Segalowitz, N. & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J.F. Kroll & A.M.B. De Groot, (Eds.), Handbook of bilingualism: Psycholinguistic approaches. (pp. 371-388) Oxford, UK: Oxford University Press.

Norman Segalowitz
Jan Hulstijn

18

Automaticity in Bilingualism and Second Language Learning

ABSTRACT In this chapter, we examine automaticity in light of the role it might play in second language acquisition and in bilingual functioning. We review various theoretical and operational definitions of automaticity, considering their respective strengths, limitations, and challenges they present to researchers studying automaticity in the context of bilingualism. Studies are reviewed regarding automaticity in grammar acquisition and in lexical access and the connection between automaticity and attention in second language acquisition. The implications of automaticity for second language instruction are also discussed. It is argued that automaticity needs to be carefully defined operationally and always viewed in the larger context of how the control system operates in the acquisition and performance of complex skills.

This chapter examines the role played by automaticity and closely related psychological constructs in bilingual functioning and in second language (L2) acquisition. Why devote a chapter to automaticity in a book about bilingualism? As we show, nearly all psychological approaches to bilingualism and L2 acquisition emphasize the importance, alongside many other factors of course, of *frequency* and practice-based *repetition* involving the mapping of L2 experiences onto their underlying cognitive representations. The prime psychological construct invoked for understanding frequency effects and how repetition leads to improvement in L2 skill (or any skill for that matter) is *automaticity*.

We begin our discussion, therefore, with a review of what automaticity is (and is not) from a general cognitive perspective on skill acquisition, placing automaticity into its larger context and examining different current approaches to its study. This is followed by a discussion of several areas in bilingualism research in which automaticity has received special emphasis. Next, we consider some of the pedagogical implications of research on automaticity for language instruction. Finally, we conclude with a brief discussion of future directions for research on automaticity and bilingualism.

Automaticity

Automaticity has been given both loose theoretical definitions and a number of highly specific operational definitions. Generally, *automaticity* refers to the absence of attentional control in the execution of a cognitive activity, with attentional control understood to imply the involvement, among other things, of intention, possibly awareness, and the consumption of cognitive resources, all in the service of dealing with limited processing capacity (Kahneman, 1973). Some have also associated parallel processing with automatic processing and serial processing with nonautomatic or attention-based processing (Schneider & Shiffrin, 1977; but see Nakayama & Joseph, 1998, for a different interpretation).

A good example of what theorists have in mind by this distinction is what happens when a skilled reader of English recognizes a single letter, say, the letter *A* (Posner & Boies, 1971). Simple, single-letter recognition, it is claimed, requires no conscious effort or effortful attention, is extremely rapid, and cannot be stopped or interfered with by other ongoing activities. When presented with such a stimulus, a fluent reader cannot help recognizing it. It is in this sense that letter recognition is said to be "automatic." In contrast, the recognition of a letter in the Hebrew alphabet by an L1 speaker of English who is only a novice reader of Hebrew might require considerable consciously directed effort, applied slowly over an interval much longer than it takes that same person to recognize a letter of the English alphabet. Thus, the relatively rapid, effortless, and ballistic (unstoppable) activities underlying fluent letter recognition are said to be automatic, standing in contrast to slower, effortful activities that can be interrupted or influenced by other ongoing internal processes (e.g., distractions, competing thoughts).

The characteristics of automaticity mentioned—its rapidity, effortlessness, unconscious nature, and ballistic nature—have each been separately operationalized in various ways in experimental research; some examples are reviewed next. In thinking about these examples, it is important to keep in mind that, in principle, these characteristics do not necessarily always have to bundle together (Bargh, 1992; Neumann, 1984; N. S. Segalowitz, 2003; Tzelgov, 1999). For example, Paap and Ogden (1981) presented evidence showing that fluent letter recognition may be automatic in the sense of being obligatory but nevertheless can consume resources. This illustrates one way in which automaticity does not refer to a unitary construct. It would be an error, therefore, to assume without first doing the requisite empirical research that extensive practice leading to expertise will unfailingly result in performance that has *all* the characteristics typically associated with automaticity.

This distinction between automatic and attention-based processing pervades the cognitive psychological literature on skill acquisition (Ackerman, 1988, 1989; Anderson, 1983; Anderson & Lebiere, 1998; LaBerge & Samuels, 1974; Levelt, 1989; Logan; 1988; Proctor & Dutta, 1995) and is central to many treatments of L2 acquisition (DeKeyser, 2001; N. C. Ellis, 2002; Hulstijn & Hulstijn, 1984; Johnson, 1996; McLaughlin & Heredia, 1996; McLaughlin, Rossman, & McLeod, 1983; and N. S. Segalowitz, 1997, 2003). As will become evident from discussion in this chapter, the idea of automaticity is itself evolving, especially as researchers devise different ways to operationalize what they mean by it.

Broadly speaking, two general theoretical approaches have been followed in attempts to understand the place of automatization during skill development. One approach is typified by Anderson's ACT* (ACT-Star: Adaptive Control of Thought) model of skill acquisition (Anderson, 1983; Anderson & Lebiere, 1998). This approach holds that, in the early phases of skill acquisition, performance largely relies on mechanisms that are under conscious control, often involving declarative knowledge (Anderson, 1983). As the learner gains practice, sequenced components of the new skill that are repeated become routinized or "chunked," rendering them very fast and efficient and unavailable to conscious awareness. The declarative knowledge is said to become proceduralized, and the change is sometimes compared by analogy to the compilation of a computer subroutine that involves converting instructions encoded in a high-level interpreted language into lower-level machine language.

An alternative approach is Logan's (1988) instance theory of automatic processing. Logan proposed that initially performance of a to-be-mastered skill is based on a set of algorithms for executing the desired action. Each time the rule is carried out, there is a new memory trace formed corresponding to the action executed. On subsequent occasions, there is a race between an algorithmic process that constructs the appropriate response and a retrieval process that searches memory for the information needed to perform the action. With increasing practice, more and more representations of the response are stored in memory, so eventually retrieval is accomplished faster than is execution of the algorithm. Logan's theory thus holds that automatization in skill acquisition involves a shift from rule-based to memory-based performance. Logan's theory is able to account very well for the power law (Newell & Rosenbloom, 1981) property of skilled performance, which refers to the frequent observation that response latency decreases as a function of the number of practice instances raised to some power (Logan, 1992).

Theoretical Perspectives and Empirical Studies on Automaticity in Bilingualism

Empirical studies addressing questions about automaticity and bilingualism can be viewed from various perspectives. We review studies that examined (a) automaticity as a characteristic of proficiency, (b) automaticity as a factor in grammar rule acquisition, (c) the relation between automaticity and attention, and (d) bilingualism as a testing ground for learning more about automaticity.

Studies of Automaticity in Second Language Proficiency

The theoretical starting point for studies in this category is the view that accords a central place to automaticity in the development of a complex skill, especially the "compilation" of production units as advocated by Anderson (1983; Anderson & Lebiere, 1998), Ackerman (1988, 1989), and others.

For a specific example related to language skill, consider Levelt's (1989, 1999) "blueprint" of the cognitive architecture of the system responsible for fluent, intentional speech. One component of this architecture is an executive system that he called the conceptualizer, a functional level responsible for generating preverbal messages and monitoring the speaking activity. The work of the conceptualizer is largely an attention-based process in which the speaker is able to attend to the communicative event as it unfolds. Levelt admitted that much of the work of the conceptualizer may become directly or automatically available to the speaker because of a lifetime of practice at generating messages, monitoring the situation for turn-taking cues, selecting sociolinguistically appropriate forms of speech, and so on. Yet, Levelt pointed out that speakers can nevertheless easily attend to these aspects of speech if necessary. In this sense, these aspects of the work of the conceptualizer are not "informationally encapsulated" (nonmodular); that is, they do not function as relatively autonomous cognitive processing modules and so are not automatic.

