attentional resources fully awares and in the spotlight fashion is condition *sine qua non* of learning.

Although – based on common sense as well as research to date (cited in both chapters of Part I) – it is difficult to deny the importance of conscious, selective attention in language learning, the present book argues for a middle-of-the-road position, closer to the weaker version of the Noticing Hypothesis, in accordance with which intensive attention is mandatory in some cases and merely facilitative in others. Such an understanding of learning can accommodate both intensive focal as well as extensive, global attention with its different pools activated with varying degrees of awareness. All this means that in learning selective and conscious attention needs to be paid in the spotlight manner but there are a number of reservations to be kept in mind. First of all, noticing does not always happen in a 0-1 fashion: as in the case of real spotlights, there are degrees of light and shade, closer to the centre of vision – understood both literally and metaphorically – or neighbouring on its peripheries. In language learning this will mean that while paying focal attention to certain aspects of language, the learner is likely to simultaneously record other features present in the input. Secondly, just as the spotlight can be moved from place to place, the learner can focus and refocus. Such attention shifts are either caused by conscious registration of incoming stimuli or induced without the subject's full awareness. Similarly in language learning, one will have mastered certain areas of grammar due to clues present in the input (called input enhancement) or deductive teacher-fronted instruction; yet, one may as well spot them owing to the considerable frequency of certain structures in the language material that is encountered. In the latter case, while the very act of noticing may be

conscious, the paths leading to the final illumination will most certainly go through territories of the unconscious, outside the main stage of events (as in Baars' Global Theatre metaphor, discussed thoroughly in Chapter 1). In fact, as will be shown in this part of the book, in human processing of input the selective, the linear and the conscious are in constant interplay with the global, the parallel and the unconscious, as argued by dual-process theories. Such interactive metal operations are also the case in language learning, as will be shown based on multiple examples of phonological, morphological, syntactic and semantic processing. All this is argued for in Chapter 1 in relation to both the native tongue and the second/foreign language in late bilinguals. It is also made clear that, selective and conscious though it has to be at lower levels, language processing becomes more parallel, conscious/-unconscious as the learner develops and automatises their language skills.

Such dual-process specificity of *FOCUS* ON FORM is also subscribed to throughout Chapter 2, which looks at how the teacher may draw the learner's attention to formal aspects of language in ways that are particularly conducive to learning. Research to date (cf. Doughty and Williams 2004; Ellis 2002 and other works) shows that the best result comes if focus on form, explicit or implicit, is combined with communicative exposure to authentic language. That such a combine is effective pedagogically is hardly surprising if we remember the focal/-global, intensive/extensive, conscious/unconscious processing modes, relate the former characteristic of each pair to instruction. What is also important is that these two types of processing are mutually reinforcing rather than mutually exclusive, which cannot be without its influence on the success of the said approach.

At the same time, however, it will be too simplistic to assume that the focal, selective, intensive and conscious instruction will only lead to conscious declarative learning while the exposure will result in global extensive processing and subsequent not-fully-conscious intake. In classroom practice the exposure component combined with a multiplicity of teaching options will lead to different learning modes, explicit or implicit (both discussed in Chapter 2). This is the reason why the present work chooses to see language instruction as a kind of continuum of options rather than in terms of binary oppositions, interaction/instruction (Ellis 2002) or non-interventionist/interventionist (Pawlak 2006). Another reason for such an understanding of language instruction is that we cannot really divide instructional approaches into purely interventionist or non-interventionist; we rather place them on a less-to-more interventionist scale. Chapter 2 argues for such an interpretation of language teaching options, reflecting on which formal aspects of language are best taught/learned in each of the possible modes, especially in the case of the advanced language learner, who is of primary interest to the present work.

CHAPTER ONE
ON THE IMPORTANCE OF NOTICING: ATTENTION

By definition, attention is a cognitive process of selectively concentrating on one thing while ignoring the others**.** For the present book, this mental sieve is of interest in relation to language learning, with special regard to Schmidt's Noticing Hypothesis (Schmidt 1990 and later works), in the light of which intake is the part of input on which the learner has *selectively concentrated*, while whatever was filtered out or *ignored* cannot be considered learned. However, in order to be able to critically reconsider Schmidt's postulate and either subscribe to it or choose its weaker version, we need to look at attention as such, with special regard to its selectivity and the consciousness of the process.

Both said aspects of attention are far from one-dimensional: there are a number of issues to be considered in relation to each of them. In the case of selectivity, we have to investigate its timing and properties as well as the related issues of parallel as opposed to linear one-track processing; contemplating consciousness we need to reflect upon whether such a cognitive filtering and focusing process has to involve only conscious mental states. The present chapter starts by clarifying the main concepts while looking at attention both diachronically and synchronically.

Moving along the former line, Section 1 traces the interest in attention back to the 1960s, examining such attentional correlates as: processes including sensory activation as well as the problem of the selectivity of these processes; the relationship between attention and the affective domain; attention and other behavioural or cognitive processes such as simultaneous performance of different actions or memory; and finally, the question of consciousness. The section concludes on a more synchronic note, with the presentation of the contemporary views on attention, with special regard to the role of working memory in the process of cognitively filtering and registering incoming stimuli

Section 2 discusses the question of consciousness, presenting different approaches to the issue from the perspective of the so-called easy and difficult problem of consciousness (Chalmers 2002). Various stances, motivated by both neuroscience and language philosophy are discussed in this part.

In both of these parts the relations between attention and language learning and processing are initially considered, to be fully dwelled upon in Section 3, which relates both aspects of attention – selectivity and consciousness – firstly to the native tongue phenomena, and then to language learning, with special regard to language instruction and the afore-mentioned Noticing Hypothesis.

1. Selectivity of attention

The idea of the selective intake of information is intuitively sound. No specialist psychological knowledge is necessary to realise that being constantly surrounded by numerous stimuli, we are aware of only a limited number of them. It is also common knowledge that it is impossible to be equally attentive to different simultaneously performed actions or different pieces of information processed at the same time. In such cases, selectivity is necessary to minimise processing investment in order to avoid cognitive overload and, consequently, error; or the risk of incohesion of the incoming data and the resulting inability to make a decision. Research to date has concentrated on the mechanisms of such selectivity as well as their timing. The latter refers to the attentional switch-on, the exact moment information processing changes qualitatively: from unconscious to conscious; from parallel to sequential.

Analysing the evolution of views on the above-mentioned issues, Czyżewska (1991) points out that the development of knowledge has been towards granting more importance to the unconscious. The following sections take the reader – chronologically – through a number of studies on attention, hopefully demonstrating Czyżewska's claim. The present work also tries to make it clear that attention is selective on both levels of processing: sensory, carried out in the primary cortex as well as representational, in the higher cortex. What is also important for learning in general – and language learning in particular – is that attentional selectivity is a factor in all three stages of the acquisition of knowledge: input, central processing and output.

1.1. Attentional selectivity: main research paths to date

It was the selectivity to perceptual input that was first proposed as the defining characteristic of attention. In the light of one of such working models of attention – Broadbent's Filter Model of Selective Attention (Broadbent 1954, 1958; in Czyżewska 1991, Kolańczyk 1992 and Eysenck 2004) – a virtually unlimited number of sensory stimuli, received via different modalities, enter the module of very short term memory (VSTM), where they are subjected to unconscious parallel processing. It is in VSTM that the process of selection takes place: based solely on the sensory analysis, only the sensory input of significant perceptual salience is filtered for further processing. This is also the moment of a major qualitative change: once filtered, the so-far unconsciously processed input is subject to mental operations that are conscious and sequential.

Broadbent justifies his theory based on two experiments: his own split-span number experiment (Broadbent 1954; in Czyżewska 1991, Kolańczyk 1992 and Eysenck 2004) and Cherry's two-texts experiment (Cherry 1953 cited in Broadbent 1958). In both experiments, the testees were subjected to a dichotic listening

task, in which each ear was a separate channel for one of two different inputs presented simultaneously. In Broadbent's number experiment, the testees had two lists of numbers presented to them. While the attending ear received numbers such as 1, 2 and 3, numbers 4, 5, and 6 were played to the unattending ear. Immediately after the presentation, the testees were asked to retrieve the numbers they remembered. Even though the numbers were presented in pairs – 1-4, 2-5 and 3-6 – the retrieved sequence was always 1, 2, 3, 4, 5 and 6. In Cherry's experiment in turn, the listening material contained a press article read by a man (the attending ear) and another text read by a different speaker (the other ear). Both texts differed in their middle part and the modifications in the material presented to the unattending ear included: the introduction of another speaker (a woman); playing the text backwards; or replacing individual words with tones. The testees were asked to concentrate on the listening material played to the attending ear. After the experiment it turned out that the text presented to the unattending reception channel had not been processed semantically: while the testees were almost certain the presentation contained actual human speech, they were unsure about the language of the presentation and totally unaware of its content. What was registered in the to-be-ignored passage, however, were the change of the speaker (man/woman) and the series of tones replacing words. The results of the two experiments were, according to Broadbent (1954, 1958), demonstrative of the mechanism of selective attention and its timing. What the testees were able to retrieve and the order of the remembered elements showed that parallel processing may take place during the initial not-yet-conscious analysis of the sensory properties of the incoming stimuli. However, once the process became conscious, it became serial and dealt with a limited amount of input. This is what Broadbent called a bottleneck effect: an inflexible, definitive narrowing down of the information channel filtering in only a part of input for further conscious processing.

