PSYCHOLOGICAL MECHANISMS
UNDERLYING SECOND LANGUAGE FLUENCY

RICHARD SCHMIDT
University of Hawai'i

Fluency in a second language is considered important by both learners and teachers, but is not well understood. This paper describes what is known about second language fluency and describes a number of psychological learning mechanisms that might explain how fluency develops. These include the mechanisms underlying the contrast between automatic and controlled processing, the learning mechanisms postulated within Anderson's ACT* theory of cognition, Bialystok's conception of the control dimension of language development, the notion of restructuring, recent proposals for the redefinition of automaticity as retrieval from memory, both instance and strength versions, and chunking theories. The paper concludes with some suggestions for research into the development of second language fluency itself that can fill gaps in existing knowledge and reduce our dependence on other fields for explanatory principles, while contributing simultaneously to discussion of the mechanisms responsible for skill development in general.

The Phenomenon of Fluency

"Fluent" and "fluency" are frequently used as non-technical terms and have a number of meanings that should be sorted out. With respect to native language fluency, Fillmore (1979) identified four different things we might have in mind when identifying someone as a particularly fluent speaker. First, we might be thinking of a speaker who easily fills time with talk, a fast talker (Kuiper & Tillis, 1986) such as a disk jockey, a sports announcer, or a conversational partner who hardly lets us get a word in edgewise. Second, we might have the quality of speech more in mind than quantity and might mean by a fluent speaker one whose speech is coherent, complex and dense (Fillmore gives Noam Chomsky as an example of this kind of fluent speaker). We might also consider someone to be especially fluent if they always seem to know the appropriate thing to say in a wide variety of contexts, identifying fluency with pragmatic and affective skills. Finally, Fillmore points out, we may focus on

speakers with exceptional control over the aesthetic functions of language, including creativity and imagination, punning, joking, the creation of metaphors, and so on.

When considering fluency with respect to nonnative learners of a language, we might also have any of these things in mind. We often use fluency as a rough synonym for global ability (Lennon, 1990), as when someone says "My friend speaks four languages fluently." In such cases we usually mean only that our friend speaks those languages well, without meaning to differentiate among the ways that "well" could be specified. In this paper, I will not be concerned with such a global conception of fluency, which appears to differ little from the concept of "proficiency," but will restrict the term in two ways:

(1) A contrast is often made in the second language literature between fluency and accuracy, knowledge, or developmental stage (Brumfit, 1984). It seems natural to say about some learners that "X really knows quite a lot of English, but doesn't speak it fluently" or that "Y speaks fluently but not very grammatically." An extreme case would be the speaker of a pidginized interlanguage, largely unanalyzed and agrammatic with respect to the target language, who speaks that variety in a fluid rather than a halting manner (Schumann, 1990). Such a speaker is not fluent under a global proficiency definition, but can be called fluent if we identify fluency with the processing of language in real time, rather than with language as the object of knowledge. It is this conception of fluency as a primarily temporal phenomenon that I will take as the basic definition for foreign and second language learning.

(2) Although there is a substantial body of literature on factors related to fluency in receptive processes (Segalowitz, 1991), I will also restrict the discussion of fluency to the productive processes involved in the planning and delivery of speech. It is certainly possible to speak of fluent listeners, readers and writers as well as speakers. Speed and ease of processing are probably common components of fluency across modalities, but other implied contrasts (if any) between "fluent" and simply "good" or "proficient" listeners, readers or writers are less clear.[1]

My own preferred label for fluency in speech production is **automatic procedural skill** (Carlson, Sullivan, & Schneider, 1989). Fluent speech is automatic, not requiring much attention or effort, and is characterized by the fact that "the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently" (Lennon, 1990, p.391). Nonfluent speech is effortful and requires a great deal of attention, so that nonfluent speakers exhibit many hesitations and other manifestations of groping for words and attempting to combine them into utterances. Fluency depends on procedural knowledge (Færch & Kasper, 1984), knowing how to do something, rather than declarative knowledge, or knowledge about something. Finally, if a distinction is to be made between procedural knowledge and procedural skill (many writers have written about the contrast with declarative knowledge, but none have been particularly concerned with this particular contrast), I prefer to identify fluency with skill rather than knowledge, again emphasizing the performance aspect of actually doing something in real time rather than the knowledge of how something is to be done.

Comments by learners provide support for a conception of second language fluency as a performance phenomenon with particular emphasis on its temporal aspects (Lennon, 1989), as do proposals concerning the empirical correlates of fluency. Möhle (1984) has suggested speech rate, the length and positioning of silent pauses, the length of fluent speech runs between pauses, the frequency and distribution of filled pauses, and the frequency of repetitions and self-corrections as possible measures of fluency. While I take a speaker-based perspective on the concept of fluency, assuming that fluency rests upon a definable bundle of production processes, there is also evidence that these factors influence hearer-based impressions of fluency. In two recent studies, Lennon (1990) and Riggenbach (in press) had native speaking judges rate non-native speech samples for fluency and then investigated quantifiable

---

[1] Sincoff and Sternberg (1987) have argued that whereas reading and listening draw primarily on "verbal comprehension," speaking and writing require "verbal fluency." Comprehension and fluency are seen as different though related abilities. I am grateful to an anonymous *SSLA* reviewer for calling this reference to my attention.

performance features in those samples that might function as objective indicators of oral fluency. Lennon analyzed the speech of four female West German university students, advanced learners of English, all of whom were judged to have improved in fluency between the beginning and the end of six months' residence in Britain. Twelve variables were assessed, and statistically significant improvement across subjects was found for three of them: faster speech rate, fewer filled pauses per t-unit, and fewer t-units followed by pause. Lennon also noted that there were individual differences among the four subjects, indicating that the perception of fluency on the part of listeners may not always be based on the same speech characteristics. Self-corrections proved to be a poor fluency indicator across subjects, but other variables that did not show statistically significant changes in this limited study deserve further investigation with larger subject samples as possible indicators of fluency. Using a cross-sectional design, Riggenbach investigated the speech of six Chinese learners of English, three of whom were rated "very fluent" and three of whom were rated "very non-fluent" by 12 judges. Riggenbach compared these learners on 19 variables, including hesitation and repair phenomena, speech rate, and a number of interactive discourse measures. Few significant differences were found, but the fluent and non-fluent learners were significantly different with respect to speech rate and the number of unfilled pauses. Like Lennon, Riggenbach found that repair frequency was not a a factor influencing fluency judgments, and also stresses the fact that individual learners had different fluency profiles. One learner in particular was more like the fluent speakers in terms of rate and amount of speech but was rated non-fluent because of the agrammaticality of her speech, suggesting that hearer-based impressions of fluency are holistic, influenced by considerations of accuracy as well as by the temporal, performance aspects on which I will focus in this paper.

Psychological Learning Mechanisms

Identifying second language fluency with automatic procedural skill provides a label that accords reasonably well with nontechnical conceptions of the phenomenon and with some empirically identifiable components of speech

judged to be fluent, but says nothing about how fluency develops. However, a number of learning mechanisms for the development of cognitive skills in general have been proposed in the psychological literature, several of which have been cited in the second language literature as plausible explanations for the development of fluency. These include the mechanisms underlying the notion of automaticity as developed by Shiffrin and Schneider (1977) and applied to L2 learning by Levelt (1977) and by McLaughlin, Rossman, and McLeod (1983); the mechanisms of proceduralization, composition, generalization, discrimination and strengthening proposed in Anderson's ACT* theory of cognition (Anderson, 1982, 1983); Bialystok's dimension of "control" (Bialystok, 1990a, 1990b); the notion of restructuring as developed by Cheng (1985) in psychology and applied to L2 learning by McLaughlin (1990); recent proposals for the redefinition of automaticity as retrieval from memory, in both instance theory (Logan, 1988a, 1991; Logan & Stadler, 1991) and associative strength theories (MacKay, 1982; Schneider, 1985); and chunking theories (Newell, 1990; Servan-Schreiber & Anderson, 1990). I will discuss each of these proposals in turn, first presenting and explaining the theoretical constructs and the proposed learning mechanisms, next considering how the theory has been or could be applied in the field of SLA, and finally providing an evaluation of the proposed mechanisms in terms of both their current status within cognitive psychology and their relevance to our understanding of second language fluency and how it develops.

Automatic and controlled processing

The distinction between controlled and automatic processing, originally conceived as a dichotomy between two qualitatively different forms of processing, has been a major topic in psychology. The theoretical contrast parallels the easily confirmable subjective experience that some skills and mental activities seem to require our full attention, while others seem to require little or no attention or effort. The most important properties of automatic processing are generally considered to be that it is (1) fast and efficient, (2) effortless, (3) not limited by short term memory capacity, (4) not under voluntary control, (5) difficult to modify or inhibit, and (6) unavailable

to introspection. Automatic processes typically occur in well-practiced tasks, and are held to be responsible for skilled performance and most of the details of cognitive processing. In contrast, controlled processing is (1) slow and inefficient, (2) effortful, (3) limited by the capacity of short term memory, (4) largely under subject control, (5) flexible, and (6) at least partly accessible to introspection. Controlled processing serves such functions as maintaining goals in working memory and applying general procedures to new circumstances, and it typically occurs in novel and inconsistent processing tasks. The development of skilled behavior involves a shift with practice from controlled to automatic processing. Novices of all kinds, including beginning second language learners, must pay careful attention to every step in the procedure, while experts do not. For further discussion of the basic contrasts between controlled and automatic processing, readers are referred to Hasher & Zacks (1979), LaBerge (1981), Logan (1991), Logan and Stadler (1991), Posner and Snyder (1975), Schneider (1985), Schneider and Detweiler (1988), Schneider, Dumais, and Shiffrin (1984), and Shiffrin and Schneider (1977).

