CHAPTER TWO

Attentional control of lexical processing pathways during word recognition and reading

David A. Balota
Department of Psychology, Washington University, USA

Stephen T. Paul
Psychology Department, Mississippi State University, USA

Daniel H. Spieler
Department of Psychology, Stanford University, USA

The focus of the present chapter is on the primary meaning bearing element in reading, i.e. the word. There are certainly many different aspects of words that play crucial roles during word recognition (see Balota, 1994 and Henderson, 1982 for reviews). In the present chapter, we have decided primarily to emphasize research contributing to our understanding of five principal factors that have been shown to modulate word identification performance. Specifically, we will review some of the current word recognition research examining the influences of orthography, phonology, and meaning, along with syntactic and discourse-level context effects. The first three factors are important in that each factor has been shown to affect processing of words both in isolation and in linguistic contexts, and each has yielded impressive amounts of data and controversy. The latter two factors are somewhat different in that discourse-based syntactic and semantic information do not contribute to isolated word recognition, but, of course, are fundamental in our use of words in the vast majority of language processing contexts.
In our discussion of each of the five factors, we will expand on some methodological and stimulus issues that we believe are crucial to our understanding of research in the area. For example, as outlined by Rayner and Pollatsek (1989), there are issues associated with the experimental paradigms involved in word recognition research that may not generalize to normal reading. Obviously, we rarely find ourselves reading with the tachistoscopic stroboscope frequently employed in word recognition research. Nor is it usual to have each word of text masked, or presented via rapid serial visual presentation (RSVP) methods. In addition, unless the text is particularly dense with unfamiliar or misspelled words, we don't normally find ourselves performing lexical decisions as we progress through paragraphs of text in a book. We will emphasize in the present chapter that perhaps the most critical difference between isolated word recognition and normal reading is the direction of attention to relatively distinct processing pathways.

The present chapter involves three sections. In the first section, we shall provide an overview of the processing pathway approach to lexical processing, which will be the unifying theme across aspects of this chapter. The second section will provide a brief overview of some of the theoretical approaches to word recognition and hence provide a foundation for interpreting the empirical literature that will be reviewed. The third and major section will provide a review of the five previously mentioned factors, with special emphasis on the manner in which the processing pathway perspective will help elucidate our understanding of the manner in which these factors influence word processing across a variety of tasks.

**ATTENTIONAL SELECTION AND LEXICAL PROCESSING**

One of the intriguing aspects of lexical processing is the multiplicity of internal representations and processes which may be engaged when a reader encounters a string of letters. For example, the deep semantic form of analysis which a reader of a classic Russian novel emphasizes is quite different from the usual mixture of orthographic and lexical information required when the same individual attempts to solve a crossword puzzle. Although there is clearly overlap in the nature of cognitive operations involved in each of these situations, the fact that a single individual may excel in such distinct situations highlights the importance of the highly flexible nature of the human language processing system.

When considering the types of codes that are available for stimuli, there are few stimuli that have such diverse codes as words. For example, a visually presented word can be queried at a number of quite distinct levels: Does it have the letters B and S (orthography)? Does it rhyme with the word SAVE (phonology)? Does it represent an animate object (semantics)? Is it a noun (syntax)? We believe that one of the key features of the human language parser is its flexibility in engaging each of these processing pathways based on the current task demands. We would argue it is precisely this flexibility that needs to be taken into consideration in developing adequate models of word recognition.

Unfortunately, the flexible nature of processing has not always played a crucial role in theories of lexical processing. A theory that incorporates processing flexibility would seem to be necessary in order to provide a complete picture of word processing in humans. Part of the problem may be the implicit notion that experiments are performed primarily to elucidate underlying cognitive architectures. The problem with the architecture metaphor is that it lends an image of a relatively static set of constraints that remain constant within and across experimental paradigms. In fact, in reviewing the work in word recognition, it appears that a fundamental goal has been to identify which codes or pathways are obligatory processed upon lexical presentation. Consider, for example, the work on the processing of word meaning. Researchers have argued from semantic priming studies (Neder, 1977) and Stroop studies (Stroop, 1935) that meaning is automatically accessed when a word is visually presented. Thus, there appears to be automatic and autonomous access of meaning. However, as we shall see, even this fundamental observation about lexical processing can come under the influence of attentional selection and such results appear to depend upon the processing pathways that subjects select as a function of task demands. Indeed, we will entertain the possibility that there are no obligatory processing pathways engaged for words, and that the available evidence suggests that the cognitive system has the remarkable ability to select distinct processing pathways in response to a given task demand, thereby minimizing the role of alternative unselected pathways. The ability to shift between reading a novel and solving a crossword puzzle or analysing a word for meaning versus orthography: phonology are a few examples of such flexibility. With that in mind, perhaps a better metaphor in which to coach experimental studies of word recognition is a quest for the specification of a cognitive "toolbox" in which, for any given task, only a subset of the available tools are brought to bear. Thus, in the spirit of using "the right tool for the right task", subtle changes in experimental context may lead individuals to employ different cognitive tools to accomplishing the goals of a given task.