It was the bottleneck – inflexible and definitive – understanding of the attention switch-on that aroused criticism and led to research, whose results and interpretations showed that the initial unconscious phase may be longer and more significant than Broadbent thought. Among others, Moray's (1959) cocktail party experiment showed that while the unattending ear may not register neutral words, the channel is definitely sensitive to the listener's name, as often is the case when people socialise at parties and are oblivious to other conversations going on around unless they spot their name mentioned in one of them. This suggests that while unconscious of the incoming stimuli, we can actually process them not only for physical properties – a voice changing from male to female; words replaced by tones (Broadbent's 1958) – but also, at least to a certain extent, semantically.

The findings were confirmed by three other studies. The first of them, a text-processing experiment, was carried out by Gray and Wedderburn (1960). Formally similar to Broadbent's number test, it was more sophisticated semantically: the text chosen for the experiment was broken down into two

complementary parts and played to the testees dichotically, through the attending and the unattending ears. When asked to recall the text, the participants were unable to render it verbatim or near-verbatim but, significantly, could retrieve the content of the combined message. The second experiment demonstrated unattended semantic processing of visual input (Neisser 1964): the testees who were asked to locate a given letter (e.g.: Z) were able to perform the task much faster, if the neighbouring distractor letters did not share the visual properties of the searched graphic symbol; in other words, the more salient/unique a given stimulus was, the easier it was to spot. In turn, in his 1969 study, which was a visually adapted replica of Moray's cocktail party, Neisser demonstrated that while concentrating on a text of a specified colour, the testees retained very little from fragments in other colours which were unattended to unless they contained the name of the test participant. All this shows that the unconsciously processed material is not entirely filtered out, especially if, to use Kolańczyk's words, it is "important to the listener or organised in a meaningful way" (1992: 81; my own translation).

As was pointed out earlier, all the above-quoted studies and their findings led to a major reconsideration of the problem of the selectivity of attention. New models of selective attention were put forward, the most important of which were the ones proposed by Treisman (1960, 1964), Norman (1968), Shiffrin and Schneider (1977), Deutsch and Deutsch (1963), Posner (1980) and Eriksen and St.James (1986). All of them, as can be seen from their presentation below, saw selection as carried out much later than Broadbent had proposed, following and not preceding the semantic analysis of the input.

Revising Broadbent's idea of the nature and timing of the selective mechanisms of attention, Treisman (1960, 1964) suggested that unattended stimuli are not filtered out at the threshold of consciousness but continue to be processed extensively and may influence the conceptualisation of the attended input. This happens because, as Treisman speculated, Broadbent's attentional bottleneck may not be as inflexible as he had claimed; nor does the final transition from unconscious parallel to conscious serial appear so early in the course of input processing (at the level of physical cues, as claimed by Broadbent 1954, 1958). According to Treisman, the process of selection is gradual and hierarchical: in the case of language input processing, it starts at the level of visual or auditory cues and proceeds through syllables, words, phrases, sentences towards the analysis of discourse. As a result, Treisman's model accounts for attentional attenuation of unattended stimuli rather than their filtering out in the bottleneck-like, definitive fashion.

What is crucial to point out here – as it will be of importance (and, consequently, revisited) later in this chapter – is the particular applicability of Treisman's intuitions about selectivity to the role of attention in the processing of linguistic input. This is because it coincides with somewhat later studies on the

mental lexicon (Aitchison 1994, Singleton 2000), research into parsing as well as memory vs. rule-based processing and understanding complex sentences or sentential ambiguity (cf. Pinker 1994 and 1999, among others). All this amounts to an assertion that we process the physical input parallelly *while* receiving it, and not only linearly after *having* received it. This happens because the sensory (auditory or visual) linguistic input carries an array of various potential candidates (words, phrases) for processing meeting – or not – our initial expectations about what is going to be heard or read. That is why, in order to understand a given utterance, we need to *bottleneck* or attenuate this input by successfully discarding dead-ending processing routes. This happens in the course of the parallel processing of the incoming data along the three pillars – phonological, syntactic and conceptual – of Jackendoff's (2002) parallel architecture. Within this architecture three parallelly operating processors of language as a system – phonological, syntactic and conceptual – generate phonological, syntactic and conceptual structures which interact with each other communicating via interfaces. As a result, as proposed by Jackendoff (2002) and recycled by Truscott and Sharwood-Smith (2004), language input is analysed parallelly by the three subsystems, which are in constant interplay: the sensory sound/image input is reprocessed in the phonological loop; simultaneously, the conceptual analysis is carried out, and this is frequently done way ahead of the structural analysis; and the syntactic module usually overproduces candidates for the structural interpretation of the message, these candidates, needless to say, being processed parallelly. The closer we get to the final selection – phonological (or graphic), syntactic, semantic – of input, the tighter the bottleneck becomes.

Going back to attentional selectivity studies, another research effort worth quoting is Shiffrin and Schneider's (1977) experiment, which corroborated the earlier findings (Neisser 1964; the letter test) that the unattended message is attentionally registered if it contains an element standing out from its surroundings, for example a figure in a group of letters. The results point to the importance of categorisation as the filtering mechanism, the consequence of which is particularly visible if the context categories – like the figures and letters used in the experiment – are clearly homogenous and internally coherent. In this respect Shiffrin and Schneider's model is similar to the one proposed by Deutsch and Deutsch (1963), who claim that all incoming stimuli are processed by comparison with the data retrieved from the long-term term memory, semantic processing included. All this is done subconsciously up to the moment of meaning recognition, so the categorisation belongs to the parallel, unfiltered domain of mental operations; in order to recognise what is meaningful, the attention filter weighs all of the content stimuli parallelly, selecting those of a considerably high value. Once singled out, these stimuli themselves become part of the selection process, filtering out all other, less important or less salient pieces of information. In effect, we can speak of a shadowing effect – a stimulus of a very high value has

the potential of outweighing the rest of input to the point of its total negligence. In the case of visual input studies such a phenomenon is known as the Stroop effect and is manifested by the testees' inability to semantically process a word like *red* if the word itself is printed in another colour (Stroop 1935; cf. also Deutsch Lezak et al. 2004). In language-related areas, the studies of Shiffrin and Schneider as well as by Deutsch and Deutsch conjure up issues such as the role of working memory as the worktable for language processing; the interplay between the declarative and the procedural memory (cf. Anderson's ACT model; Anderson 1980 and later works); or the gestalt appraisal of linguistic data which is carried out prior to the analytic processing of this data. Phenomena similar to the Stroop effect, in turn, can be identified in the case of clutter on the working memory worktable, where stronger stimuli overshadow and off-focus the weaker ones. Such strong attention attractors are, for example, emotions – including language anxiety – that are the winners among equals, the brain being the captive audience of the body (Damasio 1994: 159-160). All the issues will be considered in detail later in this part of the book.

Finally, we look at two models – Posner's (1980) and Eriksen and St. James's (1986) – together, as the two see attentional selectivity in terms of similar metaphors.

According to Posner (1980), attention is best understood based on the *spotlight* metaphor: looking at a certain scene we perceive a very small area very clearly while the rest of the scene remains in a kind of shadow or darkness. As our motivation or task demands change, we can move the spotlight around the scene bringing out its different elements and regions from the shadow/darkness. All this seems to endorse the idea of attention as flexible and its selectivity as an ongoing rather than a fixed process. What, however, is equally – if not more – important is that, according to Posner (1980), there is a phenomenon called *covert* attention: the ability to move the attentional spotlight without actually moving one's sight towards the singled-out item. This, according to Posner, means that there are actually two attentional systems: endogenous, controlled by intention and involved in the processing of central cues; and exogenous, in charge of automatic attention shifts towards peripheral cues.

Similarly, in the light Eriksen and St. James's (1986) *zoom lens model* of attention, we notice an item and concentrate on it. However, instead of changing the focus of attention altogether, as in Posner's model, we can increase or decrease the focal area depending on the demands of the task, considering trade-off effects: greater focusing limits scope while broadening the scope results in blurring the vision. In other words, we can either attend to a limited area intensively or deal with a larger number of input facets in a more extensive way. What is particularly interesting about these zoom adjustments is that – unlike spotlighting, which implies a switch from the unconscious to the awares – zooming in and out is a fluctuating rather than a definitive process: it may vary

while we are on the task with changing demands of the task in question without the switch-on/switch-off quality of the process. This means, as Awh and Pashler (2000) argue, that even if focused – regardless of the lens adjustment – our attention can still, to a certain degree and under certain conditions, process stimuli from regions of the field of vision which are non-adjacent to the area we are currently focusing on, this phenomenon being known as split attention. This proposal is reminiscent of a claim made by Kahneman et al. (1992: 176) to the effect that: "attention to any property of an object causes even irrelevant properties of the object to be attended". This happens – to revisit the issue of "certain conditions" mentioned above – because the division of attention is facilitated by the fact that all of the attributes belonging to the same object are related in a coherent, uniting way – they are part of one scene, field of vision or context. If we venture to reach out, as in the case of previous models, towards language processing, we can say that combining the intensive and the extensive, the flexibility and the degrees of focusing of the spotlight and the zoom lens is highly reminiscent of the instruction plus exposure combine proposed by focus-on-form pedagogy. That is why we will return to these issues where instructed learning is discussed at a greater length.