The three remaining components of the cognitive architecture proposed by Levelt (1989, 1999) are the formulator, responsible for grammatical and phonological encoding; the articulator, responsible for the neuromuscular execution of the phonetic plan; and the speech-comprehension system, responsible for providing the conceptualizer with information required for self-monitoring (see La Heij, chapter 14, and Costa, chapter 15, this volume). In contrast to the conceptualizer, the remaining three components are held to be largely automatic in the sense of being informationally encapsulated (modular). For example, selection of eligible grammatical argument categories permitted by a verb that has been selected, individual word retrieval, or phonological encoding and articulation all occur in an automatic fashion (although Levelt did allow for marginal forms of executive control even here; Levelt, 1989, p. 22).

As for speaking in L2, the issue then becomes whether the scope of automaticity in L2 is the same as in L1. Several authors have followed up on this. Pienemann (1998) proposed a processability theory using Levelt's model as a starting point for a theory of how procedural, and hence automatized, skills develop in the L2 learner (see Pienemann, Di Biase, Kawaguchi, & Håkansson, chapter 7, this volume). De Bot (1992) also examined how Levelt's model could be applied to the L2 context.

The implications for bilingual development of these ways of thinking about automaticity are clear; later phases of L2 learning must involve automatic processing to a far greater degree than do earlier phases, and this automatic processing will play a significant role in distinguishing fluent from nonfluent abilities. Such an idea may have important consequences for both theory and practice because a role for automaticity in acquiring L2 proficiency could be taken as support for a general skills approach to language learning. Such a skills approach (e.g., Johnson, 1996; McLeod & McLaughlin, 1986) stands in contrast to approaches emphasizing the fine tuning of underlying competence parameters believed to constitute an innate universal grammar (e.g., White, 1989, 1996). Thus, a focus on automaticity in bilingualism may be of significant interest insofar as it has the potential to link a theory of bilingual ability to the broader and more general psychological literature on cognitive skill development.

An early demonstration of the importance of automaticity for understanding L2 proficiency was provided by Favreau and Segalowitz (1983). They compared two groups of relatively fluent L2 readers. One group, here referred to as the stronger bilinguals, comprised people able to read L2 and L1 texts equally fast to achieve the same level of understanding. The other group is referred to here as the weaker bilinguals, although it must be remembered that people in this group still were very strong in an absolute sense. Participants in the weaker group, unlike those in the stronger group, read the L2 more slowly than L1 to achieve comparable levels of comprehension in the two languages. The goal of the study was to see whether the stronger bilinguals possessed more automatic single-word recognition skills than the weaker bilinguals.

To investigate this, Favreau and Segalowitz adapted Neely's (1977) operational definition of automaticity as a ballistic or unstoppable process. This involved a primed lexical decision task in which participants saw a priming stimulus followed by a target stimulus. The prime was either a category name (e.g., FRUIT) or a meaningless string of symbols. The prime signaled the onset of the upcoming target. The target was either a word naming an exemplar from the prime category (e.g., APPLE), an exemplar from another category (e.g., TABLE), or a nonword. The subject had to judge the word/nonword status (lexical decision) of the target. In some conditions, the participants were trained to expect a prime word like FRUIT to be followed by a semantically *unrelated* target word such as TABLE.

Like Neely, Favreau, and Segalowitz (1983) found that, once participants were suitably trained, they showed appropriate facilitation and inhibition effects in L1. For example, with a long interval (1,150 ms) between prime and target, a prime like FRUIT facilitated lexical decision to an *expected* but semantically unrelated target like TABLE or CHAIR, relative to the neutral prime condition. In contrast, a prime like FRUIT inhibited responses to an *unexpected* yet semantically related target like APPLE that was occasionally presented on surprise trials. On the other hand, when the prime–target interval was short (200 ms), lexical decision on these surprise trials was facilitated, indicating that the subject could not suppress the activation of semantically related concepts (APPLE, ORANGE, BANANA, etc., by FRUIT) even though instructions and training indicated that such targets were not predicted by the prime. In this way, the experiment demonstrated the ballistic nature of word meaning activation.

Favreau and Segalowitz (1983) found that in L2 only the stronger group showed this form of automaticity. Interestingly, the evidence also indicated that the weaker bilinguals did not process stimuli more slowly but only less automatically. This research illustrates the important point that subtle cognitive processing differences can exist between groups of relatively highly skilled L2 users (all bilinguals in this study were able to read mature texts to full comprehension), in which for some people (e.g., the stronger bilinguals) certain underlying processes operated in a ballistic fashion; for others, they did not.

Segalowitz and Segalowitz later proposed a somewhat different approach to the study of automaticity (N. S. Segalowitz & Segalowitz, 1993; S. J. Segalowitz, Segalowitz, & Wood, 1998). Practice and experience with a language typically lead to faster processing, which is commonly reflected in various ways, including faster lexical decision times, faster rates of speaking and reading, and better ability to process rapid speech. N. S. Segalowitz (2000; N. S. Segalowitz & Segalowitz, 1993) pointed out, however, that if the construct of automatic processing is to have explanatory value, then—to avoid circularity—the term *automatic* should be more than a synonym for *fast*.

As a consequence, there is a need to distinguish operationally between the following two situations, each involving a contrast between fast and slow performance. The first is Situation A, in which the faster performance is simply caused by a difference in the run-time speeds of the processes underlying performance and not some difference in the selection of which processes are involved or in the way processes interact with each other. In this case, there is no need to invoke the idea of *automatic processing*, defined now to mean more than fast processing to avoid circularity, to explain the difference in performance.

In contrast, there is Situation B, in which faster performance is caused by more than just a difference in the speed of underlying processes. Here, the difference may lie in the way underlying processes are organized, such as when L2 visual word recognition proceeds directly from the printed stimulus to meaning activation without first passing through a stage of phonological recoding or translation into L1. Or, instead, the difference might lie in the internal organization of a given process without necessarily involving the elimination of one or more stages of processing. Such differences could lead, for example, to more ballistic processing, more parallel processing, and so on, resulting in significantly faster and more efficient performance.

N. S. Segalowitz and Segalowitz (1993) described the fast–slow contrast of Situation A as a case reflecting simple speed-up and the fast–slow contrast in Situation B as a case for which the difference can more appropriately be attributed to automaticity. They proposed that when attempting to determine if automaticity underlies a given case of fast responding, an attempt should be made to reject the null hypothesis that the performance could be caused by merely generalized speed-up.

N. S. Segalowitz and S. J. Segalowitz (1993; S. J. Segalowitz et al., 1998) proposed a way to test—and therefore potentially reject—the speed-up null hypothesis. They argued that, when faster processing is caused only by generalized speed-up of the processes underlying performance, the standard deviation of the reaction time should drop proportional to the reduction in the reaction time. This idea can be understood at an intuitive level by considering the following metaphor. Suppose a videotaped recording of a person making a cup of tea on 50 different occasions is viewed. Each component of the action—putting the water on to boil, pouring the hot water into a cup, inserting the tea bag, and so on—will take a particular length of time. A mean execution time and a standard deviation for this mean can be calculated across the 50 repetitions both for the global action of "making tea" and for each component of this event. Suppose now a new videotape is created by rerecording the original at twice the normal speed. On the new tape, the entire event will appear to be executed in half the time with half the original standard deviation overall; moreover, the mean duration of each component and the standard deviation associated with each component will also be reduced by exactly half. This situation corresponds to what N. S. Segalowitz and Segalowitz (1993) argued to be the null case of generalized speed-up; performance becomes faster because the underlying component processes are executed more quickly and for no other reason. (Of course, this account makes a number of simplifying assumptions about the brain, including that the component processes are organized serially only. They probably are not. However, the scenario described would apply to both the nonoverlapping aspects of the underlying components and to those that are organized serially, which together determine the total time of execution.)

Suppose now we are shown still another videotape in which the mean time for the global action of making tea is again half the original mean time, but the standard deviation for the 50 repetitions is far less than half the original standard deviation. This tape cannot have been produced simply by rerecording the original at twice the normal speed. Instead, there must have been some change in the way the activity of making tea had been carried out, such that some of the slower and more variable components of the action sequence had been dropped or replaced by faster, less-variable components. In other words, there must have been a change that involved more than simple speed-up, namely, some form of restructuring of the underlying processes.

