THEORETICAL AND EMPIRICAL ISSUES IN THE STUDY OF IMPLICIT AND EXPLICIT SECOND-LANGUAGE LEARNING

Introduction

Jan H. Hulstijn
University of Amsterdam

There are good theoretical and educational reasons to place matters of implicit and explicit learning high on the agenda for SLA research. As for theoretical motivations, perhaps the most central issue in SLA theory construction in need of explanation is the differential success in one’s first language (L1) and in one’s second language (L2). Although acquisition of an L1 results in full mastery of the language (provided that children are exposed to sufficient quantities of input and do not suffer from mental disabilities), learners of an L2—even after many years of L2 exposure—differ widely in level of attainment. How can we explain universal success in the case of L1 acquisition and differential success in the case of L2 acquisition? Among the many explanations that have been proposed, including brain maturation and brain adaptation processes (critical period), access to Universal Grammar, L1 interference, and sociopsychological factors (see Hyltenstam & Abrahamsson, 2003, for a review), one finds explanations that involve the notions of implicit and explicit learning. Scholars working in different disciplines, in different theoretical schools, and sometimes using different terminology have argued that L1 acquisition (or at least the acquisition of L1 grammar) relies principally on processes of what we might now call implicit learning, whereas the acquisition of an L2 often relies on both implicit and explicit learning (Bley-Vroman, 1991; DeKeyser, 2003; N. Ellis, this issue; R. Ellis, 2004; Krashen, 1981; Reber & Allen, 2000).

I am grateful to Rod Ellis for his thoughtful comments on previous versions of this text.

Address correspondence to: Jan H. Hulstijn, Amsterdam Center for Language and Communication, Faculty of Humanities, University of Amsterdam, 134 Spuistraat, 1012 VB Amsterdam, Netherlands; e-mail: hulstijn@uva.nl.
As concerns educational motivations, the extent to which implicit and explicit learning can be shown to explain the differential success of SLA is likely to determine their relevance for L2 instruction. Curriculum planners, material designers, teachers, and learners all have a vested interest in knowing in which linguistic domains L2 learning might best benefit from implicit or explicit learning modes.

DEFINITIONS

To put the papers brought together in this thematic issue into perspective, I will make some introductory remarks concerning key concepts. Following Schmidt (1994a), I will distinguish between implicit and explicit memory, implicit and explicit knowledge, implicit and explicit learning, implicit and explicit instruction, inductive and deductive learning, and incidental and intentional learning. The definitions given here are in line with most, but not all, of the literature. As it is more practical to define the implicit member of a pair in relation to the explicit member in some cases, I will define the latter first.

Implicit and Explicit Memory

Explicit and implicit memory is memory of a past event with or without conscious awareness, respectively (Schacter, 1987). The constructs are usually operationalized in terms of behavior exhibited in an information-retrieval task. The tasks are distinguished operationally by the instructions given to the participants. On explicit memory tasks, participants are explicitly asked to recall past events or to recognize previously studied events. On implicit memory tasks, no reference is made to past events; participants are simply asked to perform the task as accurately and quickly as possible. In general, studies have found that participants’ responses are affected by the absence or presence of previous events (“priming”), with no awareness of the effect on the part of the participants. Evidence for the dissociation between explicit and implicit memory comes from experiments with patients suffering from retrograde amnesia who perform very poorly on explicit tasks but almost equally well as normal subjects on implicit tasks (Jacoby, 1983; Schacter; Roediger, 1990; for a discussion, see Buchner & Wippich, 1998).