In conclusion to the review of the selectivity studies, we can say that, in spite of their differences in how they conceptualise attention or their understanding of the quality and quantity of filtering mechanisms, all of the above proposals share two basic assumptions:

1. there are different kinds of paying attention, intensive and extensive, overt and covert;
2. these parallelly processed different foci of attention can be of varying intensity as our attention fluctuates and shifts.

These assumptions are discussed in the following sub-sections.

1.1.1. Focusing in controlled vs. automated processing

The necessity of distinguishing between controlled and automated processes was put forward in the already quoted Shiffrin and Schneider (1977). They argued that while controlled processing requires attention and does not offer a lot of focusing capacity for any competing input, there are no such limitations on automated processes, which are performed without attention. This, however, means that automated actions are rather inflexible: they are not subject to modification as *control* also means the potential to adjust one's performance when on task.

A slightly different view is presented in Posner (1982): attention, or – as he also calls it – the central executive (cf. Baddeley and Hitch's 1974; model of working memory discussed later in this chapter) is in charge of both controlled and automated processes. There is a kind of interplay between these two types of foci, resulting from the fact that the brain is a system of systems (cf. Damasio 1994, discussed later in this chapter). Consequently, on the one hand, the automated processes consume a part of attentional resources allocated to controlled processes (Posner 1982). On the other hand, controlled processes will be, to a certain extent, automated and unconscious. The latter is because even new tasks, which require attentional focusing, are at the same time wired to background knowledge, previous similar tasks, etc. Such factors, which may be called contextual in the broad sense of the word, will influence attention-focusing processes by automatically providing conceptual blueprints within which the controlled processes can be accommodated. This influence will be positive or negative, as proved in an experiment carried out by Kahneman and Henik (1977; for details see Eysenck 2004), in the sense of speeding up the controlled processing or inhibiting it, if what is automatically provided is a mismatch.

A similar model proposing a certain gradation of automaticity and control is Norman and Shallice's (1986) schema activation, which distinguishes between:

– fully controlled actions directed by a supervisory system similar to the central executive known from Baddeley and Hitch's (1974; see also Baddeley 1986 and 2000) working memory model. This supervisory system (Norman and Shallice 1986; see also Shallice and Burgess 1996) is activated in the case of tasks which require control: novel and incoherently structured (Kolańczyk 1992). The system is in charge of constructing new schemas to underlie task execution, monitoring those novel schemas for error and evaluating them for potential future use;
– partly automated actions, without conscious control yet involving what can be called *contention scheduling*, which is a mechanism used to choose between two conflicting schemas competing to blueprint an action; such a selection is motivated by contextual clues;
– fully automatic actions, controlled by schemas or our plans of actions, requiring hardly any conscious awareness of what is being done.

In the light of all of the above considerations, we can say that while the division into automated and controlled processing is a fact, it is not decidedly bipolar, automaticity and control being gradual rather than final. It is also important to note that while automatic processes are not particularly attention-consuming, they do, to a certain extent, compete for attentional resources because both controlled and automatic processes are supervised by a kind of central executive. This supervisory attentional resource, in addition to coordinating the input, monitors it

for its adherence to schemas, exercising less control when familiar action plans are activated as a blueprint for the currently perceived stimuli; or more control when the input is novel, incoherent and in need of new schemas to be constructed and monitored. All this brings up the question of intensive and extensive attention.

1.1.2. Extensive and intensive attention

According to Kolańczyk (1992), the intensity of attention can be understood in two ways:

1. how much we decide to concentrate on the task and, consequently, how many resources (=attention, effort; Wickens 1984) we are ready to invest;

and

2. how well the object of attention is singled out from the background/context.

Extensive attention, in turn, amounts to the holistic processing of data, often in a relaxed way, with paratelic – not goal-oriented – motivation towards the task (Kolańczyk 1992).

Paying attention extensively to a broader area of interest – often carried out in addition to intensively concentrating on one item only – re-introduces the phenomenon of *divided attention*. Broadbent (1954, 1958), as mentioned earlier in this chapter, discarded the idea altogether, arguing that we should rather be speaking about shifts of (intensive) attention between different tasks. Contrary to his assumption – and in agreement with what has already been written about controlled and automated actions – it is possible to be on a number of tasks at the same time, depending on their novelty/familiarity as well as where on the novice-expert continuum the performer can be placed. As a consequence, a novel task and/or the one which is performed by a novice will necessitate a greater intensity of attention; in turn, an expert will be able to extensively attend to a number of operations especially if they require a standard, well-known course of action. Additional factors in the area of attentional intensity will include the similarity of the parallelly processed tasks and their relative difficulty (Eysenck and Keane 2000, Eysenck 2004). In the case of task similarity, two or more actions performed together benefit from different modalities; consequently, while we can read and listen to music, simultaneous reading and watching television is considerably more demanding. When it comes to task difficulty, attention is more efficiently divided between less demanding tasks; going back to the reading/listening example, it is less problematic to listen to music while engrossed in a novel than when one is on an exam-preparation task. Summing up research into divided attention, Eysenck

(2004: 205) states that: "two dissimilar, highly practised and simple tasks can typically be performed well together, whereas two similar, novel and complex tasks cannot".

The intensity and extensity of attention can also be understood in a way proposed by Kentridge et al. (1999). Similarly to Posner's (1980) concept of endogenous/exogenous attention described earlier in this chapter, they argued that attentional resources can be item-oriented or field-oriented. In the first case we can speak of attention as volitional and selective, amounting to focusing on a particular item or its particular feature. In the latter case, attention will have the quality of a more general vigilance, alertness or readiness to be attracted to a certain conspicuous stimulus coming from the field.

A contemporary model which seems to offer an interesting combination of the elements of intensive and extensive attention is the one presented by Arvidson (2004, 2006). Based on the work of Gurwitsch (1964 cited in Arvidson 2004, 2006), rooted in phenomenology and Gestalt psychology, Arvidson's model proposes that we perceive all incoming stimuli in their context. We are aware of this context but to a different degree than we are aware of the main focus of our attention. The immediate context, in turn, is placed within a much larger situational framework to which attention is paid only marginally. That is why, when talking about attention, Arvidson (2006: 1) proposes working from the inside of the attention circle towards the outer, more peripheral layers, the "outer shells", as he calls them. Consequently, we will be moving from the *theme*, the centrally located dimension of the main focus; through the *thematic context* of attention, our consciousness of the immediate circumstances surrounding the object of our thematic (focal) attention; towards the very vague peripheral attention, the consciousness of the outer word *halo*. As demonstrated below, each layer on the attentional tripartite model operates according to different organisational principles.

In order to be the focus of attention, the *theme* has to be singled out from the background and consolidated as a unit on the basis of the "gestalt-coherence principle" (Arvidson 2006: 3). According to this principle, everything that is noticed about a given theme – clear, vague, complete or not – contributes to the perception of the theme as central against the backgrounded context (a *thematic* dog in the *contextual* yard, in Arvidson's example; Arvidson 2006). Central as it is, the main theme is also prone to shiftiness and jumpiness, to use Arvidson's terminology: any part of the central focus perceived as a whole may result in us refocusing our attention. This means that the above-mentioned dog (as a whole) may be backgrounded if we notice his/her wounded leg; the dog and his/her leg will, however, remain coherently linked[1].

---

1 A similar point was made earlier in this chapter following the quote from Kahneman et al. (1992: 176): "attention to any property of an object causes even irrelevant properties of the object to be attended" on condition that all the properties are related in a coherent, uniting way – they are part of one scene, field of vision or context.

The *thematic context*, in turn, is subject to relevancy rules: we notice everything that presents itself as relevant to the theme (Arvidson 2006: 5). While it is true, as mentioned above, that the theme segregates itself from the background, it remains part of it in the sense that it continues to be relevant to its surroundings as an organising factor. Arvidson (2006: 5) calls this *unity by relevancy* and explains it as the inner coherence of the context, whose value goes beyond just "being there", and the theme, "a central gestalt" in the network of gestalts which are all part of the same action or essive field. The closer a given gestalt is to the theme, the more important it appears; consequently, the context as a whole has what Arvidson calls a "gradation intensity" (2006: 6), this quality of gradual transition making the context itself flexible and prone to transformation, fluctuation and shifting. Modifying the dog example slightly, we can stipulate that the yard as a context for the dog will be subject to different kinds of processing activities performed in relation to the dog itself and in relation to the dog's aforementioned wound. In the latter case we may want to refocus our context-oriented attention looking for sharp objects which, when located, will evoke – although not equally clearly – sharp objects as a class with their properties and possible accompanying incidents. Whether, as a result, our contextual attention will transform into focal/thematic attention depends on how important the original theme – the dog – still is.