In the context of the present chapter, the "cognitive tools" are psychological pathways (see also Posner, 1978) which are devoted to the processing of particular forms of information. These processing pathways might be analogous to neural pathways, such as the ventral and dorsal visual
pathways, which are devoted to the processing of different forms of visual information (Ungerleider & Mishkin, 1982). Indeed, part of the motivation behind the emphasis on processing pathways is the involvement of different regions of the brain in processing different dimensions of a single stimulus. For example, in processing words, Petersen, Fox, Posner, Mintun, and Raichle (1990) have demonstrated through the use of positron emission tomography that vastly distinct areas of the brain are engaged when participants watch visually presented words (recipient), read aloud visually presented words (temporal), and generate verbs to nouns (frontal). The role of attention is to modulate processing along multiple pathways in a manner consistent with optimal performance of a task. The degree to which a pathway contributes to performance is continuously valued and dependent upon task instructions, strategies the subject might adopt, stimulus contexts, and a host of other factors.

A processing pathway approach emphasizes two points which should be kept in mind throughout the present chapter (and possibly others in this volume). The first point concerns the frequently discrepant and often contradictory findings in the literature. A particular experimental design, combined with the stimuli used and the instructions given to individual participants, will result in the individual choosing some subset of available processing pathways and weighing the information from these pathways in a particular fashion. The combined contribution of information from each of these pathways will yield some pattern of performance. The highly flexible control which the attentional system has on the implicated pathways suggests that a different set of instructions, different lexical processing task, or even changes in list configuration could lead to a differential involvement of processing pathways, and hence, a different pattern of performance. Obviously, careful experimental design and task analyses must be applied to reveal differential involvement of processing pathways and to interpret the results of particular experiments.

The second point which we feel deserves emphasis is that theories and models of language processing cannot afford to ignore the central role played by attention. Researchers in the area of visual word recognition have formulated a number of elegant models of word processing in which attentional influences on processing have typically been ignored or have fallen outside the purview of the model. Clearly such simplifications by omission are to some extent necessary in any research endeavor. However, neglect of the role of attentional processing may not always serve to simplify the theoretical questions. In particular, such neglect runs the risk of asking comparatively inflexible models to account for data from highly flexible cognitive systems. In this light, we believe that an adequate model of visual word recognition must reflect the type of attentional selection that occurs during reading. Therefore, we believe that it is paramount to keep in mind the goal of developing a model of lexical processing in which task demands direct attentional selection to meaning level information.

Of course, in order to appreciate the role of attentional selection in lexical processing, we must first review the elegant theoretical and empirical work that has been developed in the lexical processing literature. Thus, we now turn to a brief overview of some of the most theoretical perspectives on lexical processing.

OVERVIEW OF THEORETICAL PERSPECTIVES

In lieu of attempting to provide an overview of the rich theory that has developed in word recognition research, we will simply highlight a few theoretical perspectives that provide a foundation for the later discussion of the empirical literature. The goal here is simply to mention a few of the approaches to word processing, not to provide a detailed evaluation of any single model. In pursuit of this goal, we have selected Forster’s serial search model (1956), the Seidenberg and McClelland (1980) parallel distributed processing (PDP) model, and Coltheart’s (1978) dual-route model. At the outset, one should note that a goal of this modelling endeavour is to develop a task-independent model of word recognition. It is precisely this assumption that we believe will need to be substantially modified by the constraining influence of attentional selection.

