Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal
Hulstijn, J.H.; de Graaff, R.

Published in:
AILA Review

Citation for published version (APA):
UNDER WHAT CONDITIONS DOES EXPLICIT KNOWLEDGE OF A SECOND LANGUAGE FACILITATE THE ACQUISITION OF IMPLICIT KNOWLEDGE? A RESEARCH PROPOSAL

Jan H. Hulstijn & Rick de Graaff
Vrije Universiteit Amsterdam

Abstract
This paper proposes nine hypotheses for empirical research aiming to assess under which conditions explicit grammar instruction helps the acquisition of implicit knowledge of a second language. The results of empirical studies, designed on the basis of these hypotheses, may replace the unqualified interface and noninterface positions by qualified positions claiming that explicit instruction facilitates the acquisition of implicit knowledge under specific conditions. These conditions depend on factors such as linguistic domain, complexity, scope and reliability, and semantic or formal redundancy of the target structures. Furthermore, they depend on factors such as rule versus item learning and reception versus production. Finally, they depend on moderating factors such as type of instruction, fluency of performance, and learner characteristics.

1. Introduction
In this paper we take an empirical approach to the question to what extent the acquisition of implicit knowledge, as the product and goal of second language (L2) learning, can be facilitated by explicit learning. We use the term "explicit learning" in the sense defined by Schmidt (this issue), as learning with awareness at the point of learning. According to Schmidt, explicit learning may vary from spontaneous rule discovery by the learner on the one hand to explicit instruction on the other hand, i.e. guidance on the part of the teacher through input enhancement or presentation and explanation of grammar rules.

Some researchers take the stance that explicit grammar instruction is of little help and that L2 acquisition must be fostered by giving L2 learners large amounts of "comprehensible input" (e.g. Krashen, 1981, 1982, 1985). Others recognize the prime importance of large amounts of comprehensible input and involvement of L2 learners in communicative activities, but take the view that, under certain conditions, explicit grammar instruction, in one way or another, can be of considerable help (e.g. Bley-Vroman, 1988; Doughty, 1991; Ellis, 1990, 1993a, 1993b; Larsen-Freeman, 1991; Larsen-Freeman & Long, 1991; Lighthowen & Spada, 1993; Schmidt, 1988, 1990, 1993; Rutherford & Sharwood Smith, 1985; Sharwood Smith, 1993; Terrell, 1991; VanPatten & Cadiemo, 1993a, 1993b).

Our perspective is a programmatic one: assuming that the answer to the question of whether explicit grammar instruction facilitates the acquisition of implicit knowledge will not be an unqualified "never" nor an unqualified "always", how can we lay out a programme of research which aims at assessing when, i.e. under which conditions, explicit grammar instruction helps?

This paper is set up as follows. Section 2 lists our assumptions, and summarizes and discusses various theoretical positions: Krashen's (1981, 1982, 1985) noninterface position, Anderson's (1980, 1982) strong interface position, and Ellis' (1993a) weak interface hypothesis. Section 3 provides a number of testable hypotheses, based on various theoretical

1 We would like to thank the following people for their comments on an earlier version of this paper: Rod Ellis, Leo van Lier, Richard Schmidt, and Bill VanPatten.
and pedagogical considerations. Section 4 sums up some moderating factors which have to be taken into account when designing investigations to test any of the hypotheses mentioned in section 3. Section 5 discusses the relative merits of "natural" versus (semi)artificial experiments. In section 6, we give, by way of example, a brief description of a study in progress, conducted by ourselves. In section 7 we draw our conclusions.

2. Theoretical positions

Let us summarize the assumptions underlying most current theories of L1 and L2 acquisition. We sum these assumptions up, couched in our own terminology:

(1) Fluent language proficiency of native speakers of any age, is based on "implicit knowledge" (I-knowledge) of the grammar of their L1.

(2) Most native speakers have little "explicit knowledge" (E-knowledge) of their L1, although some E-knowledge may be learned in school.

(3) Almost nothing is known about the process of the acquisition of I-knowledge of the L1. One thing is certain, however, caretakers do not teach young children grammar rules. Thus, I-knowledge is not acquired through the acquisition of E-knowledge.

(4) Although almost nothing is known about the process of L1 acquisition, we may assume that children's L1 acquisition necessarily involves some sort of "paying attention" to language forms, to the meanings expressed by forms, and to form-meaning relations (Levelt, Sinclair & Jarvella, 1978). Such focusing of attention may often be caused by error-detecting devices (Clark & Hecht, 1983; Marshall & Norton, 1978).

