Language Learning 54:2, June 2004, pp. 227–275

The Definition and Measurement of L2 Explicit Knowledge

Rod Ellis
University of Auckland, New Zealand

A number of theories of second language (L2) acquisition acknowledge a role for explicit L2 knowledge. However, the testing of these theories remains problematic because of the lack of a widely accepted means for measuring L2 explicit knowledge. This article seeks to address this lacuna by examining L2 explicit knowledge from two perspectives. First, it considers explicit knowledge as a construct. How can explicit knowledge be defined? How does it differ from other constructs such as *L2 proficiency* and *language aptitude*? Second, the article considers how L2 explicit knowledge can be measured. It critically reviews some of the ways in which explicit knowledge has been operationalized in second language acquisition research and discusses some of the instruments that have been used to measure L2 explicit knowledge. It concludes with some guidelines for investigating explicit knowledge as analyzed knowledge and as metalanguage.

I would like to thank Catherine Elder and Alan Davies for their helpful comments on an earlier version of this article. I also appreciated the constructive reports from four anonymous reviewers.

Correspondence concerning this article should be addressed to Rod Ellis, Department of Applied Language Studies and Linguistics, University of Auckland, Private Bag 92019, Auckland, New Zealand. Internet: r.ellis@auckland.ac.nz

The purpose of this article is to examine explicit knowledge of a second language (L2) from two perspectives. First, I want to consider L2 explicit knowledge as a *construct*. How do we define L2 explicit knowledge? How does it differ from other constructs such as *L2 proficiency* and *language aptitude*? Second, I shall review some of the ways in which explicit knowledge has been operationalized for study (i.e., how it is measured) and, on the basis of this, advance some guidelines for the development of measuring instruments. However, the article does not seek to provide a comprehensive review of theory and research addressing L2 explicit knowledge; rather it is concerned with advancing a number of grounded proposals for measuring this construct.

The need to define and measure explicit knowledge arises from the importance attached to this type of knowledge in some theories of L2 acquisition (e.g., Bialystok, 1994; R. Ellis, 1994a; Hulstijn, 2002; Krashen, 1981). These theories make contradictory claims about the role of explicit knowledge in L2 acquisition. Thus, for example, where Ellis sees explicit knowledge as contributing indirectly to the acquisition of implicit knowledge by facilitating attention to form in the input, Krashen sees it as playing a role only in L2 production through monitoring. Other researchers (e.g., DeKeyser, 1998; Sharwood-Smith, 1981) argue for a direct interface between explicit and implicit knowledge. Yet other researchers (e.g., Cummins, 1983) see explicit knowledge as a component of L2 proficiency: It is needed to engage effectively in context-free language use. One of the difficulties in testing these various claims is the absence of an agreed-upon measure of explicit knowledge. Thus, although this article is not directly concerned with the controversies themselves, it will seek to make it possible to address them more effectively than has been the case to date by proposing instruments and procedures for assessing learners' explicit knowledge.

Defining Explicit Knowledge

The psychological and second language acquisition (SLA) literature is full of terms such as *language awareness, metalinguistic*

phenomena / awareness / abilities / performance, analyzed knowledge, conscious knowledge, declarative knowledge / rules / memory, learned knowledge and *explicit knowledge*. These terms overlap in ways that are not clear, and any attempt to tease apart the subtle differences in reference is beyond the scope of this article. Clearly, though, there is a need to establish a term or terms that can be used in some consistent fashion to refer to this area of mental representation. I shall use the term *explicit knowledge* to refer to that knowledge of language about which users are consciously aware. A working definition of explicit knowledge is as follows:

> the conscious awareness of what a language or language in general consists of and/or of the roles that it plays in human life.

To put it more simply, explicit knowledge is knowledge *about* language and *about* the uses to which language can be put. The problem is, as Birdsong (1989) points out, that explicit knowledge is manifest in a "complex and contingent set of behaviors that defy simplistic assumptions and explanations" (p. 49). In the discussion that follows, an attempt will be made to penetrate this complexity, first by considering what explicit knowledge is *not* and then by specifying what it *is*.

First, explicit knowledge is not an "attitude." Gombert (1992) defines *metalinguistics* as "a reflexive attitude with regard to language objects and their manipulation" (p. 1). Explicit knowledge can be viewed as the *outcome* of such an attitude but is distinct from it. Second, it is not a "practice" or an "activity." Valtin (1979) defines *metalinguistic phenomena* as "all the practices and activities concerning language and language judgment which are not themselves a part (or very closely tied to) production and comprehension processes." What a person knows explicitly is distinct from the actual uses to which he or she puts this knowledge: We can distinguish a tool as an object from our use of that tool. Unfortunately, this has not always been done. Paradis (1994), for example, defines explicit

knowledge as knowledge that individuals "are capable of representing to themselves and of verbalising on demand" (p. 394). Such a definition confuses the ability to verbalize knowledge with the knowledge itself. Of course, because explicit knowledge is a mental phenomenon and because we cannot directly access such a phenomenon, the only way we have of examining it is through activities that involve its application. Indeed, it is precisely the relationship between explicit knowledge as a mental phenomenon and its manifestation in human activity that this article seeks to examine. Nevertheless, we need to maintain a theoretical distinction between explicit knowledge itself and the use of this knowledge. Third, explicit knowledge is not a pedagogic construct (cf. the use of the term *language awareness* to refer to the British educational movement directed at increasing students' understanding of language and how it works), although pedagogic constructs may be derived from a theory of explicit knowledge and its role in language acquisition (R. Ellis, 1993, 1994a).

In short, I am proposing that explicit knowledge of language needs to be viewed as part of declarative memory, which is distinct from, although certainly related to, the processes through which this memory is constructed (i.e., learning) and from the uses to which it is put (i.e., performance). In the subsequent discussion I will pursue the definition of explicit knowledge first by examining how it differs from implicit knowledge and then by identifying a number of key characteristics of explicit knowledge. I will conclude this section by addressing the content of learners' explicit knowledge (i.e., what learners develop explicit knowledge of).

Implicit and Explicit Knowledge Compared

Explicit knowledge cannot be defined without reference to implicit knowledge. Thus, the key question becomes: To what extent and in what ways are explicit and implicit linguistic knowledge distinct? Claims regarding the distinctiveness of

explicit and implicit knowledge have been advanced by developmental psycholinguists, by cognitive psychologists, and by SLA theorists.

Developmental psycholinguists have shown that children, irrespective of the language they speak, develop a metalinguistic understanding of language during middle childhood (i.e., after basic linguistic competence has been acquired). Tunman and Herriman (1984), for example, propose that at a certain point in their development (5 years old or later), children cease using their language to just produce, comprehend, edit, or repair utterances and begin to treat it as an object of thought. They demonstrate awareness about language. Birdsong (1989) reviews studies that show how children vary considerably in their metalinguistic awareness, depending in particular on their literacy skills, whereas they do not differ significantly in the acquisition of basic linguistic competence (i.e., their implicit knowledge). This suggests that the kind of knowledge that involves metalingual awareness is distinct from the kind of knowledge that underlies everyday language use. As Tunman and Herriman point out, the crucial difference lies in the kind of processes involved. In the case of normal language use, production and comprehension processes require little or no attention and are executed very rapidly. In the case of operations involving explicit knowledge, conscious control needs to be exerted. In short, whereas implicit knowledge is available for automatic use, explicit knowledge typically involves controlled processes, although, as we shall see, in the opinion of some (e.g., DeKeyser, 2003), it may be possible to proceduralize explicit knowledge to the point that it cannot be easily distinguished from implicit knowledge.

Just as developmental psycholinguists view explicit knowledge as ontogenetically later than implicit knowledge, so cognitive psychologists such as Reber (1989) view it as phylogenetically later. Reber comments that "consciousness evolutionarily speaking is a late arrival on the mental scene" (p. 230). He argues that "primitive unconscious processes" of

the kinds found in animals enable individuals to learn about the world in very basic ways. They are, he says, "automatic and ineluctable." They contrast with "sophisticated unconscious" knowledge, which is abstract and automatized but which is available to consciousness through reflection. Explicit knowledge derives from reflection on this knowledge but is disassociated from it. Its function is as an executive system that exerts, in some degree, control over the operation of the cognitive unconscious.

The separateness of explicit and implicit knowledge is forcibly argued by Paradis (1994). Drawing on the work of Cohen (1991), Paradis postulates that the two types of knowledge reside in neuroanatomically distinct systems. Explicit memory is stored diffusely over large areas of the tertiary cortex and involves the limbic system; implicit memory is "linked to the cortical processors through which it was acquired" (p. 397) and does not involve the limbic system. The two memory systems are also susceptible to selective impairment. Paradis cites evidence to suggest that bilinguals who have learned their L2 formally (and therefore can be assumed to possess substantial explicit knowledge) may lose the ability to use their first language (L1) in the case of aphasia while maintaining the ability to speak haltingly in the L2. Paradis recognizes that the two systems can interact but disputes that that there is transfer of knowledge from one system to the other. He argues that practice cannot convert explicit knowledge into implicit, because the two types of representation are separate.

This latter claim is controversial, however. Skill-building theories, such as that proposed by Anderson (1983), allow for declarative knowledge to evolve into procedural knowledge through practice (see also DeKeyser, 1998, 2003), with intervening stages of explicitness. Also, in a carefully argued discussion of implicit and explicit knowledge, Dienes and Perner (1999) claim that the distinction represents a continuum rather than a dichotomy, a position they see supported by Karmiloff-Smith's (1992) account of how implicit linguistic knowledge becomes progressively more explicit in children. They propose that the degree of explicitness of a proposition depends on the extent to which a number of

elements are present. If their proposal is applied to statements about language (they illustrate it in terms of general propositions of the kind *This is a cat*), a fully explicit statement of a grammatical rule for the relative pronoun *that* might take the form of:

> I know that in the grammar of English *that* is a relative pronoun that can refer to both animate and inanimate nouns.