According to this approach, if it is believed that practice and experience have produced some cognitive change other than generalized speed-up—restructuring, more ballistic processing, reduced reliance on decision processes, and so on—then one should try to reject the null hypothesis represented by generalized speed-up. N. S. Segalowitz and Segalowitz (1993) proposed that if faster performance reflects more change than is accounted for by speed-up, then the standard deviation should change by a greater proportion than that seen in the reaction time. Put another way, if the coefficients of variability—the ratio of the standard deviation to the mean reaction time for each individual—remain the same while reaction times become faster (that is, both standard deviation and reaction time change by the same proportion), then there will be no grounds for rejecting the speed-up null hypothesis, and there will not be a significant correlation between reaction time and coefficient of variability across subjects. If, on the other hand, the coefficient of variability is significantly reduced as reaction time becomes faster, then the null hypothesis can be rejected, and a claim can be made that there has been a change—which N. S. Segalowitz and Segalowitz (1993) called automatization—that must reflect a different recruitment or organization of underlying mechanisms. In this case, as the reaction time reduces, so does the coefficient of variability, and there will be a significant correlation between the two. (See Wingfield, Goodglass, & Lindfield, 1997, for a different approach to dissociating speed of processing from automaticity.)

N. S. Segalowitz and Segalowitz (1993) collected lexical decision data from adults who varied in ability in L2 English or L2 French (S. J. Segalowitz et al., 1998). The results were consistent with their approach for distinguishing automaticity from speed-up. They found that coefficient of variability varied with reaction time in those conditions for which faster responding was logically expected to reflect a change involving more than just speed-up. They also found that coefficient of variability did not vary when faster responding was expected to reflect only speed-up (see also N. S. Segalowitz, Poulsen, & Segalowitz, 1999). These results are interesting for two reasons. Methodologically, they demonstrate how to move beyond merely speculating that an observed case of increased performance speed reflects a higher level of automaticity; it is now possible to assess the degree to which this performance is not solely attributable to generalized speed-up. On a theoretical level, this research demonstrated that higher levels of L2 proficiency, unlike lower levels of L2 proficiency, are associated with more than just differences in processing speed.

Caution must be taken, of course, when using the coefficient of variability analysis just described. Failure to reject the generalized speed-up hypothesis carries with it the usual caveats concerning failure to reject the null hypothesis; it is always wise, therefore, to have convergent evidence to support a generalized speed-up account to conclude from failure to reject that speed-up is what actually occurred. Also, the method of analysis proposed by N. S. Segalowitz and Segalowitz (1993) does not address the many interesting questions that could be asked regarding the kind of change that has taken place when analysis supports a claim for automatization; it only allows concluding that something other than generalized speed-up occurred.

Further research is always required to pinpoint the exact nature of the change; however, analysis of the coefficient of variability may again be useful in that follow-up research. For example, suppose the results of a study indicated that L2 word recognition became faster after some particular form of training, and that a generalized speed-up explanation can be rejected by the coefficient of variability analysis. Follow-up research using a design permitting a coefficient of variability analysis could be useful for looking into whether performance improved because, say, perception of orthographic redundancies (knowledge of spelling pattern frequencies) or phonological recoding had become more automatic.

Automaticity and Grammar Rule Acquisition

Perhaps one of the most hotly debated issues in the field of foreign or L2 learning concerns the learning and subsequent use of explicit grammar rules. Currently, there are three main theoretical positions on this issue, commonly referred to as the strong interface, weak interface, and no interface positions (R. Ellis, 1993; Larsen, Freeman, & Long, 1991, p. 324). Adherents of the strong interface position claim that explicit, declarative knowledge can be transformed or converted into implicit knowledge through practice, as proposed in Anderson's skill acquisition theory (Anderson, 1983; Anderson & Lebiere, 1998). According to the weak interface position, explicit, declarative knowledge may somehow, in a way not yet properly understood, facilitate the acquisition of implicit, procedural knowledge. The no interface position denies a causal role of explicit knowledge in the acquisition of implicit knowledge. In the area of language pedagogy, Krashen (1981) is perhaps the best-known proponent of the no interface position. For a discussion of the theoretical issues involved in the three positions, see the work of R. Ellis (2000), Hulstijn (2002), and Paradis (1994).

Little empirical research has been conducted to test claims made on the basis of these three positions in relation to issues of automatization. Robinson and Ha (1993) and Robinson (1997) investigated the learning of the so-called dative alternation rule of English by adult speakers of Japanese, Korean, and French (Robinson & Ha, 1993) and Japanese (Robinson, 1997) in a single learning session lasting not longer than 30 min. (*Dative alternation* refers to the fact that, for some monosyllabic verbs in English, the indirect object form can alternate with the direct object form, as in "She gave the book to the boy" and "She gave the boy the book," whereas some bisyllabic verbs only allow the indirect form, as in "She donated the painting to the museum.") In these studies, automaticity was defined as reaction time patterns conforming to the power law. Participants in the 1993 study were presented with the dative alternation rule. Subsequently, in the training phase, they were shown 36 sentences, one at a time. They had to indicate whether the sentence did or did not conform to the rule just presented. Feedback was given on the correctness of each response. There were 8 sentences in the training set. One sentence was presented eight times, one sentence seven times, one six times, and so on, and the 36 sentences were presented in random order. In a subsequent transfer test, participants performed the same task, this time with 32 sentences, 8 of which were identical to the ones used in the training set. Reaction times of responses to old sentences, which had been presented in the previous training phase, were faster than those to new sentences. However, no evidence was found for the hypothesis, based on Logan's instance theory (1988), that reaction times would be faster for sentences presented more often in the training phase than for sentences presented less often.

In interpreting the complex findings of this study, it is important to bear in mind that it was concerned with the application of a rule, explained in advance, in a metalinguistic task (grammaticality judgment) rather than in a functional listening, reading, speaking, or writing task, and that the training phase comprised only 36 trials. We concur with DeKeyser's (2001) interpretation that "neither rule application nor instance retrieval was at work, but a similarity-based item retrieval process" (pp. 142–143). The pattern of results in the 1997 study, which adopted a more complex design and addressed other issues in addition to Logan's instance theory, was similar to that of the 1993 study regarding instance learning. Again, no gradual improvement as a function of number of previous item presentations was found. In summary, the two Robinson studies did not provide evidence for automatization as operationally defined.

Healy and her coworkers (Bourne, Healy, Parker, & Rickard, 1999; Healy, Barshi, Crutcher, et al., 1998) investigated the acquisition of easy and difficult rules by adult native speakers of English. The easy rule required pronunciation of the article *the* as *thuh* or *thee* when preceding nouns beginning with a consonant or a vowel, respectively. The difficult rule required judging the order of letters in meaningless three-letter sequences, such as the invalid LMV and the valid PRQ sequences. PRQ is valid because it can be rearranged to correspond to a sequential string in the alphabet (PQR), whereas LMV cannot. Participants in both experiments were presented with well- and ill-formed stimuli. They judged the stimuli's well-formedness and received feedback on the correctness of their responses. Participants also reported whether their responses were based on a guess, on a rule, on memory of the instance, or on other strategies. In both experiments, response accuracy rose to around 95%, and latencies dropped over the course of 30 learning blocks. Healy et al. (1998) reported that:

> Although all subjects [in the difficult-rule experiment] guessed initially, many subjects soon discovered and started using the rule. However, by block 6, rule use began to give way to an instance strategy so that by the end of 30 blocks of practice, subjects exhibited the instance-based strategy almost exclusively. (p. 26)

In the easy-rule experiment, 40% rule use was reported initially, suggesting that some participants, not surprisingly, were familiar with the *thuh/thee* rule from the start. In Block 30, participants reported using the rule 65% of the time. An interesting finding was that, in the case of the easy rule, rule use resulted in faster response latencies than did use of the instance strategy, whereas the reverse pattern was obtained in the case of the difficult rule. In interpreting these results, one has to bear in mind, as in the case of the Robinson studies reported above, that participants were engaged in a metacognitive judgment task rather than in a speech production task requiring the application of the rules.