The most frequently cited theory of the development of automatic processing is that of Shiffrin and Schneider (1977), based on experiments involving the detection of target letter stimuli presented in a field of distractors. Comparing the development of this skill when target mappings were varied (presumably calling on controlled processing, because subjects had to activate a different memory set relating particular letters to "yes" and "no" responses on every trial) with conditions in which mappings were held constant, Shiffrin and Schneider found that only the latter condition led to automatic processing with practice. Performance under the consistent mapping condition exhibited two of the defining characteristics of automaticity: once a task is automatized, attentional resources are freed to perform other tasks concurrently (e.g.carrying on a conversation while performing the visual search task), and automatic processes also occur even when subjects consciously try to prevent them. In the Shiffrin and Schneider model, controlled processing utilizes temporary sequences of nodes activated under attentional control. Automatic processing involves sequences of nodes in memory that nearly always become active in response to a particular input configuration. Both the associative links between stimulus and response and

the links among steps of the response process can be automatized, so that attention is required for neither the initiation of a fully automatic response nor its completion.

The mechanism considered responsible for the development of automatic responses is **strengthening** of the connections among nodes, as a result of repeated exposure and rehearsal (association learning). This model does not assume that the structure of the response is modified in any way, only that it runs off more rapidly. Since strengthening is a purely a process-improvement mechanism, it is not adequate (nor is it intended) as an explanation for cases in which development is manifested by new or modified responses to a particular input configuration.

Shiffrin and Schneider's view of the contrast between controlled and automatic processing has a number of possible implications for second language learning, several of which were identified by McLaughlin, Rossman and McLeod (1983):

1. Complex skills such as those involved in language learning are learned and become automatic only after the earlier use of controlled processes.

2. Since speaking is a complex cognitive task with hierarchical task structure (involving discoursal, pragmatic, syntactic and lexical choices), each component requires more or less attention depending on how well-learned it is. The development of new skills is possible only when other task demands are minimized.

McLaughlin, Rossman and McLeod were concerned with introducing an information-processing perspective to the second language field in general terms, not with the particular problem of fluency. They identified the task of understanding second language learning in information processing terms with the need to formulate a component skills analysis in which the processing skills that make up the task of learning a second language are identified, with particular attention to those skills contributing to individual variation and overall success. However, to the extent that McLaughlin, Rossman and McLeod were concerned with fluency, they linked fluency with automatic processing, reporting that one difference between fluent and nonfluent bilinguals is the degree of automatization of lexical processing and citing studies showing that fluent second language learners exhibit some of the

hallmarks of automatic processing (speed and the reallocation of attention) more than novice learners.

The automatization of second language processes is a useful concept for understanding second language fluency. The point has been made frequently that speech is possible at a normal rate only when most of the procedures involved have been automatized (de Bot, in press; Levelt , 1979, 1989; Rehbein, 1987; Sajavaara , 1987). Practice seems to be the necessary condition for fluency in a second language, and this is given a theoretical justification in models of automatization. Such characteristics of L2 fluency as speech rate and the length of fluent runs between pauses may reflect automaticity fairly directly, while other aspects of fluency may reflect the fact that virtually all complex tasks require a mixture of automatic and controlled processes, usually organized in a systematic network or hierarchy (Levelt, 1977, 1989; Schneider, Dumais, & Shiffrin, 1984).

However, a number of objections have been raised against the concept of automaticity as formulated in the Shiffrin and Schneider model. These include theoretical and empirical questions concerning the single-capacity view of attention which underlies the theory (Kahneman & Treisman, 1984; Wickens, 1984), disagreement with the claim that automatic processing is free of attentional limitations (Cheng, 1985; Cohen, Dunbar, & McClelland, 1990), challenges to the notion that automatization reflects the withdrawal of attention (Hirst, Spelke, Reaves, Caharack, & Neisser, 1980; Logan, 1988a), and competing views of the contrast between automatic and controlled processing as a dichotomy or a continuum (Strayer & Kramer, 1990). Schneider and Detweiler (1988) have proposed a revised model in which automatization is viewed as a gradual, continuous transition through five identifiable phases: controlled comparison from buffered memory (fully controlled processing), context-maintained controlled comparison, goal-state-maintained controlled comparison, controlled assist of automatic processing, and fully automatic processing. The five stage model suggests that even for completely fluent native speakers, some processes may remain at the stages of controlled comparison or controlled assist of automatic processing. Examples might be subject-verb agreement, especially when subject and verb are separated by

intervening linguistic material, the *who/whom* distinction, and many other normative aspects of language. Unfortunately, Schneider and Detweiler provide precise definitions of the phases of automatization only with reference to specific laboratory tasks, so extensions to natural language performance can only be metaphorical.

Shiffrin and Schneider were more concerned with describing the characteristics of controlled and automatic processing than with how it develops, and their discussion of learning mechanisms is under-developed. Most of the theories to be discussed in the remainder of this section can be seen as attempts to flesh out the basic contrast between automatic and controlled processing with more detailed specification of the learning mechanisms responsible for automatization.

ACT*

John Anderson has proposed a multi-stage, multi-mechanism theory of the acquisition of cognitive skills (Anderson, 1980, 1981, 1982, 1983, 1989), referred to as ACT* (pronounced "act-star"). According to Anderson, the first stage of skill development relies on declarative (propositional) knowledge. Facts about how a skill is performed are maintained in working memory and are used interpretively by general-purpose productions (IF-THEN rule-statements), which are flexible but which carry heavy costs in terms of time and working memory space. The declarative stage of skill in Anderson's theory is equivalent in most respects to controlled processing in the Shiffrin and Schneider theory. A straightforward example is the case of classroom learning in which a student is told a rule of the second language, for example, a rule concerning tense inflection, and then carries out a drill requiring that a number of verbs be inflected using the rule. This would be an equally good example of Krashen's notion of the conscious application of a rule of language (Krashen, 1981), although Anderson's model is incompatible with Krashen's position that there is no interface between conscious learning and subconscious acquisition (Krashen, 1985). Since Anderson maintains that knowledge in a new domain *always* starts out in declarative form and is used interpretively, this initial stage must also encompass examples in which learners produce linguistic forms by self-discovered rules of thumb or by analogy with known

forms. If working memory and consciousness are assumed to be roughly equivalent (Baars, 1988; Kihlstrom, 1984; Schmidt , 1990), then language learners and users ought to be able to report many of the details of this declarative stage.

The second stage of skill acquisition is a procedural stage, in which knowledge is directly embedded in procedures for performing the skill. The procedural stage of skill acquisition is roughly equivalent to automatic processing in the Shiffrin and Schneider model, but Anderson offers more detail concerning how it may develop. Anderson outlines two general processes in the development of procedural knowledge: knowledge compilation, by which the skill moves from the declarative to the procedural stage, and tuning, through which productions become more selective in their range of applications. Five learning mechanisms are proposed to explain these changes.

The mechanisms of knowledge compilation are **composition** and **proceduralization**. Composition refers to the collapsing of sequences of productions into macroproductions, prepackaged sequences, or chunks (Miller, 1956, 1958). One does not normally remember a social security number as 018305267, for example, but as 018-30-5267. Telephone and credit card account numbers are other examples in which long numerical sequences are pre-chunked by their issuers for easier processing. Proceduralization, the other mechanism of knowledge compilation, refers to the embedding of factual knowledge into productions so that the products of frequently executed productions can be retrieved directly from memory and declarative knowledge does not need to be activated in working memory for their execution. It is not uncommon for declarative knowledge either to be lost (this assumes that memory traces decay, which is controversial) or simply to be no longer retrievable after proceduralization is complete. We can often drive or walk familiar routes more accurately than we can give directions to others on how to do so; skilled typists may have better kinesthetic than verbal control over spelling; and foreign language learners may be able to retrieve the products of frequently used productions even when they cannot remember anything about the declarative knowledge that was presumably used to guide such

productions initially (Sharwood Smith, 1981).

The mechanisms that contribute to the fine tuning of procedural knowledge in Anderson's theory are **generalization, discrimination**, and **strengthening**. Both generalization, by which rules become broader in scope, and discrimination, which narrows the scope of rule application (Anderson, 1982; Klahr, 1984; Reese, 1989), have obvious examples in language learning. The mechanism of strengthening, by which better rules are strengthened and poorer rules are weakened, leading eventually to rule replacement, appears equally relevant to second language learning. It should be noted that the mechanism of strengthening proposed in Anderson's ACT* model is somewhat different from the mechanism of strengthening discussed above with reference to the Shiffrin and Schneider theory of controlled and automatic processing, since strengthening in Anderson's model affects only the likelihood of a procedure or rule being selected, not the strength of associations among the elements of a response or procedure.

Anderson has offered a detailed and powerful theory of cognition in general which has considerable appeal as a model of second language skill. Hulstijn (1990) points out that the more sub-procedures get subsumed into overall procedures (by the composition mechanism), "the more language use can be said to take place fluently and automatically, requiring less attention" (Hulstijn, 1990, p.32). However, the acquisition of language skills is not limited to the speeding up of the same procedures originally formed from declarative knowledge, but also the establishment of new procedures which reorganize previously acquired rules and procedures. Anderson's model can account for this (whereas the Shiffrin and Schneider model of automatization cannot) because it is a self-modifying production system that includes mechanisms that change productions (generalization and discrimination), rather than only mechanisms responsible for the running off of productions (Reese, 1989). The Anderson model thus relates to broader concerns than the development of fluency in a narrow sense, but even if we maintain our limited concept of fluency in terms of the temporal aspects of skill, Anderson's mechanisms all seem to contribute to fluency, though in rather different ways. Composition results in processing speedup through the unitary application of procedures. Schneider and Detweiler (1988) point out that under certain postulated fixed

times for the encoding, decoding and transmission of information, a copy typist using sequential transmission (e.g. transmitting the letter pattern THE to the fingers as T, H, and E) would average 50 wpm. By transmitting the same information in parallel (transmitting T,H, and E simultaneously across the visual-motor loop), the same typist would average 60 wpm, and using chunk transmission and decoding (transmitting the chunk THE from a single visual module to a single motor module and then decoding the motor chunk into its components) the same typist would average 100 wpm. Proceduralization also contributes to speed of processing, because working memory demands are reduced and the system can simultaneously perform other tasks that make demands on working memory. It is harder to make predictions concerning the temporal effects of the mechanisms of generalization and discrimination, and these are both included in the theory primarily because of their effects on accuracy and development (changes in the representation of linguistic knowledge rather than access to that knowledge), but both may also produce speedup through what Anderson calls "algorithmic improvement" (Anderson, 1982, p. 398). Strengthening also has a major influence on timing because it reduces the time it takes for a production rule to be selected.