Such a statement includes all of the key elements that Dienes and Perner claim must be present for a statement to be fully explicit: It identifies a property (the word *that*), it specifies a predication of the property ("a relative pronoun that can refer to both animate and inanimate nouns"), it locates the property in the grammar of English, and it includes a marker of factuality ("I know"). It is easy to see how less explicit (more implicit) versions of this proposition could be formulated:

> I used the word *that*.
> *That* is a relative pronoun.
> *That* is a relative pronoun in English.

Such a view of explicit knowledge relates to performance (i.e., the extent to which a statement is explicit), but, as we noted earlier, explicit knowledge needs to be distinguished from the uses to which it is put. Thus the theory fails to demonstrate that explicit/implicit linguistic memory itself constitutes a continuum. Although we will certainly need to acknowledge that learners' uses of explicit knowledge may reflect different degrees of consciousness and also that the extent to which learners are able to articulate their explicit knowledge will be variable (i.e., more or less explicit), Dienes and Perner's arguments do not obligate us to accept that the mental representations of implicit/explicit knowledge are themselves continuous rather than dichotomous.

In SLA, Krashen (1981) has also argued for the separateness of implicit ("acquired") and explicit ("learned") knowledge, suggesting that explicit knowledge is only available to monitor the output derived from implicit knowledge. Like Paradis, he adopts a

noninterface position; there is no direct transfer of knowledge from explicit to implicit. Krashen's position has come under considerable attack, but he has continued to maintain it (see, for example, Krashen, 1994) and has received support from Schwartz (1993). Schmidt (1994) has disputed Krashen's view that the acquisition of implicit knowledge is an entirely unconscious process but he endorses the implicit/explicit distinction at the level of *knowledge*. Schmidt makes the point that "implicit and explicit *learning* and implicit and explicit *knowledge* are related but distinct concepts that need to be separated" (p. 20). Schmidt appears to view implicit/explicit knowledge as continuous rather than dichotomous but does not elaborate on this point.

My own view is that we would do better to view the two types of knowledge as dichotomous. Adopting a connectionist account of implicit linguistic knowledge as an elaborate interconnected network (e.g., N. Ellis, 1994a), it is not easy to see how knowledge as weighted content (i.e., as a set of neural pathways of greater and lesser strength) can be anything other than separate from knowledge of linguistic facts. Thus, where representation is concerned I would argue, along with Paradis, Krashen, and Ellis and contrary to Dienes and Perner and Schmidt, for the separateness of the two types of knowledge. Again, though, it must be noted that in arguing for such a position, I do not wish to imply that implicit knowledge cannot be rendered explicit or vice versa. Also, as Bialystok (1982) has argued, in performing different tasks, learners are likely to draw differentially on both knowledge sources, a point that has obvious implications for the measurement of learners' linguistic knowledge.

The following picture emerges from the above discussion. It should be noted, however, that all the points constitute assumptions, supported by the literature to a greater or lesser extent but in need of further empirical validation.

1. Explicit knowledge appears phylogenetically and ontogenetically later than implicit knowledge, and it involves different access mechanisms.

2. Explicit knowledge is neurologically distinct from implicit knowledge.

3. The question of whether the two types of knowledge are to be seen as dichotomous or continuous is a matter of controversy, but neurological evidence and current connectionist models of linguistic knowledge point to a dichotomy.

4. The question of the separateness of the representation of the two types of knowledge is independent from the question of whether the processes of implicit and explicit learning are similar or different. This remains a controversial issue. It is likely, however, that learning processes and knowledge types are correlated to some degree, at least.

5. Although there is controversy regarding the interface of explicit and implicit knowledge at the level of learning, there is wide acceptance that they interact at the level of performance.

We will now turn to a more detailed examination of what is meant by explicit L2 knowledge. To address this question it is necessary to consider two separate questions: (a) What are the key characteristics of L2 explicit knowledge? and (b) What is the content of L2 learners' explicit knowledge?

Key Characteristics of Explicit Knowledge

1. *Explicit knowledge is conscious*.[1] That is, in contrast to implicit knowledge, which is entirely tacit, learners know what they know; they are consciously aware of some aspect or feature of the L2. *Conscious awareness*, however, must be distinguished from *intuitive awareness*. Karmiloff-Smith (1979) has noted with reference to the study of child language development that a distinction needs to be made between *metalinguistic* data (e.g., What is a word? Is *the* a word? How many words are there in this sentence? Is it funny to say *x*?) and what may be termed *epilinguistic* data

(e.g., tapping the child's awareness of the implicit grammatical rules he is using such as gender concord, use of one article in preference to another, anaphoric reference). Intuitive awareness or what Karmiloff-Smith calls "epilinguistic behaviour" is evident in the ability to recognize instantly that a sentence is ungrammatical. Conscious awareness or what Karmiloff-Smith calls "metalinguistic behaviour" is evident when learners can cognize why a sentence is ungrammatical. For example, faced with a sentence like

*The policeman explained Wong the law.

a learner may know intuitively that there is something ungrammatical and may even be able to identify the part of the sentence in which the error occurs but may have no conscious awareness of the rule that is being broken. Such a learner has implicit but no explicit knowledge of the feature, dative alternation, in question. Another learner, however, may understand that the sentence is ungrammatical because the verb *explain* cannot be followed by an indirect object without *to*.

2. *Explicit knowledge is declarative*. That is, it is comprised of facts about the L2. These facts concern both rule-based knowledge and knowledge of fragments and exemplars (Eichenbaum, 1997). Explicit knowledge of an L2 is no different from encyclopedic knowledge of any other kind. I know, declaratively, that the Normans invaded England in 1066. Similarly, I know that verbs like *explain* require an indirect object with *to* and, further, that the indirect object usually follows the direct object. I also know that *How do you do*? serves as a formulaic chunk for greeting someone. These facts are only loosely connected; they do not constitute a "system" in the same way that the implicit knowledge of proficient L2 users does. To a considerable degree, declarative facts about the L2 may be stored separately and can be easily accessed as separate units of information.

3. *L2 learners' declarative rules are often imprecise and inaccurate*. For example, a learner who responded to the ungrammatical sentence above with the comment "You can't use a proper noun after *explain*" clearly has some explicit understanding of what makes the sentence ungrammatical but equally clearly does not have a very accurate notion. Also, as Clapham (2001) shows, learners may be able to recognize a specific metalingual term in one sentence or one language but not in another.

4. *The development of a learner's explicit knowledge can take place on two planes*. It can grow in breadth as the learner accumulates more declarative facts about the language. It can also advance in terms of depth as the learner refines existing explicit knowledge, making it more precise and accurate and applying it more consistently across different contexts and languages. For example, a "deep" understanding of what makes the above sentence ungrammatical might include an awareness that Latinate verbs like *explain* require a different pattern of dative alternation from that of Anglo-Saxon verbs like *give* and that there are various semantic restrictions on dative alternation (Pinker, 1989). A number of studies (e.g., Butler, 2002; Green & Hecht, 1992; Sorace, 1985) have reported a relationship between the quality of learners' explicit knowledge and their overall proficiency. This relationship, however, cannot be interpreted as demonstrating that explicit knowledge promotes the development of implicit knowledge, as it is equally possible that explicit knowledge develops as a by-product of extending implicit knowledge, as suggested by Bialystok (1994). Bialystok, like a number of other researchers, argues that "language that is explicit does not become implicit" (p. 567).

5. *Explicit knowledge is generally accessible through controlled processing*. This contrasts with the automatic processing that characterizes the use of implicit knowledge.

One of the widely commented on uses of explicit knowledge is to edit or monitor production, a process that is possible only in those types of language use that allow learners sufficient time to access the relevant declarative facts. For this reason, explicit knowledge may not be readily available in spontaneous language use where there is little opportunity for careful on-line planning. R. Ellis and Yuan (in press), for example, demonstrated that L2 learners' grammatical accuracy was significantly greater in both oral and written narrative tasks if they were given time to plan on-line than if they were pressured to produce rapidly. One explanation of this finding is that the learners were able to monitor their productions in the careful on-line planning condition using their explicit knowledge.

It is possible, however, that some learners are able to proceduralize their explicit knowledge and thus access it for rapid on-line processing in much the same way as they access implicit knowledge. DeKeyser (2003) suggests that proceduralized explicit knowledge can be considered "functionally equivalent" to implicit knowledge. Hulstijn (2002), however, is doubtful, arguing that although practice "may speed up the execution of algorithmic rules to some extent" (p. 211), it is necessary to distinguish the automatization of implicit and explicit knowledge, and that what appears to be the automatization of explicit knowledge through practice may in fact entail the separate development of implicit knowledge. N. Ellis (1994a) suggests how this might come about; he proposes that sequences produced through the application of declarative rules can come to be performed automatically if the sequences are sufficiently practiced. That is, it is not the rules themselves that become implicit, but rather the sequences of language that the rules are used to construct. This is clearly a crucial issue for the measurement of the two types of knowledge and will be considered again later.

6. *Any language task that a learner finds difficult may naturally result in an attempt to exploit explicit knowledge*. In terms of sociocultural theory (Lantolf, 2000), explicit knowledge might be viewed as a tool that learners use to achieve self-control in linguistically demanding situations. In accordance with this theory, explicit knowledge might manifest itself through the private speech that learners use to grapple with a problem. In support of such a contention, it can be noted that when learners are asked to make and justify grammaticality judgments in a think-aloud or dyadic problem-solving task, they typically try to access declarative information to help them do so if they feel unable or lacking in sufficient confidence to make a judgment intuitively (R. Ellis, 1991a; Goss, Ying-Hua, & Lantolf, 1994).