N. C. Ellis and Schmidt (1997) and DeKeyser (1997) investigated how adult, literate native speakers of English acquired some rules of grammar of an artificial language in a computer-controlled laboratory setting. These experiments were limited to the written mode for input and output; listening and speaking were not involved. In all three studies, the participants had to learn patterns in an artificial language that were analogous to grammatical rules. Participants in the first experiment reported in the study of N. C. Ellis and Schmidt (1997) had to learn plural forms in an artificial language (e.g., *bupoon* for the plural of the artificial word *poon*, meaning plane), some of which conformed to frequency criteria that made them "regular" plurals, whereas others did not and hence were "irregular." Participants studied the artificial language names given to 20 picture stimuli in 15 sessions of 1 hour sessions and spanning up to 15 days. Participants in the second experiment were shown meaningless artificial language sentences for a period of 75 min. Participants in DeKeyser's study (1997) were shown artificial language sentences with pictures illustrating their meaning in 22 sessions of an hour or less and spread over an 11-week period. The exposure-learning regimes in these studies differed somewhat, but they had in common that both accuracy and reaction times of participants' responses were measured during the learning. All three studies showed an increase in response accuracy for stimuli conforming to the appropriate grammatical patterns and a concomitant decrease in latency over the course of trials and sessions following a power law of learning. The authors interpreted these results as evidence for automatization of grammar learning. The main focus of these studies, however, was on the issues of implicit versus explicit learning and top-down learning by rule versus bottom-up learning by association and analogy.

N. C. Ellis and Schmidt (1997) argued that the findings of their studies can be accounted for by a simple associative learning mechanism even in the case of the acquisition of regular rule-governed forms. DeKeyser (1997) found that performance in both comprehension practice and production practice followed the same power function learning curve, but that acquisition was skill specific, showing little transfer from comprehension to production and vice versa. DeKeyser argued that L2 rules can be learned in much the same way as learning in other cognitive domains and can be accounted for by Anderson's model of skill acquisition, according to which declarative knowledge, with practice, turns into procedural knowledge.

One of the crucial issues in the debate between proponents of the strong, weak, and no interface positions is concerned with the meaning of the expression *turn into* (*transform* is used as a synonym in this debate) when it is claimed by some and denied by others that explicit knowledge can turn into (or transform into) implicit knowledge. Does this mean that explicit knowledge undergoes a metamorphosis such that, eventually, explicit knowledge has ceased to exist and that, "in its place" implicit knowledge has arisen?

Such a view implicitly rests on the idea that first there is an area in the brain where explicit knowledge resides, and furthermore that, during the process of proceduralization, implicit knowledge is formed, settling itself in the same area, forcing explicit knowledge to dissolve. However, such a strong view of transformation is not supported by brain research. Brain research suggests that declarative knowledge resides in the medial temporal lobe, including the hippocampus, whereas implicit knowledge is distributed over the neocortex (Paradis, 1994; Squire & Knowlton, 2000; Ullman, 2001). Viewed from this neurophysiological perspective, the strong interface position in the L2 acquisition field should be taken to mean that explicit knowledge forms a prerequisite for implicit knowledge to come into existence rather than the claim that explicit knowledge transforms into implicit knowledge.

The evidence of the studies reviewed in this section are consistent with Willingham's (1998) position that, already in the initial phases of learning, implicit knowledge is spontaneously formed, and that explicit processes are simply not used any longer in later phases. The practical relevance of the interface issue remains great: Of course, language teachers and language learners alike want to know to what extent knowledge of grammar rules may foster or hinder the attainment of fluency in language use. In terms of theoretical explanations, however, the interface issue is likely to form part of the much broader neurocognitive issue of explicit and implicit cognition. Empirical evidence may come not only from behavioral data (such as response time and response variability, presented elsewhere in this chapter) but also from neurophysiological data (such as event-related potential and neuroimaging).

Automaticity and Attention in Second Language Proficiency

The research reviewed so far attempted to integrate the concept of automaticity with theories of L2 proficiency and L2 grammar learning. N. S. Segalowitz (1997, 2000) proposed, however, that automaticity addresses just one component of a larger set of issues underlying "cognitive fluency" or processing efficiency that is responsible for the linguistic fluency or proficiency (the rapidity, fluidity, and accuracy) observed in a bilingual individual. Besides automaticity, there is a complementary, nonautomatic aspect involving attention-based processes that are also required for fluent language use. These operate in a close fashion with more automatic processes to determine the overall underlying efficiency of L2 functioning.

Such attention-based processes include focusing on (directing awareness to) the language itself while learning it, such as the noticing and focus-on-form skills that may be necessary for successful learning (Doughty & Williams, 1998; Lightbown & Spada, 1990; Robinson, 1995; Schmidt, 2001). Selective attention is also involved in fluency insofar as the ability to focus on the speech stream as a channel of communication under noisy conditions or focus selectively on phonological cues carrying sociolinguistic messages or on cues to turn taking and the like, as pointed out by Levelt (1989) (see also Eviatar, 1998, and Fischler, 1998, for more on selective attention and language). Finally, there is the attention-directing function of language itself, in which language is used to shape the way a listener or reader builds a mental representation of the message conveyed. This attention-directing function is believed by cognitive linguists to be central to the communicative purpose of language (Langacker, 1987; Talmy, 1996, 2000).

N. S. Segalowitz and Frenkiel-Fishman (in press) found that L2 skills reflecting attention-directing functions of language were significantly related to levels of automaticity of single-word recognition as indexed by the coefficient of variability measure described here. This study involved an attention-shifting task adapted from Rogers and Monsell's (1995) alternating runs paradigm. The stimuli were time adverbials and conjunctions, both good examples of words that serve to direct a person's attention in particular ways while building a mental representation of a message's meaning. Time adverbials direct the listener/reader on how elements of a mental representation should be foregrounded or backgrounded with respect to time. Conjunctions convey the need to form particular links between elements of a mental representation.

Participants were given two tasks (N. S. Segalowitz & Frenkiel-Fishman, in press). In one they had to judge the meaning of a target word belonging to the time adverbial stimulus set. The other task required them to judge a conjunction. For example, in the time adverbial task, subjects judged whether a word (*soon*, *later*, etc.) referred to a moment in time relatively close to or relatively far from the present moment (as an illustration, compare the meanings of "I'll do it *soon*" versus "I'll do it *later*"). In the conjunction task, subjects judged whether a word (*because*, *despite*, etc.) normally indicates the presence or absence of a causal link between the clauses it conjoins (e.g., compare "John passed the exam *because* he studied all night" versus "John passed the exam *despite* partying all night").

In the N. S. Segalowitz and Frenkiel-Fishman (in press) experiment, on each trial either a time adverbial or a conjunction appeared in one of four spatial locations on a screen. This location indicated which task (time adverbial or conjunction judgment) was to be performed. As in the work of Rogers and Monsell (1995), the tasks alternated in a predictable manner according to the sequence "...adverbial adverbial conjunction conjunction..." and thereby requiring a repeat of a given task and a switch to the alternate task on every second trial. This design provided a measure of the switch cost, that is, the cost in response time to switch from one task to the other, compared to repeating a task. Participants performed the experiment in separate L2 and L1 blocks, thus providing a measure of switch cost in each language.

In a separate part of the study (N. S. Segalowitz & Frenkiel-Fishman, in press), subjects' ability to process L2 word meaning was indexed in terms of the coefficient of variability of latency (as discussed in the section Studies of Automaticity in Second Language Proficiency) in a classification task in which nouns were judged as referring to living or nonliving objects. Here also, L1 measures were used as baseline. The results indicated that the switch cost in L2 was significantly correlated with the coefficient of variability of reaction time in the classification task after taking into account performance on the same tasks in L1. The results were interpreted as indicating that attention-focusing skill is related to proficiency as indexed by switch cost and coefficient of variability of reaction time respectively.