Implicit and Explicit Knowledge

Explicit and implicit knowledge differ in the extent to which one has or has not (respectively) an awareness of the regularities underlying the information one has knowledge of, and to what extent one can or cannot (respectively) verbalize these regularities (Anderson & Lebiere, 1998, p. 5; Bialystok, 1982; R. Ellis, 2004). Explicit and implicit knowledge are often associated with effort-
ful and automatic processing, respectively (Hasher & Zacks, 1979; Segalowitz, 2003; Segalowitz & Hulstijn, 2005). Declarative knowledge is sometimes used as a synonym for explicit knowledge (Anderson & Lebiere). Knowledge is declarative when subjects can explicitly declare or verbalize their knowledge. Episodic knowledge (Tulving, 1983) is knowing “when and where.” (Episodic memory is the behavioral manifestation of episodic knowledge. They are often used as virtually synonymous.) Episodic knowledge is a form of autobiographical memory. L2 learners sometimes have episodic knowledge of new, recently encountered, L2 words or expressions. This episodic knowledge might also be regarded as a form of explicit knowledge.

In terms of brain activation, long-term memory of explicit, declarative, and episodic facts—including many aspects of vocabulary knowledge—has been claimed to reside in various areas of the neocortex (especially the frontal and temporal lobes but also the parietal and occipital lobes). During initial stages of its acquisition (i.e., during the consolidation phase), disparate cortical sites associated with a memory are bound together by the hippocampus in the medial temporal lobe, which is part of the limbic system. The consolidation phase might last several weeks or months (Byrnes, 2001, p. 71; Eichenbaum, 2001; Meeter & Murre, 2004; Squire & Knowlton, 2000; Ullman, 2001, 2004).

Implicit knowledge also resides in various regions of the neocortex (especially the basal forebrain, striatum, amygdala, and cerebellum) but is not subserved by the hippocampus (Byrnes, p. 71; Paradis, 1994; Reber, Allen, & Reber, 1999; Ullman, 2001, 2004). Implicit and Explicit Learning

Of all key concepts dealt with in this introduction, explicit and implicit learning are the two for which the least consensus exists. As I will explain subsequently in more detail, there are several reasons for this lack of consensus. For the moment, let me give a definition, notwithstanding the issues that continue to be debated. Explicit learning is input processing with the conscious intention to find out whether the input information contains regularities and, if so, to work out the concepts and rules with which these regularities can be captured. Implicit learning is input processing without such an intention, taking place unconsciously. According to Reber et al. (1999):

Implicit learning (a) operates largely independent of awareness, (b) is subsumed by neuroanatomical structures distinct from those that serve explicit, declarative processes, (c) yields memorial representations that can be either abstract or concrete, (d) is a relatively robust system that survives psychological, psychiatric, and neuroanatomical injury, (e) shows relatively little interindividual variability, and (f) is relatively unaffected by ontogenetic factors. (p. 504)

Note that learning is often defined with reference to the nature of the knowledge learned. Explicit and implicit learning is then the learning of explicit and
implicit knowledge, respectively. In the pedagogical literature, explicit and implicit L2 learning are sometimes rather loosely defined as learning with or without the aid of grammar rules, respectively.

**Implicit and Explicit Instruction**

Instruction is explicit or implicit when learners do or do not receive information concerning rules underlying the input, respectively (R. Ellis, 1994, p. 642; Norris & Ortega, 2000).

**Inductive and Deductive Learning**

Deductive learning takes place when rules are presented before examples are provided; inductive learning takes place when examples are given before rules are presented (DeKeyser, 1995, p. 380). The terms *deductive* and *inductive* learning are used in an instructional context. By definition, deductive and inductive learning are part of explicit instruction because the correct rule is always given at some point.

**Incidental and Intentional Learning**

Intentional learning refers to the learning mode in which participants are informed, prior to their engagement in a learning task, that they will be tested afterward on their retention of a particular type of information. Incidental learning refers to the mode in which participants are not forewarned of an upcoming retention test for a particular type of information. Incidental learning is also given a more general definition—not limited to experimental situations—as the unintentional picking up of information (see Hulstijn, 2003, for a detailed discussion).