It may be possible to relate specific mechanisms from Anderson's theory to specific aspects of second language fluency (such as Lennon's empirical correlates of perceived fluency, discussed above), although this has not yet been attempted. Anderson has illustrated the operation of his theory with respect to L1 acquisition phenomena (Anderson, 1980), but the examples presented have more to do with the types of rules that function in acquisition than with fluency factors.

There are disadvantages as well as strengths in multimechanism theories. Bialystok has objected to Anderson's model on the grounds that it conflates the representation of knowledge with access to that knowledge (Bialystok & Bouchard Ryan, 1985), two dimensions that should be kept separate for both theoretical and methodological reasons. In addition, as Carlson and Schneider (1990) have noted, ACT* mechanisms often compete with each other. Composition speeds processing, but at the same time the complexity of the composed productions slows production, negating the effect of composition

(Carlson & Schneider, 1990). The postulated existence of both cooperating and competing mechanisms makes the theory difficult to falsify.

Executive control

Whereas McLaughlin, Rossman and McLeod have drawn the attention of second language theorists to the Shiffrin and Schneider contrast between automatic and controlled processing and Hulstijn has drawn upon Anderson's model to illuminate the development of procedural skill in a second language, Bialystok has carried the discussion of fluency in other directions. For some time, Bialystok has advocated a model of second language development that rests upon a two dimensional framework (Bialystok, 1982, 1985, 1990a, 1990b; Bialystok & Bouchard Ryan, 1985; Bialystok & Sharwood Smith, 1985). The first of these dimensions, that of **analysis**, has to do with the ways in which linguistic knowledge is represented cognitively and the ways in which representations change in the course of linguistic development. This component involves the progressive development of a knowledge system which is initially implicit in the mind of the learner but which gradually becomes both more explicit and more organized formally, a crucial development for advanced language skills such as literacy. (A similar view of first language development has been advanced by Karmiloff-Smith, 1986.) In my view (and apparently also in Bialystok's), the analysis dimension of the model is not relevant to an understanding of second language fluency, so it will not be discussed further here.

The second dimension of Bialystok's model, **control**, concerns the use of linguistic knowledge and is assigned the task of accounting for access to linguistic knowledge (whether analyzed or unanalyzed) and describing the cognitive demands that language tasks place upon learners. The characterization of this dimension has evolved in Bialystok's thinking. In the 1982 version, it was labeled the "automatic factor," and the achievement of automatic access to the information represented by the analyzed dimension was seen as the essence of development on this dimension (Bialystok, 1982: 183). In more recent versions, the concept of control has been broadened beyond that of automaticity, and refers to the ability to select, coordinate and integrate relevant information in real time, the key to which is the "ability to

intentionally focus attention on relevant parts of a problem to arrive at a solution" (Bialystok & Mitterer, 1987, p. 148). For Bialystok, fluency is an outcome of development along the dimension of control:

> Since control of processing is constrained by real time, effective control processes confer the impression of fluency or automaticity upon performance. ...*fluency is considered to be an emergent property of high levels of control*. Skilled selective attention, that is, creates a performance that appears automatic and effortless. (Bialystok, 1990a, p.125, emphasis mine)

For Bialystok, the direction of development is from low to high levels of control. Hulstijn (1990) has found this objectionable, commenting that development must proceed from high control to low control, but this disagreement is largely a question of terminology and perspective. Hulstijn's argument is the traditional one that the developmental path towards fluency for each particular procedure is from controlled processing (requiring attentional supervision) to automatic processing (either requiring no attentional control or less attentional control). However, if the focus is on control processes themselves, including not only the selective allocation of attention but also more specific control processes such as rehearsal, search, planning (Crookes, 1989), monitoring (Morrison & Low, 1983), and decision making of all kinds (Shiffrin & Schneider, 1977), then it is reasonable to speak of development in the direction of higher (i.e. increasingly skillful) levels of control over a skill or its components. "Higher" levels of control may also refer to the level of organization to which attention is directed.

In addition, at least a partial shift in focus away from automaticity towards efficient self-regulation as an essential characteristic of fluency is justified by the fact that skilled performance requires a balance between the speed of automatic processing and the goal-directedness of controlled processing. Stillings, Feinstein, Garfield, Rissland, Rosenbaum, Weisler and Baker-Ward (1987) have pointed out that "a system that acted only by allowing the currently most active automatic procedure to carry through to completion without any influence by goals would be incoherently impulsive" (Stillings et

al., 1987, p. 59). A great deal of empirical data suggest that automatic processes (indeed, all cognitive processes) are subject to attentional control to some degree (Cheng, 1985; Cohen, Dunbar & McClelland, 1990). Fluency can validly be described as the control of mostly automatic processes by selective attention in the service of intentional goals (Bialystok's point), although determination of the mechanisms by which automatic and controlled, goal-directed behaviors are coordinated remains a difficult problem (Phillips & Hughes, 1988).

This model also has a number of weaknesses in its ability to provide an explanation for second language fluency. Following Jackendoff (1987), Bialystok (1990b) has argued that automaticity is epiphenomenal. However, while automaticity may not be the only explanation for L2 fluency, it or some substitute must be part of any such explanation, and Bialystok has acknowledged that automaticity must be incorporated "somehow and somewhere" within the dimension of control (Ellen Bialystok, personal communication, July 17, 1991). The model might be improved by reference to psychological theories that have attempted to describe the details of executive control structures while recognizing that intentional control by itself is too slow and unwieldy to provide the precision and timing needed to perform skilled acts and that one reason automatic processes are important is that they can be harnessed to provide information relevant to a person's goals (Cohen, Dunbar, & McClelland, 1990; Logan, 1988a; Norman & Shallice, 1986; Phillips & Hughes, 1988; Reason, 1984; Schneider & Detweiler, 1988; Shallice, 1978). Levelt's model of speech production (Levelt 1989) is equally relevant here. Levelt argues that speaking is usually an intentional activity, serving purposes that speakers want to realize, and is thus under executive control, but in order for control to be allocated where it is needed (primarily at the conceptual level), virtually all low level components (including the selection of grammatical structures, retrieval of lexical items and the formulation of articulatory plans) must be largely automatic if fluent speech is to result (Levelt,1989, pp. 20-22).

It may be possible to elaborate a theory in which it is made explicit how skilled selective attention develops and produces the impression of both automaticity and fluency, but Bialystok has not yet provided this level of detail. From the perspective of the concerns raised in this paper, the major weakness of the model is that it contains no learning mechanisms. Selective

attention, identified as the main processing mechanism for the dimension of control, is said to develop as the result of age, experience, practice, and bilingualism, but there is no attempt at an explanation of how this comes about.

Restructuring

Cheng (1985) introduced the notion of **restructuring** (a process-switching mechanism) as an alternative to that of automaticity (process-improvement) for the explanation of skilled performance. Cheng argued that the results of the Shiffrin and Schneider experiments did not require an explanation in terms of a dichotomy between controlled and automatic processes, because other explanations were readily available:

> In particular, [improved performance] can be due to a restructuring of the task components so that they are coordinated, integrated, or reorganized into new perceptual, cognitive, or motor units, thereby allowing the procedure involving the old components to be replaced by a more efficient procedure involving the new components. (Cheng, 1985, p.414)

Cheng used a simple analogy to illustrate her point. One way to find the sum of ten 2's is to perform nine addition operations. But anyone who has learned the multiplication table can solve the same problem by looking up the entry for 2 X 10 in memory. Remembering the answer is faster and more efficient than performing nine addition operations, but the gain in efficiency is not accomplished by a speed-up of the addition process, or even through an automatic process of multiplication (no operation is performed other than retrieval from memory). Cheng claimed that consistent and variable mapping tasks, such as those carried out by Shiffrin and Schneider, cannot distinguish between shifts in modes of processing and the restructuring of task procedures, making automaticity a suspect concept. Schneider and Shiffrin (1985) responded to Cheng that restructuring was indeed a factor in improvement in such experiments but was insufficient to explain a number of key findings, for which the concept of automatization was still required.

For second language learning, McLaughlin (1990) has taken a position

similar to Schneider and Shiffrin's, that both automaticity and restructuring are required concepts, citing Cheng's definition of restructuring as the coordination, integration and reorganization of task components resulting in more efficient procedures. McLaughlin cites discontinuities in linguistic development as the primary evidence for the interplay of automaticity and restructuring, discussing numerous cases of U-shaped behavior in which correct forms that have become automatic are replaced by over-regularized forms based on qualitative representational changes (restructuring) before the correct forms reappear.

McLaughlin's and Cheng's conceptions of restructuring are alike in being concerned with shifts in the strategies used to carry out a cognitive skill and in defining restructuring in terms of changes in the organization of task components. However, there is an equally important difference between their interpretations of the concept that becomes apparent if one considers the specific examples of restructuring that are presented and discussed. Cheng emphasized the choice of more efficient strategies as an alternative to automatization, giving the shift from a series of addition processes to memory look-up as a basic example. McLaughlin's examples of restructuring in language learning all involve a shift away from memory and exemplar-based strategies, such as an early reliance on memorized formulas, to strategies based on more abstract, rule-based representations.[2] Following Karmiloff-Smith

[2] It could be argued that the concept of restructuring is incompatible with that of automaticity, if restructuring is viewed as development from implicit to explicit knowledge (from memory-based procedures to rule-based procedures) and automatization is viewed as development from explicit to implicit knowledge (Michael Long, personal communication). McLaughlin protects himself from this criticism by scrupulously avoiding any characterization of the controlled:automatic distinction in terms of knowledge types, but the conceptual problems involved in many discussions of explicit (or analyzed) and implicit (or unanalyzed) knowledge are evident in a dispute between Hulstijn and Bialystok. Hulstijn (1990) criticized Bialystok's model on the grounds that it permits development in only one direction, from unanalyzed to analyzed knowledge, arguing that the existence of second language learners who use explicit grammar rules as the starting point for the establishment of automatic routines provides sufficient evidence that language learning need not start with unanalyzed knowledge. Bialystok responded that language learning must indeed start with unanalyzed knowledge, that it is simply not possible for mental representations to become less analyzed (Bialystok, 1990b). One can only agree with Bialystok that knowledge, *qua* knowledge, cannot become less analyzed, while recognizing that Hulstijn's point has less to do with

(1986), McLaughlin suggests that restructuring occurs when learners go beyond success: "once the procedures at any phase become automatized, consolidated, and function efficiently, learners step up to a 'metaprocedural' level, which generates representational change and restructuring" (McLaughlin, 1990, p.120). The shift from memory-based to rule-based representations is not motivated by processing considerations, and the establishment of complex internal representations is unlikely to result in faster, more efficient processing in the ways in which Cheng's prototypical strategy shifts do.