Although it is beyond the scope of this chapter to discuss further the role of attention-based processes in L2 functioning and in fluency acquisition (see Schmidt, 2001), it is important to keep in mind that automaticity cannot really be talked about without also talking about attention. Automaticity is recognized only by virtue of its contrast to nonautomatic or less-automatic (attention-based) modes of processing.

Moreover, there are automatic modes of processing that are fully integrated within nonautomatic modes, and it is impossible to fully tease them apart. To illustrate, consider the relatively "simple" case of reading a sentence in L2. One has to process letters, words, and syntactic patterns and integrate all this into the ongoing construction of a representation of the meaning of the sentence and of the larger text. Reading will, of course, be fluent to the extent that many of the mechanisms involved are ballistic and do not consume resources better used for other purposes. However, such a need for automaticity can be identified at *all* levels of processing, from relatively "low-level" letter recognition to aspects of relatively "high-level" attention focusing (see also Tzelgov, Henik, & Leiser, 1990, for a similar point).

This tight relationship between automatic and attention-based mechanisms raises the following interesting question that has yet to be addressed empirically: Do the attention-based and automatic components of proficiency develop independently? If so, can such development account for individual differences in learning success in a given learning context (e.g., a classroom, study abroad, immersion, etc.)? If not, there are at least three alternatives to consider: (a) Do attention-based language skills require a threshold level of automatic processing before they can develop? (b) Does the acquisition of automatic processing abilities require some critical level of supporting attention-based mechanisms in place? (c) Should automatic and attention-based processing be conceived as mutually dependent? These questions have important practical value in addition to theoretical interest because the answers may point in particular directions regarding the most effective way to organize L2 learning experiences.

Studies Using Bilinguals to Investigate Automaticity

The studies reviewed in the previous sections directly addressed questions about the role played by automaticity in bilingualism. Next, we review several related examples of research that made use of the automatic and nonautomatic characteristics of bilingualism to study automaticity itself and related constructs in addition to contributing directly to an understanding of bilingualism as such.

One interesting study in this category is Meuter and Allport's (1999) study of attention. Meuter and Allport were interested in the processes responsible for the shift cost or slowed response time observed when subjects have to perform tasks in two different languages. In their study, bilinguals named numerals shown on the screen using L1 or L2 in a paradigm in which the language of response was cued by color.

Meuter and Allport (1999) found that the cost associated with switching to L1 (that is, the slowing of the L1 response observed after having just responded in L2 compared to having just responded in L1) was greater than the cost associated with switching to L2. This effect is paradoxical because normally it would be expected to be easier to switch to the stronger L1 than to switch to the weaker L2. In fact, however, the authors had predicted this paradoxical effect from their theory of the nature of switch costs. They believe that the cost observed on a given switch trial reflects the need to overcome inhibition activated on the immediately preceding trial. Thus, on a switch trial involving an L2 response, the bilingual has to suppress or inhibit the automatic activation of the competing stimulus name in L1 to respond correctly in L2. If, however, the switch trial requires an L1 response, the bilingual has to do two things: cancel the inhibition to responding in L1 that was activated on the previous trial and overcome any persisting inhibition from that trial.

Meuter and Allport's results were consistent with the idea that there is automatic activation of L1 representations in L1 naming tasks, whereas there is little or no automatic activation of L2 representations in an L2 naming task; this L1 activation may be difficult to overcome when competition between the languages is important. Presumably, their paradigm could be adapted to quantify this automatic activation when it is useful to measure an individual's degree of balance between L1 and L2 in terms of automatic processing and attention flexibility. (See also chapter 17, by Meuter, this volume.)

Bialystok (2001) reported a series of highly original studies on the possible cognitive benefits associated with early bilingualism (also, see Bialystok's chapter 20 in this volume). She investigated what happens when both languages are automatically activated and are always in competition because the individual is growing up bilingual. She compared bilingual children learning their two languages at the same time with monolingual children and found that, in certain nonlinguistic domains, children with strong L2 abilities outperformed monolingual children. The results were consistent with the following idea: Because almost every waking moment involves dealing with language-based cognitive demands, a bilingual with two (or more) equally strong languages at his or her command continually has to inhibit competition from the currently not-to-be-used language(s), competition that arises from the automatic activation of language representations elicited by ongoing thoughts and by stimuli in the environment. For the young bilingual child, this may constitute intensive training of frontal inhibitory systems, training that normally does not occur to the same degree for monolinguals. If correct, this view would then suggest that the automatic activation of language-based representations can, in a bilingual child, have far-reaching consequences by providing sustained training of inhibitory systems that are required even for nonlinguistic cognitive activity (such as those documented by Bialystok). This idea merits further investigation, especially through studies using more direct measures of the automatic nature of language activation and suppression.

Automaticity has also been studied in bilinguals with a view to understand the nature of lexical access. Tzelgov, Henik, Sneg, and Baruch (1996), for example, exploited certain automatic aspects of reading in bilinguals to understand further the nature of lexical access in skilled readers. Some theories of skilled reading hold that readers access meaning from print automatically in a process that is mediated by preassembled phonological representations (Van Orden, 1987) developed during earlier phases of skill acquisition. Other theories suggest that automatization in reading skill acquisition involves a shift from dependence on assembled phonological representations to direct access of meaning from visual input (Waters, Seidenberg, & Bruck, 1984). The mediated access approach characterizes automatic processing as making use of activity-specific "precompiled" productions, as proposed by Anderson (1983) in his process-based ACT* approach.

In contrast, the direct access approach characterizes automatic processing as a memory-based, single-step retrieval process, similar to Logan's (1988) memory-based instance theory of automaticity (i.e., a shift from algorithmic to instance retrieval). Tzelgov et al. examined bilingual readers in a Hebrew-English version of the Stroop paradigm (Stroop, 1935). They used cross-script homophones, such as Hebrew color words written in the Latin (English) alphabet (e.g., *adom*), and English color words written in Hebrew letters that, when sounded out, sound like English words.

Consider now the case in which *adom* is written in green ink, and the correct response is therefore "green." According to the mediated access approach, if the subject is a Hebrew speaker and a skilled reader of English, he or she will automatically access via a phonological route the concept of red because */adom/* in Hebrew means red. According to the direct access approach, however, the phonologically based link between */adom/* and the concept red will be bypassed.

In a series of experiments with Hebrew-English bilinguals, Tzelgov et al. studied whether the automatic processing underlying skilled L2 reading made use of the phonological route (and hence precompiled productions) or the direct route (and hence instance retrieval). They reported finding a strong cross-script Stroop effect, particularly when the stimulus was a transliteration of a color name in Hebrew, the subject's L1 (*adom* activating *red*). The results thus supported the first model described above, namely, that unintentional automatic processing in reading involves precompiled phonological productions and not retrieval of stored instances.

Tzelgov et al. (1996) argued, on the basis of these results and others they obtained, that there is evidence for two different, coexisting forms of automaticity, one involving activity-specific precompiled productions and the other the development of a database for memory retrieval in the execution of the skill in question. The results were also interpreted as support for the asymmetric model of bilingual memory proposed by Kroll (e.g., Kroll & Stewart, 1994) because the Stroop effects were themselves asymmetrical as a function of which language was L1 (see Kroll & Tokowicz, chapter 26, and Dijkstra, chapter 9, this volume, for related discussion).

In sum, it can be seen from these studies that bilingualism can provide a particularly useful situation for studying cognitive mechanisms not only as they relate to L2 processes, but also as they relate to basic, more general cognitive issues such as automaticity.