(5) As in the case of L1, a fluent use of the L2 is based on I-knowledge of that language. The behavioral correlate of this I-knowledge is fluency in language use.

(6) Fluent language use is characterized by nonserial (= parallel) processing. Not much is known about the process which takes I-knowledge as its input and has fluent behaviour as its output. However, current psycholinguistic theories claim that this process is characterized by parallel processing. Speaking of word recognition in reading, Rayner & Pollatsck (1989:108) conclude that "parallel processing is the hallmark of automatic processing". This may be true not only for word recognition, but also for many other productive and receptive language skills (Levelt, 1989: 20-22).

(7) Fluent, automatic language use is a matter of parallel, incremental processing (Levelt, 1989: 24), and certainly not a rapid serial application of rules. It is, for example, not a matter of running extremely fast through a flow chart, making all the individual steps, taking into account all the conditions specified. Parallel processing creates the phenomenological experience of not being aware of the way in which information is being processed and how these processes are monitored.

One of the perennial issues of contention in L2 acquisition research and in L2 pedagogy is the following one:

Although I-knowledge of a L1 is not acquired through the acquisition of E-knowledge (assumption 3), it is an open question whether the acquisition of E-knowledge of a L2 can facilitate the acquisition of I-knowledge. In other words, to what extent can the goal of I-knowledge be reached more efficiently (i.e. with fewer difficulties and/or in less time) via the teaching and learning of explicit rules?

Basically, there are two answers to this question:

(1) I-knowledge of an L2 cannot come into existence as the direct result of E-knowledge, nor can it indirectly be influenced by E-knowledge of that language. This position has become known as the "noninterface position".

(2) The acquisition of I-knowledge of an L2 can be driven or influenced by the learning and teaching of E-knowledge. This is the "interface position".

The next subsections consider these positions in more detail.
2.1. The noninterface position

One of the best known proponents of the noninterface position is Krashen (1981, 1982, 1985). According to Krashen, there is a fundamental difference between "acquisition" (of I-knowledge) and "learning" (of E-knowledge). Learning does not facilitate acquisition. L2-utterances are generated on the basis of I-knowledge. E-knowledge only plays a role in monitoring utterances thus generated. Grammar instruction may result in E-knowledge but cannot result in, nor affect I-knowledge. For that reason, grammar instruction is of little or no importance for becoming fluent in the L2. I-knowledge can only be acquired by receiving large amounts of "comprehensible input". That is about all the noninterface position has to say about the mechanics of the actual process of acquiring I-knowledge. In this respect, Krashen's position reflects Chomsky's much quoted statement (1970), viz. "that we should probably try to create a rich linguistic environment for the intuitive heuristics that the normal human automatically possesses".

Krashen's noninterface position certainly is a well defendable position for L2 pedagogy. But from an empirical perspective it is a fruitless position, as long as the distinction between acquisition and learning has not been operationalized (Hulstijn & Hulstijn, 1984; Larsen-Freeman & Long, 1991: 224-225, 245-249; McLaughlin 1978, 1987 Ch. 2).^2

2.2. The interface position

The interface position comes in two forms (Larsen-Freeman & Long, 1991:324). First, that E-knowledge becomes I-knowledge through practice. Second, that E-knowledge cannot become I-knowledge but that it aids the acquisition of I-knowledge. Thus, the interface position sees E-knowledge as either necessary or instrumental in acquiring I-knowledge. We will discuss each form in turn.

2.2.1. A strong interface position

According to what Ellis (1993a) has called the "strong interface position", I-knowledge emerges from E-knowledge directly, i.e. through practice. This position is derived from the literature on skill acquisition in cognitive psychology, in particular theories of controlled and automatic processing. One of the best known of these theories is Anderson's ACT* theory (Anderson, 1982). According to Anderson, declarative, propositional knowledge is converted into procedural knowledge by processes of compilation, tuning and restructuring. The original declarative knowledge need not always be replaced by procedural knowledge: "Sometimes the two forms of knowledge can coexist side by side, as when we can speak a foreign language fluently and still remember many rules of grammar. However, it is the procedural, not the declarative knowledge that governs the skilled performance" (Anderson 1980:226; see also recommendation 7 in Schmidt's contribution to this issue).^4

For many years, the strong interface position was the accepted view of most practitioners in the field of L2 pedagogy. A skilled and fluent command of the L2 was (and, by many, still is) viewed as the automatization of the application of explicit grammar rules. It should be noticed, however, that most educationalists interpreted the notion of automatization as rapid serial execution rather than as parallel processing.

For a proposal to make the acquisition/learning distinction testable, see Schwartz, 1993:155.