Finally, there is the *outer world halo*, which, as Arvidson (2006: 7) proposes "is irrelevant to the theme but is presented nonetheless". Irrelevant, however, according to Arvidson, does not mean dispensable. The halo consists of three ever-present domains: the stream of consciousness or "streaming in attending" (Arvidson 2006: 7), embodied existence and the perceptual or environing world. To refer back to the dog example, the thematic focus on the animal and its immediate context will be accompanied by the marginal awareness of the passing of time (streaming), the fact that the person watching the dog is sitting or standing (embodiment) and the fact that cars are passing outside the fence surrounding the yard (the environing world). What Arvidson points out is that the marginal consciousness has its inner/outer structure too: there is gradation of the parts of the halo adhering to the thematic context and those more remote from it.

Summing it all up, in the light of the considerations of attentional selectivity, historical and contemporary presented in Section 1.1, referring to both controlled vs. automated processing as well as intensive vis à vis extensive attention, we can assume that there is, as Eysenck (2004) puts it, a grain of truth in all three theories: the central processor theory, the bottleneck theory and the divided attention theory, also called the separate pools theory. Consequently, it may be a good idea to combine the three kinds of findings into a coherent construct of attention as multifaceted processing controlled by a kind of a supervisory resource. There have been a few attempts at translating the assumptions presented above into working models, the most convincing of which is the proposal to link

attention to working memory. Two such models are presented in the following section, the one put forward by Cowan et al. (2005), and Baars' Global Workspace Theory (Baars 1997a, 1997b and 2002; Baars and Franklin 2003 and 2007).

1.2. Working memory and attention

Before presenting the theoretical models that accommodate the concept of central executive and different pools of attention based on the idea of the convergence of attention and working memory (WM), the very construct of WM proposed by Baddeley and Hitch (1974) and later amended by Baddleley (1986 and 2000) will be briefly described.

As Cowan et al. (2005: 42) put it, "[w]orking memory (WM) is the set of mental processes holding limited information in a temporarily accessible state in service of cognition" (Cowan et al. 2005: 42). It was exactly this kind of approach – the "in service of cognition" idea – that led to the reconsideration of the notion of memory, originally referred to in terms of three stores: sensory, short-term and long-term memory. Baddeley and Hitch (1974), the proponents of the working memory construct, put forward a claim that a module has to exist – later often referred to as a kind of worktable (cf. Stevick 1999, among others) – which is in charge not only of storage but also of retrieval and processing. Baddeley and Hitch saw such module as tripartite, its components being:

– the phonological loop, which stores information in the form of speech (=phonologically);
– the visuo-spatial sketchpad, in charge of coding of the visual/spatial input;
– and, most importantly, the central executive, in charge of both processing tracks, which is modality free but whose capacity is limited.

The three-component model was amended by Baddeley (2000), who added a fourth element – the episodic buffer – which enables multimodal temporary storage fed by and mediating between both the two original subsystems (the loop and the sketchpad) on the one hand and long-term memory on the other. The buffer itself, however, is separate from LTM and serves as its lead-in, a kind of modelling space, working, as it seems, through new episode/known episode comparisons and categorisation, similarly to the model of Deutsch and Deutsch (1963). Unlike it, however, retrieval from the episode memory buffer proposed by Baddeley (2000) is possible based on conscious awareness.

The Baddeley and Hitch's WM model and the earlier-quoted Eysenck's (2004) proposal for a comprehensive model of attention have a number of convergence zones. First of all, the central executive – which was also linked to attention by Posner (1982; cf. earlier in the present chapter) – fits into the concept

of the supervisory resource. Its limited capacity, in turn, is the most likely underlying cause for the bottleneck effect on tasks in which "rehearsal and grouping processes are prevented, allowing a clearer estimate of how many separate chunks of information the focus of attention circumscribes at once" (Cowan et al. 2005: 43). Finally, the two ancillary subsystems, the phonological loop and the visuo-spatial sketchpad explain the idea of separate pools of attention and the fact that we can simultaneously attend to two different tasks if each of them utilises a different modality subcomponent.

The tendency to find parallels between working memory and attention is very strong in contemporary research, one of the bents boiling down to amending the traditional concept of WM capacity, understood in terms of storage-and-processing measures and control over them, by highlighting either storing with its speed of retrieval or processing with its scope-of-attention measures. Cowan et al. (2005) argue in favour of this, claiming that the traditional interpretation of working memory in terms of a united concept of storage and processing can be difficult because these measures may be subject to individual differences (IDs). As a result, some people may exercise better control over storing and some over processing. What is more, in addition to IDs, it is the task specificity that will be of importance to these two measures, as some tasks require more storage capacity and some – a greater processing effort. Considering the latter we have to admit that somebody may be able to perform two or more tasks simultaneously because of their high WM capacity or because of the fact that the tasks in question are well automatised, comparatively easy or engage different modalities.

Based on what was written earlier about attention, we can note that both the processing-and-storage dexterity (IDs) of working memory as well as its task-related specificity mentioned above call for a certain attention-related module/area on the WM worktable. Such a general, amodal attention resource which is a subcomponent of WM, shared between storage and processing, if necessary, is included in the working memory model of Cowan et al. (2005). Such an attentional capacity within working memory is, according to its proponents, characterised by 4 kinds of findings (Cowan et al. 2005: 49):

1. There is a limit in the capacity of the focus of attention.
2. This limit varies between individuals.
3. Measures of this capacity are theoretically and empirically related to storage-and-processing measures of WM.
4. The common variance between these measures is related to intellectual aptitude measures.

As for the first assertion, Cowan (2001) in his earlier work speaks of 3-5 chunks of information that can be subject to simultaneous processing. This, as noted by Bierwiaczonek (2010; personal communication), is an assertion which is intui-

tively sound in relation to language processing. The predicate-argument structure of a sentence – the subject, the verb, its (maximum) two objects and its (optional) adverbial – is elegantly accommodated within the 3-5 chunk range, much more neatly than within the earlier-proposed (Miller 1956) short-term memory capacity for processing information, in which the magical number 7 (5-9) was put forward. In turn, the fact that both Cowan and Miller suggest range of 3-5/5-9 implies that scores in this area differ between testees (assertion 2). As a result, it seems reasonable to argue that attentional capacity *is* a relevant individual difference, in general and in language acquisition/learning. In the light of assertion 3, it stands to reason that the total capacity of working memory will be the result of a multiplication between the number of chunks of information that can be put together on the worktable (memory capacity) and the effectiveness with which they can be effectively manipulated (attentional capacity). The result of this multiplication will be the most satisfying if both capacities are high. It can also be said that low memory capacity and high attentional capacity and high memory capacity and low attentional capacity will lead to comparable scores, an acknowledgment which, among others, lies at the explanation of different types of intellectual aptitude (4), a part of which is language aptitude, an issue to be revisited in Section 3 of the present chapter, where attention-related phenomena are discussed in connection with language learning.

Another development on Baddeley and Hitch's model of working memory (Baddeley and Hitch 1974; Baddeley 1986 and 2000) which incorporates the idea of a certain central/supervisory resource is Baars' Global Workspace Theory (Baars 1997a, 1997b, 2002, Baars and Franklin 2003, 2007). The model proposes a relationship between working memory and consciousness, which is motivated by Baars' (1997a and 1997b) intuition that consciousness has to be related to the limited capacity aspects of brain: we can be consciously involved with only one flow of information. At the same time, however, consciousness creates widespread access to sources of knowledge which are mostly unconscious and whose joint capacity is practically unlimited. While the issue of consciousness per se will be returned to in Section 2, here we will concentrate on the important role to be played by attention in the Global Workspace model: to explain the conscious gateway offering global access, Baars (1997b) speaks of "the spotlight of attention" in "the theatre of consciousness".

How exactly should this "theatre of consciousness" metaphor be understood? And how important is the attentional spotlight? The answers to these questions can only be clear if based on the Global Workspace Theory (Baars 1997b), in the light of which there are five important elements of this theatre: the stage, the spotlight, the actors, the behind-the-scenes context and the audience.

Based on the early model by Baddeley and Hitch (1974), Baars proposes two dimensions of *the stage of working memory*, corresponding respectively to the phonological loop and the visuo-spatial sketchpad: the inner speech, practically

ongoing and difficult to stop for more than a few seconds; and the visual component, constantly underlying our spatial processing of the world. In Baars and Franklin (2003, 2007) a third dimension is introduced. Following the already-mentioned Baddeley's (2000) amendment of the WM model, the episodic buffer is included in the WM stage of the consciousness theatre, a kind of filter between working memory and the long-term episodic memory. According to Baars and Franklin (2003), while the interaction between working memory and the long-term store, carried out via the inner speech and the visual component, is mostly unconscious, the episodic buffer is a conscious go-between whatever happens at the stage and the long-term memory.