It would be interesting to see whether knowledge restructuring in second language learning results in U-shaped behavior with respect to fluency indicators such as speech rate and pause distribution as well as accuracy measures. Although restructuring may result in "algorithmic improvement" in Anderson's sense, this is a difficult notion to pin down in any specific way, and the general principle that it is more efficient to remember than to compute (to be discussed in the next section) suggests that fluency as well as accuracy may follow a U-shaped curve, declining when restructured procedures are introduced and increasing again as the new procedures are routinized so that they can be drawn directly from memory. This has not been McLaughlin's concern, since he has identified restructuring with developmental stages and changes in linguistic representations. His position appears similar to earlier versions of Bialystok's model (Bialystok, 1982), that there are two partially independent though interacting dimensions to second language learning, one having to do with the development of fluency, attributed to automaticity, the other (McLaughlin's "restructuring," Bialystok's "analysis") having to do with the evolution of increasingly abstract representations of knowledge. Once restructuring has been re-defined in terms of representational changes, the concept gains relevance as a mechanism underlying the development of linguistic competence but has less relevance as a possible mechanism underlying the development of fluency.

---

knowledge than with the issue of whether such knowledge is accessed and used in production, which is an empirical question rather than a logical one.

Instance theory

In Cheng's discussion of restructuring, direct retrieval from memory was presented as an example of a strategy shift resulting in more efficient processing, and such strategy shifts were seen as suggesting that automaticity may be epiphenomenal. Logan (1985, 1988a, 1988b, 1990, 1991; Logan & Klapp, 1991; Logan & Stadler, 1991) has developed a theory that recognizes automaticity as a real phenomenon to be accounted for, while arguing that that the principles determining the establishment and retrieval of memories provide a more satisfactory accounting of the properties of automaticity than do traditional accounts based on notions of attention and resource limitations.

Logan's theory rests upon certain assumptions about how memories are established and retrieved:

1. Encoding into memory is an obligatory, unavoidable consequence of attention, although the quality of the encoding depends on the quality and quantity of attention. Not all contextual details are represented in the memory trace; subjects only encode what they pay attention to.

2. Retrieval from memory is an obligatory, unavoidable consequence of attention, although retrieval may not be easy and is not always successful—practice and the match between context at encoding and context at retrieval are crucial.

3. Each encounter with a stimulus is encoded, stored, and retrieved separately.

In Logan's instance theory, the learning mechanism responsible for automaticity is **memory-retrieval**, or more precisely, the accumulation of separate episodic traces with experience that produces a gradual transition from algorithmic (rule-based) processing to memory-based processing:

> The theory assumes that novices begin with a general algorithm that is sufficient to perform the task. As they gain experience, they learn specific solutions to specific problems, which they retrieve when they encounter the same problems again. Then, they can respond with the solution retrieved from memory, or the one computed by the algorithm. At some point, they may gain enough experience to respond with a solution from memory on

every trial and abandon the algorithm entirely. At that point, their performance is automatic.[3] (Logan, 1988a, p.493)

Instance theory contrasts sharply with process-based models of the development of skill, including both the Shiffrin and Schneider (1977) model, which assumes that underlying processes do not change but simply run off more rapidly when automatized, and Anderson's ACT*, which attributes speedup to reductions in the amount of resources required (through proceduralization and algorithmic improvement) or the number of steps to be executed (through composition). In Logan's view, the underlying process (the algorithm) does not change and remains available to handle examples that have not been encountered before. The theory assumes that application of the algorithm does not speed up significantly or require less attention. Use of the algorithm is simply replaced over time by a much more efficient process, single-step memory retrieval. Output is determined by a race between the algorithm and memory retrieval, and memory retrieval becomes faster and faster as instances accumulate. Memory-retrieval dominates eventually, because the greater the number of instances in the race, the greater the probability that one of them will finish before the algorithm.

Instance theory derives support from a number of experiments involving lexical decision tasks and alphabet arithmetic (Logan, 1988a). In the lexical decision experiment, most interesting because of the use of linguistic stimuli, subjects were presented with letter strings and were required to indicate as quickly as possible whether or not each was an English word. With practice, reaction times decreased substantially over blocks for the specific items practiced, while there was little evidence of a general practice effect which improved performance for new items, suggesting that subjects remembered their previous encounters with individual words and nonwords and apparently ruling out process-based theories that explain automatization in terms of general procedures that deal with new stimuli as effectively as old ones.

---

[3] Although Logan considers automatization to have been achieved when performance relies entirely on retrieval from memory, in a sense automatization is never complete, because each additional instance continues to have an effect on memory (Logan, 1988a).

Logan (1991) has provided arguments based on instance theory to account for the frequently cited properties of automaticity. Automatic processing is autonomous because memory encoding and retrieval are obligatory consequences of attention. It is harder to control than algorithm-based processing because the rapid finishing time of memory retrieval allows less time for an act of control to take effect. It is effortless when sufficient instances have been stored in memory to ensure easy retrieval, fast because memory retrieval is a single step process, and inaccessible to consciousness because memory retrieval does not have constituent processes that might be introspectable.

Instance theory has not been discussed in the SLA literature, but its implications for our understanding of second language fluency are provocative, partly because they are strikingly different from conventional wisdom. Learners and teachers often assume that fluency rests upon "internalized rules," perhaps without giving much thought to how such rules might operate but generally assuming that some process of linguistic construction is going on out of reach of awareness. Linguistically based theories of speech production also assume that rules are operative as processes in some sense, for example the process of "formulating" in the detailed model of speech production proposed by Levelt (1989). Pinker (1991) argues that both memory-based and rule-based processes play a role in speech production, but that rule-based production prevails whenever there is a sufficiently productive rule available; memory-based processing applies only to linguistic exceptions. In the spreading-activation model of sentence production proposed by Dell (1986, 1989), generative rules define sequences of categorically defined slots, which are then filled by items that have been retrieved from a lexical network. In this model, all inflected words (*boys, singing, brought*) and many derived words (*genuineness, clearly*) are also created by rule governed processes, assembled from morphemes during production (Dell, 1986).

Instance theory suggests that morphological rules are not manipulated in the production of morphologically complex words unless the words are first-time creations or have been produced so infrequently that memory traces are insufficient for direct retrieval. Instance theory also suggests that the retrieval of past solutions from memory may include complex forms at higher levels

than the word, including phrases and both completely formulaic and partly open clause structures. If memory is the basis of fluency, the conventional wisdom regarding the status of formulaic speech as a peripheral phenomenon in second language learning is therefore in serious need of revision. Instance theory provides theoretical support for the position of Pawley and Syder (1983), who argued that nativelike fluency is possible only because native speakers have memorized hundreds of thousands of morphologically complex lexical items, a considerable proportion of which consists of lexicalized sentence stems. For SLA, Widdowson (1989) has argued that the ability to use linguistic knowledge for communication depends less on the use of rules to assemble utterances from scratch than it does on having available a stock of partially pre-assembled patterns and formulaic frameworks, using rules primarily to make those adjustments to formulae that are necessary according to contextual demands. Gatbonton and Segalowitz (1988) and Arevart and Nation (1991) have advocated a number of classroom teaching techniques in harmony with instance theory, procedures designed to promote the automatization of specific utterances and utterance frames rather than the routinization of structures or rules.

The major weakness of instance theory is that it cannot account for studies showing evidence of transfer of training or generalization in skill learning. Carlson, Sullivan, and Schneider (1989) have reported the results from experiments in which subjects practiced judgments about digital logic gates for over 8,000 trials. Results indicated that the organization of component processes and use of working memory remained constant while the speed of component processes increased and attentional load decreased. Logan's own experiments have also produced some evidence of algorithm speed-up (Logan, 1988a), and Logan & Stadler (1991) have reported on a series of memory search experiments in which some evidence was found for a category comparison strategy (Hintzman, 1986), beyond the effects that could be produced by instance-based learning.

Strength theories

Logan's claim that automatization relies upon instance learning, in which each encounter with a stimulus is represented separately in memory, is the

most controversial aspect of the theory. Other automaticity-as-memory theories rely on strengthening to account for retrieval from memory (Cohen, Dunbar, & McClelland, 1990; MacKay, 1982; Schneider, 1985; Schneider & Detweiler, 1988). In strength theories, response representations do not accumulate, but connections between stimulus and response become progressively stronger with practice. Schneider has proposed a theory of automatization in visual search (Schneider, 1985; Schneider & Detweiler, 1988) that includes two kinds of learning, priority learning (responsible for response selection) and associative learning (responsible for the internal structure of responses). Within a connectionist-control architecture involving neural-like units at the micro level of processing, a common mechanism, proportional strengthening, is theorized to underlie both associative and priority learning. In a similar fashion, Cohen, Dunbar and McClelland (1990) have attempted to provide an explicit (although so far partial) account of the mechanisms underlying automaticity that can explain both its gradual development with practice and its relation to selective attention. The role of attention in this account is to select among competing processes on the basis of task instructions, and attention is conceived of as an additional source of input modulating interactions occurring at the intersections of pathways.