Instructional Implications

Questions about what role, if any, automaticity plays in L2 acquisition and proficiency will naturally have implications for how to optimize language instruction. Here, the central instructional question is the following: Once language learners have been exposed to new linguistic information, what must they do to be able to achieve later automatic access to that information? As pointed out, the functional, communicative use of language involves the simultaneous manipulation of many linguistic elements at different levels, ranging from the higher levels of content and discourse organization to the lower levels of processing speech sounds and letters (in oral and written communication, respectively). Given the fact that humans have a limited capacity for information processing, it is obvious that language users cannot pay attention to all information at all linguistic levels simultaneously to the same high degree. In most communicative situations, the processing of information at the higher levels—that is, information concerning the content and the course of the communication—consumes much of this limited capacity.

VanPatten (1990), for example, reported a study indicating that, in the early stages of L2 acquisition, learners find it difficult to focus both on message content and various aspects of form (verb form, grammatical functors). Because of the novelty of most communicative acts, the processing of information at the higher levels can hardly be automatized. What can be automatized to a large extent, however, is the processing of information at the intermediate levels of the retrieval of words to express personal thoughts; processing at the lower levels of the planning of the morphosyntactic, phonological, and phonetic aspects of the utterance; and the execution of the planned part of an utterance with the aid of speech organs accompanied by appropriate gestures.

This is what happens during the many years of L1 acquisition and what has turned most adults into fluent speakers of their native tongue. It is therefore important, in the case of L2 instruction, to devise tasks that do not require the allocation of much attention to the higher levels of information, allowing learners to pay attention to information at particular lower levels standing in need of automatization.

The basic principle underlying such tasks is repetition (N. C. Ellis, 2002), but as we discuss in the following sections, simple repetition as such cannot be the whole answer. To the extent learning is promoted through repetition (e.g., in the case of the automatization of the word-by-word understanding of speech), learners should listen to materials that do not contain very many unfamiliar words, preferably several times. Similarly, in the case of reading, learners should be given linguistically "easy" (but yet authentic or quasi-authentic) texts to allow them to increase their reading speed. Teachers and learners alike should make a principled distinction between two types of listening and reading activities: exposure to materials containing new linguistic elements for the purpose of acquiring new knowledge and exposure to materials containing familiar elements for the purpose of automatization.

Research on the Training of Lexical Access

Hulstijn (2001, pp. 283–286) discussed various pedagogical approaches designed, based on the ideas presented in the preceding section, to enhance automatic word recognition. Empirical research on the impact of training for automaticity on subsequent reading and writing skills has only just begun (e.g., Schoonen et al., 2003; Van Gelderen et al., 2004). In a study involving 281 high school students in Grade 8 in the Netherlands, Van Gelderen et al. and Schoonen et al. investigated the relative contribution of three sources of linguistic cognition on reading and writing in Dutch as an L1 and English as an L2 (after approximately 250 hours of instruction). The dependent variables were reading and writing both in L1 (Dutch) and L2 (English). The predictor variables fell into three categories: (a) knowledge of language, measured with tests of receptive vocabulary, grammar, and spelling, in both L1 and L2; (b) speed of access to knowledge of language, measured with computer-based tests of word recognition, lexical retrieval, sentence verification, and sentence building, in both L1 and L2; and (c) metacognitive knowledge, assessed with a questionnaire pertaining to knowledge of text characteristics and strategies of reading and writing in L1 and L2. In analyses using structural equation modeling, significant correlations were found between speed measures and measures for reading and writing skills.

Stronger correlations were found (Van Gelderen et al., 2004) between predictor variables and reading and writing in the case of L2 than in the case of L1. However, no variance in L1 and L2 reading or writing performance was uniquely accounted for by the speed measures when the knowledge of language and metacognitive knowledge measures were also entered into the regression analysis. One plausible interpretation of these findings is that most of these low-intermediate L2 learners could already access their L2 knowledge sufficiently fast to allow processing of semantic and pragmatic information at the text level.

Finally, we review a study that attempted to increase L2 automaticity through explicit training using the latency coefficient of variability as an operational definition of automaticity. Akamatsu (2001) trained 46 Japanese university students, who had at least 6 years of prior instruction in English, in seven weekly sessions to recognize 150 English words quickly. In each session, students had to draw separator lines as quickly and as accurately as possible between words that had been printed with no interword spaces. Before and after training, students took a computer-controlled word recognition test. This test comprised 50 nonwords and 50 high-frequency and 50 low-frequency words that had been part of the training set. Both accuracy and reaction time on correct trials improved significantly from pretest to posttest. More interesting, individuals' latency coefficient of variability and reaction time were highly and significantly correlated in the processing of low-frequency words, but not of high-frequency words, both before and after training. The author speculated that students in this study had already passed the automatization phase for high-frequency words. Training of these words had only resulted in speed-up, whereas training with the low-frequency words had produced a qualitative change, reflecting automatization. The results of this study (and those of N. S. Segalowitz, Watson, & Segalowitz, 1995, reviewed in the section Automaticity and Communicative Approaches to Teaching) support the idea that training activities of a relatively short duration can bring about a qualitative change in the processing of lexical information, indicating a gain in efficiency that reflects more than simple speed-up.

Clearly, research has only begun on the important practical question of how to enhance automatic processing to promote L2 proficiency. Such research is in its infancy because researchers have only just started to identify the learning issues involved and have only recently developed practical performance measures for operationalizing automaticity. There are, of course, ways of bringing L2 materials to the learner and to create repetition conditions in a manner that could promote automaticity beyond those reviewed earlier. For example, some authors have pointed out that, for any task aimed at helping learners gain fluency in oral production, it is essential to provide learners ample time to plan ahead (Robinson, 2001b). In a review of the literature on factors affecting cognitive complexity of L2 production tasks, Skehan and Foster (2001) claimed that "there is considerable agreement that complexity and fluency are enhanced by pre-task planning" (p. 201). They also pointed out that one of the things speakers do when they are given time to plan their oral production well ahead of execution is to bring into working memory elements from long-term memory perceived to be relevant to the task at hand.

Automaticity and Communicative Approaches to Teaching

A fundamental question remains, however, about how to best promote automatization through repetition in real learning situations outside the laboratory. In answering this question, we have to take into account the extensive and essentially negative experience with so-called pattern drills of the audiolingual method in the 1960s and 1970s. This method was used to help L2 learners improve their production skills in language laboratories through the use of equipment to listen to audio recordings and by making recordings of their own speech (Rivers, 1967).

One of the main reasons why many of these drills failed to bring about the desired effect on spontaneous language use is that they required learners to focus on grammatical forms almost exclusively. Many drills did not force learners to process information at the higher levels of discourse. These methods gave way to what are called communicative language teaching (CLT) methods that stress the importance of meaningful communication as part of the learning process. Unfortunately, most CLT methods do not provide sufficient repetition to promote automatization. This is because the openness of typical CLT communication activities cannot guarantee there will be the necessary opportunities for repeating and rehearing language input; efforts by teachers to supplement communicative activities with special repetition exercises are largely unsuccessful for the same reasons earlier audiolingual drill methods were (Johnson, 1996, especially pp. 171–172).

Thus, teachers are faced with the following dilemma: Typical methods that provide the repetition necessary for automaticity to develop ultimately fail to promote learning because of the highly decontextualized nature of the repeated material; at the same time, typical communicative methods that provide opportunities to fully contextualize learning through meaningful communication fail to provide the repetition necessary for automatization. Can this dilemma be overcome?

Gatbonton has addressed this problem by proposing an analysis of L2 learning that focuses on repetition leading to automaticity within highly contextualized learning (Gatbonton & Segalowitz, 1988; in press) and by providing sets of systematically constructed materials for practical applications based on this approach (Gatbonton, 1994). First, Gatbonton agreed with others that a fundamental step in early L2 learning is the automatization of formulaic utterances—chunks of language that are routinized even in the speech of native speakers (N. C. Ellis, 1997, 2002; Pawley & Syder, 1983; Wray, 2002; see chapters in Schmitt, 2004). Second, she advocated selecting the utterances to be automatized from among those expressions and utterance frames useful for a variety of communicative purposes. Third, she proposed ways to create activities systematically that are genuinely communicative (i.e., for which the communication meets a genuinely felt psychological need for information) and the activity is inherently repetitive (i.e., the activities involve, in a way that feels natural, the need to report information to many people, one by one). Thus, Gatbonton advocated repetition to promote automaticity of basic communicative utterances within a context that requires the learner to coordinate these learning activities with the control of attention, decision making, and other higher level aspects of language processing. She called this process *creative automatization* to reflect the idea that the learner achieves automatization through repetition of acts involving the creation of communicatively valuable utterances.