7. *Explicit knowledge is potentially verbalizable*. Earlier I noted that explicit knowledge exists independently of whether it can be verbalized. Nevertheless, because it is declarative in nature, it is stable. Butler (2002), for example, found that the learners she studied (adult Japanese learners of English) were generally able to provide some kind of explanation for their choice of articles in a cloze task, although they had difficulty in giving specific reasons. Dienes and Perner's (1999) theory of implicit/explicit knowledge may best be seen as an account of the degree of explicitness of verbalizations of explicit knowledge.

It is important to recognize, however, that verbalizing a rule or feature need not entail the use of metalanguage. As James and Garrett (1992) point out, language can be talked about in a "standard received language" or a "non-technical" one. Thus, the error in the sentence above might be explained nontechnically by saying, "You can't say, 'explain Wong.' You've got to say, 'to Wong' after 'explain.'" Alternatively, the explanation might call on extensive metalanguage; for example, "In the case of dative

alternation, there are some verbs like *explain* that require the indirect object to be realized as a prepositional phrase rather than as a noun phrase." Although metalanguage is not an essential component of explicit knowledge, it would seem to be closely related. It is possible that an increase in the depth of explicit knowledge will occur hand in hand with the acquisition of more metalanguage, if only because access to linguistic labels may help sharpen understanding of linguistic constructs.

8. *Explicit knowledge is learnable*. Implicit knowledge is also learnable, but there would appear to be relatively universal constraints on the ability of adult learners to fully learn an L2 implicitly given that only a very few learners achieve native-speaker proficiency. In contrast, as Bialystok (1994) points out, "explicit knowledge can be learned at any age" (p. 566). The constraints that exist on learners' ability to learn explicit facts about a language are of a different order, probably relating to individual differences in the analytical skills needed to memorize, induce, or deduce those facts. However, with careful instruction it may be possible to teach many learners a very substantial amount of declarative information about a language, although this is also controversial. Krashen (1982), for example, argues that, in general, only simple and portable rules, such as English third-person *–s*, can be readily learned as explicit knowledge.[2]

The Content of a Learner's Explicit Knowledge

What kind of knowledge do learners develop a declarative understanding of? It is important to make an initial distinction between knowledge of *language* and knowledge of *a language*. As James and Garrett (1992) have noted, the study of language awareness can be directed at language in a general sense—the domain of linguistics—or at a specific language. With the

exception of linguists, however, it is not to be expected that speakers/learners of a language will develop awareness of the universal features of language (what in the generative school of linguistics is referred to as Universal Grammar). They will be limited to language-specific details. My concern here is only with explicit knowledge of language-specific details.

Learners can potentially gain an explicit knowledge of any aspect of a language and of its use. It is important to emphasize this, because, in general, explicit knowledge has been discussed primarily in relation to grammar. I will briefly consider some of the aspects of a language that have figured in studies of explicit knowledge.

Pronunciation. Developmental psychologists such as Nesdale, Herriman, and Tunmer (1984) and Gombert (1992) have paid considerable attention to children's awareness of phonological features. The features they have studied include the difference between linguistic and nonlinguistic sounds, phonemic segmentation, and syllable structure. In contrast, L2 researchers have paid scant attention to learners' explicit knowledge of pronunciation. This is probably because it is assumed that conscious awareness of aspects of language such as phonemic segmentation arises in the process of developing L1 literacy and transfers automatically to the L2. However, there are some aspects of pronunciation that individuals may not become very aware of during the course of L1 development or that may not readily transfer to the L2. Newman and White (1999) found that native speakers of New Zealand English were not always able to identify the syllable that carried the main stress in a word and, in this respect, performed less well than Japanese learners of English. In contrast, they found that the Japanese learners were much less successful in counting the syllables in English words like *kites* and *caravan* than the native speakers. It would seem, therefore, that L2 learners' explicit understanding of at least some aspects of pronunciation warrants attention. These aspects might include syllable structure and, perhaps, contrasts between the learners' L1 and the L2 sound systems.

Vocabulary. There is a wide consensus among cognitive psychologists (see, for example, N. Ellis, 1994b) that whereas knowledge of the *form* of a word and of word collocations is largely implicit (e.g., we know "by feel" whether a word is pronounced or spelled correctly and which word goes with which), knowledge of the *meaning* of a word is explicit (e.g., we are consciously aware what a word means and can give a definition of it). Word meanings, then, probably constitute the largest single area in a learner's explicit knowledge. In addition, of course, individuals may develop a conscious awareness of the form of at least some words—problematic spellings or the differences between American and British spellings, for example—and of some collocations (e.g., idioms that they have intentionally learned).

Developmental psychologists have investigated children's conscious awareness of a number of aspects of vocabulary: words as signifiers (i.e., differentiating between a word and its referent), lexical segmentation (i.e., the ability to segment a sentence into its constituent words), and the metalinguistic term *word*. Again, these aspects of explicit knowledge have not attracted the attention of L2 researchers, probably because it is assumed that such knowledge transfers directly to the L2. In fact, with the obvious exception of word meanings, explicit knowledge of vocabulary appears to have been entirely neglected in L2 research.

Grammar. Discussion and studies of explicit knowledge have largely focused on grammar. This reflects the centrality of grammar in such fields as linguistics and language teaching and also, perhaps, the fact that grammar, in contrast to pronunciation and vocabulary, is more amenable to conscious reflection and manipulation (Odlin, 1989). There is also, of course, a rich and well-established metalanguage for talking about grammar.

The term *grammar* covers a lot of ground. As we have already noted, it is important to distinguish knowledge as awareness of grammatical rules/features (i.e., analyzed knowledge) from knowledge of the metalanguage for talking about

these. It is also useful to distinguish knowledge of forms that constitute *items* (e.g., consciously knowing that the preposition *in* is used with the verb *interest*) and morphological and syntactical *rules* (e.g., consciously knowing that one-syllable adjectives form the comparative by adding *–er*).

Studies of L2 metalanguage (e.g., Alderson, Clapham, & Steel, 1997; Butler, 2002) have focused on such areas of grammar as parts of speech, the use of English articles, grammatical functions (i.e., subject, object, predicate, etc.), and rules that figure in standard pedagogic grammars of the L2 (e.g., use *avoir* rather than *etre* to form the passé composé of *aller*). These studies have attempted to examine both learners' conscious awareness of grammatical forms and their knowledge of metalanguage. It is interesting to note that, to the best of my knowledge, there has been no attempt to investigate learners' conscious understanding of the grammatical contrasts between their L1 and the L2 they are learning.

Pragmatic features. Just as learners can have an explicit knowledge of linguistic aspects of the L2, so too they can develop a conscious understanding of pragmatic aspects. Wolfson (1983) makes the point forcibly that pragmatic rules generally lie below the level of conscious awareness. However, some parents make explicit attempts to teach children a number of standard formulas for expressive speech acts (Clark & Clark, 1977), and L2 learners can also develop explicit knowledge of these features through instruction or through reflection.

Developmental psychologists have shown an interest in this area of explicit knowledge. Aspects that have been investigated include the explicitness of a verbal message (i.e., whether a message contains sufficient information for its comprehension by an addressee), sociolinguistic appropriateness (i.e., the relationship between linguistic choice and addressee), illocutionary acts (e.g., strategies for realizing requests), and linguistic humor (e.g., the manipulation of ambiguity for humorous effects). L2 researchers have not specifically set out to investigate explicit knowledge of L2 pragmatic features. However, many of the

instruments that have been used to investigate learners' knowledge of illocutionary acts, such as the discourse completion questionnaire (see Kasper & Dahl, 1991), are arguably more likely to tap explicit than implicit knowledge.

Sociocritical features. Critical discourse analysis (Fairclough, 1985) has as its goal the identification of social values that have become "naturalized" through use and thereby opaque to language users. In effect, then, its purpose is to make explicit the way in which language implicitly encodes social structures and political positions. The assumption is that individuals are typically unaware of how language works in these respects but can be made aware. Here, again, then we find the essential dichotomy of implicit and explicit knowledge of language. Sociocritical features include the marked use of linguistic forms (e.g., lexical items, pronouns, or pragmalinguistic formulas) to achieve a particular social purpose (e.g., persuade the addressee to buy a product), to establish a particular identity (e.g., to signal that someone is an outsider) and to enact and affirm a particular power relationship (e.g., to assert one's right to something on the basis of one's class, gender or ethnicity). To the best of my knowledge, however, there have been no studies of L2 learners' explicit understanding of the sociocritical uses of the L2. One reason for this might be the difficulty of operationalizing sociocritical awareness in learners for empirical study, as this depends on the particular ideological orientation of the individual.

Summary

In this section I have worked toward an extended definition of L2 explicit knowledge, which can now be summarized as follows:

> Explicit L2 knowledge is the declarative and often anomalous knowledge of the phonological, lexical, grammatical, pragmatic, and sociocritical features of an L2 together with the metalanguage for labeling this

> knowledge. It is held consciously and is learnable and verbalizable. It is typically accessed through controlled processing when L2 learners experience some kind of linguistic difficulty in the use of the L2. Learners vary in the breadth and depth of their L2 explicit knowledge.

In the following section I will examine how explicit knowledge has been investigated empirically before turning to consider some of the different ways of measuring it.

Studies of L2 Explicit Knowledge

My concern in this section is not with the results of studies that have investigated L2 explicit knowledge but with (a) what aspects of explicit knowledge they have examined and (b) how they have measured these aspects. As such, this section serves as a bridge between the foregoing definition of L2 explicit knowledge and the following section, which examines more schematically different ways of measuring it. It aims to provide the reader with an account of how researchers to date have operationalized the construct.