Connectionist models are associative memory strength theories. McClelland & Rumelhart (1985) and Cohen, Dunbar, and McClelland (1990) have argued that the problem of determining what counts as an instance in learning can be overcome through the use of distributed representations, in which memory for events is encoded neither as discrete instances nor as a single connection between a generic stimulus and a generic response, but in the strengths of a set of connections involving different units that are used to provide overlapping but nevertheless distinct representations. Within a connectionist model of language production, Stemberger (1985) proposed a resolution of an apparent paradox: morphologically complex words are analyzed by speakers of the language (i.e. their constituent parts are understood) and can be and often are produced by rule, but such forms are nonetheless lexicalized and need not be produced by rule. Stemberger suggested that analyzed units are organized for production in networks of

shared representations, by which grammatical morphemes are directly accessed by many different words. Shared representations can be acquired in two ways: complex forms with idiosyncratic meanings are learned as units, and high-frequency complex forms are knitted together in the process of automatization.

Although connectionist models have attracted a great deal of attention in the fields of both first and second language acquisition recently (Gasser, 1190; MacWhinney, 1989; Pinker & Mehler, 1988; Rumelhart & McClelland, 1986; Schmidt, 1988; Sokolik, 1990), there are as yet no published connectionist accounts dealing specifically with language fluency. However, a pre-connectionist interactive strength model proposed by MacKay (1982) deals specifically with speech production and addresses the question of flexibility in the development of fluency. In contrast to the standard view that automatic processes are less flexible than controlled processes, MacKay argues that fluency and flexibility are positively correlated and that increased practice leads to increased transfer to functionally equivalent actions. In MacKay's model, consistent practice strengthens associations among nodes in hierarchical networks (for speech, these consist of propositional, conceptual, syntactic, lexical, syllable, phonemic, and muscle movement nodes) and activating a node at any level of the network primes connected nodes. For all multilevel skills, automaticity varies with the level under consideration. Learning takes place at the level of abstraction that can benefit most from practice and in proportion to the learning that has already taken place at lower levels in the hierarchy. For adult native speakers of a language, fluency gains take place almost exclusively at the level of the propositional-conceptual system (for example, becoming more fluent with practice in discussing new and difficult concepts encountered in a new academic discipline), because associations at the phonological and muscle movement systems have been practiced for a lifetime and are already at maximal strength. The model predicts flexibility (transfer) whenever lower level processes have already become automatic. Only partial imperfect transfer can be expected when writing with the unaccustomed hand, for example, because the component muscle movements have received little practice, but the theory predicts a high level of transfer in the case of a bi-instrumental musician who practices a piece

on one instrument and then plays it on another.

MacKay's model was supported by a number of experiments in which German-English bilinguals translated some sentences from English to German and others from German to English during training and then demonstrated nearly perfect transfer of skill when translating the same sentences in the opposite direction. Kramer, Strayer, and Buckley (1990) specifically contrasted the predictions of Logan's instance theory with those of MacKay's model, reporting the results of two studies that examined the development and transfer of automatic processing in rule-based memory search tasks. For these tasks, high positive transfer occurred despite replacement of the exemplars of the memory set rules, suggesting that learning was not specific to the instances encountered during training.

Hierarchical strength models make some interesting predictions regarding the development of L2 fluency. All such models are sensitive to the number of specific examples (lower level components) practiced in connection with higher level tasks and are consistent with the claim that a substantial amount of experience with item-learning must take place before system learning begins (Cruttendon, 1981; Nattinger, 1990). For SLA, a number of studies have reported that rule application is sensitive to particular lexical items that are frequent and well-practiced (Abraham, 1984; Schmidt & Frota, 1986). Schmidt & Frota (1986) reported on a beginning learner of Portuguese who managed to achieve better than 80% accuracy for the choice of aspectual verb forms, but who achieved this through lexical learning (each specific verb assigned a consistent aspectual choice) rather than through a productive rule that applied equally to all verbs. However, interactive strength theories predict that transfer and generalization of skill do take place once a sufficient number of specific items have been stored and practiced. The relevant SLA studies mentioned above have only reported lexical effects on syntactic accuracy; it would be worth investigating whether second language fluency also depends on specific examples in the early stages (with faster response times for well practiced examples than for less practiced items) but eventually comes to be process-based (with no differences between response times for well practiced and new items), as MacKay's model predicts. On the other

hand, a possible weakness of this particular model is its emphasis on linguistically defined levels of production and the claim that fluency is achieved bottom-up, which is reminiscent of the audiolingual claim that phonological habits must be established first and that manipulation of the conceptual level in free expression must be delayed until all lower-level processes are well established. This seems falsified by the familiar case of the foreign graduate student at an English medium university who is adept at manipulation of the conceptual level (for example, in writing) without having achieved fluency at the lower component levels of speech production.

Chunking theories

The ACT* theory (discussed in an earlier section) proposed a total of five learning mechanisms for the development of procedural skill. Anderson (1986) subsequently claimed that the knowledge compilation mechanisms of proceduralization and composition were sufficient to produce the inductive learning for which he had previously used separate mechanisms of generalization, discrimination and strengthening. Most recently, Servan-Schreiber and Anderson (1990) have proposed a theory with a single mechanism, **competitive chunking** (equivalent to composition in the earlier account), to account for the empirical results of studies in which subjects learned artificial grammars. Servan-Schreiber and Anderson trained subjects on exemplar sentences generated by a miniature artificial grammar, finding strong support for the hypothesis that the primary mechanism responsible for learning was chunking and that grammatical discrimination after training was based on the degree to which representations of new strings could be built from the collection of learned chunks. The theory assumes that productions will be composed or chunked if they follow each other and are linked by goal settings. A computer simulation of the chunking theory reproduced the experimental results. This was an important demonstration, because it is commonly claimed that the result of exposure to the exemplars of an artificial grammars is the unconscious (unintentional and unaware) abstraction of the underlying rules of the system (Reber, 1967; Reber, Allen & Regan, 1985).

Newell and Rosenbloom (Newell, 1990; Newell & Rosenbloom, 1981; Rosenbloom & Newell, 1987) have proposed a theory in which chunking has

even more importance as a learning mechanism. The attempt is an ambitious one: to construct a unified theory of cognition that posits a single set of mechanisms that operate to produce the full range of human cognition, including problem solving, decision making, routine action, memory, learning, skill, perception, motor behavior, language, motivation, emotion, and imagination. Newell and Rosenbloom focus on the power law of practice, which appears to hold over the full range of human tasks, from purely perceptual skills such as target detection to purely mental tasks such as working out geometric proofs. The power law (sometimes referred to as the log-linear law) refers to the fact that a plot of the logarithm of the time to perform a task against the log of the amount of practice approximates a straight line (Anderson, 1982; Carlson, Sullivan, & Schneider, 1989; Logan, 1990; Rosenbloom & Newell, 1987). The ubiquity of the power law argues that there may be a common underlying mechanism responsible for improvement in skill on all tasks. Rosenbloom and Newell (1990) argue that chunking is the common learning mechanism. Using an artificial intelligence system called Soar, which uses production system architecture, Newell (1990) has been able to demonstrate that the chunking mechanism can account for the log-linear law in the learning of a large number of cognitive skills.

Language learning was not among the cognitive skills simulated within the Soar system (partly because Newell felt that the level of controversy over competing views in linguistics made it impossible to establish a set of commonly agreed upon facts of learning as the basis for simulation). Extensions to the problem of fluency in second language learning must be somewhat speculative, but there is some evidence for chunking in SLA. Rescorla and Okuda (1987) analyzed data from the first six months of acquisition of English by a Japanese speaking child, reporting that Atsuko was able to produce a large number of creative and novel referential sentences, but accomplished this by using a small number of patterns or modules (rather than by composing varied sentences from a large stock of single words), building up longer units by chunking the smaller modular components. The fact that the Soar model is hierarchical is appropriate for modeling speech production, in which an utterance may consist of higher level chunking into clauses and

phrases and lower level chunking into words and phonemes (Rosenbloom & Newell, 1987). At higher levels of analysis, task descriptions, plans, explanations and life stories have also been shown to be tree-structured or hierarchically chunked (Linde, 1987). Soar generates chunks that are based on lower-level patterns, higher-level (more abstract) patterns, or mixtures, such as chunks created from one higher-level pattern and one more primitive pattern. This suggests a model for representing the ways in which creative and routine elements may vary in fluent speech—for example, when formulaic utterances fill slots within a larger discourse pattern (Coulmas, 1981; Hatch, Flashner & Hunt, 1986; Hatch & Hawkins, 1987) or when formulaic frames themselves have open slots (Hakuta, 1974; Pawley & Syder, 1983).

General Discussion

The power law of practice is the most frequently cited characteristic of skill development by experimental psychologists, and the ability to account for the log-linear function is often considered the most important test of any theory of cognitive skill learning. Three of the proposals for the development of automatic procedural skill discussed in this paper fail this test. The original Shiffrin and Schneider dichotomy between automatic and controlled processing, restructuring, and the notion of executive control as incorporated in Bialystok's two-dimensional model do not include sufficiently well-specified mechanisms to make the relevant predictions. The remaining four proposals pass this particular test. Anderson has presented a detailed argument that ACT* mechanisms can account for the power law, because the interaction of strength dynamics and the characteristics of working memory produce a power function which overrides the predicted exponential speedup of algorithmic improvement (Anderson, 1982). Other theories predict the power law more directly, including the chunking model of Newell and Rosenbloom, as discussed above. Instance theory accounts quantitatively for the power-function speed-up of mean reaction times observed in the skill acquisition, as well as the power-function of standard deviations of reaction times (Logan, 1988a). Strength theories also predict the power law, because if strengthening formulas are mapped onto reaction time, speed of responding can be shown to

decrease as a power function of practice (Logan, 1991; Cohen, Dunbar, & McClelland, 1990).

One might choose among available theories on other grounds. If the objective is to find the richest set of metaphors for describing the development of second language fluency, then ACT*, with five mechanisms, may be ideal. But if these mechanisms are not all necessary, we may prefer economy of description, choosing among theories that rely on a single mechanism (memory retrieval, strengthening, or chunking) as long as they account for the relevant facts.[4]

For SLA, the crucial question must be which theory and which mechanisms best fit the facts of second language fluency and the way in which it develops. The following general observations concerning L2 fluency seem uncontroversial:

(1) Speaking is a complex task that requires processing at many different levels more or less simultaneously (Rehbein, 1987). Planning and uttering are partly accomplished in cycles, but, as Levelt has pointed out, if the language processors could not work in parallel, then "speaking would be more like playing chess: an overt move now and then, but mostly silent processing" (Levelt, 1989, p. 27). To the beginning, nonfluent second language learner, speaking sometimes does seem to require as much thought and effort as planning a chess move.