N. S. Segalowitz et al. (1995) provided some preliminary experimental support for this creative automatization proposal in a study using a single case design. In this study, a Greek-speaking psychology student who spoke English as an L2 participated in a psychology tutorial over a 3-week period in which a single article from a psychology journal was analyzed from several different perspectives. Throughout the 3-week period the student performed lexical decision tasks involving a large number of words, including keywords from the studied article and control words matched for frequency but not appearing in the article. The results showed that a measure analogous to the coefficient of variability of the lexical decision reaction time (coefficient of variability could not be used directly because this was a single-subject study) improved significantly for the words contained in the studied article but not for the control words. This result is consistent with the idea that natural and communicatively meaningful activities inherently repetitive can improve automaticity of lexical access. What is not known is how enduring such improved automaticity of lexical access is.

This would be an important question for future research to address.

Summary and Future Directions

We conclude this review by addressing two questions. First, is it possible to have a general theory of automaticity that will apply in a useful way to phenomena of bilingualism? Second, what future directions ought L2 acquisition research on automaticity take?

A General Theory of Automaticity

We have seen that automaticity figures prominently in most accounts of L2 acquisition and proficiency development, just as it does in most accounts of skill acquisition. Nevertheless, the usefulness of studying automaticity cannot to be taken for granted (see, e.g., Pashler's reservations [1998, pp. 357–382]). A major stumbling block to a general theory of automaticity is that the term has either been used in a very broad sense, without clear operational definition, or else has been defined narrowly but in different ways by different authors (e.g., in terms of ballistic processing; as a shift from serial to parallel processing; as restructuring resulting in a significant change in latency coefficient of variability; as latency patterns reflecting the power law). These are exactly the same problems that confront L2 acquisition researchers. They are attempting to distinguish between explicit and implicit learning processes, to understand when awareness is and is not useful in learning, to find ways of determining when language functioning is proceeding in an autonomous versus a monitored manner, and to understand the conditions under which autonomous processing might be acquired and enhanced (N. C. Ellis, 1994; Hulstijn, 2002; Robinson, 2001a). Thus, both cognitive psychologists interested in automaticity in general and L2 acquisition researchers interested specifically in how languages are learned face the common challenge of having to tease apart a complex of deeply intertwined issues. Is progress being made on this, or are we moving around in circles?

We think there is reason for optimism. One interesting example of potential progress in this area was provided by LaBerge (1997, 2000b) in his triangular circuit theory of attention. This theory identifies particular neural circuits as underlying so-called attention-based and automatic phenomena. These involve neuron clusters in the posterior and anterior cortex for the perception of objects, their attributes, and the organization and execution of action plans; excitatory neurons in the thalamic nuclei that, by virtue of their wide cortical distribution, can selectively enhance cortical activity; and frontal cortex circuitry responsible for control. The linking of these sites forms what LaBerge referred to as "a triangular circuit of attention." Awareness of an object is said to occur when an attentional circuit for that object becomes linked to an attentional circuit related to the self, such as a self-attended representation of a person's spatial or temporal location in relation to the attended object. LaBerge separated automatic processing from attention-based processing in terms of the presence or absence of activity in these triangular circuits.

This theory has generated considerable discussion. For example, Tzelgov (1999), basing his work on LaBerge's theory, proposed that automaticity be used to refer to cases when there is activation of a triangular circuit not involving a self-attended circuit (see also LaBerge, 2000a, for commentary on this). By explicitly proposing neural correlates of attention-based and automatic phenomena, LaBerge raised the bar in the way we talk about automaticity. It is hoped that in time the multiple criteria that have up to now complicated discussion about automaticity will become more precisely defined and distinguishable from one another in terms of underlying neural mechanisms, whether in terms of LaBerge's theory or some other neurobiological approach to attention.

Even prior to the emergence of neural theories of automaticity, there has been a growing consensus that the common element in most automatic phenomena is *ballisticity*, the unstoppable execution of a process once triggered (Bargh, 1992; Favreau & Segalowitz, 1983; Neumann, 1984; Tzelgov, 1999). Although clearly still a work in progress, it appears that it may become possible to provide an account that integrates the neural and behavioral evidence for ballistic processing, thereby allowing more rigorous specification of the relation between automatic and other closely related phenomena.

Future Second Language Research on Automaticity

The developments identified in this chapter should make it possible to address basic questions in L2 acquisition in ways not before possible, using neurophysiological measures (such as event-related potential and neuroimaging) as well as behavioral measures (such as reaction times of responses elicited in a variety of single and dual tasks). Can we monitor the degree of automatic (ballistic) processing in L2 learners at different stages of acquisition? Can we do so for specific aspects of L2 cognition? The current storage-versus-computation debate in linguistics (cf. Nooteboom, Weerman, & Wijnen, 2002) concerning the division of labor between the lexicon (containing chunks of ready-made, stored linguistic information) and the grammar (containing procedures for computing or parsing remaining linguistic information, in language production and reception, respectively) may be highly relevant for the questions of (a) which linguistic phenomena are amenable to automatization and (b) to what extent knowledge of grammar rules can foster or hinder automatization. Perhaps, the success of L2 acquisition that results in increasingly fluent behavior resides, at least partly, in greater availability of ever-larger, preassembled linguistic units and the reduced need to compute information.

As we point out in this chapter, there is reason to believe that it is especially at the intermediate levels of syntactic, morphological, and phonological encoding/decoding, as well as at the lower levels of articulation and perception of acoustic or orthographic signals, that component processes can become automatic to a large extent. Nevertheless, under certain circumstances, the language user can consciously monitor the outcome of these processes and, for instance, decide to repair an error.

A further question that remains to be studied concerns the relationship between the ability to mobilize attentional resources (e.g., noticing) and L2 acquisition. Is noticing a cognitive prerequisite for attaining fluency? If so, for which linguistic phenomena and at which levels of processing might this be the case? How do neurobiological mechanisms of attention and automatic processing determine the cognitive efficiency that underlies high levels of language proficiency? What are the most effective ways to promote proficiency in terms of changing the way attention and automatic processes operate? Can neuroimaging techniques be used to monitor such change (see especially chapters 23 by Hull & Vaid and 24 by Abutalebi, Cappi, & Perani in this volume)?

The acquisition of, and functioning in, an L2 provide paradigmatic examples of the challenges facing cognitive scientists interested in how people acquire the ability to perform complex skills. The availability of the new techniques and the new ways of conceptualizing the issues reviewed here promise to bring important insights to this area.

Acknowledgments

We thank Laura Collin, Elizabeth Gatbonton, Randall Halter, Patsy Lightbown, and Irene O'Brien for helpful comments on earlier versions of this chapter. Support for this chapter came from a grant to Norman Segalowitz from the Natural Sciences and Engineering Research Council of Canada.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*, 288–318.

Ackerman, P. L. (1989). Individual differences and skill acquisition. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 165–217). New York: Freeman.

Akamatsu, N. (2001, February). *Effects of training in word recognition on automatization of word-recognition processing of EFL learners.* Paper presented at the 2001 annual conference of the American Association for Applied Linguistics, St. Louis, MO.

Anderson, J. R. (1983). *The architecture of cognition.* Mahwah, NJ: Erlbaum.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought.* Mahwah, NJ: Erlbaum.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology, 105*, 181–199.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition.* New York: Cambridge University Press.

Bourne, L. E., Jr., Healy, A. F., Parker, J. T., & Rickard, T. C. (1999). The strategic basis of performance in binary classification tasks: Strategy choices and strategy transitions. *Journal of Memory and Language, 41*, 223–252.

De Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics, 13*, 1–24.