As the contents of the working memory stage – by which we mean the current input, processing and output – may be within the conscious grasp as well as they may fade away and come back again, Baars (1997b and later works) introduces the concept of *the spotlight of attention*. Working memory is described as fleeting (Baars 1997b and later works), which is the best characteristic of the come-into-light-fade-away-come-into-light-again phenomenon of spotlight in the theatre. In relation to attention it means that the attentional spotlight can be guided – both voluntarily and spontaneously – singling out, for conscious processing, different elements on the scene which Baars (1997b and later works) calls *the actors*, trying to get into the bright spot. As a result, on the stage of working memory there is an ongoing competition between sensations, thoughts and images that want to get into the spotlight. The more important an actor is the more actively it will compete for attentional resources.

However, the theatre of consciousness – like any theatre – is not limited to the stage. There are two additional elements: the out-of-stage context, including the director of the play, and the audience.

As Baars (1997) points out, a lot of attentional selection – metaphorically understood as the fleeting of the spotlight from one element on the scene to another – is unconscious and spontaneous. This is because it is motivated from behind the scenes by factors such as beliefs or past memories which make *the out-of-stage context*. There is also the director – which is how Baars (1997) understands the executive brain functions, the goal-driven system guiding human WM – who is also behind the scenes. We rarely have access to reasons – there is scientific proof that our minds play dice (Klarreich 2001), making inexplicable decisions on the spur of the moment – which is why we have to assume that the director seated off stage is not always conscious, surprising as it may sound.

Finally, *the audience* in the theatre metaphor stands for long-term memory and automated productions.

In its above-described shape, the theatre of consciousness performs 9 functions (Baars 1997b): 1) the adaptation and learning function: the more new information we encounter the more conscious involvement is needed; 2) the definitional and contextualising function: every conscious experience is shaped by

contextual unconscious factors (like prior ideas about a given phenomenon, now out of the stage); 3) access to a self system: self is the observant and the controller of conscious experience; it adds the subjective feel to the ongoing events; 4) the prioritising and access control functions: events are consciously related to higher-level goals (an example is smoking related to health in social campaigns); 5) the recruitment and control function: conscious goals can recruit unconscious routines for the execution of new goals; 6) the decision-making and executive function: the theatre is controlled from outside; however, an actor on the scene can incite the controller into noticing an issue and dealing with it; 7) the error detection and editing function: unconsciously detected error breaks through to consciousness; 8) the reflective and self-monitoring function: we reflect upon our own functioning through inner speech and imagery; 9) optimising the trade-off between organisation and flexibility: what is in the spotlight is consciously controlled but the spotlight itself can be moved around freely.

The model presented in Baars (1997b) was confirmed by neuroimaging studies carried out in subsequent years (Baars 2002; Baars and Franklin 2003, Baars and Franklin 2007) leading, once again, to the following observations (Baars and Franklin 2003: 166-167):

– the brain can be viewed as a collection of distributed specialised networks;
– consciousness is associated with a global workspace in the brain – a fleeting memory capacity whose focal contents are widely distributed (broadcast) to many unconscious specialised networks;
– this workspace can also integrate many competing and integrating networks;
– some unconscious contents called contexts shape conscious content;
– sometime such contexts work together jointly constraining conscious events;
– motives and emotions can be viewed as goal contexts (in this way they belong to the central executive domain);
– executive functions work as hierarchies of goal contexts.

All of this is also in agreement with contemporary views on working memory, which, if referred to as a *space*, is no longer seen as a two/three-dimensional enclosed area (a blackboard; a desk) but is rather "a transient pattern of activation of elements within long-term memory stores" (Miyake and Shah 1999 cited in Truscott and Sharwood-Smith 2004: 3).

The Global Workspace Theory model presented above is also, to a large extent, similar to somewhat earlier considerations offered in LeDoux (1996) and much earlier claims put forward by Lashley (1951). Both of these authors emphasise the fact that we have access to only a fraction of information processing in our brain, namely the result of these operations. The rest is unconscious, yet not in the Freudian, ominous, dynamic, emotionally-loaded and motivated-by-repressed-memories way but *cognitively* unconscious. The term

itself – the cognitive unconscious – was put forward by Kihlstrom (1987 but see also Kihlstrom 1984), and is simply applied to all of the workings of the human mind which are inaccessible to us. If we agree that the final result of the analysis is conscious while the analysis itself remains in the shadow of the cognitive unconscious, the comparison between this idea and the theatre of the Global Workspace with its spot-lighted stage and the rest of the theatre in darkness is well justified.

In the light of what has been written so far, we can draw a number of interim conclusions about *attentional selectivity*. First of all, we need to emphasise that attention is subject to bottleneck-like effects. As a result, the number of information-processing tasks we can intellectually handle at the same time is not unlimited. It, however, would be far-fetched to state that attention is undivided, and that the filtering process, once in progress, is one-directional and finite. The key words in the description of attention are *fleeting* (Baars 1997b; see also later works) and *flexible*, the latter manifesting itself in attenuation of attention – which is a term preferred to filtering – its graduity and hierarchality (Treisman 1960 and 1964). As a result, we have to assume that attentional processes will be subject to fluctuation in two different senses. On the one hand, a number of information chunks will be processed parallelly, the attentional focus shifting from one to another depending on current – and constantly changing – task demands. All this happens very much along the lines of the spotlight metaphors proposed by Baars as well as – much earlier – by Posner (1980), the legitimate intuitive interpretation of *the spotlight* being the one of *attentional control*. On the other hand, attention will be on the move in terms of its spectrum, in accordance with Eriksen and St. James's (1986) *zoom lens model*, transforming flexibly – again motivated by task demands – from intensive to extensive. This transitional propensity of attention is quite adequately included in Arvidson's (2004, 2006) three-level model of attention in which the fluctuation of focus between the theme, the thematic context and the halo of the outer world is based in the internal gestalt-like coherence between the three layers of the model, which is a continuum with a number of flexible seamlessly-crossed boundaries. At the same time however, we have to remember that the gestalt nature of attentional processes has one more dimension: the very same wholeness that allows for the flexibility of these processes calls for a central supervisory unit, a kind of central executive in control of the whole process. This demand is satisfactorily fulfilled by the concept of linking attention to working memory, as proposed by Baars (1997b and later works) as well as Cowan et al. (2005).

All of the above can best be summed up in the words of Eysenck, who concludes that (2004: 207-209):

– there is considerable evidence for the existence of multiple [attentional] resources;
– [at the same time] there is also evidence for the central processor on the one hand and, on the other, for the bottleneck effect – frequent but not omnipresent – which amounts to serial execution of tasks in dual-task – and, as it seems, multiple-task – performance;
– as a result, the amount of dual task interference depends on the extent to which two tasks share common resources; simultaneously, there is often some disruption to performance even though two tasks make use of separate pools of resources, which depends, among others, on individual differences between task performers.

That is why it is important to remember about a number of constraints on attention. Those mentioned so far have been: modality, with the assertion that divided attention is possible on the condition that competing inputs come from different modalities; individual differences, especially in the area of the capacity of working memory, with the acknowledgement that such differences may be the result of varying sub-capacities in the areas of both storage and processing; the degree to which certain operations are automatised; and last but not least, where the individual performing the tasks can be placed within the novice-expert continuum.

Finally, it has to be pointed out that there is one more attention-related factor whose importance is still subject to debate, namely, the role of conscious and unconscious processes, emphasised by numerous theoretical models, Baars' Global Workspace Theory, among others. What is consciousness and how can it be understood and related to attention? Are attention and consciousness synonymous? Is their relation as straightforward as it seems on the basis of the quoted research, which generally treats attentive as conscious? All these questions are considered in Section 2.

2. The problem of consciousness

According to a number of researchers, there is a difference between attention and consciousness. Baars (1997a, 1997b and later works) argues that attention is a window to consciousness, the difference between the two being a matter of selecting an item/chunk – be it a person, an object, an event, etc. – from the background (attention) and becoming aware of this event (consciousness). Eysenck and Keane (2000: 119) explain this attention/consciousness variance as the difference between looking and seeing; listening and hearing; choosing a channel on a TV and actually watching what appears on the screen. Alternatively, it will be the difference between implicit and explicit perception, to use a term utilised in Chun and Wolfe (2001).

In fact, there is ample research to date which sees attention and consciousness as independent. Auksztulewicz (2007) quotes a number of studies which prove that consciousness and attention are not fully convergent; we can only talk about an overlap between some attentional processes and some types of consciousness. The most interesting model of such an interplay between attention and consciousness is the one presented in Lamme (2003), who, contrary to Baars' claim that attention is a gateway to consciousness, sees consciousness as being the gateway to both attended and unattended intakes, out of which only the first are subject to reportability[2]. Based on the analysis of both psychological/theoretical and neurobiological arguments, Lamme (2003) claims that conscious experiences – like attention – are selective; in neural terms, however, they are two kinds of cerebral activities.

In the light of the multiplicity of consciousness-related phenomena only just referred to above, it has to be stated that discussing in detail all of the intricacies of the consciousness/attention mutual relationship is definitely beyond the scope of the present work. Yet, as has been shown in the course of the present chapter, the word *conscious* appears in numerous attention-related contexts. That is why, while not venturing to offer a full explication of the problem, the following subsections attempt to clarify the very notion of consciousness, so that we can specify the meaning of the very frequently used phrase of "conscious attention".