**THREE FACTORS THAT CAUSE CONFUSION**

Not everyone will agree with these definitions. Especially in the case of the two labels that form the umbrella for this special issue—implicit and explicit learning—there exist many definitions. SLA-oriented reviews of the vast literature on implicit and explicit learning have been provided by DeKeyser (2003), N. Ellis (1994, this issue), Paradis (1994), and Schmidt (1994a, 1994b, 2001). Rather than attempting to summarize the information conveyed in these review papers, I will first focus on three factors that must be taken into account in theories of explicit and implicit L2 learning and then argue that it is especially the first of these (concerned with the regularity or irregularity of the linguistic phenomenon to be learned and closely related to the differences between symbolic and subsymbolic accounts of language knowledge) that
might have caused confusion in the explicit-implicit debate. Let us first consider the following two learning tasks.

In task A, the input information can easily be described in symbolic terms with noncompeting rules operating on nonoverlapping categories or features. In task B, the input information cannot be described with noncompeting, large-scope rules but only with fuzzy categories resulting from the competition between many cues, differing in availability, strength, and validity (in terms of the competition model of Bates & MacWhinney, 1989). Explicit learning—defined as input processing with the intention to find out whether the data can be described with rules and, if so, to discover the rules—is possible in task A provided that the rules are not too complex but is bound to fail in task B. Implicit learning—defined as an absence of the intentions just mentioned—is a possibility in task A. However, to the extent that the data are presented in a way that their underlying structure becomes salient and to the extent that learners have acquired (in school and elsewhere over time) metalinguistic knowledge and metacognitive problem solving strategies that facilitate the discovery of the underlying regularities, input processing in task A might spontaneously invoke an explicit learning mode.\(^3\) What this example aims to illustrate is that definitions of learning—whether implicit or explicit—as a process (how) can easily become contaminated with the object of learning (what). Different views on the object of learning will easily lead to different views on its process. Thus, when making claims about the effect or feasibility of implicit and explicit learning modes, we must take into account the possible interaction of at least the following three factors: (a) the regularity and complexity of the system underlying the data (see N. Ellis, this issue; Williams, this issue); (b) the frequency and salience with which any underlying regularity of the data is represented in the input to which learners are exposed (see N. Ellis, this issue; Williams, this issue); and (c) learners’ individual differences in knowledge, skills, and information processing styles, which might be beneficial or detrimental to discovering underlying regularities (Robinson, this issue; see also Reber & Allen, 2000; Reber, Walkenfeld, & Hernstadt, 1991).

The interaction of these factors might determine how learners process and categorize the input data to which they are exposed. The psychological literature on input categorization might be highly relevant for the study of L2 learning. Learners’ categorization of the input data might take place (a) by computing similarities to previously encountered exemplars, (b) by computing similarities to prototypes, (c) by computing the frequency of relevant features, (d) by applying rules, or (e) by a combination of these mechanisms (see Ashby & Maddox, 1998, for an overview; and Zaki, Nosofsky, Stanton, & Cohen, 2003, for a recent contribution to an ongoing debate).

In conclusion, a systematic investigation into the ways in which the three factors mentioned in this section might interact under explicit and implicit learning modes should be part of our research agenda, which calls for collaboration between linguists and psychologists.
ATTENTION

The three factors mentioned in the previous section also play a role, I believe, in discussions of whether implicit learning requires attention, consciousness, and awareness (DeKeyser, 2003; N. Ellis 1994, this issue; Schmidt, 1994a). In the field of SLA, it has now become customary to accept the noticing hypothesis of Schmidt (2001), which claims that at least some attention to and awareness of elements of the surface structure of utterances in the input is necessary for learning to take place. There is currently a debate in cognitive psychology over the question of whether implicit learning runs independently of attention or might rely on the same type of attentional mediation that is often considered to govern explicit learning processes. In the introduction to a volume dedicated to the controversial role of attention in implicit learning, Jiménez (2003) suggested that the analysis of the effects of attention on implicit and explicit learning could “tell us something very important about which kind of regularities our cognitive systems are prepared to capture immediately, and which other contingencies can be grasped exclusively by relying on a series of strategic, resource-demanding, and conscious recoding operations” (pp. 6–7). Hence, the question of attention still stands as an essential topic. However, its investigation must take into account the factors mentioned in the previous section.