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

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

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge, U.K.: Cambridge University Press.

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

Ellis, N. C. (1997). Vocabulary acquisition: word structure, collocation, word-class, and meaning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 122–139). Cambridge, U.K.: Cambridge University Press.

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

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition, 19*, 145–171.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly, 27*, 91–113.

Ellis, R. (2000, September). *The representation and measurement of L2 explicit knowledge*. Paper presented at the conference on Language in the Mind, National University of Singapore.

Eviatar, Z. (1998). Attention as a psychological entity and its effects on language and communication. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of neurolinguistics* (pp. 275–287). New York: Academic Press.

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory & Cognition, 11*, 565–574.

Fischler, I. (1998). Attention and language. In R. Parasuraman (Ed.), *The attentive brain* (pp. 381–399). Cambridge, MA: MIT Press.

Gatbonton, E. (1994). *Bridge to fluency: Speaking*. Scarborough, Ontario: Prentice Hall Canada.

Gatbonton, E., & Segalowitz, N. S. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly, 22*, 473–492.

Gatbonton, E., & Segalowitz, N. S. (in press). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review*.

Healy, A. F., Barshi, I., Crutcher, R. J., et al. (1998). Toward the improvement of training in foreign languages. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Foreign language learning* (pp. 3–53). Mahwah, NJ: Erlbaum.

Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge, U.K.: Cambridge University Press.

Hulstijn, J. H. (2002). Towards a unified account of the representation, processing, and acquisition of second-language knowledge. *Second Language Research, 18*, 193–223.

Hulstijn, J. H., & Hulstijn, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning, 34*, 23–43.

Johnson, K. (1996). *Language teaching and skill learning*. Oxford, U.K.: Blackwell.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford, U.K.: Pergamon Press.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

LaBerge, D. (1997). Attention, awareness, and the triangular circuit. *Consciousness and Cognition, 6*, 149–181.

LaBerge, D. (2000a). Clarifying the triangular circuit theory of attention and its relations to awareness: Replies to seven commentaries. *Psyche, 6*. Retrieved February 20, 2002, from http://psyche.cs.monash.edu.au/v6/psyche-6-06-laberge.html

LaBerge, D. (2000b). Networks of attention. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 711–724). Cambridge, MA: MIT Press.

LaBerge, D., & Samuels, J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323.

Langacker, R. W. (1987). *Foundations of cognitive grammar, Vol. 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Larsen Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. London: Longman.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83–122). Oxford, U.K.: Oxford University Press.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on

second language learning. *Studies in Second Language Acquisition, 12*, 429–448.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492–527.
Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 883–914.
McLaughlin, B., & Heredia, R. (1996). Information-processing approaches to research on second language acquisition and use. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 213–228). New York: Academic Press.
McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information processing perspective. *Language Learning, 33*, 135–158.
McLeod, B., & McLaughlin, B. (1986). Restructuring or automaticity? Reading in a second language. *Language Learning, 36*, 109–123.
Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*, 25–40.
Nakayama, K., & Joseph, J. S. (1998). Attention, pattern recognition, and pop-out in visual search. In R. Parasuraman (Ed.), *The attentive brain* (pp. 279–298). Cambridge, MA: MIT Press.
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226–254.
Neumann, O. (1984). Automatic processing: A review of recent findings and a plea for an old theory. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 255–293). New York: Springer-Verlag.
Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.
Nooteboom, C., Weerman, F., & Wijnen, F. (Eds.). (2002). *Storage and computation in the language faculty*. Dordrecht, The Netherlands: Kluwer.
Paap, K. R., & Ogden, W. C. (1981). Letter encoding is an obligatory but capacity-demanding operation. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 518–527.
Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In Ellis, N. (Ed.), *Implicit and explicit learning of languages* (pp. 393–419). London: Academic Press.
Pashler, H. E. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). New York: Longman.
Pienemann, M. (1998). Developmental dynamics in L1 and L2 acquisition: Processability theory and generative entrenchment. *Bilingualism: Language and Cognition, 1*, 1–20.
Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review, 78*, 391–408.
Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance*. London: Sage.
Rivers, W. M. (1967). *Teaching foreign-language skills*. Chicago: University of Chicago Press.
Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*, 283–331.
Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition, 19*, 223–247.
Robinson, P. (Ed.). (2001a). *Cognition and second language instruction*. Cambridge, U.K.: Cambridge University Press.
Robinson, P. (2001b). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge, U.K.: Cambridge University Press.
Robinson, P., & Ha, M. (1993). Instance theory and second language rule learning under explicit conditions. *Studies in Second Language Acquisition, 15*, 413–438.
Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207–231.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge, U.K.: Cambridge University Press.
Schmitt, N. (Ed.). (2004). *Formulaic sequences*. Amsterdam: Benjamins.
Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: 1. Detection, search and attention. *Psychological Review, 84*, 1–66.
Schoonen, R., Van Gelderen, A., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., et al.

(2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language Learning, 53*, 165–202.

Segalowitz, N. S. (1997). Individual differences in second language acquisition. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 85–112). Mahwah, NJ: Erlbaum.

Segalowitz, N. S. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 200–219). Ann Arbor: University of Michigan Press.

Segalowitz, N. S. (2003). Automaticity and second languages. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Oxford, U.K.: Blackwell.

Segalowitz, N. S., & Frenkiel-Fishman, S. (in press). Attention control and ability level in a complex cognitive skill: Attention shifting and second language proficiency. *Memory & Cognition*.

Segalowitz, N. S., Poulsen, C., & Segalowitz, S. J. (1999). RT coefficient of variation is differentially sensitive to executive control involvement in an attention switching task. *Brain and Cognition, 38*, 255–258.

Segalowitz, N. S., Watson, V., & Segalowitz, S. J. (1995). Vocabulary skill: Single case assessment of automaticity of word recognition in a timed lexical decision task. *Second Language Research, 11*, 121–136.

Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics, 14*, 369–385.

Segalowitz, S. J., Segalowitz, N. S., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics, 19*, 53–67.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183–205). Cambridge, U.K.: Cambridge University Press.

Squire, L. R., & Knowlton, B. J. (2000). The medial temporal lobe, the hippocampus, and the memory systems of the brain. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 765–779). Cambridge, MA: MIT Press.

Stroop, J. R. (1935). Studies of interference in serial and verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Talmy, L. (1996). The windowing of attention. In M. Shibatani & S. A. Thompson (Eds.), *Grammatical constructions* (pp. 235–287). Oxford, U.K.: Oxford University Press.

Talmy, L. (2000). *Toward a cognitive semantics* (Vols. 1 & 2). Cambridge, MA: MIT Press.

Tzelgov, J. (1999). Automaticity and processing without awareness. *Psyche, 5*. Retrieved February 20, 2002, from http://psyche.cs.monash.edu.au/v5/psyche-5-05-tzelgov.html

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 760–771.

Tzelgov, J., Henik, A., Sneg, R., & Baruch, O. (1996). Unintentional word reading via the phonological route: The Stroop effect with cross-script homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 336–349.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*, 105–122.

Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., et al. (2004). Linguistic knowledge, processing speed and metacognitive knowledge in first- and second-language reading comprehension: A componential analysis. *Journal of Educational Psychology, 96*, 19–30.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound and reading. *Memory & Cognition, 15*, 181–198.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12*, 287–301.

Waters, G., Seidenberg, M. S., & Bruck, M. (1984). Children and adults use of spelling-sound information in three reading tasks. *Memory & Cognition, 12*, 293–305.

White, L. (1989). *Universal grammar and second language acquisition* (Vol. 9). Amsterdam: Benjamins.

White, L. (1996). Universal grammar and second language acquisition: Current trends and new directions. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 85–120). San Diego, CA: Academic Press.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105*, 558–584.

Wingfield, A., Goodglass, H., & Lindfield, K. C. (1997). Separating speed from automaticity in a patient with focal brain atrophy. *Psychological Science, 8*, 247–249.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, U.K.: Cambridge University Press.