(2) It has seemed to most theorists that working memory limitations and the speed at which speech is processed in relation to the size of the knowledge base underlying it (de Bot, in press) necessitate some concept of automatic

---

[4] One may also choose among theories on the grounds of the attractiveness of their pedagogical implications. These are beyond the scope of this paper, but different theories of skill development often do have different practical implications. While the traditional view of automaticity stresses the necessity of awareness, attention and controlled practice in the early stages if automaticity is to be achieved (O'Malley, Chamot, & Walker, 1987), instance theory takes a different view of the importance of attention (based on assumptions about memory encoding and retrieval) and suggests that extended practice is not necessary in principle for automatization. What is necessary is having the required knowledge in memory in sufficient strength that it can be retrieved, and this can sometimes be achieved by rote memorization if the number of facts is small enough (Logan & Klapp, 1991). Gregg (1984) has cited anecdotal evidence from his learning of Japanese that seems to constitute a case for almost immediate automatic performance after memorization and very brief practice.

processing as the basis for fluency. However, it is better to consider components of second language production rather than the task as a whole as automatic or nonautomatic. Certain aspects of speaking may become automatic (e.g. phonological processing) while others remain nonautomatic (conceptual processing in particular).

(3) Second language fluency develops through practice in a gradual fashion, in harmony with the view that the controlled-automatic distinction should be viewed as a continuum rather than a dichotomy. The power law of practice has not been empirically demonstrated for any of the components of second language fluency, but seems a plausible extension of the observation that fluency improvement is more easily observable in the early stages of active second language use, with a gradual slowing of improvement over time until asymptote is reached.

None of these observations provides grounds for choosing among memory-based, strengthening, chunking, or multi-mechanism theories. However, there are other aspects of fluent speech which are controversial and which mirror controversies in the psychological literature on skill development. The most basic disagreement concerns the terminal behavior that any theory must explain. Leeson defines fluency as "the ability of the speaker to produce indefinitely many sentences conforming to the phonological, syntactical and semantic exigencies of a given natural language on the basis of a finite exposure to a finite corpus of that language" (Leeson, 1975, p. 136). But Pawley and Syder believe that "memorized sentences and phrases are the normal building blocks of fluent spoken discourse, and at the same time, that they provide models for the creation of many (partly) new sequences which are memorable and in their turn enter the stock of familiar usages" (Pawley & Syder, 1983, p. 208). The question is not whether fluent native and nonnative speakers can produce novel utterances. They can and do, as do nonfluent learners. The important questions are whether novel utterances are ever produced as fluently as familiar utterances, without attention or effort, and if this does come about, how and when is it accomplished in the course of development?

This strikes to the heart of the major point of disagreement among psychological theories of skill development, the relative importance of well-

practiced, specific items, instances or exemplars for the development of skilled performance, as opposed to improvement in performance attributed to the increasingly skillful application of abstract rules or algorithms, uninfluenced by specific instantiations of the rules (Singley & Anderson, 1989). Fluency gains are attributed to the influence of specific examples by the mechanisms of proceduralization and composition in the ACT* theory, restructuring as discussed by Cheng, instance theory, and the competitive chunking model of Servan-Schreiber and Anderson, none of which predict any increase in fluency through generalization to new examples. The generalization and strengthening mechanisms of ACT* and strength theories as a general class do predict skill-improvement at more abstract levels (Roitblat, 1988), although in connectionist models rule-like performance is never independent from the exemplars in the knowledge base (Sokolik, 1990). The Newell and Rosenbloom chunking mechanism also produces process-improvement at the level of rules when implemented in Soar architecture, which uses symbols and abstractions and produces implicit generalizations (Newell, 1990). It should be noted, however, that there is no currently viable psychological theory that rests upon generalization as the *sole* mechanism of skill development. There is therefore little theoretical support from psychology for the common belief that the development of fluency in a second language is almost exclusively a matter of the increasingly skillful application of rules.[5]

One of the major claims of researchers who have investigated pausal phenomena in native speech (Butterworth, 1975, 1989; Chafe, 1980; Goldman-Eisler, 1964, 1968; Henderson, Goldman-Eisler, & Skarbek, 1966; Levelt, 1989) is that the distribution of pausing and speaking in speech reflects an alternation between phases in which hesitant speech is due to attentional preoccupation with macroplanning while stretches of fluent speech with little pausing reflect skilled microplanning which does not require much attention. Goldman-Eisler (1964) characterized these speech production cycles in terms of

---

[5] As commonly practiced, the technique of pattern practice rests on the assumption that short training sessions with a small number of exemplars, each of which is typically practiced once, will lead to fluency based on automatic rule application. The theories reviewed in this paper suggest that unless such practice is very extensive (introducing the boredom factor), neither the specific examples practiced nor the general rule will be available subsequently for fluent use.

alternation between newly organized speech and old, well organized speech consisting of learned sequences and ready-made phrases, with pausal phenomena more likely to occur in newly created speech. These ideas have been taken up by the Kassel group (Dechert, 1980, 1983: Dechert, Möhle & Raupach, 1984; Dechert & Raupach, 1980a, 1980b, 1987; Raupach, 1980, 1984; Rehbein, 1987) and applied to a corpus of L1 and L2 samples of speech production in German, French and English. One major difference between fluent and nonfluent second language learners that emerges across these studies is that the nonfluent learner's pauses, false starts and other signs of hesitation reflect the need to focus attention on the lower levels of planning, whereas fluent learners act more like native speakers in exhibiting hesitation primarily as a reflection of integration and macroplanning.

The role of formulaic speech and chunk learning in these developments is less clear. Within the field of SLA, formulaic speech has been investigated primarily in early stage learning, motivated by the question of whether or not formulaic speech provides a point of initial entry to the productive linguistic system (Bohn, 1986; Bolander, 1989; Hakuta, 1974; Krashen & Scarcella, 1978; Schmidt, 1983; Schmidt & Frota, 1986; Wong Fillmore, 1979), with some attention to the role of formulaicity in pragmatic competence (Blum-Kulka, 1989; Manes & Wolfson, 1981; Yorio, 1980). There has been comparatively little investigation of the possible role of memorized language as a mechanism for the production of fluent speech which continues to operate in competition with productive rules, under a tacit assumption that once forms have been analyzed and some productive use established then all subsequent appearances of the forms in question can be taken as productive synthesis by rule-application. It is just as possible that, once created, new forms are subsequently stored and pulled from memory for subsequent use. Even errors, normally considered the best evidence for rule-based productivity, may be stored and retrieved as wholes by learners rather than being committed each time as a creative act (Platt & MacWhinney, 1983; Schmidt & Frota, 1986).

Reliance on formulaic speech has often been considered a strategy to outperform one's competence and, occasionally, as a crutch to be used in the absence of fluency as well (Rehbein 1987), a strategy used to compensate for the lack of automatic processing ability. However, Dechert (1983) has

proposed that learners must develop lexicalized "islands of reliability" that become the basis for the search processes necessary for the course of planning and executing less formulaic speech. Raupach (1984) argues that fluency is heavily dependent upon stored chunks, and suggests that second language learners may go through several identifiable stages in the development of fluency: (1) in the earliest stages, most planning activities take place within filled and unfilled pauses; (2) with the adoption of new forms of hesitating such as drawled syllables, speech planning activities take place in different places and in connection with different islands of reliability; (3) new organizers are acquired that lead to a preferred set of formulaic schemata (comparable to Pawley and Syder's lexicalized sentence stems); and (4) formulaic islands of reliability eventually become integral parts of longer speech stretches.

Suggestions for Research

The empirical facts about automatization are well enough established with respect to tasks that are amenable to careful laboratory control that connectionist and other computational simulations offer the potential for resolving disputes concerning the mechanisms assumed to underlie automatization. Yet it is unsettling to realize that the mechanisms made available by psychological theorizing for understanding second language fluency derive primarily from the study of skill in such tasks as typing, the detection of target letters in fields of distractors, judgments about digital logic gates, alphabet arithmetic, and computer simulations of the same tasks, tasks which cannot be assumed to rely necessarily on the same learning mechanisms as speaking a second language.

Whether the investigation of second language fluency itself can contribute to identification of the mechanisms underlying it is an unanswerable question at this point, because there has been relatively research on the topic. As Crookes (1991) has pointed out, the study of second language performance has received less attention than that of second language competence, and within the language processing field, studies of second language comprehension greatly outnumber studies of production. Studies of second language fluency

constitute an even smaller subset of production studies, and not many of these have been concerned with how fluency develops. Further descriptive studies are justified along the lines of Lennon (1990) and Riggenbach (in press), in order to establish reliable and valid quantitative measures of fluency. The multiple case-study approach used in these studies is particularly recommended, because (a) extensive data is required for each language learner in order to be able to identify such factors as idiosyncratic formulaic utterances and their evolution over time (Raupach, 1984), and (b) there are likely to be differences among learners in the ways in which they cope with the demands of processing language in real time (Peters, 1983; Skehan, 1991; Wong Fillmore 1979). In addition, studies that attempt to define the empirical correlates of perceived fluency should probably be balanced by introspective reports by learners. Language learners cannot introspect the microstructure of automatic processes but can say what the focus of their attention is at a particular time, can be assumed to be consciousness of underdetermined choice points in the flow of action (Baars, 1988), and may have something to contribute to an understanding of the interplay of automatic and controlled processing in fluent and nonfluent second language speech.