In doing so, we have to remember that defining consciousness is far from easy as there is no agreement on what consciousness is. It "poses the most baffling problems in the science of the mind. There is nothing that we know more intimately than conscious experience, but there is nothing that is harder to explain" (Chalmers 1994: 200). Additionally, there are few things as multi-faceted as consciousness. In the light of the two above-mentioned constraints, trying to pinpoint consciousness as a phenomenon seems to be mission impossible or at least not easily accomplished. Considering the problem of consciousness is likely to send us back to Leibnitz and Newton and their interest in the physics of perception as well as to inevitably bring up the homunculus argument: the question of the Cartesian theatre and the little audience watching the scene of events. The following sections of the present chapter will look at some of these issues, considering them at least to a certain extent. Finally, there seems to be a terminological problem to solve: in noticing-related debates, consciousness is used interchangeably with awareness. The question which should be resolved before a more profound discussion is started is if these two are actually synonymous. Based on the distinction between the easy and hard problems of consciousness (to be discussed later in this section) the answer is "no": *awareness*

---

2 Lamme's model will be discussed in more detail in relation to Schmidt's Noticing Hypothesis

is the easy part of the consciousness issue: it relates to mental functions and their underlying mechanisms which are objective and physical (Chalmers 2002), such as focusing attention together with the ability to discriminate and integrate information, report our mental states etc. The label *consciousness*, in turn is frequently reserved for the hard-problem phenomenon, that of the subjective experience which, at least according to some researchers, goes beyond the scientifically detectable. As a result, the hard problem amounts to explaining: why human beings have phenomenal experience together with issues related to such experience including awareness of sensory input and qualia; the question of philosophical zombies; and subjectivity of experience or phenomenal natures. What should also be kept in mind is that there are different approaches to the problem: philosophical, neurobiological, linguistic and so on. The present section looks at awareness and consciousness, as well as the easy and hard problems of consciousness, respectively, examining them from two – out of many – different stances: Chalmers' (1994, 1996, 2002) dualism and Dennett's (Dennett 1991, 1993 and 1997; Dennett and Akins 2008) physicalism.

2.1. The easy problems of consciousness

According to Chalmers (1994: 200-201), the easy problems of consciousness – also called A-consciousness (access consciousness; Block 1997) – boil down to explaining the following phenomena:

- the ability to discriminate, categorise, and react to environmental stimuli
- the integration of information by a cognitive system
- the reportability of mental states
- the ability of a system to access its own internal states
- the focus of attention
- the deliberate control of behaviour
- the difference between wakefulness and sleep

Consequently, being aware will amount to a number of easily differentiated states and actions, such as: reacting to a certain input in a way that might be called intentional or deliberate; having access to one's state of mind and its dynamics; being able to verbalise the accompanying sensations to describe both the current mental state as well as possible changes resulting from the accommodation of new information in the existing architecture of knowledge, which Gut (2009: 188) calls the "solidification of thoughts in a form of ... language"; as well as being aware that one is focusing on the incoming stimuli. The key word to explaining these states and actions is *control* because, as Block (1997) puts it, awareness does not amount to availability of certain stimuli alone, active control of thought and behaviour is indispensible.

All of the above phenomena can be – and have been – explained scientifically based on studies of cognitive and neurophysiologic models of a number of mechanisms: access and reportability, which are the mechanisms in charge of retrieving information about internal states and making it available for verbal report; integration of information, responsible for which are the mechanisms which consolidate and process incoming data; etc. The models that offer explanations of these mechanisms include, among others, the following theories (based on Chalmers 1994): Dennett's (Dennett 1991; Dennett and Akins 2008) multiple drafts theory, in the light of which numerous processes in the brain integrate into the final perception of the experienced event; the already-quoted Baars' (1988, 2003) Global Workspace Theory of consciousness, whose main idea is that of a central processor containing consciousness; and Crick and Koch's (1990, 1994) neurobiological theory of consciousness relating the phenomenon in question to neuronal oscillations in the cerebral cortex; the final model will be endorsed by ideas put forward by two other brain scientists, Damasio (1999) and LeDoux (2002).

Dennett's (1991 and later works) theory of multiple drafts is a model developed in response to the so-called Cartesian materialism, in the light of which there is a central consciousness centre in the brain where the results of our experience are presented to the inner self or homunculus. Dennett's escape from the Cartesian theatre, as explicated in Dennett and Akins (2008), amounts to:

– breaking up the work supposedly done by the homunculus and distributing it, temporarily and spatially, to a number of lesser but more *specialised agencies* in the brain;
– the effects of the *work* not having to be re-analysed or stored in memory
– all the processes involved being *parallel*; the final draft is the result of preparing and discarding a number of interim drafts
– the *impossibility of precisely timing* when the human being becomes conscious of the experience.

The multiple draft process was subsequently renamed in Dennett's later works as *fame in the brain* (cf. 1996, among others), and seen as a kind of competition (not all can be *famous*), in which some subsystems in our brain are quicker in processing what we experience; other subsystems take their time, discarding their initial drafts and taking up different ones. This often results in us deciding to choose one option but actually choosing another; screaming with fear and laughing at this fear at the same time; performing an action, like changing gear while driving, only to become aware of what has happened subsequently.

What is important in relation to the problem of consciousness is that whether or not some events become *famous* in Dennett's understanding of the term is a matter of their ability to draw attention to incoming visual or auditory stimuli and

give prominence to concurrent content fixations in the brain[3]. Events demonstrating such an ability are called probes and are characterised by their subsequent recollectability and reportability. Probes are what we are aware of as "the ability to report a content is conclusive evidence of consciousness" (Dennett and Akins 2008: 4321).

In turn, Baars' (1988, 1997a and 1997b, 2003) Global Workspace Theory of consciousness – already described in some detail in the section of the present chapter devoted to models of attention related to working memory – is based on a similar idea of the integrative function of consciousness: the brain being a kind of web, "a massive parallel distributed system" (Baars 2003: 1), there is a "fleeting memory" enabling access between otherwise separate brain functions. The present chapter has already offered a comprehensive description of the theatre; what remains to be added is that Baars (2003) claims that consciousness is the primary agent in such a global access function because it acts as a gateway to a number of brain functions. According to Baars, conscious perception opens the route, among others, to working memory, unconscious perception offering much more limited processing possibilities; conscious events enable all kinds of learning: explicit, implicit, episodic and skill learning; and, finally, consciousness is behind attentional selectivity.

Clarifying the idea of the above-mentioned agent consciousness active in the Global Workspace, Baars (1988, 2003) uses the theatre metaphor. Yet, contrary to the Cartesian theatre idea, we are not dealing with a homunculus consciousness but rather with "the bright spot on the stage of immediate memory" – working memory also called "extended consciousness" (Baars 2003: 7) – directed there by a spotlight of attention under central executive guidance. "The rest of the theatre is dark and unconscious" (Baars 2003: 3). Poetic as it may sound, the model is well grounded neurologically: in the case of sensory consciousness, the "bright spot" amounts to the activation of visual, auditory, etc. cortex resulting in inner imagery or speech. In this form, the input is forwarded to the "decentralised audience of expert networks sitting in the darkened theatre" (Baars 2003: 3). In such a way consciousness performs its primary function: it sets the multiple unconscious networks in motion, often in a state of competition, while coordinating and integrating their activities. All this happens under dual cerebral control of the frontal executive cortex on the one hand and areas of, as Baars (2003a: 4) calls them, automatic interrupt control: the brain stem, pain systems or the amygdale. The coordination, integration and control exercised by consciousness in the theatre – like Dennett's fame model – are the basis of subsequent reportability of the supervised events. In other words, in the theatre of consciousness, the performance starts in the spotlight (=the gateway), continues on stage (=within awareness) and out of it (=unawares) and returns to the

---

3 Very much like Baars' actors; cf. earlier in the present chapter

stage (or at least some of the mental operations do) to be reportable when spotlighted (=within consciousness again).

Finally, Crick and Koch's neurobiological theory of consciousness (Crick and Koch 1990, 1994 and 2002) is an attempt to prove that "the best approach to the problem of explaining consciousness is to concentrate on finding what is known as the neural correlates of consciousness – the processes in the brain that are most directly responsible for consciousness" (Crick and Koch 2002: 94). Such an approach amounts to treating the hard problem of consciousness– the one of the subjective experience (cf. the next subsection) – as an easy problem by breaking it down into a number of questions such as: the reason why we experience things at all; the factors underlying particular experiences; the fact that some aspects of our conscious experience are not reportable to others. Crick and Koch offer – as they put it – "an answer to the last question and a suggestion to the first two" (2002: 94): explicit neuronal representation.

In order to understand the nature of such representation, Koch and Crick claim, we need to be aware that everything we perceive is represented in our brain in a semi-hierarchical way. The representations appear in the primary and higher cortex, respectively: implicitly in the former, as the firing neurons generate a kind of general idea, lines and edges (differences between colours as they used to be represented on black-and-white TV); and explicitly, more particularly in the latter, including all of the specificities of a given item like the perspective, point of view, specific colour or the effects of light and shadow[4]. Crick and Koch claim that it is the latter, explicit representations that are the actual neural correlates of the subjective experience.