SYMBOLIC AND SUBSYMBOLIC ACCOUNTS OF REGULARITY AND IRREGULARITY IN LANGUAGE

Factor one, which concerns regularity and complexity in form-meaning relationships of the input data, perhaps contributes the most to the lack of consensus in the literature on implicit and explicit learning. Natural languages are particularly intriguing insofar as they are multifaceted phenomena that defy simple definitions. A machine language could consist of (a) a set of well-defined form-meaning units (lexical items) and (b) a set of syntactic rules, with which grammatically well-formed and semantically unambiguous strings of units can be formed. Natural languages, however, are characterized by the absence of a one-to-one relationship between form and meaning, from the morpheme level all the way up to the text level. Both the lexicon and the grammar of natural languages contain, on the one hand, too many irregular form-meaning phenomena to allow a comprehensive characterization by means of rules operating on categories and, on the other hand, too many regular form-meaning phenomena to represent them simply as a large unstructured set of items. Thus, on the one hand, grammars can be seen as governed by abstract principles with great generality (e.g., the structure-dependency principle, the projection principle, and the subjacency principle in generative linguistics), but, on the other hand, grammar rules without exceptions hardly exist (Givón, 1999). Theories of the representation of linguistic knowledge must reflect this
competition between regularity and irregularity in one way or another. Theories of the symbolist school attempt to do this with rules that operate on categories (the grammar) and an inventory of categories and category members (items). Theories of the subsymbolic school represent knowledge in the form of small, meaningless units interconnected in an enormous network that is in a state of permanent flux, the activation and inhibition of internode connections resulting from verbal communication (Daelemans & de Smedt, 1996; Hulstijn, 2002; see also Williams, this issue).

It is useful to keep the symbolic and subsymbolic approaches to knowledge representation in mind when reading the literature on implicit and explicit language learning. Of course, when starting to think about implicit and explicit learning, one would begin, as did Schmidt (1994a), to distinguish between the product and the process of learning. One could then distinguish between implicit and explicit knowledge and between implicit and explicit learning and then argue that—in principle—these are orthogonal dimensions. However, in the view that linguistic cognition is at least partially subsymbolic, for an individual to have explicit knowledge of the architecture of that subsymbolic portion of his or her cognition would amount to being a neurocognitive scientist having advanced skills in programming parallel-distributed processing systems. Explicit learning—defined as an intentional effort to uncover the rules of the system underlying the input data—must then be seen as a fruitless process that is bound to fail. On the other hand, in the view that linguistic cognition is symbolic (at least partially), it would make sense to argue that an individual can have either implicit or explicit knowledge of the rules and categories of the system and, hence, that the adoption of an explicit learning mode is, in principle, a viable option.4

When we look at this issue from a less principled but more practical perspective, a picture emerges characteristic of L2 pedagogical grammars, L2 learners, and L2 classrooms all over the world: Although natural languages can only insufficiently be described with categorical rules, teachers and textbooks provide learners with categorical rules (“rules of thumb,” as Krashen (1981, p. 114) called them). L2 learners, to the extent that they know these rules, can be said to have explicit, metalinguistic knowledge of the L2, which might coexist with implicit knowledge (Hulstijn, 2002; Macnamara, 1973).

In conclusion, given that there are different views regarding the type of architecture to best represent linguistic knowledge, we should not be surprised to see definitions of implicit and explicit learning (the process) that are influenced by views on knowledge representation (the object or product of learning).