Such research should not just be exploratory in nature, but needs to identify gaps in our current understanding and attempt to fill them. The controversy in psychological theorizing concerning the relative importance of memory retrieval and speeded computation in skill development raises questions that should be investigated with respect to the development of second language fluency. The acquisition of general purpose procedures for the use of language is sometimes assumed by psychologists to be the main evidence against theories emphasizing memory for specific instances. In fact, the prevailing practice in SLA has been to assume rule productivity as soon as it is possible to do so (as soon as there is evidence of the analysis of chunks or productive use), and what is needed is more careful investigation of the interplay between routine and creative speech and the relationship of this to fluency development. There are many ways in which this can be examined. The simplest would be to study the ability of learners to produce target language constructions fluently, comparing their performance on items that have been practiced with their performance on the same rules with different

exemplars. Such research could go beyond Logan's findings from lexical decision and alphabet arithmetic tasks by comparing the contribution of algorithm speed-up and the retrieval of specific instances on such varied tasks as verb inflection (it seems plausible to hypothesize that fluent speakers of Spanish may have memorized all the high frequency forms of most verbs) and relative clause formation (it is intuitively unlikely that speakers of English could have memorized all the relative clauses they have ever produced).

The problem of the lack of experimental control in second language studies cannot be completely overcome, but a research paradigm introduced by Hulstijn (1989a, 1989b) can help. Hulstijn has been involved in a series of investigations adopting twin experiments, combining a natural L2 learning experiment with a semi-artificial experiment using artificial input inserted into verbal materials of a natural language. For example, the learning of Dutch function words and word order by second language learners (the study of which has high validity but limited reliability) can be compared with the learning of artificial functional morphemes and word order by native speakers of Dutch when these artificial elements are inserted into Dutch sentences (the study of which would have limited validity but high reliability). Hulstijn has not used the paired experiment paradigm to investigate fluency issues, but it could easily be done, for example with respect to the kind of studies mentioned above comparing learner performance on previously encountered examples with their performance on new realizations of the same rules. An additional desirable modification of research methodology in the investigation of fluency is greater use of reaction time instead of accuracy as a basic tool of measurement, since most accounts of the development of skill make clearer predictions concerning processing speed than they do concerning accuracy.

As yet, we know little about how second language fluency may vary under different task demands, in particular interactional contexts, and with respect to particular topics, or how fluency develops over time for specific tasks and across tasks for younger and older learners, with and without instruction. One goal of future research should be to investigate such aspects of second language fluency development directly, in order to reduce the degree to which our understanding of the mechanisms underlying fluency is dependent on theories from other fields. At the present time, this dependence

is almost absolute. It may turn out that such research can also contribute to the discussion of current issues raised by general theories of skill development.

Author's address for correspondence:

Richard Schmidt
Department of English as a Second Language
University of Hawai'i
1890 East-West Road
Honolulu, HI 96822

References

Abraham, R. (1984). Patterns in the use of the present tense third person singular -s by university level ESL speakers. *TESOL Quarterly*, 18, 55–69.

Anderson, J.R. (1980). *Cognitive psychology and its implications*. San Francisco: Freeman.

Anderson, J.R. (1981). A theory of language acquisition based on general learning mechanisms. *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, 97–103.

Anderson, J.R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J.R. (1986). Knowledge compilation: The general learning mechanism. In R.S. Michalski, J.G. Carbonell, & T.M. Mitchell (Eds), *Machine learning: An artificial intelligence approach: Vol. 2* (pp. 289–310). Los Altos, CA: Morgan Kaufman.

Anderson, J.R. (1989). Practice, working memory, and the ACT* theory of skill acquisition: a comment on Carlson, Sullivan, and Schneider. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 527–530.

Arevart, S., & Nation, P. (1991). Fluency improvement in a second language. *RELC Journal*, 22, 84–94.

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

Bialystok, E. (1982). On the relationship between knowing and using linguistic forms. *Applied Linguistics*, 3, 181–206.

Bialystok, E. (1985). The compatibility of teaching and learning strategies. *Applied Linguistics*, 6, 181–206.

Bialystok, E. (1990a). *Communication strategies: a psychological analysis of second-language use*. Oxford: Basil Blackwell.

Bialystok, E. (1990b). The dangers of dichotomy: a reply to Hulstijn. *Applied Linguistics*, 11, 46–51.

Bialystok, E., & Bouchard Ryan, E. (1985). A metacognitive framework for the development of of first and second language skills. In D. Forrest-Pressley, G. MacKinnon, & T. Waller (Eds.), *Metacognition, cognition, and human performance, Vol. 1* (pp. 207–252). New York: Academic Press.

Bialystok, E., & Mitterer, J. (1987). Metalinguistic differences among three kinds of readers. *Journal of Educational Psychology*, 79, 147–53.

Bialystok, E., & Sharwood Smith, M. (1985). Interlanguage is not a state of mind: An evaluation of the construct for second-language acquisition. *Applied Linguistics*, 6, 101–117.

Blum-Kulka, S. (1989). Playing it safe: the role of conventionality in indirectness. In S. Blum-Kulka, J. House & G. Kasper (Eds.), *Cross-cultural pragmatics* (pp. 37–70). Norwood, NJ: Ablex.

Bohn, O-S. (1986). Formulas, frame structures, and stereotypes in early syntactic development: Some new evidence from L2 acquisition. *Linguistics*, 24, 185–202.

Bolander, M. (1989). Prefabs, patterns and rules in interaction? Formulaic speech in adult learners' L2 Swedish. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the life span: Aspects of acquisition, maturity, and loss* (pp. 73–86). Cambridge: Cambridge University Press.

Brumfit, C. (1984). *Communicative methodology in language teaching: The roles of fluency and accuracy.* Cambridge: Cambridge University Press.

Butterworth, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research*, 4, 75–87.

Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process*, (pp. 108–135). Cambridge, MA: Bradford/MIT Press

Carlson, R.A., & Schneider, W. (1990). Practice effects and composition: A reply to Anderson. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 531–533

Carlson, R.A., Sullivan, M., & Schneider, W. (1989). Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 517–526.

Chafe, W.L. (1980). Some reasons for hesitating. In Dechert and Raupach (Eds.), *Temporal variables in speech* (pp. 169–180). The Hague: Mouton.

Cheng, P.W. (1985). Restructuring versus automaticity: alternative accounts of skill acquisition. *Psychological Review*, 92, 414–423.

Cohen, J.D., Dunbar, K., & McClelland, J.L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review*, 97, 332–361.

Coulmas, F. (Ed.) (1981). *Conversational routine: Explorations in standardized communication situations and prepatterned speech*. The Hague: Mouton.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11, 367–383.

Crookes, G. (1991). Second language speech production research: A methodologically-oriented review. *Studies in Second Language Acquisition*, 13, 113–132.

Cruttenden, A. (1981). Item-learning and system learning. *Journal of Psycholinguistic Research*, 10, 79–88.

de Bot, K. (in press). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics*.

Dechert, H.W. (1980). Pauses and intonation as indicators of verbal planning in second-language speech productions: Two examples from a case study. In H.W. Dechert & M. Raupach (Eds), *Temporal variables in speech* (pp. 271–285). The Hague: Mouton.

Dechert, H.W. (1983). How a story is done in a second language. In C. Færch & G. Kasper (Eds), *Strategies in interlanguage communication*, (pp. 175–198). London: Longman.

Dechert, H.W., Möhle, D., & Raupach, M. (Eds.). (1984). *Second language productions*. Tübingen: Narr.

Dechert, H.W., & Raupach, M. (Eds.). (1980a). *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler*. The Hague: Mouton.

Dechert, H.W., & Raupach, M. (Eds.). (1980b). *Towards a cross-linguistic assessment of speech production*. Frankfurt a.M.: Lang.

Dechert, H.W., & Raupach, M. (Eds). (1987). *Psycholinguistic models of production*. Norwood, NJ: Ablex.

Dell, G.S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.

Dell, G.S. (1989). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 136–165). Cambridge, MA: Bradford/MIT Press.

Færch, C., & Kasper, G. (1984). Pragmatic knowledge: Rules and procedures. *Applied Linguistics*, 5, 214–225.

Fillmore, C.J. (1979). On fluency. In C. Fillmore, D. Kempler, & W.S-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85–101). New York: Academic Press.

Gasser, M. (1990). Connectionism and universals of second language acquisition. *Studies in Second Language Acquisition*, 12, 179–199.

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

Goldman-Eisler, F. (1964). Discussion and further comments. In E.H. Lenneberg (Ed.), *New Directions in the study of language* (pp. 109–131). Cambridge, MA: MIT Press.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* New York: Academic Press.

Gregg, K.R. (1984). Krashen's monitor and Occam's razor. *Applied Lingusitics*, 5, 79–100.

Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning*, 24, 287–298.

Hasher, L., & Zacks, R.T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356–388.

Hatch, E., Flashner, V., & Hunt, L. (1986). The experience model and language teaching. In R.R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 5–22). Rowley, MA: Newbury House.

Hatch, E., & Hawkins, B. (1987). Second-language acquisition: An experiential approach. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics, Vol.2: Reading, writing and language learning* (pp. 241–283). Cambridge: Cambridge University Press.

Henderson, A., Goldman-Eisler, F., & Skarbek, A. (1966). Sequential temporal patterns in spontaneous speech. *Language and Speech*, 9, 207–216.

Hintzman, D.L. (1986). "Schema abstraction" in a multiple-trace model. *Psychological Review*, 93, 411–428.

Hirst, W., Spelke, E.S., Reaves, C.C., Caharack, G., & Neisser, U. (1980). Dividing attention with alternation or automaticity. *Journal of Experimental Psychology: General*, 109, 98–117.

Hulstijn, J. (1989a). Experiments with semi-artificial input in second language research. In B. Hammarberg (Ed.), *Language learning and learner language. Scandinavian Working Papers on Bilingualism* (The University of Stockholm, Centre for Research on Bilingualism), 8, 28–40.

Hulstijn, J. (1989b). Implicit and incidental second language learning: experiments in the processing of natural and partly artificial input. In H.W. Dechert & M. Raupach (Eds.), *Interlingual processes* (pp. 49–73). Tübingen: Narr.

Hulstijn, J. (1990). A comparison between the information-processing and the analysis/control approaches to language learning. *Applied Linguistics*, 11, 30–45.

Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge, MA: MIT Press.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. Davies (Eds.,) *Varieties of attention* (pp. 29–61). Orlando, Fla: Academic Press.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: evidence from conscious access: evidence from children's metalinguistic and repair data. *Cognition*, 23, 95–147.

Kihlstrom, J. (1984). Conscious, subconscious, unconscious: A cognitive perspective. In K. Bowers & D. Meichenbaum (Eds.), *The unconscious reconsidered* (pp. 149–211). New York: Wiley.