This proposal – with is hierarchality and mappings between primary and higher cerebral processes – is very much in accord with the model of consciousness put forward by Damasio (1999), who distinguishes between emotions, our feelings of these emotions and the sense that it is our self feeling this emotion. He argues that the first two belong to the realm of core consciousness while the latter is demonstrative of higher-reason, extended consciousness[5]. Core consciousness is activated on perception of a certain object as a result of which, as Damasio (1999) suggests, our brain construes a non-verbal message of how our mental state is affected by the processing of this object. In other words, as a result of a certain

---

4 As Crick and Koch emphasise, brain damage affecting the neurons in charge of explicit item representation – and not defective receptors in the eye – results in dysfunctions such as prosopagnosia (the inability to consciously recognise a familiar item if this item is the face of someone we know or a certain known colour).

5 Koch (2004), who compares Damasio's core and extended consciousness to Block's (1997) phenomenal and acess consciousness (terms used throughout the present chapter), writes: "Core consciousness is all about here and now, while extended consciousness requires a sense of self – the self-referential aspect that, for many people, epitomizes consciousness – and of the past and the anticipated future" (2004: 15).

emotion, which is reflected in our physiological responses (such as, for example, elevated blood pressure), a feeling – which Damasio describes as a "sensory pattern" or "image" (1999: 55) is generated. Neither of these two, as pointed out by Damasio, has to be conscious. It is not until the "[c]omplex customized plans of response are formulated in conscious reason and may be executed as behaviour" (Damasio 1999: 55) that the extended consciousness is reached. The transition – or mapping – between the two levels of consciousness implies that the first-order (or proto-self; 1999: 154) feelings are transformed into memories (or become part of the conscious autobiographical self; 1999: 173). What is important is that – as in Crick and Koch's model (2002) – the extended consciousness has to be the substrate, or correlate, of the core consciousness because, as Damasio observes, the former does not exist without the latter.

This correlation between the two types of consciousness vis à vis their above-described differences brings back the problem of the nature of subjective experience and its reportability. Crick and Koch's (2002) claim that for us to be able to verbalise the subjective experience of a given item – be it somebody's face seen at dawn or a certain colour – explicit information represented in the higher visual cortex, the correlate of the phenomenal experience registered by the primary cortex, has to be further transferred to the motor cortex. As a result, we may not be able to fully report the nature of the inner feel accompanying the perception, but, according to Crick and Koch (2002), we can certainly see and report the difference between one subjective experience and another because of the specific encoding-reencoding going on between these two cortical areas. Addressing the other two questions – why we consciously experience at all and what underlies specific experiences – Crick and Koch suggest that there are ways in which neurons that explicitly encode an item can "convey the meaning" of this item "to the rest of the brain" (2002: 95). Such neuronal ability is related to what they call "a neuron's projective field" (2002: 95), a synaptic pattern explicitly coding a certain concept. In other words, a familiar face represented explicitly in the higher visual cortex will be linked via the white brain to a corresponding area in the motor cortex, where the name of the person is represented and ready for uttering; to auditory cortex and the sound of the person's voice; to the emotional brain and all of the memories of this person; etc.

The above-mentioned idea of consciousness as the elusive *self* located in the white brain is endorsed by another neurobiologist – LeDoux (2002). He argues that what makes every self unique are synapses with their unique activation patterns. In other words we are what we remember and how[6] (in terms of the connection strength as well as the content of these memories). Explicating the mechanism, LeDoux (2002: 303) writes:

---

6 An idea similar to Schacter's (1996, 2001) view of memory-motivated self.

> We all have the same brain systems, and the number of neurons in each brain system is more or less the same in each of us as well. However, the particular way those neurons are connected is distinct, and that uniqueness, in short, is what makes us who we are.

Reviewing all the three cognitive perspectives on consciousness discussed above, it is important to point out that all of them share certain characteristics, the most important being breaking up the work done in the brain and assigning it to the subsystems parallelly processing the data. In this respect all these theories seem highly convergent with other widely accepted models of cerebral organisation and performance: Damasio's (1994) idea of the brain as a system of systems or the brain as a parallel processor, the core concept of the PDP (Parallel Distributed Processing) theory put forward by Rumelhart and McClelland (1986). The following paragraphs look at these similarities.

First of all, Dennett's model of multiple drafts with its idea of breaking up mental work and sharing it between special cerebral agencies as well as Crick and Koch's concept of a neuron's projective field are reminiscent of Damasio's idea of brain which is not "a single, contiguous map, but rather an interaction and coordination of signals in separate maps" (1994: 66), a system of systems. As Damasio puts it:

> We can now say with confidence that there are no single "centres" for vision, or language, or for that matter, reason or social behaviour. There are "systems" made up of several interconnected brain units … dedicated to relatively separable operations that constitute the basis of mental functions … What determines the contribution of a given brain unit to the operation of the system to which it belongs is not just the structure of the unit but also its place in the system. (1994: 15)

As a result of such compositionality of brain representations, the images we store in our memories are what Damasio (1994: 102-103) calls "dispositional representations", neuronal firing patterns which enable a momentary reconstruction of an image – for example a memory of a certain person – based on the joined activity of the individual assemblies of neurons storing the person's voice, their profile at dawn, their giggle, their freckled nose, etc. Such ensembles of smaller neuronal systems are called "convergence zones" (Damasio 1994).

Additionally, Dennett's multiple draft and Crick and Koch's projection field concepts will be considerably convergent with the core claim of PDP (Rumelhart and McClelland 1986): that information about individual entities like people, objects or situations is not stored as a single memory trace but in the form of a number of interconnected units any of which can trigger the retrieval of the whole entity.

There is also a perceivable analogy between the way Damasio (1994) specifies neural correlates of mind and Crick and Koch's (1990, 1994, 2002) attempt to define consciousness via explicit representations in the higher cortex. We can speculate that what Damasio means by "the place in the system" (cf. the

quotation on the previous page; Damasio 1994: 15) is the relative importance and, consequently, activation potential of a given higher cortical area. This potential will determine which of the interconnected neural subsystems and the representations they encode will participate in the inner feel of a given perception in a given moment. Equally important, to quote Damasio once more, will be the context in which the perception takes place as "the physiological operations that we call mind … mental phenomena can be fully understood only in the context of an organism interacting in an environment" (1994: xvii).

In search of other conceptual similarities, we can note that Dennett's proposal that the final draft of input is the result of multiple drafting and redrafting on the level of individual perception and that these processes involve parallel processing goes hand in hand with PDP, in the light of which the human mind consists of a number of elementary units making up the neuronal network and all mental processes involve parallel (rather than sequential) interactions – excitatory or inhibitatory – of these units (Rumelhart and McClelland 1986). The drafting-redrafting concept also agrees with what Damasio (1994) writes about brain operations in macro-scale such as recall and learning:

> The brain's systems and circuits, as well as the operations they perform, depend on the patterns of connections among neurons and on the strength of synapses constituting the connections (108).
> Since different experiences cause synaptic strengths to vary within and across many neural systems, experience shapes the design of circuits. ... Some circuits are remodelled over and over throughout the life span, according to the changes an organism undergoes (112)

What is emphasised is the cyclicality of the process of remodelling – cf. Dennett's redrafting – as well as the process of strengthening neuronal connections. This goes hand in hand with PDP's claim that learning consists of strengthening connections between the interconnected units of the overall memory trace.

Finally, Baars' concept of a "fleeting memory" enabling access between otherwise separate brain functions, though of different cerebral origin, is reminiscent of Damasio's (1994: 182-183) "convergence zones" – assemblies of neurons in the prefrontal cortices in charge of item identification based on putting together all kinds of incoming information – and their content – "dispositional representations for the appropriately categorised and unique contingencies of our life experience" (Damasio 1994: 182-183). The latter are not representations of items per se but rather means for reconstructing such representations in recall, patterns of neuronal activity in the convergence-zone assemblies.

What remains to be said as a form of sum-up is that when it comes to offering solutions to the easy problems of consciousness listed at the beginning of this section – the above-mentioned theories deal with a number of them: Dennett's multiple drafts theory addresses reportability of mental states; Baars' (1988, 2003)

Global Workspace theory of consciousness, deals with the issues of information integration and reportability; finally Crick and Koch's (1990, 1994) neurobiological theory of consciousness refers to integration – or binding – of information as well as the reportability of explicit representations . However, if we were to quote Chalmers (1994: 202), the conclusion might be that if "these phenomena were all there was to consciousness, then consciousness would not be much of a problem".