INDIVIDUAL DIFFERENCES IN IMPLICIT AND EXPLICIT LEARNING

Robinson (this issue) examines the influence of individual differences in language learning aptitude, intelligence, and working memory on rule and instance
learning for learners in what he calls incidental and implicit conditions, finding areas of both similarity and difference. Robinson concluded, on the basis of his results, that individual differences are influential on both instance-based and rule-based implicit and incidental learning. Furthermore, this study suggested that these two learning conditions and input stimuli draw on related but separable learning processes. Investigations like the one conducted by Robinson have the potential to substantially deepen our understanding of L1 and L2 learning. Most of the SLA literature has treated individual differences as additional, mediating variables rather than as intrinsically associated to the fundamental issue of the learnability of language. If language aptitude, intelligence, and working memory can be conceptually related to the constructs of implicit and explicit learning and knowledge, their status might change from peripheral and correlational to central and causal. However, to make this enterprise successful, the rather general notions of aptitude and intelligence will need to be broken down into their components. The role of differences in working memory in implicit and explicit learning is also addressed by N. Ellis (this issue). Furthermore, according to R. Ellis (personal communication, December 26, 2003), individual differences in learner orientation must be separated from developmental factors that influence learners’ ability to process specific information as implicit or explicit knowledge.

THEORY CONSTRUCTION AND DEVELOPMENT: TOP-DOWN AND BOTTOM-UP

The titles and abstracts of the papers in this thematic issue speak for themselves. There is no need to summarize their contents here. However, it is appropriate to end this introduction with some remarks on potential pitfalls concerning theory construction and development in the years to come. There are lessons to be learned from history.

Almost 20 years ago, Krashen (1981) formulated his Monitor Theory of SLA, centered around two central constructs: acquisition and learning. One of the criticisms leveled against Monitor Theory was that it failed to provide precise definitions of its two main constructs (acquisition and learning), thus preventing researchers from operationalizing them and putting the hypotheses of Monitor Theory to the test (Gregg, 1984; McLaughlin, 1978). In principle, this criticism was correct. However, as Jordan (2004) noticed (referring to Nicola, 1991), when we look at the history of science, it appears that many scientific theories, in their formative years, used poorly defined constructs in rather vague laws. Newton, when he launched his theory that explained why the Moon travels around the Earth, was not able to provide a precise definition of the proposed construct of gravity. Thus, initially, his theory violated the principle of falsifiability. Only in later years were Newton’s followers able to develop the construct of gravity conceptually and empirically. Unfortunately, Krashen’s constructs of acquisition and learning have fared less well. Since the
launch of Monitor Theory, which, at the time, had a high potential for success and fruitfulness, nobody has been able to give definitions of the notions of acquisition and learning that would render Monitor Theory testable (with the exception, perhaps, of Schwartz (1993) and Paradis (1994), although Paradis treated acquisition and learning as implicit and explicit knowledge, respectively). To avoid a similar fate in the case of the notions of implicit and explicit learning, with which we are concerned here, it is crucial that we make simultaneous progress on two fronts: theory development and empirical testing. If we continue to focus on the conceptual and speculative aspects of theory construction, neglecting measurement issues, theories of implicit and explicit L2 learning will not survive. For example, the potentially important issue of an interface between implicit and explicit knowledge (or, in Krashen’s terms, between acquired and learned knowledge), known in the literature as the strong interface, weak interface, and no interface positions (R. Ellis, 1993; Hulstijn & De Graaff, 1994; see N. Ellis, this issue, for a fresh look at this issue), might then become untestable (as suggested by Hulstijn, 2002). On the other hand, however, the history of science also shows that theories, and the definitions of their key constructs, are adapted as new empirical findings are being produced. In summary, progress in our understanding of SLA is best served by a cycle of top-down theoretical and bottom-up empirical work, avoiding not only overly strong forms of (exclusively concept-driven) rationalism but also overly strong forms of (exclusively data-driven) empiricism.