Forster’s serial search model

In general, serial search models (e.g., Becker, 1979, 1982; Norris, 1986; Paap, Newcombe, McDonald, & Schwartz, 1982; Tall & Hambly, 1986) propose that, based on a preliminary visual analysis of a stimulus, an assortment of lexical possibilities becomes available. The input stimulus is compared with each member of a candidate set, one at a time, until a match is found. The search set is typically assumed to be organized so that more frequent words are checked before less frequent words. Thus, search models have an easy way to handle a principal finding in word recognition literature, i.e. low-frequency relatively uncommon words are processed more slowly and less accurately than high-frequency relatively more common words (see Balota & Chumbley, 1984, 1985, 1990; Balota & Spieler, 1999; Monsell, Dowle, & Haggard, 1989, for a discussion of the role of word-frequency in word recognition tasks).

Forster’s (1976) autonomous search model includes a number of distinct access biases to the master lexicon. These access biases correspond to orthographic codes (reading), phonological codes (speech perception), and semantic/syntactic codes (speech production). Most of the work by
Forster and colleagues has dealt with the orthographic access file. The notion is that when a word is visually presented, an orthographically defined bin is created and then searched according to frequency of occurrence until a match is found. Then, based on a pointer from the matched item, the reader can access the plethora of lexical information available with that item in the mental dictionary.

Although much of Forster’s work has dealt with the orthographic access file, other search models, such as Becker’s (1979) verification model, nicely demonstrate why one might wish to hypothesize other access routines, such as one based on semantic access files. Becker used the serial search framework to account for another principal finding in word recognition, the semantic priming effect. The semantic priming effect refers to the finding that readers are faster and more accurate to process a visually presented word (DOG) when it follows a semantically related word (CAT) compared to when it follows a semantically unrelated word (PIN). Becker simply argues that in addition to an orthographically defined search bin, subjects also have a semantically defined bin, which is a generated set of target candidates made available based on the meaning of the prime item. These related candidates are compared with the stimulus item before unrelated items, and hence, one finds a semantic priming effect.

The emphasis in Forster’s original search model is on an autonomous word recognition device that is relatively uninfluenced by attentional control. In fact, such a model was quite consistent with modular views of lexical processing, in which visual lexical presentation drives the search through orthographic bins independent of attentional control. However, the primary data base that was used to test aspects of this model where speeded naming and lexical decision performance, both tasks that place relatively minimal load on meaning processing, at least compared to reading comprehension. Moreover, work by Glanzer and Ehrenreich (1979) and Becker (1980), both advocates of the serial search framework, have demonstrated that the two fundamental aspects of word recognition, word frequency effects and semantic priming effects, can be modulated by probability manipulations across different lists. Interestingly both Glanzer and Ehrenreich and Becker have suggested that list probability manipulations influence the strategic control of different types of search processes. Thus, these studies are quite consistent with the processing pathway approach in suggesting that attentional control can modulate the manner in which access files are searched.

Coltheart’s (1978) dual-route model

This model is motivated by a logical analysis of the problem facing a reader of English. Namely, that while there exists a degree of consistency of mapping spelling to sound correspondences (e.g. MINT) there are many words which violate these “standard” spelling to sound rules (e.g. PINT). In order to address such problems, Coltheart (1978) suggested that there may be two routes through which a word may be named (see Coltheart, Curtis, Atkins, & Haller, 1993, for a more recent computational version). One route entails a direct look-up of the pronunciation in the lexicon (i.e. the lexical route). Presumably, any word that has been learned by the reader is stored in memory along with its correct pronunciation. Whenever the word is subsequently encountered during reading, the visual form is used to access the lexical entry for the word. Once accessed, the correct pronunciation of the word may be retrieved directly from memory. This direct route is frequency modulated and is quite similar to Morton’s (1970) logogen model. The second route employs the pronunciation based on general letter-to-sound rules (i.e. the assembled route). Use of this route is necessary at least for cases in which unfamiliar or new words are encountered, as well as to account for the relative ease with which non-words (pronounceable letter strings such as BLANT and PLATAMARG) may be pronounced (see, however, Mervel, 1980 for an alternative view of non-word naming). Because this route makes use of fairly consistent rules for translating groups of letters into sounds, it will deliver incorrect pronunciations for words that do not comply with these spelling-to-sound rules (e.g. "tint" as a rhyme for "lint"). In this way, the model nicely accounts for the frequency by regularity interaction (e.g. Sedenberg, Waters, Barnes, & Tannenhaus, 1994). Specifically, low-frequency words that have inconsistent spelling to sound correspondences (e.g. PINT) produce slowed response latencies compared to low-frequency words that have consistent spelling-to-sound correspondences (e.g. LINK). Because high-frequency words have a relatively fast lexical route, there is relatively little competition from the slower sublexical route and so one finds little or no consistency effect for high frequency stimuli (e.g. SAME vs. HAVE).