2.2. The hard problem of consciousness – Chalmers' dualism vs. Dennett's physicalism

What is problematic – and a bone of contention, as will be shown in the present section – is the hard problem of consciousness: the one of experience, the subjective aspect of information processing. As Chalmers (1994) puts it, we can easily explain how our cognitive apparatuses engage in information intake and processing, which means that we clarify issues such as our ability to see colours, feel tastes, and hear sounds. What posits a problem is what makes this experience subject-specific: how and why we see colours as deep or pale; how and why tastes and sounds can lead to rich inner states, etc. To take this even further, the question is of organisms being conscious of what it is like to be themselves; of these organisms' mental states being conscious in the sense that they know what it is like to be in a certain state; and of phenomenal consciousness or qualia, which amount to the way things seem to us (cf. among other, Chalmers 1994, 1996, 2002 and Block 1997). What is additionally important is that the unique character of this kind of consciousness makes it qualitatively different from the functional/access, the easy problem type of consciousness presented in the previous section.

This is exactly where the disagreement arises. To start with, the already quoted Crick and Koch's neurobiological theory of consciousness (Crick and Koch 1990, 1994, 2002) states that there is no such thing as the hard problem of consciousness. All of the questions related to it, such as the question of subjective experience, can be explained on the basis of scientifically detectable – and "easy" – neural correlates of consciousness; a similar stance is presented in the synaptic self put forward by LeDoux (2002). The very qualia, the way things seem to us – as described in the previous section – are the result of neuronal activity – neuronal connections (LeDoux) – even though their full reportability is not possible.

Similar scepticism towards consciousness as a hard problem is expressed by Dennett (Dennett 1991, 1993, 1997 Dennett and Akins 2008), who claims that qualia do not exist, or, as he puts it, if qualia are something that we know about the objects of perception, something real, they are real in no special way, because:

> whenever someone experiences something as being one way rather than another, this is true in virtue of some property of something happening in them at the time, but these properties are so unlike the properties traditionally imputed to consciousness that it

> would be grossly misleading to call any of them the long-sought qualia. Qualia are supposed to be *special* properties, in some hard-to-define way. My claim – which can only come into focus as we proceed – is that conscious experience has *no* properties that are special in *any* of the ways qualia have been supposed to be special.
>
> (Dennett 1993: http://www.tufts.edu/as/cogstud/papers/quinqual.htm)

By stating the above Dennett, as he puts it himself, does not deny the existence of conscious experience; he simply claims that this experience can be explained in terms of its easily defined properties. Addressing the functional/phenomenal duality of consciousness, he inclines towards a quantitative rather than qualitative (cf. Chalmers and Block earlier in this section) distinction between the two sides of the consciousness coin (Dennett 1997: 417): he suggests richness of content *and* degree of influence for the phenomenal/experiential and functional/access consciousness, respectively. He refuses to see these two as separate phenomena, arguing that Block himself, in his own consideration of the duality problem (Block 1997), could not satisfactorily prove that "in the normal run of things"[7] (Dennett 1997: 417) these two types of consciousness can actually exist separately.

Chalmers (2002) discards the above-mentioned argument based, as he claims, on a methodological flaw in Dennett's line of reasoning and the other cognitively motivated approaches (described in Section 2.1, devoted to the easy problem of consciousness). Cognitive research methodology, as Chalmers claims, while very effective in explaining functions and their underlying mechanisms, is of little use when it comes to investigating the origins and the quality of the very "inner feel" (Chalmers 1994, 1996, 2002), the subjective experience accompanying the performance of the above-mentioned functions and related mechanisms. According to Chalmers, while Dennett's (1991, 1993) multiple drafts theory as well as the other two already-mentioned theories of consciousness: Baars' (1988) Global Workspace Theory and Crick and Koch's neurobiological theory (Crick and Koch 1990; Crick 1994) offer some insights into: discrimination, categorisation and integration of input; its reportability and accessibility; as well as attention-related phenomena of selectivity and control, they leave the "the inner feel" unexplained. Theories from other fields of science – neuroscience (nonlinear dynamics and non-algorithmic processing; Penrose 1994) or quantum mechanics (consciousness

---

7 Dennett (1997) addresses Block's (1997) argument of blindsight, in the light of which patients suffering from damages to visual cortex can still see in their minds' eyes. Block argues that such patients have access consciousness without phenomenal consciousness; according to Dennett they just have impoverished content with a simultaneous high degree of influence of the higher cortex able to compansate for the deficiencies in the primary cortex. We may observe that a similar higher-cortex influence is exerted in the case of prosopagnosia, yet in this case it is the higher-cortex deficiency which overcomes the uncomromised intake by the primary cortex.

arising from quantum-physical processes taking place in neuronal protein structures; Hameroff et al. 1994) – seem equally deficient to Chalmers (1994, 1996). Finally, there are also the so-called "mysterians" (Chalmers 2002: 92), who claim consciousness can never be understood and explained. To quote Chalmers (1994: 201), the main flaw of all these attempts boils down to the fact that we know that subjective experience "arises from a physical basis, but we have no good explanation of why and how it so arises".

As for the so-far unsatisfactory solutions offered to the hard problems of consciousness, the underlying cause of their failure to pinpoint the how and why of subjectivity is, according to Chalmers (1994, 1996, 2002), their reductionism, which amounts to trying to explain consciousness in terms simpler than the phenomenon itself. Where these cognitively-based approaches seem to have failed, a non-reductionist stance is offered. The point of departure for such a theory is seeing experience as fundamental, an axiom taken for granted and, consequently, unanalysable. What follows are the principles of structural coherence and organisational invariance as well as the double-aspect theory of information whose aim is to enable the leap (Chalmers 2002) across Levine's (1999) explanatory gap between physical processes and consciousness. All three principles are briefly discussed below, following Chalmers (1994, 1996 and 2002).

The principle of structural coherence refers to a certain isomorphism between consciousness and awareness, the former representing the mysterious inner feel and the latter applicable to various functional phenomena, accessible and reportable, discussed in the section devoted to the easy problem of consciousness. In the light of this principle, the conscious experience, even if impossible to analyse, will in no way be unrelated to the cognitive representation of the incoming sensory information. As Chalmers (1994, 1996) puts it, every instance of conscious experience leaves a trace of corresponding controllable and verbalisable information in the functional system; and vice versa: every trace in the system is a proof of conscious experience, these two as if mirroring each other (Chalmers 2002). While it is true that we still do not fully understand certain properties of experience because of its intrinsic subjective nature, Chalmers (2002) claims that we can still see through it into awareness-related substrates or correlates. To put it in a simpler yet metaphoric way, we cannot catch the main culprit but we know they were at the crime scene, as we can clearly see and analyse the footprints. This logically relates to the principle of organisational invariance, which states that if two different systems exhibit the same kind of awareness-related traces – the same neural substrates; the same functional architecture – they are demonstrative of the same kind of conscious experience. Finally, there is the double-aspect theory of information which stems from the above acknowledgement of the isomorphism between awareness and consciousness, or, to use Chalmers' (1994, 1996) terms, between the physically embodied

information spaces and experiential (phenomenal) information spaces, where structural differences between the latter are the result of differences between the former. Logically inevitable is the conclusion that information has two aspects: physical and phenomenal. These two aspects being linked and mutually dependent, it seems obvious that activation of the former implies the emergence of the latter.

How convincing Chalmers' argument is and whether his three principles, supporting the qualitative difference between the two types of consciousness, are actually significantly different from the quantitative variance stance adopted by cognitive approaches to consciousness is a problem that goes beyond the scope of the present work; as a result, the dualism-physicalism debate is not going to be resolved here. Yet, keeping in mind that any consideration of consciousness means taking into account the two levels of the emergent information space, the phenomenal and the functional, is crucial as the point of departure for two important acknowledgements.

First of all, it has to be pointed out that it is the latter, functional dimension of consciousness that is of much greater importance and use to the present argument. To start with, the easy problem of consciousness is where the very issue of attention as a cognitive phenomenon can be accommodated. Since the present work is written from the cognitive perspective, such a cognitively-grounded approach to consciousness is a most natural point of departure. As a consequence, the discussion on noticing will, to a certain, manageable extent, concentrate on the subjectivity of experience and phenomenal consciousness, particularly in our discussion of Schmidt's Noticing Hypothesis; it will, however, deal much more thoroughly with discrimination, categorisation and integration of input; accessibility and reportability of mental states; different foci of attention; and the degree of control. These latter issues are going to be discussed later in this chapter. Different foci of attention will be considered in relation to strong and weak versions of the noticing hypothesis; accessibility, reportability and control of mental states are going to come back with problems of implicit and explicit learning in Chapter 2. Finally – and most importantly – it is necessary to point out that from this point on the term *consciousness* will be used interchangeably with *awareness,* both in reference to the functional dimension of the phenomenon in question.

At the same time, however, it will be kept in mind throughout the subsequent parts of the argument that learning of any kind amounts to behaviour regulation and there are two – the immediate (phenomenal) and the intermediate (functional) – systems (Obuchowski 1967, Tomaszewski 1975 and Kolańczyk 1999) that deal with such change. Kolańczyk looks at them through their functions – which are, respectively, labelled adaptive and transgressive (Obuchowski 1967; 1993) or reactive and purposive (Tomaszewski 1975) – and the type of memory they involve – procedural for the immediate behaviour regulation system and declarative for the intermediate system. What is important to remember is that the