Klahr, D. (1984). Transition processes in quantitative development. In R.J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 101–139). New York: Freeman.

Kramer, A.F., Strayer, D.L., & Buckley, J. (1990). Development and transfer of automatic processing. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 505–522.

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

Krashen, S.D. (1985). *The input hypothesis: Issues and implications.* London: Longman.

Krashen, S.D., & Scarcella, R. (1978). On routines and patterns in language acquisition and performance. *Language Learning*, 28, 283–300.

Kuiper, K., & Tillis, F. (1986). The chant of the tobacco auctioneer. *American Speech*, 60, 141–149.

LaBerge, D. (1981). Automatic information processing: a review. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 173–185). Hillsdale, NJ: Erlbaum.

Leeson, R. (1975). *Fluency and language teaching.* London: Longman.

Lennon, P. (1989). Introspection and intentionality in advanced second-language learners. *Language Learning*, 39, 375–396).

Lennon, P. (1990). Investigating fluency in EFL: a quantitative approach. *Language Learning*, 40, 387–417.

Levelt, W.J.M. (1977). Skill theory and language teaching. *Studies in Second Language Acquisition*, 1, 53–70.

Levelt, W. (1989). *Speaking.* Cambridge, MA: Bradford/MIT

Linde, C. (1987). The life story: a temporally discontinuous discourse type. In Dechert & Raupach (Eds.), *Psycholinguistic models of production* (pp. 189–206). Norwood, NJ: Ablex.

Logan, G.D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology*, 39, 367–386.

Logan, G.D. (1988a). Toward an instance theory of automatization. *Psychological Review*, 95, 492–527.

Logan, G.D. (1988b). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors*, 30, 583–598.

Logan, G.D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, 22, 1–35.

Logan, G.D. (1991). Automaticity and memory. In W.E. Hockley & S. Lewandowski (Eds.), *Relating theory and data* (pp. 347–367). Hillsdale, NJ: Erlbaum.

Logan, G.D., & Klapp, S.T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 179–195.

Logan, G.D., & Stadler, M.A. (1991). Mechanisms of performance improvement in consistent mapping memory search: Automaticity or strategy shift? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 478–496.

MacKay, D.G. (1982). The problems of flexibility, fluency, and speed-accuracy tradeoff in skilled behavior. *Psychological Review*, 89, 483–506.

MacWhinney, B. (1989). Competition and connectionism. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 422–457). Cambridge: Cambridge University Press.

Manes, J., & Wolfson, N. (1981). The compliment formula. In F. Coulmas (Ed.), *Conversational routine: explorations in standardized communication* (pp. 115–132), The Hague: Mouton.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11, 113–128.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning*, 33, 135–157.

McClelland, J., & Rumelhart, D.E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, 114, 159–188.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.

Miller, G.A. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology*, 56, 485–491.

Möhle, D. (1984). A comparison of the second language speech production of different native speakers. In H.W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 26–49). Tübingen: Narr.

Morrison, D.M., & Low, G. (1983). Monitoring and the second language learner. In J.C. Richards and R.W. Schmidt (Eds.), *Language and communication* (pp. 228–250). London: Longman.

Nattinger, J.R. (1990). Prefabricated speech for language learning. In B. VanPatten & J.F. Lee (Eds.), *Second language acquisition, foreign language learning* (pp. 198–206). Clevedon, England: Multilingual Matters.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

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.

Norman, D. (1986). Reflections on cognition and parallel distributed processing. In J. McClelland, D. Rumelhart, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 2: Psychological and biological models* (pp. 531–546). Cambridge, MA: MIT Press.

Norman, D., & Shallice, T. (1986). Attention to action: willed and automatic control of behavior. In R. Davidson, G. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory, Vol. 4* (pp. 1–18). New York: Plenum.

O'Malley, J.M., Chamot, A.U., & Walker, C. (1987). Some applications of cognitive theory to second language acquisition. *Studies in Second Language Acquisition, 9,* 287–306.

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). London: Longman.

Peters, A.M. (1983). *The units of language acquisition.* Cambridge: Cambridge University Press.

Phillips, J.G., & Hughes, B.G. (1988). Internal consistency of the concept of automaticity. In A.M. Colley & J.R. Beech (Eds.), *Cognition and action in skilled behaviour* (pp. 317–331). Amsterdam: Elsevier (North-Holland).

Pinker, S. (1991). Rules of language. *Science, 253,* 530–535.

Pinker, S., & Mehler,J. (Eds.). (1988). *Connection and symbols.* Cambridge, MA: MIT Press.

Platt, C.B., & MacWhinney, B. (1983). Error assimilation as a mechanism in language learning. *Journal of Child Language*, 10, 401–414.

Posner, M.I., & Snyder, C.R.R. (1975). Attention and cognitive control. In R.L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Raupach, M. (1980). Temporal variables in first and second language speech production. In Dechert & Raupach (Eds.), *Temporal variables in speech*. (pp. 263–270). The Hague: Mouton.

Raupach, M. (1984). Formulae in second language production. In H. Dechert, D. Möhle & M. Raupach (Eds.), *Second language productions* (pp. 114–137). Tübingen: Narr.

Reason, J. (1984). Lapses of attention in everyday life. In R. Parasuraman & D. Davies (Eds.), *Varieties of attention* (pp. 515–549). Orlando, FL: Academic Press.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.

Reber, A.S., Allen, R., & Regan, S. (1985). Syntactical learning and judgment: Still unconscious and still abstract. *Journal of Experimental Psychology: General*, 114, 17–24.

Reese, H.W. (1989). Rules and rule-governance: Cognitive and behavioristic views. In S.C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 3–84). New York: Plenum.

Rehbein, J. (1987). On fluency in second language speech. In H.W. Dechert and M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 97–105) Tübingen: Narr.

Rescorla, L., & Okuda, S. (1987). Modular patterns in second language acquisition. *Applied Psycholinguistics*, 8, 281–308.

Riggenbach, H. (in press). Towards an understanding of fluency: A microanalysis of nonnative speaker conversation. *Discourse Processes*.

Roitblat, H.L. (1988). A cognitive action theory of learning. In J. Delacourt & J.C.S. Levy (Eds.), *Systems with learning and memory abilities* (pp. 13–26). Amsterdam: Elsevier (North-Holland).

Rosenbloom, P., & Newell, A. (1987). Learning by chunking: A production system model of practice. In D. Klahr, P. Langley & R. Neches (Eds.), *Production system models of learning and development* (pp. 221–286). Cambridge, MA: MIT Press.

Rumelhart, D.E., & McClelland, J.L. (1986). On learning the past tenses of English verbs. In J.L. McClelland, D.E. Rumelhart, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition, Vol.2: Psychological and biological models* (pp. 216–271). Cambridge, MA: MIT Press.

Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In Dechert & Raupach (Eds.), *Psycholinguistic models of production* (pp. 45–75). Norwood, NJ: Ablex.

Schmidt, R. (1983). Interaction, acculturation and the acquisition of communicative competence. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 137–174). Rowley, MA: Newbury House.

Schmidt, R. (1988). The potential of parallel distributed processing for S.L.A. theory and research. *University of Hawai'i Working Papers in ESL, 7*(1), 55–66.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,* 129–158)

Schmidt, R., & Frota, S.N. (1986). Developing basic conversational ability in a second language: a case study of an adult learner of Portuguese. In R.R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–322). Rowley, MA: Newbury House.

Schneider, W. (1985). Toward a model of attention and the development of automatic processing. In M. Posner & O. Maris (Eds.), *Attention and performance XI* (pp. 475–492). Hillsdale, NJ: Erlbaum.

Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. *Human Factors*, 30, 539–566.

Schneider, W., Dumais, S., & Shiffrin, R. (1984). Automatic and control processing and attention. In R. Parasuraman & D. Davies (Eds.), Varieties of Attention (pp. 1–27). Orlando, FL: Academic Press.

Schneider, W., & Shiffrin, R. (1985). Categorization (restructuring) and automatization: two separable factors. *Psychological Review*, 92, 424–428.

Schumann, J.H. (1990). Extending the scope of the acculturation/pidginization model to include cognition. *TESOL Quarterly*, 24, 667–684.

Segalowitz, N. (1991). Does advanced skill in a second language reduce automaticity in the first language? *Language Learning*, 41, 59–83.

Servan-Schreiber, E., & Anderson, J.R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 592–608,

Shallice, T. (1978). The dominant action system: an information-processing approach to consciousness. In K. Pope and J. Singer (Eds.), *The stream of consciousness: Scientific investigations into the flow of human experience* (pp. 117–157). New York: Plenum Press.

Sharwood Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics*, 7, 239–256.

Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing II: Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–90.

Sincoff, J.B., & Sternberg, R.J. (1987). Two faces of verbal ability. *Intelligence*, 11, 263–276.

Singley, M.K., & Anderson, J.R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Learning*, 13, 275–298.

Sokolik, M.E. (1990). Learning without rules: PDP and a resolution of the adult language learning paradox. *TESOL Quarterly*, 24, 685–695.

Stemberger, J.P. (1985). An interactive activation model of language production. In A.W. Ellis (Ed.), *Progress in the psychology of language, Vol. 1* (pp. 143–185). Hillsdale, NJ: Erlbaum.

Stillings, N.A., Feinstein, M.H., Garfield, J.L., Rissland, E.L., Rosenbaum, D.A., Weisler, S.E., & Baker-Ward, L. (1987). *Cognitive science.* Cambridge, MA: Bradford /MIT Press.

Strayer, D.L., & Kramer, A.F. (1990). An analysis of memory-based theories of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 291–304.

Wickens, C.D. (1984). Processing resources in attention. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 63–102). New York: Academic Press.

Widdowson, H.G. (1989). Knowledge of language and ability for use. *Applied Linguistics*, 10, 128–137.

Wong Fillmore, L. (1979). Individual differences in second language acquisition. In C.J. Fillmore, D. Kempler & W.S-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 203–228). New York: Academic Press.

Yorio, C.A. (1980). Conventionalized language forms and the development of communicative competence. *TESOL Quarterly*, 15, 433–442.