One of the more powerful lines of support for the dual-route model has come from studies of acquired dyslexias, wherein there is evidence of a double dissociation between the two processing pathways. Specifically, for one type of acquired dyslexic (surface dyslexia) there appears to be a breakdown in the lexical processing pathway. Hence, these individuals are fine at pronouncing non-words and regularly spelled words. However, when confronted with an irregularly spelled word these individuals are likely to regularise it, i.e. pronounce BROAD such that it rhymes with the non-word BRODE (see Marshall & Newcombe, 1980; Shallice, Warrington, & McCarthy, 1983). On the other hand, there is a second class of acquired dyslexics, phonological dyslexics, who appear to have an intact lexical route but an impaired phonological route. These individuals can
REVIEW OF THE LITERATURE

Communication channels are the medium through which people can send and receive messages. In order to ensure effective communication, it is important to understand the factors that influence the process. The model of communication proposed by Shannon and Weaver (1949) is widely accepted in the field. This model consists of five components: source, encoder, channel, decoder, and receiver. Each component plays a crucial role in the process of communication.

In the context of computer networks, the model of communication is often referred to as the Berenson model (1997). This model includes additional components such as the medium access control (MAC) layer and the physical layer. The MAC layer is responsible for managing the access to the shared medium, while the physical layer deals with the actual transmission of data over the network.

Another important aspect of communication in computer networks is cryptography. Cryptography is the practice of securing communication from unauthorized access. There are two main types of cryptography: symmetric and asymmetric. Symmetric cryptography involves using the same key for both encryption and decryption, while asymmetric cryptography uses different keys for each process.

In conclusion, understanding the process of communication is crucial for effective communication in computer networks. The models of communication proposed by Shannon and Weaver (1949) and Berenson (1997) provide a framework for analyzing and designing communication systems.
certain issue in our discussion of this hierarchy.

3. ATTENTIONAL CONTROL DURING VISUAL RECOGNITION
2 ATTENTION CONTROL DURING WIDE RECOGNITION
2. ATTENTION CONTROL DURING WORD RECOGNITION

The process of recognizing words is a complex cognitive task that involves multiple stages of attention and information processing. In the context of word recognition, the focus is on the ability to extract and interpret meaningful information from the visual or auditory input. This process is essential for reading, writing, and speaking, and it is influenced by various factors, including the frequency of the word, its semantic and syntactic context, and the individual's knowledge and experience.

For instance, the recognition of a word like "attention" can be facilitated by the activation of related concepts and the integration of contextual information. This process is supported by a network of neural connections that are activated during word recognition, and it is influenced by the individual's attentional resources and the task demands.

In summary, the process of word recognition is a dynamic and interactive process that involves the coordination of various cognitive processes, including attention, memory, and language processing. This process is essential for effective communication and learning, and it is a critical component of human cognition.
The problem of control of dynamical systems is one of the most important problems in science and technology. The solution to this problem is based on the development of methods and algorithms that allow for the stabilization of the system in a desired state. The effectiveness of these methods depends on the accuracy and reliability of the model used. In this paper, we consider the problem of control of a dynamical system described by the differential equations:

\[ \dot{x} = f(x, u) \]

where \( x \) is the state vector, \( u \) is the control input, and \( f \) is a nonlinear function.

The goal of control is to find a control law \( u(t) \) that stabilizes the system at a desired equilibrium point. This problem can be solved using various methods, such as the Lyapunov method, dynamic programming, and optimal control theory.

In this section, we will focus on the use of the Lyapunov method, which is based on the construction of a Lyapunov function that has a negative definite derivative along the trajectories of the system. This function is used to prove the stability of the system and to design a control law that guarantees the desired performance.