Table 1 summarizes the studies. It should be noted that this table includes only studies that expressly set out to investigate explicit knowledge (i.e., it does not include studies that set out to investigate L2 learning in general using instruments that may have incidentally afforded measures of explicit knowledge). In fact, there have been relatively few studies that have had the study of L2 learners' explicit knowledge as their main purpose, whereas there have been many more studies that may have addressed this aspect in the context of studying L2 acquisition. For example, a very large number of SLA studies have made use of grammaticality judgment tasks (see R. Ellis, 1991a, for an early review of these) as a way of measuring L2 learners' knowledge. These will not be considered here unless they were designed specifically to measure explicit knowledge.

Table 1 also shows that the research to date has focused almost exclusively on explicit knowledge of L2 grammar. Only

Table 1

Survey of studies of L2 explicit knowledge

Study	Participants	Aspects Studied	Instruments
Bialystok (1979)	Three groups of English-speaking learners of French ($N = 185$).	Grammar; adjectives, pronouns, verbs.	Grammaticality judgment test; sentences read twice on a tape; three conditions: (a) judge whether a sentence is grammatical, (b) indicate the ungrammatical part of speech, and (c) indicate from a list of rules which rule has been violated. All items given in a timed and untimed format.
Sorace (1985)	17 English-speaking learners of Italian (9 beginners and 8 intermediate learners).	Grammar; six structures (perfect; imperfect; indirect pronouns with *piacere* and *dire*; auxiliary in compound tenses; past participle agreement).	Grammaticality judgment test; indicate errors, correct errors, state the grammatical rule.
Masny (1987)	74 francophone students in advanced courses in English as a second language.	Grammar; three syntactic categories (pronoun, relative clause, and concord).	Grammaticality judgment test; judge grammaticality of sentences, correct deviant sentences.

Green and Hecht (1992)	300 German-speaking learners of English from German secondary schools and a university with between 3 and 12 years of exposure to formal instruction. 50 native speakers.	Grammar; content based on errors learners had made in two communicative tasks and on grammar points taught as part of syllabus.	Grammaticality explanation test; (a) correct the sentences and (b) offer explanations of the rules that would enable someone else to correct the errors. Twelve sentences in the test.
Alderson et al. (1997)	599 university students of foreign languages in British universities.	English and French grammar; knowledge of metalanguage for parts of speech; a range of grammatical features such as adverbs, pronouns, verb tenses, articles, negatives, verb complements.	Metalinguistic Assessment Test for English and French; identification of named parts of speech; grammaticality judgment test requiring correction and rule statement.
Han and Ellis (1998)	48 adult learners enrolled in a university intensive English program; level was intermediate-plus.	Grammar; verb complements.	Grammaticality judgment test; unspeeded judgments. Oral rule statement in an interview used to construct a metalingual comments score.
Elder et al. (1999)	334 university students of foreign languages at the University of Melbourne.	Same as Alderson et al. (1997).	Same as Alderson et al. (1997).

Table 1 (continued)

Survey of studies of L2 explicit knowledge

Study	Participants	Aspects Studied	Instruments
Newman and White (1999)	87 1st-year linguistics students (61 New Zealand educated and 11 Japanese).	Pronunciation (syllable structure and stress placement); parts of speech; grammar rules.	Identifying the number of syllables in four words; indicating which syllable carries the main stress; identifying named part of speech; stating a rule of English grammar.
Clapham (2001)	Same as Alderson et al. (1997).	Same as Alderson et al. (1997).	Participants' ability to identify metalingual terms and to use them varied according to whether the the sentence was simple/complex and according to the language.
Hu (2002)	64 Chinese learners of English (upper-intermediate level).	Prototypical and peripheral rules for six English grammatical structures, for example, use of the simple present tense to refer to habitual events (prototypical) or to future events (peripheral).	Participants' verbal explanations of the rules; error correction task completed without time pressure; participants' judgment of the prototypicality of the rules.
Butler (2002)	80 mixed-proficiency Japanese learners of English; 20 native speakers.	English articles (a, the, zero).	Participants' verbalizations of their reasons for choosing article forms in a cloze task.

Newman and White (1999) investigated another area of language: pronunciation. This emphasis on grammar contrasts quite markedly with studies of L1 explicit knowledge in children, which have addressed pronunciation, vocabulary, and pragmatic aspects as well as grammar. All of the studies in Table 1 have attempted to investigate explicit knowledge as conscious awareness. Several studies (Alderson et al. 1997; Clapham, 2001; Elder, Warren, Hajek, Manwaring, & Davies, 1999; Han & Ellis, 1998; Newman & White, 1999) have addressed learners' knowledge of metalanguage. A number of studies have also examined learners' ability to verbalize rules explicitly (Butler, 2002; Green & Hecht, 1992; Hu, 2002). All the studies except Han and Ellis (1998) and Butler (2002), both of which investigated a single area of English grammar, investigated a range of grammatical features. In general, the features chosen were those for which there were readily available pedagogic rules, although Green and Hecht (1992) deliberately included features for which simple explanations were not readily available. Also, interestingly, Hu (2002) examined learners' knowledge of simple (prototypical) and difficult (peripheral) rules for the same grammatical forms. All the studies except Clapham and Butler investigated breadth of explicit knowledge. Clapham's study shows the importance of investigating depth (e.g., the extent to which participants can identify metalingual terms in a variety of sentence types and in different languages).

It is clear that the favored method of investigating L2 explicit knowledge as conscious awareness is the grammaticality judgment task. All the studies except Newman and White (1999) and Butler (2002) made use of some version of such an instrument. However, they varied somewhat in the specific operations required of learners. Most of the studies required learners to perform three operations: (a) identification of the ungrammatical sentences, (b) correction of the errors, and (c) provision of rules. Green and Hecht's (1992) study omitted (a), as all the sentences on their test were ungrammatical. Masny's (1987) study omitted (c). Several of the studies acknowledge the difficulty of deriving

a quantitative measure from (c), given that learners vary in the degree of precision and accuracy with which they are able to state a rule. Han and Ellis (1998) attempted to develop a scale for assessing learners' metalingual comments. Bialystok (1979) overcame the problem by testing learners' receptive knowledge of rules. Interestingly, none of the studies invited learners to report what kind of knowledge they were using to judge individual sentences. Nor did any of the studies attempt to measure the certainty with which learners made their judgments.

Another method for measuring explicit knowledge that figures in these studies is a test of learners' ability to identify named grammatical features. Such a test, used by Alderson et al. (1997), Elder et al. (1999), and Newman and White (1999), measures receptive knowledge of grammatical metalanguage. Newman and White (1999) used two tests that do not figure in the other studies: a test of the ability to state the number of syllables in words and to indicate which syllable in a word receives the main stress and a test requiring participants to state any explicit rule of grammar they can remember. Only one study (Alderson et al.) reports on the reliability of the instruments used. The alpha coefficient of reliability for the various tests used in this study ranged between an acceptable .73 and .86. The general failure to consider the reliability of the use of the testing instruments is considered further later in the article.

Ways of Measuring Explicit L2 Knowledge

In this section I will consider in greater depth three types of measure that figured in the studies of L2 explicit/metalingual knowledge listed in Table 1: language aptitude tests, grammaticality judgment tests (GJTs), and tests of metalanguage. In addition I will examine the use of verbal reports as a methodological device for arriving at measures of explicit knowledge.

Language Aptitude Tests

The ability to reflect on language and extract abstract information has been viewed as one aspect of *language aptitude*. For example, the widely used Words in Sentences Test from the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959) claims to measure language learners' ability "to understand the function of words and phrases in sentences" (p. 1). The test asks learners to find the word in a sentence that has the same function as a key word underlined in another sentence. Carroll (1965) refers to the abilities that this test is designed to tap as "grammatical sensitivity" and "inductive language learning ability." Pimsleur (1966) also saw "the ability to reason analytically about verbal materials" as intrinsic to what he called "verbal intelligence" (p. 8). Likewise, the Paired Associates Test of the MLAT can be seen as a test of explicit vocabulary learning. Thus, there are grounds for viewing language aptitude, or, at least, the components relating to grammatical analysis and vocabulary learning, as the essential ability underlying the development of explicit knowledge. In other words, measuring learners' ability to analyze sentences grammatically or memorize the meanings of words may provide an indication of the extent of their explicit knowledge.

What support is there for such a position? Krashen (1981) was among the first to suggest that language aptitude related only to explicit and not to implicit knowledge. He proposed that Carroll's inductive ability and grammatical sensitivity and Pimsleur's verbal intelligence relate "directly to, or reflect, conscious language learning, the Monitor" (p. 21). DeKeyser (2003) adopts a similar position. He argues that language aptitude is a factor only with older learners, who employ explicit learning processes. It does not figure in children, who rely on implicit learning processes. DeKeyser claims that the shift from implicit to explicit learning processes in older learners explains the differences that have been observed regarding rate and ultimate level of achievement between adults and children. Thus, for both

Krashen and DeKeyser, language aptitude as traditionally measured is de facto a measure of learners' ability to learn explicit knowledge. From this perspective, if one assumes that learners learn what they set out to learn (a big assumption), then the Words in Sentences and Paired Associates Tests of the MLAT might be seen as measures of explicit knowledge. That is, learners with the kind of aptitude needed to acquire explicit knowledge are assumed to have more such knowledge than learners lacking in this aptitude.