For the moment, it appears that we should first be concerned with the empirical side of implicit and explicit learning and knowledge. To reach this goal, the papers in this issue make an important contribution. R. Ellis presents concrete proposals of how to operationalize implicit and explicit knowledge by means of various tests. We might disagree with his operationalizations or with his interpretation of the outcome of his factor analyses, but, regardless of future empirical or theoretical work, Ellis’s paper signals a crucial moment in rendering theories of implicit and explicit knowledge and learning testable. It is not unlikely that the SLA field will now enter a phase marked by questions of validity, reminiscent of the debate concerning the definition and testing of the notion of intelligence. We could soon witness discussions of construct-definition claims such as “Implicit knowledge of a L2 is what task X measures.” I would welcome such discussions as part and parcel of normal science.

The advent of technologies with which we can look into the brain (such as the measurement of event-related brain potentials, or ERPs) will add depth to the measurement issue, as we might now be able to compare the data obtained with behavioral measures (such as the tasks used in the contributions by R. Ellis, De Jong, Robinson, and Williams) with those elicited with neurophysiological measures, as in the contribution of Tokowitz and MacWhinney. The latter authors argue that comparing ERP and behavioral data might provide a sensitive method for disentangling implicit and explicit knowledge. At the same time, Tokowitz and MacWhinney’s study shows that there are difficulties to
overcome with respect to data elicitation techniques and statistical analysis in ERP studies comparing the processing of L1 and L2 stimuli. R. Ellis sees a need for cross-validating behavioral measures with neurophysiological measures. Meanwhile, his study attempted to explore to what extent it is possible to distinguish implicit from explicit knowledge on the basis of behavioral measures. Thus, the present papers by R. Ellis and Tokowitz and MacWhinney already provide empirical data for the debate alluded to in the previous paragraph.

The papers in this issue illustrate how the construction and development of a theory of implicit and explicit learning and knowledge that attempts to explain the differential success of L1 and L2 acquisition can move forward on both the conceptual and the empirical fronts, avoiding the pitfalls of extreme rationalism and extreme empiricism while producing ideas and data on which we can base further research.

NOTES

1. Ullman (2001, 2004) did not speak of explicit and implicit memory or knowledge but of declarative and procedural memory/knowledge. The declarative system, depending on medial temporal lobe structures including the hippocampal region, subserves knowledge of facts and events including word-specific knowledge (the mental lexicon). The procedural system, depending on "frontal/basal-ganglia circuits, with a likely role for portions of parietal cortex, superior temporal cortex, and the cerebellum" (Ullman, 2004, p. 238), subserves procedural memory, supporting the learning and execution of motor and cognitive skills, especially those involving sequences, including the rule-governed combination of lexical items into complex representations (the mental grammar). An essential feature of Ullman’s declarative/procedural model is its specific claims concerning the lexical and grammatical types of linguistic knowledge being subserved by the declarative and procedural systems, respectively.

2. The definitions of explicit and implicit instructional treatment, given by Norris and Ortega (2000), actually comprise, respectively, a combination of what is here teased apart as implicit instruction and implicit learning, on the one hand, and explicit instruction and explicit learning on the other.

3. The debate in the psychological literature concerning the question of whether implicit memory does or does not result from a process of implicit learning is plagued, in my opinion, by the same confusion. In Reber’s implicit learning experiments (Reber & Allen, 2000; Reber et al., 1991, 1999), the stimuli could be represented by noncompeting, categorical rules representing a miniature artificial grammar (as in task A). This was not the case, however, in the implicit-memory experiments of Jacoby (1983) and Schacter (1987). These experiments (as in task B) were concerned with the recall or recognition of individual words or pictures, whose interrelationship (if there was any) could not be represented with categorical, syntactic rules.

4. For a discussion of the question of how regularities and irregularities are distributed in the syntax and lexicon of a mental grammar, see, for instance, Jackendoff (2002, p. 57 and Chapter 6). For some fascinating speculations as to how differential representations of symbolic and subsymbolic knowledge might interface in early and later interlanguages, see Towell (2003).

REFERENCES