Although Anderson (1980, 1982) does not use the term "parallel processing", which was introduced some years later in cognitive psychology (e.g. Rumelhart & McClelland, 1987), his notion of procedural knowledge appears to be more akin to parallel processing than to rapid serial processing.

Anderson appears to have changed his mind on the issue whether LI acquisition comes about through the proceduralization of declarative knowledge. In the third edition of his popular textbook "Cognitive psychology and its implications" (1990), Anderson writes: "With respect to the issue of whether language is really a system different from other human cognitive systems, it is a fair summary to say that the jury is still out" (p.356), and "Children also appear not to be much helped by direct instruction but rather learn language more implicitly" (p. 325).
2.2.2. A weak interface position
Ellis (1993a) rejects the strong interface position while replacing it by what he calls a "weak interface position". To illustrate all three positions (noninterface, strong interface, and weak interface), Ellis presents the following figure, using labels from both Anderson's skill acquisition theory (1982) and Bialystok's two-dimensional model of language proficiency (see e.g. Bialystok, 1990a, pp. 118-129; 1991).

<table>
<thead>
<tr>
<th>Declarative</th>
<th>Procedural</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type A</strong></td>
<td><strong>Type B</strong></td>
</tr>
<tr>
<td>Explicit</td>
<td>Conscious knowledge of L2 items</td>
</tr>
<tr>
<td></td>
<td>Conscious knowledge of learning, production, and communication strategies. The learner can use explicit knowledge easily and rapidly.</td>
</tr>
</tbody>
</table>

| **Type C**  | **Type D** |
| Implicit    | Intuitive knowledge of L2 items |
|             | Ability to employ learning, production, and communication strategies automatically. The learner can use intuitive knowledge fluently. |

Source: Ellis, 1993a, p. 94.

**Figure 1: The difference between explicit/implicit and declarative/procedural knowledge**

Ellis refers to the vertical dimension with the labels "knowledge dimension" and "analysis". For the horizontal dimension he uses the labels "process dimension", "control" and "automatic" (p.94-95).

According to Ellis, the noninterface position claims that Type D knowledge derives from Type C knowledge only. The strong interface claims that Type A knowledge can be converted into Type D knowledge through practice. According to Ellis' own weak interface position, Type A knowledge may develop into Type C knowledge, provided learners are ready to accommodate the new knowledge into their interlanguage systems. Type D knowledge can develop from Type C knowledge through formal practice or natural communication.

It seems to us that Ellis' model comprises an incompatible combination of Anderson's theory postulating a transition from declarative to procedural knowledge, and Bialystok's dimensions of Analysis and Control. According to Bialystok (1988: 40; 1990b: 47), E-knowledge cannot be converted into I-knowledge, as Ellis wants it in his weak interface position, and automaticity is epiphenomenal. However, it must be acknowledged that Bialystok's earlier work (1978) allows for E-knowledge to be converted into I-knowledge. Ellis (1994; and personal communication 31 March, 1994) points out that his 1993 model is based on Bialystok's earlier work.

Thus, if the horizontal dimension cannot be conceived as a matter of gaining automaticity, there is no difference anymore between Type C and Type D knowledge. Having I-knowledge of
grammar amounts to being able to use that knowledge fluently (cf. section 2, assumption 5), although, evidently, fluency may be affected by performance factors.\(^5\)

2.3. The interface position as assumption for empirical research

Although we do not wish to commit ourselves to the strong or weak interface position as defined by Ellis (1993), we do not dismiss the interface idea altogether. On the contrary, the most fruitful approach to further an understanding of L2 learning is to adopt the interface position as a framework within which hypotheses can be formulated and empirical research can be based. A discipline cannot exist of theories alone, it must advance through empirical work as well. For that reason, we contend that it is potentially more fruitful to start on the assumption of an interface position than on the basis of a noninterface position. Starting on the basis of an interface position, then, our task is to specify under what conditions E-knowledge may be instrumental in acquiring I-knowledge. In other words, assuming that E-knowledge can be the result of explicit learning and that explicit learning can (but need not) follow from explicit instruction (see recommendation 4 - 6 in Schmidt's contribution to this issue), our task is to specify under what conditions explicit instruction of grammar rules does further L2 acquisition.

3. The interface position translated into testable hypotheses

Basing empirical research on the assumption that explicit instruction (and hence that explicit learning) somehow "facilitates" the acquisition of I-knowledge, does not mean that we have no theories or other sources of knowledge to draw upon. In this section, we present six dimensions within which variables can be chosen for the formulation of testable research questions. An illustrative example is given for each dimension, followed by one or more testable hypotheses.