A number of studies have found that there is a relatively strong relationship between the Words in Sentences Test and measures of explicit knowledge. Robinson's (1995) study of the relationship between learning under different conditions (implicit, incidental, rule search, and instructed), measured by means of GJTs, and language aptitude suggests that the ability measured by the Words in Sentences Test may be more strongly associated with the kind of learning involving attention to form (explicit learning) rather than with meaning (implicit learning). Masny (1987) and Alderson et al. (1997) both report moderate statistically significant correlations between MLAT scores and grammaticality judgment scores.

Nevertheless, there are empirical grounds for disputing the view that language aptitude, including inductive grammar-learning ability, relates exclusively to explicit knowledge. Skehan argues that it is involved in the development of both explicit and implicit knowledge. Contrary to DeKeyser's claim that language aptitude is not a factor in child language acquisition, Skehan (1986, 1990) has shown that scores on tests measuring ability in grammatical analysis (administered when the participants in the study were teenagers) produced statistically significant correlations with measures of the participants' L1 acquisition based on their communicative speech when they were children. For example, the Hidden Words Test and the Matching Words Test of the Elementary Modern Language Aptitude Test both correlated significantly ($r = 0.32$ and 0.52, respectively) with mean length of utterance measured at 42 months.

Further, a number of studies (see R. Ellis, 1994b, for a survey of these) have found a clear relationship between language aptitude and measures of communicative language use, which can be hypothesized to reflect implicit knowledge. A sounder position, then, is that the ability to analyze sentences grammatically serves, to some extent at least, as a measure of explicit knowledge, but that it also measures learners' ability to learn implicitly and, therefore, their implicit knowledge. Adopting this position, Skehan (1998) argues that language aptitude is to be seen as partially isomorphous with and partially distinct from the general cognitive abilities that underpin explicit knowledge.

An alternative measure of the ability to acquire explicit knowledge might be a general intelligence test. Language aptitude tests may be measuring explicit language learning ability to the extent that they are also measuring general intelligence. Sasaki (1996) found a strong correlation between measures of intelligence and measures of language aptitude. However, the only study (Masny, 1987) that has examined the relationship between general intelligence and explicit L2 knowledge (measured by means of a GJT) failed to find such a relationship, although this study did report the expected relationship between MLAT and general intelligence.

To sum up, tests of language aptitude probably cannot provide valid measures of L2 explicit knowledge. First, they measure the ability to learn, which may or may not be correlated with learners' stored explicit knowledge. Second, there is evidence to suggest that language aptitude is involved in the development of implicit as well as explicit knowledge.

Grammaticality Judgment Tests

There is now a large literature on GJTs.[3] Since Chaudron's (1983) and R. Ellis's (1991a) review of GJT studies, there have been a number of additional studies (e.g., Bard, Robertson, & Sorace, 1996; Cowan & Hatasa, 1994; Davies & Kaplan, 1998; Gass, 1994; Goss, Ying-Hua, & Lantolf, 1994; Leow, 1996;

Mandell, 1999). These studies have been directed specifically at investigating the validity and reliability of GJTs.

The construct validity of GJTs as measures of explicit knowledge needs to be considered in relation to the specific tasks learners are asked to perform. As I noted earlier, learners have been asked to (a) identify the error in an ungrammatical sentence, (b) correct the error, and (c) state the grammatical rule that has been broken. They can also be asked to (d) indicate the degree of certainty of their judgment. In this section I will consider (a), (b), and (d), reserving (c) for the "Using Verbal Reports" section later in the article.

There are various ways in which (a) can be accomplished. Sorace (1996; Bard et al., 1996) argues the case for "magnitude estimation" (i.e., participants are asked to assign a number to the first sentence in the test to indicate its acceptability and then to assign numbers to the next sentence to indicate its acceptability in relation to the acceptability of the first sentence, and so on). Sorace argues that such a method takes account of the indeterminancy of interlanguage rules and also that it allows for the use of parametric statistics whereas other, traditional methods do not. In effect, magnitude estimation affords both a judgment and a measure of how certain a learner is about the judgment.

The key construct validity issue concerns what a GJT measures. What kind of knowledge do learners draw on when they judge the grammaticality of a sentence: explicit knowledge, implicit knowledge, or some kind of mixture of both? As Birdsong (1989) has noted, "metalinguistic data [from a GJT] are like 25-cent hot dogs: they contain meat, but a lot of other ingredients too" (p. 69). Surprisingly, this is an issue that many SLA researchers employing GJTs fail to address. Sorace (1996), however, explicitly acknowledges the problem:

> It can be a more complex task [than is the case with native-speaker judgments] to decide about the kind of

> norm consulted by learners in the process of producing a judgment, particularly in a learning environment that fosters the development of metalinguistic knowledge. It is difficult to tell whether subjects reveal what they think or what they think they should think. (p. 385)

What kind of knowledge a GJT measures may depend on whether the judgment is timed or untimed. Bialystok (1979) concludes from a study of learners' judgments that

> grammaticality decisions are made initially on an intuitive basis that may or may not be supportable by the subjects' knowledge of the relevant structures. Detailed information about errors, however, relies on information in explicit knowledge and hence requires time to produce. (p. 98)

In other words, it can be hypothesized that when learners are asked to judge the grammaticality of a sentence rapidly, they are more likely to rely on implicit knowledge, but if they are given time, they are able to gain controlled access to explicit knowledge. Sorace, too, proposes that a timed procedure is necessary to ensure that the test taps tacit rather than metalinguistic knowledge. Han and Ellis's (1998) study lends support to such a view. They found that measures derived from a timed and an untimed version of the same GJT factored out separately. In a principal-components analysis, the timed GJT loaded on the same factor as an oral production test, whereas the untimed GJT loaded on the same factor as a metalingual-comments score. Han and Ellis labeled these two factors *implicit* and *explicit L2 knowledge*, respectively. DeKeyser (2003), however, strikes a word of warning, noting that time pressure does not guarantee a measure of implicit knowledge. As I noted earlier, it is possible that some learners have developed relatively automatized explicit knowledge, which they can access even under time pressure. Also, it does not follow that learners will apply explicit knowledge if they have time to do so. They may still choose to rely on their implicit knowledge. Indeed, they may have to (or alternatively, to guess) if they do not possess the

necessary explicit knowledge to judge the grammaticality of a particular sentence. R. Ellis (1991a) and Goss et al. (1994) found that even when there was the opportunity to deliberate about a judgment, learners sometimes opted to respond immediately. At best, then, we can say that an immediate judgment is more likely to reflect implicit knowledge and a delayed judgment explicit knowledge.

There is also the question of how to operationalize timed and untimed judgments[4] (i.e., what length of time for making a judgment to give). A GJT potentially involves three principal processing operations:

1. semantic processing (i.e., understanding the meaning of a sentence)

2. noticing (i.e., searching to establish whether something is formally incorrect in the sentence)

3. reflecting (i.e., considering what is incorrect about the sentence and, possibly, why it is incorrect)

In the case of a timed test, the goal should be to allow (1) and (2) but not (3). In an untimed test, the goal should be to allow opportunity for all three processing operations to take place. Clearly, the time that learners need to perform just (1) or (2) or all three options will be highly variable, depending on their general L2 proficiency, among other factors. From a practical point of view, however, it will not be possible to design individual timed/untimed GJTs that take account of each learner's processing capacity, as this would require entirely individualized tests. Thus, researchers such as Bialystok (1979) and Han (1996) have identified a fixed time (3 s and 3.5 s, respectively) for their timed tests. These times were arrived at by trialing the GJTs to ensure that learners (in general) had just sufficient time to carry out operations (1) and (2) without time for (3). It remains a distinct possibility, however, that the fixed time allocated was either too short or too long for some learners. There is a further problem with the procedure adopted by Bialystok and Han; they allocated

the same length of time to *each* sentence in their tests, but it is probable that the time needed to perform operations (1) and (2) will also vary according to the length and/or grammatical complexity of the particular sentence to be judged. This problem is more easily addressable. A computerized GJT allows individual learners' response times for each sentence to be recorded. Following trialing of the test, average response times can be calculated for each sentence and used to decide the time to be allocated to each sentence in the timed test.[5]

Another problem related to the judgment of sentences as grammatical or ungrammatical concerns whether learners actually judge the specific structures that researchers intended them to judge or some other structures contained in the test sentences. This problem can be overcome if learners are asked to indicate or correct the errors in the sentences they have judged as ungrammatical. However, it is not clear whether this enhances the validity of a GJT as a measure of explicit knowledge. In studies of L1 metalingual knowledge, the ability to repair sentences at a relatively early age (4 years) is seen as reflecting a tacit knowledge of the rules of language rather than conscious awareness. According to Gombert (1992), it reflects "episyntactic" rather than "metasyntactic" behavior. Later, of course, children do draw on conscious knowledge to correct ungrammatical sentences, but clearly the ability to perform such an operation using tacit knowledge does not disappear. Time can again be expected to be a crucial factor; asked to indicate or correct an error on-line, L2 learners can be expected to rely more on their implicit knowledge, whereas if they are given time, they will have the opportunity to access their explicit knowledge. However, with the exception of Hu (2002), I know of no studies that have examined the effect of time pressure on learners' corrections in a GJT.