The Lyapunov function is defined as:

\[ V(x) = \frac{1}{2} x^T P x \]

where \( P \) is a positive definite matrix. The derivative of the Lyapunov function is:

\[ \dot{V}(x) = x^T P \dot{x} \]

Using the differential equation \( \dot{x} = f(x, u) \), we can rewrite the derivative of the Lyapunov function as:

\[ \dot{V}(x) = x^T P (f(x, u)) \]

The control law can be designed by minimizing \( \dot{V}(x) \) with respect to \( u \). This leads to the following control law:

\[ u = -K x \]

where \( K \) is a gain matrix.

The control law is designed such that the derivative of the Lyapunov function is negative definite, which guarantees the stability of the system.

In summary, the control of dynamical systems is a complex problem that requires the development of advanced methods and algorithms. The Lyapunov method is a powerful tool that can be used to solve this problem. By constructing a Lyapunov function and minimizing its derivative, we can design a control law that stabilizes the system at a desired equilibrium point.
The figure illustrates the process of visual word recognition, which begins with the presentation of a visual stimulus. The visual stimulus is then processed by the visual cortex, which extracts features and responds to the stimulus. These responses are then passed to the visual word-form area, which is responsible for the initial processing of visual words. This area then sends information to the visual word-form area, which is responsible for the more detailed processing of visual words. This process continues until the final decision is made about the identity of the visual word.
The influence of emotional context on word recognition was a focus of research in the 1980s. It was found that the emotional content of a word can affect its processing and recognition. For example, negative words may be recognized more quickly and with less effort than positive words. This effect has been attributed to the automatic activation of emotional processes in the brain, which can influence the way we process information.

In a study published in the journal *Cognition* in 1986, researchers found that participants were more likely to correctly identify words that were associated with positive emotions than with negative emotions. This effect was observed even when the words were presented in a neutral context. The researchers suggested that the emotional content of a word can influence the way we process information by activating specific neural networks in the brain.

The results of this study highlight the importance of considering the emotional context in which words are presented, as it can affect the way we process and recognize them. This has implications for fields such as psychology, linguistics, and education, where understanding how emotional context influences language processing is crucial.
The weather forecast is often referred to as the "planned ignorance" of the future. For example, the weather forecast for tomorrow may predict clear skies, but if it rains, the forecast was actually "right" because it was the most likely outcome. Similarly, if it doesn't rain, the forecast was "wrong" because it was not the most likely outcome. This highlights the nature of prediction and the limitations of our ability to accurately predict the future.

In the context of weather forecasting, the forecast is based on historical data and current conditions. However, the accuracy of the forecast is limited by the inherent uncertainty in the system being predicted. This is known as the "uncertainty principle" in physics, which states that the more precisely we know the position of an object, the less precisely we can know its momentum, and vice versa. Similarly, the more accurately we can predict the weather, the less accurately we can predict the location of a specific object at a specific time.

The forecast is also limited by the finite amount of information available. For example, if we were to measure the temperature at every possible location on Earth, we would still be limited by the fact that we cannot measure the temperature at the same time everywhere. This means that the forecast is always based on a subset of the available information, which introduces additional uncertainty.

In summary, the weather forecast is a good example of the limitations of prediction and the inherent uncertainty in complex systems. While we can make accurate predictions under certain conditions, the uncertainty principle and the finite amount of information available mean that we can never make perfectly accurate predictions of the future.
in which attention is remodeled into processes that can be used in subsequent

In conclusion, we should like to outline a few key insights from this research:

1. The focus on the role of attention in memory and learning is critical for understanding how we process information.
2. The development of attention networks is an important aspect of cognitive development.

Conclusions

Based on our findings, we propose that further research should be conducted on the role of attention in various cognitive processes. This could include exploring the mechanisms underlying attentional control and how they are influenced by factors such as age, gender, and educational background.

Further research could also help to identify effective interventions for improving attentional control, particularly in children and adolescents who may be experiencing difficulties in this area. By better understanding the role of attention in the brain, we can develop targeted strategies to improve attentional control and ultimately enhance cognitive performance.
REFERENCES
2 ATENIONAL CONTROL DURING WORD RECOGNITION