From the outset, however, it must be emphasized that each hypothesis embodies a simplification: the reader has to add the phrase "ceteris paribus" to each hypothesis, as it were. In reality, however, all other things aren't equal. A factor mentioned in one hypothesis may well interact with a factor mentioned in another hypothesis. Indeed, it is likely that cases of interaction constitute more interesting issues to be pursued than the straight ones covered by the hypotheses below.

3.1. Linguistic domain

The first dimension within which one must specify the interface position is linguistic. In which linguistic domains can explicit instruction help the acquisition of I-knowledge? In the domain of morpho-phonology? In the subdomain of lexical, flectional morphology only (e.g. verb

---

5 Ellis (personal communication, 31st March 1994) comments as follows: "You may still wish to argue that there is no distinction between Type C and Type D knowledge and that 'fluency is an epiphenomenon of I-knowledge'. I think, however, that there is an empirical basis for claiming a difference - in the study of interlanguage variability. How do we explain variability within the learner's vernacular style (i.e. the style that can be considered to favour use of I-knowledge)? One possibility is that the learner attempts to edit his/her performance using E-knowledge. Such an explanation is not entirely adequate, however. Learners find it easier to control (access for production) 'old' I-knowledge than 'new' I-knowledge, as reflected, for example, in the forms they supply when performing in this style. Thus, form x (for example, the target form) may be supplied more readily and more consistently than form y (for example, a non-target form) in the vernacular style, in both cases the learner not being 'aware' of which form is used. A good example might be the kind of variability that is evident in the use of copula (see my article in Applied Linguistics 1988). Learners begin with zero copula and then learn to supply the copula. Their control over copula can be seen to increase over time, partly in accordance with linguistic context. I agree, however, that the case for distinguishing Type C and Type D knowledge has to be made more carefully than in my TQ93 article. I think, though, that the distinction can be operationalized, although more easily in a longitudinal study than in the kind of experimental study you have in mind."
conjugation, noun declension)? In syntax? In certain limited areas within syntax only (e.g. word order phenomena related to L1/L2 differences at the level of D-structure, in contrast to S-structure)?

Given the current state of linguistics, one of the most important criteria to distinguish, in a principled way, between various linguistic domains, is offered by Generative Grammar. Generative grammarians distinguish between formal aspects of grammar within the domain of Universal Grammar, and domains falling outside the UG scope. Concerning the UG aspects of grammar, our hypotheses might depend on our stance on the issue of whether UG is still available to L2 learners or not. In the former case, parameter resetting is possible. In the latter case, L2 learners will depend more on general cognitive capacities to overcome the loss of the ability to acquire the L2 grammar on the basis of positive L2 input alone. If we assume that UG is still available, our hypotheses might depend on whether the L1 setting of a parameter constitutes a subset or a superset of the L2 setting. In the latter case, L2 learners have to "unlearn" the marked L1 setting, and this may require receiving negative evidence. If a parameter has two possible settings only (e.g. the Head-position parameter, which can either be head-initial or head-final), and one is found in L1 and the other in L2, the acquisition of the L2 setting may be possible on the basis of primary linguistic input (positive evidence) only. Finally, when various rules cluster under a single parameter, teaching all rules may not be necessary and even fruitless after one rule has been acquired (Rutherford, 1989; White, 1992; Schwartz, 1993; Jordens, 1993).

Here are some tentative hypotheses which could be tested on the basis of the above considerations:

* H1 The advantage of the provision of explicit instruction, in comparison with the non-provision of explicit instruction, is greater in the case of aspects falling outside the scope of UG than in the case of aspects falling inside its scope.

* H2 The advantage of explicit instruction is greater when the LI setting of a parameter forms a superset and the L2 setting a subset of a parameter, than in the reverse situation. With "is greater" in these two and all following hypotheses we mean that the acquisition of knowledge with the help of explicit instruction takes place more easily and in less time than without. We do not claim, however, that explicit instruction (and learning) will always result in a higher level of ultimate attainment. Note that, for ease of reading, we have left out the phrase "in comparison with the non provision of explicit instruction" in hypothesis 2 and all following hypotheses.

3.2. Complexity

The basic idea here is that L2 grammar rules, represented in a declarative, propositional form (as they appear in course books, for example), differ in complexity. Complexity here is not to be defined in terms of linguistic theory, but in cognitive terms. Consider the following two fabricated examples, illustrating complexity differences between two purely formal (phonological) rules and two semantic (aspectal) rules respectively.