Learners can also be asked to state how certain they are about the judgments they have made. As I noted above, magnitude estimation incorporates an indication of certainty. Also, learners' judgments can be elicited by means of multiple-choice items that

allow for an indication of certainty. Schachter (1989), for example, offered learners four choices: clearly grammatical, probably grammatical, probably ungrammatical, and clearly ungrammatical. Another way is to ask learners how certain they are *after* they have made a judgment (e.g., on a percentage scale, where 100% means totally certain and 0% means totally uncertain). Interestingly, none of the studies in Table 1 that used a GJT to measure explicit knowledge made use of this option. Certainty assessment can be used to improve the trustworthiness of a GJT. As Birdsong (1989) points out, learners, especially those of low proficiency, lack confidence in their judgments, with the result that "ratings typically hover around neutrality" (p. 86). In this respect they contrast markedly with native speakers, who cluster at the extremities of the scale. The key question for us here, however, is whether the type of knowledge that learners draw upon influences their degree of confidence. Earlier I observed that learners' explicit knowledge is often anomalous (see Sorace, 1985). This being so, learners may be less certain about judgments based on explicit knowledge and more certain about judgments based on implicit knowledge. However, again, I know of no study that has investigated such a hypothesis.

It is also possible that learners draw differentially on their implicit/explicit knowledge depending on whether the strings they are judging are grammatical or ungrammatical. Hedgcock (1993) considers this issue in some detail. He comments:

> It would be ill-advised to claim that subjects rely on *different* L2 data bases or cognitive processes in approving well-formed strings and in rejecting ungrammatical strings, although such a possibility is not entirely implausible. (p. 15)

He then goes on to suggest that "positing autonomous L2 knowledge systems . . . is an attractive way of accounting for variable performance across learners and tasks" (p. 15). One possibility, then, is that learners are more likely to draw on their implicit knowledge for judging grammatical sentences and their explicit

knowledge for judging ungrammatical sentences. Drawing on the three processing operations referred to earlier in the section, it can be hypothesized that learners typically engage in semantic processing and noticing—searching for a possible error in the case of grammatical strings—but activate reflecting once they have actually noticed an error. Reflecting may lead to attempts to activate explicit knowledge to confirm or disconfirm whether there is a deviation in the sentence. Of course, learners may wrongly notice an error in a grammatical sentence and thus activate their explicit knowledge through reflection. Thus, if the goal in using a GJT is to measure explicit knowledge, one strategy might be to identify those sentences that each learner has judged as ungrammatical. These may provide a more valid measure of the learners' explicit knowledge than either the test as a whole or the ungrammatical items on the test. Hedgcock notes that there is no research to date that has systematically investigated the modularity of learners' knowledge in relation to the grammaticality of the sentences in a test.

The reliability of the GJTs used in particular studies has also been called into question. Birdsong (1989) points to the dangers of response bias (e.g., a general tendency to judge sentences as ungrammatical). R. Ellis (1991b) reports three studies in which the same GJT was administered to L2 learners within a week. The learners changed 22.5%, 31.0%, and 45% of their judgments from one test to the other. Ellis concludes that the GJTs may have been unreliable because the learners' L2 knowledge was uncertain, causing them to draw inconsistently on a variable set of strategies for producing a judgment.[6] One, but only one, of these strategies involved using explicit knowledge in the form of pedagogical rules of varying accuracy. Ellis's interpretation has been endorsed by other studies that have investigated how learners arrive at judgments (e.g., Goss, Ying-Hua & Lantolf , 1994). Further, L2 learners have been found to employ a greater variety of strategies than native speakers when completing a GJT (Davies & Kaplan, 1998).

Of course, reliability is not an inherent characteristic of a test but is demonstrable only when the test is administered to a

particular sample of learners. Some researchers have found that GJTs can be reliable. Bard et al. (1996) demonstrate that when relative judgments using magnitude estimation are elicited, a high level of intraparticipant consistency can be observed. Gass (1994) reports a study that suggests that test-retest variability may be a function of the syntactic structures being investigated. Overall she found that the reliability coefficients for the different structures examined were high. The size of the coefficients varied according the degree of markedness of the structures (e.g., participants were more likely to reproduce a judgment with sentences involving a relative pronoun as subject than with sentences involving a genitive relative pronoun). Gass suggests that the question of reliability is related to the degree of indeterminacy of the learners' L2 knowledge, with judgments becoming progressively less reliable as knowledge becomes more indeterminate. Gass does not explain why this might be so, but it is conceivable that indeterminacy arises when learners lack implicit knowledge of a feature and as a result attempt to make a judgment on the basis of an explicit rule, which, as Sorace (1985) has suggested, can vary in precision and accuracy as a function of development.

The very considerable interest that SLA researchers have shown in GJTs is a testimony both to the potential usefulness of these instruments and to their inherent problems. What is troubling is the almost complete disregard shown by SLA researchers using GJTs to measure language acquisition for the need to demonstrate reliability. Out of 19 studies published in *Studies in Second Language Acquisition* between 1997 and 2001 that employed some kind of GJT, only 2 reported a measure of reliability for the test. Also, in many cases, these studies failed to indicate whether the test was timed or untimed. This bears out Douglas's (2001) criticism that SLA researchers have been negligent in demonstrating the validity and reliability of their instruments of measurement. However, GJTs continue to be widely used in SLA research generally and, as Table 1 shows, they remain the preferred means of investigating explicit L2 knowledge.

To conclude this section on GJTs, as Birdsong (1989), among others, has emphasized, a GJT does not provide a direct window for viewing learners' linguistic competence but rather constitutes a performance that is influenced by a range of factors. The likelihood of a GJT's providing a measure of explicit knowledge can be increased if (a) learners are given time to judge sentences and to correct ungrammatical sentences, (b) learners' responses to the ungrammatical sentences on the test (or the sentences the learner has judged as ungrammatical) are considered separately from their responses to the grammatical sentences, and (c) learners' uncertainty in judging individual sentences is taken into account. Above all, however, what is needed are careful studies that investigate GJTs as decision-making behavior with a view to determining the factors that predispose learners to draw on a particular knowledge source.

Tests of Metalanguage

Metalanguage is not an essential component of explicit knowledge. As I pointed out earlier, explicit knowledge is verbalizable, but this need not entail the use of technical metalanguage. If metalanguage is of value to learners, it is so because it contributes to the development of explicit knowledge as awareness and facilitates its access. Metalinguistic knowledge may assist learners in developing explicit knowledge that has greater precision and accuracy.

One way of obtaining a measure of learners' metalanguage is through the use of verbal reports. These are considered later in the article. Here I will consider the kind of tests used by Alderson et al. (1997), Elder et al. (1999), and Newman and White (1999). These test learners' receptive or productive knowledge of metalingual terms by asking them to identify a named grammatical feature in a sentence, label a specified feature, or select, from multiple choices, which explanation best accounts for a particular error. Bialystok (1979), for example, gave her participants a list of nine grammatical rules and asked them to

identify which of the stated rules was violated in each sentence they identified as ungrammatical. Receptive tests such as this may provide a more valid measure of a learner's knowledge of metalanguage than production tests, as, arguably, what is important to learners is not being able to label linguistic constructs, but having access to them. Somewhat surprisingly, however, receptive tests have scarcely been used. All the studies in Table 1 that attempted to measure metalingual knowledge of rules, with the exception of Bialystok's early study, did so using tasks that required learners to verbalize rules.[7]

Irrespective of whether a test measures receptive or productive knowledge, consideration needs to be given to the selection of the metalingual terms to be examined. Ideally, a comprehensive test of metalanguage would need to measure learners' knowledge of metalingual terms relating to all aspects of language. In fact, though, studies to date have concentrated almost exclusively on grammatical terms (e.g., *subject, noun, infinitive verb, past participle*), choosing items that vary in their level of technicality, and to a lesser extent phonological terms (e.g., receptive tests of knowledge of terms like *syllable* and *stress*). To the best of my knowledge, there have been no tests of lexical metalanguage (e.g., *synonym, connotative meaning*) or of metalanguage relating to the pragmatic uses of English, although Hirose and Sasaki (2000) and Schoonen et al. (2003) have attempted to measure metacognitive knowledge of expository writing.[8]

Using Verbal Reports

Verbal reports offer another way of investigating both whether learners are drawing on a conscious representation of L2 knowledge and the extent of their metalanguage. The kinds of tasks used to elicit verbal reports are of several kinds. As I noted above, some GJTs incorporate a task that requires learners to provide a verbal explanation for their judgments. Learners might also be asked to state what kind of knowledge they used

to make each judgment (e.g., feel or rule). Alternatively, an explanation task can be used, in which learners are given sentences exemplifying the correct or the incorrect use of specific grammatical features (usually underlined), which they are then asked to explain.

Collecting verbal explanations based either on sentences that learners have judged to be ungrammatical or on specified features in grammatically correct/incorrect sentences would appear, on the face of it, to provide the most valid measure of a learner's explicit L2 knowledge. Hu (2002) comments that it is generally agreed that tasks that ask learners to explain grammatical features elicit explicit information. However, he also notes that cognitive psychologists acknowledge that verbal-report tasks may not be a sensitive or exhaustive measure of explicit knowledge, because the ability to verbalize a rule is distinct from conscious awareness of the rule. Learners may possess explicit knowledge of a specific rule but fail to verbalize it satisfactorily simply because they lack the necessary skill to talk about language. Thus, collecting verbal reports should ideally be accompanied with more receptive tests of metalanguage of the kind used by Bialystok (1979).

There is a further threat to the validity of verbal reports that is more difficult to guard against. Faced with a task that requires them to explain a grammatical feature, learners may engage in analysis of the feature using the evidence provided by the sentence and thereby produce accounts that reflect their aptitude for this kind of on-line analysis rather than their pre-existing explicit knowledge. Indeed, the kinds of tasks used to elicit learner explanations resemble in some respects the kinds of tasks used to measure language aptitude. It is difficult to see, then, how a clear distinction can be drawn between measuring what learners already know and measuring their ability to conduct grammatical analysis. One possibility might be to look for evidence of learner reflection in the verbal-report protocols (e.g., in the degree of hesitancy and pausing that accompanies an explanation), but there is probably no foolproof way of

determining whether an explanation reflects preexisting explicit knowledge or aptitude for grammatical analysis.

Verbal reports in themselves do not provide a precise measure of learners' metalingual knowledge. To obtain such a measure, it is necessary to undertake some kind of analysis of the protocols obtained from verbal reports. Perhaps the simplest way of deriving a measure from the protocols is to simply score the learners' explanations as correct or incorrect. This was the procedure adopted by Hu (2002). Hu comments that he adopted a "somewhat relaxed view of what constitutes a correct rule" (p. 361) by not insisting that the explanation provided by a learner employ technical metalanguage to be considered correct. An alternative approach is to acknowledge that learners' verbalizations will vary on a continuum of accuracy and precision and attempt to develop a scale to reflect this. Han and Ellis (1998), for example, developed a rating scale for evaluating learners' metalingual comments, a generalized version of which is provided in Table 2.[9] Such a scale provides a way of assessing learners' ability to verbalize explicitly.

Table 2

A scale for rating metalingual comments (based on Han and Ellis, 1998)

Level	Description
0	The learner is unable to explain how he/she reaches a judgment.
1	The learner is able to identify verbally the element that is the source of the problem, but his/her explanation is incorrect and does not contain even very simple technical language.
2	The learner verbalizes a rule using at least some technical language, but the rule is incorrect.
3	The learner states a partly correct rule, or the learner states a correct rule that is imprecise and incomplete.
4	The learner states a correct rule fairly precisely using some technical language.
5	The learner states a completely correct rule using appropriate technical language.

Relatively few studies have asked learners to state directly what kind of knowledge (i.e., feel or rule) they have used to make a grammaticality judgment. This, of course, would be possible only on an untimed GJT. But it is precisely this kind of GJT that is best equipped to elicit explicit knowledge. Thus asking learners to confirm the kind of knowledge they used in making grammaticality judgments could help to increase the construct validity of GJTs as a measure of explicit knowledge (i.e., the researcher could discount any judgment that a learner indicated had been made on the basis of implicit knowledge).

Guidelines for Measuring Explicit Knowledge

It should be clear from the foregoing discussion that the construction of tests of explicit knowledge is highly problematic. The impossibility of designing tests that provide pure measures of explicit knowledge as analyzed knowledge must be recognized; as should also be clear from the preceding discussion, no matter what the instrument, learners will be able to use their implicit knowledge to respond to the assessment tasks. Nevertheless, as has been argued earlier, there is a clear theoretical need for a valid and reliable measure of explicit knowledge. I would like to conclude this article, therefore, with a set of guidelines for constructing tests of explicit L2 knowledge, basing these on the definition of L2 explicit knowledge provided earlier in the article and on the preceding discussion of different types of measurement. In line with the preceding claim that a distinction needs to be made between explicit knowledge as analyzed (potentially aware) knowledge and as metalanguage, separate guidelines for measuring each aspect will be provided.

Guidelines for Measuring Analyzed Knowledge

1. The measurement of analyzed knowledge (as opposed to metalanguage) should be the primary goal in testing explicit knowledge.

2. Ideally, a test should distinguish between the measurement of learners' explicit L2 knowledge and their ability to construct such knowledge for a given context (i.e., their language aptitude).

3. A general test of analyzed knowledge needs to measure learners' explicit knowledge of different aspects of the L2: phonology, vocabulary, grammar, and pragmatics. However, it may prove impossible to devise a measure of learners' sociocritical awareness in the L2, and, as a starting point, it would probably be wise to focus on the area of language in which most research on L2 explicit knowledge has been carried out to date: grammar. Where a test is intended for use in SLA research, its precise content will obviously depend on the research questions.

4. A test might also seek to measure learners' awareness of the linguistic differences between their L1 and the target language.

5. A test might seek to identify whether the degree of markedness/prototypicality of a linguistic form is a factor in learners' analyzed knowledge (i.e., whether they are more or less aware of unmarked/marked forms).

6. A test asking learners to judge the grammaticality or acceptability of target language sentences or texts is more likely to tap analyzed knowledge if it is untimed and encourages learners to deliberate carefully before making a judgment. Such a test should also ask learners to indicate the part(s) of the sentence/text they find problematic, to indicate their degree of certainty about a particular judgment, and to state what kind of knowledge (i.e., rule or feel) was used to judge each sentence. Collecting response times may serve to identify (and eliminate) sentences to which learners have responded intuitively. Sentences that learners have judged ungrammatical may provide the best measure of their explicit knowledge. GJTs need to be constructed in accordance with

the general guidelines proposed by Chaudron (1983) and Birdsong (1989, pp. 111–125) in order to avoid some of the obvious pitfalls, such as response bias.

Guidelines for Measuring Metalanguage

1. A test providing a measure of learners' knowledge of metalanguage is of secondary importance, pending studies that demonstrate that such knowledge is an important component of L2 proficiency and/or plays a role in L2 acquisition.

2. When learners are asked to provide explanations of grammatical features (e.g., as part of a GJT), steps need to be taken to ascertain to what extent learners are creating explanations on-line (i.e., measuring aptitude) rather than drawing on their preexisting explicit knowledge.

3. A test of productive metalanguage should strive to measure depth as well as breadth of knowledge. One way in which this might be achieved is by developing a scale of metalingual understanding. Also, measuring depth requires examining participants' ability to identify metalingual constructs in a variety of sentence types (e.g., complex as well as simple).

4. However, a test of metalanguage may achieve greater validity if it measures receptive rather than productive knowledge of metalanguage, as, arguably, it is learners' understanding of explicit linguistic constructs rather than their ability to articulate metalinguistic rules that is important where language acquisition and use are concerned (see note 7).

Learners' explicit knowledge cannot be measured by means of a single test but will require multiple instruments to demonstrate concurrent validity. There is an obvious need for tests of

the objective, discrete-item type, as these have the advantage that they can measure a wide range of items and are easy to mark. However, another entirely different approach based on tasks that elicit verbal reports is also needed. Such an approach will involve asking learners to perform some task (such as a grammar explanation or perhaps a translation task) that, by its nature, invites the use of explicit knowledge, and then, by analyzing the resulting protocols, identifying what kinds of explicit knowledge learners exploit and in what ways.

Given the importance of explicit L2 knowledge in a number of different theories of L2 acquisition, there is a clear need for researchers to agree on both what L2 explicit knowledge is (a problem of definition) and how it can be assessed (a problem of measurement). Only when these problems have been solved will progress in theory building and theory testing become possible.

Revised version accepted 12 August 2003

Notes

[1]Not everyone would agree with the claim that explicit knowledge is conscious. Bialystok (1994), for example, comments that "only a small portion of the knowledge that becomes explicit will ever become conscious" (p. 566). This is because what is criterial for Bialystok about explicit knowledge is that it is "analysed," and analysis need not imply consciousness. In line with this position, it might be more accurate to say that explicit knowledge can be brought to consciousness.

[2]One of the reviewers of this article disagrees with the view I have expressed here and argues, with Krashen, that most people can learn only a limited amount of explicit language. Clearly, this is an issue in need of empirical study. However, such study can be undertaken only if there is some means of measuring learners' explicit knowledge.

[3]In addition to GJTs, there are also multiple-choice discourse completion questionnaires of the kind used by Olshtain and Blum-Kulka (1985) that are similar in design (i.e., they ask learners to judge the pragmatic acceptability of sentences). I have limited the discussion here to tests involving grammaticality judgments, but the main points made apply equally to pragmatic acceptability judgments.

[4]In a timed GJT, learners are given a fixed time to judge the grammaticality of each sentence. A number of studies involving GJTs (e.g., Robinson, 1996) have adopted a different approach: They have

allowed learners to make judgments in their own time but have recorded (via a computer) the time the learners take to make each judgment. It may be possible, retrospectively, to hypothesize, by inspecting their response times, whether the learners have relied on implicit knowledge or attempted to access explicit knowledge. However, such a procedure does not result in instruments that can be claimed to constitute separate measures of implicit and explicit knowledge. Also, because there is no time pressure in such tests, learners are encouraged to act reflectively and, I would argue, are likely to do so. However, it would be useful to obtain response times in untimed tests (the purpose of which was to measure explicit knowledge), as these could be used to eliminate judgments that were made instantly and, therefore, are unlikely to have involved reflection.

[5]Some studies (e.g., Schoonen et al., 2003) have devised measures based entirely on response times. These provide a general indication of the extent to which learners are able to access their linguistic knowledge. However, they do not tell us which specific items of linguistic knowledge are implicit and explicit and it is this that concerns me in this article.

[6]Interestingly, offering learners a *not sure* option did not appear to enhance the consistency of their judgments.

[7]One reviewer of this article argues differently, asserting that "the best evidence of learners' mastery of metalanguage is their ability to use metalingual terms to label linguistic constructs correctly." However, this ignores a central problem of tests that require learners to verbalize metalingually. As an example of this problem, consider Leow (1996). Leow sought to demonstrate that GJT scores were related to other measures of oral and written L2 production (thus demonstrating the validity of GJTs). His GJT required learners to verbalize rules. He reports strong, positive correlations between the GJT scores and the oral and written production scores. But this is hardly surprising, because the GJT scores were themselves measures of learners' ability to verbalize! Also, even if a production test is designed to measure just the ability to produce the metalingual terms being assessed (i.e., does not involve extensive verbalization of rules), it may still underestimate a learner's knowledge of metalanguage. This is because, as is the case with all vocabulary, receptive knowledge always exceeds productive knowledge. For these reasons, the kind of receptive measure employed by Bialystok is to be preferred. A disadvantage of receptive tests (pointed out by the same reviewer) is that they permit guessing. However, the receptive test used by Elder et al. (1999) sidesteps this problem by asking learners to specify, from a text provided, words for a large number of metalinguistic terms, making guesswork unrewarding.

[8]Hirose and Sasaki's test asked learners to read several statements relating to such concepts as coherence, topic sentence, and conclusion and asked them to select the most appropriate one to describe English expository writing.

[9]A reviewer of this article argues that the descriptors in Han and Ellis's scale confound correctness and metalinguistic vocabulary. However, in devising a scale for evaluating learners' rule statements such a "confound" is inevitable, as the scale must necessarily consider (a) whether the stated rule is accurate or not and (b) the extent to which it demonstrates metalingual sophistication.

References

Alderson, J., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research*, *1*, 93–121.

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

Bard, E., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, *72*, 32–68.

Bialystok, E. (1979). Explicit and implicit judgments of L2 grammaticality. *Language Learning*, *29*, 81–103.

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

Bialystok, E. (1994). Representation and ways of knowing: Three issues in second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 549–569). San Diego, CA: Academic Press.

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. Berlin, Germany: Springer-Verlag.

Butler, Y. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition*, *24*, 451–480.

Carroll, J. (1965). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training, research, and education* (pp. 87–136). New York: Wiley.

Carroll, J., & Sapon, S. (1959). *Modern Language Aptitude Test—Form A*. New York: Psychological Corporation.

Chaudron, C. (1983). Research on metalinguistic judgments: A review of theory, methods and results. *Language Learning*, *33*, 343–377.

Clapham, C. (2001). The assessment of metalinguistic knowledge. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. Macnamara, & K. O'Loughlin (Eds.), *Experiencing with uncertainty: Essays in honour of Alan Davies* (pp. 31–41). Cambridge, England: Cambridge University Press.

Clark, H., & E. Clark. (1977). *Psychology of language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

Cohen, N. (1991). *Memory, amnesia and hippocampal system*. Paper presented at the Cognitive and Neuroscience Colloquium, McGill University.

Cowan, R., & Hatasa, Y. (1994). Investigating the validity and reliability of native speaker and second-language learner judgments about sentences. In E. Tarone, S. Gass, & A. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 287–302). Hillsdale, NJ: Lawrence Erlbaum.

Cummins, J. (1983). Language proficiency and academic achievement. In J. Oller (Ed.), *Issues in language testing research* (pp. 108–129). Rowley, MA: Newbury House.

Davies, W., & Kaplan, T. (1998). Native speaker vs. L2 learner grammaticality judgments. *Applied Linguistics*, *19*, 183–203.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 42–63). Cambridge, England: Cambridge University Press.

DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty and M. Long (Eds.), *Handbook of second language learning*. Oxford, England: Blackwell.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735–808.

Douglas, D. (2001). Performance consistency in second language acquisition and language testing: A conceptual gap. *Second Language Research*, *17*, 442–456.

Eichenbaum, H. (1997). Declarative memory: Insights from cognitive neurobiology. *Annual Review of Psychology*, *48*, 547–572.

Elder, C., Warren, J., Hajek, J., Manwaring, D., & Davies, A. (1999). Metalinguistic knowledge: How important is it in studying at university? *Australian Review of Applied Linguistics*, *22*, 81–95.

Ellis, N. (1994a). Introduction: Implicit and explicit language learning—An overview. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1–31). San Diego, CA: Academic Press.

Ellis, N. (1994b). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 211–282). San Diego, CA: Academic Press.

Ellis, R. (1991a). Grammaticality judgments and second language acquisition. *Studies in Second Language Acquisition*, *13*, 161–186.

Ellis, R. (1991b). Grammaticality judgments and learner variability. In H. Burmeister and P. Rounds (Eds.), *Variability in second language acquisition: Proceedings of the tenth meeting of the Second Language*

Acquisition Forum (pp. 25–60). Eugene: University of Oregon, Department of Linguistics.

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

Ellis, R. (1994a). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79–114). San Diego, CA: Academic Press.

Ellis, R. (1994b). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R., & Yuan, F. (in press). The effects of careful on-line planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language*. Amsterdam: John Benjamins.

Fairclough, N. (1985). Critical and descriptive goals in discourse analysis. *Journal of Pragmatics*, *9*, 739–763.

Gass, S. (1994). The reliability of second-language grammaticality judgments. In E. Tarone, S. Gass, & A. Cohen (Eds.), *Research methodology in second language acquisition* (pp. 303–322). Hillsdale, NJ: Erlbaum.

Gombert, J. (1992). *Metalinguistic development.* T. Pownall (Trans.). New York: Harvester Wheatsheaf.

Goss, N., Ying-Hua, Z., & Lantolf, J. (1994). Two heads may be better than one: Mental activity in second-language grammaticality judgments. In E. Tarone, S. Gass, & A. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 263–286). Mahwah, NJ: Erlbaum.

Green, P., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics*, *13*, 168–184.

Han, Y. (1996). *L2 learners' explicit knowledge of verb complement structures and its relationship to implicit knowledge*. Unpublished doctoral dissertation, Temple University, Philadelphia.

Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research*, *2*, 1–23.

Hedgcock, J. (1993). Well-formed vs. ill-formed strings in L2 metalingual tasks: Specifying features of grammaticality judgements. *Second Language Research*, *9*, 1–21.

Hirose, K., & Sasaki, M. (2000). Effects of teaching metaknowledge and journal writing on Japanese university students' EFL writing. *JALT Journal*, *22*, 69–93.

Hu, G. (2002). Psychological constraints on the utility of metalinguistic knowledge in second language production. *Studies in Second Language Acquisition*, *24*, 347–386.

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

James, C., & Garrett, P. (1992). The scope of awareness. In C. James and P. Garrett (Eds.), *Language awareness in the classroom* (pp. 3–20). London: Longman.

Karmiloff-Smith, A. (1979). Micro- and macro-developmental changes in language acquisition and other representation systems. *Cognitive Science*, *3*, 91–118.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition*, *12*, 215–247.

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

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford, England: Pergamon.

Krashen, S. (1994). The input hypothesis and its rivals. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45–77). San Diego, CA: Academic Press.

Lantolf, J. (2000). Introducing sociocultural theory. In J. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 1–26). Oxford, England: Oxford University Press.

Leow, R. (1996). Grammaticality judgment tasks and second-language development. In G. Alatis, C. Straehle, M. Ronkin, & B. Gallenberger (Eds.), *Georgetown University Round Table on Language and Linguistics, 1996* (pp. 126–139). Washington, DC: Georgetown University Press.

Mandell, O. (1999). On the reliability of grammaticality judgment tests in second language acquisition research. *Second Language Research*, *15*, 73–99.

Masny, D. (1987). The role of language and cognition in second language metalinguistic awareness. In J. Lantolf & A. Labarca (Eds.), *Research in second language learning: Focus on the classroom* (pp. 59–73). Norwood, NJ: Ablex.

Nesdale, A., Herriman, M., & Tunmer, W. (1984). Phonological awareness in children. In W. Tunman, C. Pratt, & M. Herriman (Eds.), *Metalinguistic awareness in children* (pp. 56–72). Berlin: Springer.

Newman, J., & White, L. (1999). A pilot study of language awareness at the New Zealand tertiary level. *New Zealand Language Teacher*, *25*, 41–53.

Odlin, T. (1989). *Language transfer*. Cambridge, England: Cambridge University Press.

Olshtain, E., & Blum-Kulka, S. (1985). Degree of approximation: Nonnative reactions to native speech act behavior. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 303–325). Rowley, MA: Newbury House.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393–419). San Diego, CA: Academic Press.

Pimsleur, P. (1966). Testing foreign language learning. In A. Valdman (Ed.), *Trends in language teaching* (pp. 175–214). New York: McGraw Hill.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Reber, A. (1989) Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219–235.

Robinson, P. (1995). Aptitude, awareness and the fundamental similarity of implicit and explicit second language learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Tech. Rep. No. 9, pp. 303–358). Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, *18*, 27–67.

Sasaki, M. (1996). *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses*. New York: Lang.

Schachter, J. (1989). A new look at an old classic. *Second Language Research*, *5*, 30–42.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, *11*, 11–26.

Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacoginitive knowledge. *Language Learning*, *53*, 165–202.

Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behaviour. *Studies in Second Language Acquisition*, *15*, 147–163.

Sharwood-Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics*, *2*, 159–169.

Skehan, P. (1986). Where does language aptitude come from? In P. Meara (Ed.), *Spoken language*. London: Center for Information on Language Teaching.

Skehan, P. (1990). The relationship between native and foreign language learning ability: Educational and linguistic factors. In H. Dechert (Ed.), *Current trends in European second language acquisition research* (pp. 83–106). Clevedon, England: Multilingual Matters.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, England: Oxford University Press.

Sorace, A. (1985). Metalinguistic knowledge and language use in acquisition-poor environments. *Applied Linguistics*, *6*, 239–254.

Sorace, A. (1996). The use of acceptability judgments in second language acquisition research. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 375–409). San Diego, CA: Academic Press.

Tunman, W., & Herriman, M. (1984). The development of metalinguistic awareness: A conceptual overview. In W. Tunman, C. Pratt, & M. Herriman (Eds.), *Metalinguistic awareness in children* (pp. 12–35). Berlin, Germany: Springer.

Valtin, R. (1979). *Increasing awareness of linguistic awareness in research on beginning reading and dyslexia*. Paper presented at the International Reading Research Seminar on Linguistic Awareness and Learning to Read, Victoria, British Columbia, Canada.

Wolfson, N. (1983). Rules of speaking. In J. Richards and R. Schmidt (Eds.), *Language and communication* (pp. 61–87). London: Longman.