Example 1: In language M, verbs have different endings for two aspects, perfect and non-perfect; in language N, however, verbs have different inflections for three aspects, perfective, durative, and punctual. Example 2: In language X there are two plural suffixes for nouns: the suffix -s is added to singular nouns ending on a vowel, and the suffix -os is added to nouns ending on a consonant. In language Y has the same two suffixes: -s for nouns ending on a vowel and for nouns ending on a consonant containing a front vowel in the penultimate syllable, -os for nouns ending on a consonant and containing a back vowel in the penultimate syllable. If we were to make flow charts for the derivation of inflections in languages M and N and for the derivation of the plural suffix in languages X and Y, the charts for languages M and X would

---

6 In fact, it may be impossible to define linguistic (in contrast to cognitive) complexity independently from language development and accuracy data (Hoekstra, 1990). Circularity must be avoided. Thus, it would be wrong to define "late acquired" as "complex", and define a "complex" rule as one which is "late acquired".
contain fewer steps than those for N and Y respectively. Note that degree of complexity is contingent not so much on the number of forms in a paradigm, but rather on the number (and/or the type) of criteria to be applied in order to arrive at the correct form. On such a declarative notion of complexity, we may base the following hypothesis:

* H3 The advantage of explicit instruction is greater in the case of complex rules than in the case of simple rules.

The reason is that simple formal phenomena may be salient enough in the input to be discovered by L2 learners spontaneously, without the help of explicit instruction. In the case of complex phenomena, however, explicit instruction may save learners considerable time in discovering their intricacies.\(^7\) Note that this hypothesis is concerned with the (cognitive) complexity of a rule of language, not with the complexity of the way in which the rule is explained by teacher or textbook. Of course, it is a sound pedagogical principle that explanations should be formulated in as simple as possible terms. Furthermore, declarative complexity as defined here, is not to be confused with "complicatedness" or "difficulty", experienced by learners as a result of contrast between their L1 and L2 (see also H2).

Some researchers have tried to explain the notion of cognitive complexity in terms of processing constraints (Clahsen, 1984; Pienemann, 1989; Pienemann & Johnston, 1987). This was done to explain so called "natural" acquisition orders, found among L2 learners who did not receive formal L2 instruction.\(^8\) The hypotheses of Pienemann c.s. are of an entirely different nature than the one proposed here. The former hypotheses are embedded in a developmental theory pertaining to acquisition orders. Our claim that explicit instruction helps more in the case of complex than in the case of simple rules, pertains to the ease and duration of acquisition, not to the rank order of acquisition onset times of structural features.

3.3. Scope and reliability

Many rules of language apply only probabilistically. Such "rules", or "probabilistic tendencies" can differ in scope (the number of cases covered) and reliability (the extent to which the rule holds true, cf. Bates & MacWhinney, 1989).\(^9\) Scope may be large or small and reliability may be high or low. Let us draw two (arbitrary) border lines: the scope of a rule is said to be large or small when the rule covers more or fewer than 50 cases; the reliability of a rule is said to be high or low when the rule applies in more or less than 90% of all cases. According to these definitions, there are four possibilities. Consider the following examples concerning noun gender in German (masculine, feminine and neuter), as specified by Mills (1986:33):

1. Of the ca. 15,000 singular nouns ending on -e about 13,500 (90%) are feminine.
2. Of the 15 monosyllabic nouns beginning with Kn-, 14 are masculine (93%).
3. Of the 107 monosyllabic nouns ending on -C\(\text{nasal}C\), 75 are masculine (70%).
4. Of the 45 nouns ending on -ier, 27 are neuter (60%).

\(^7\) Complexity is a likely candidate for interaction with various other factors. It is important to distinguish the notion of complexity from the notion of scope, to be explained in section 3.3. A language teacher, for example, might say that a verb conjugation rule which applies to only a small number of verbs (constituting one of a great number of verb classes), is too complex to teach with much effect. This would amount to claiming that the teaching of an explicit grammar rule does not help when this rule is too complex. In this case, however, it is the small scope rather than the degree of complexity of the phenomenon that would lead one to claim that explicit instruction does not aid the acquisition of L-knowledge. We cannot think of examples of rules, falling outside UG, being large in scope and reliability (see section 3.3.) that would be so complex that explicit instruction of them would not facilitate the acquisition of their L-knowledge. Anyway, our complexity hypothesis appears to us a well founded one and one which can and should be tested.

\(^8\) DuPlessis, Travis & White (1987) give a linguistic explanation of this acquisition order.

\(^9\) We suppose that the notion of scope and reliability can and must be distinguished to the extent that we are dealing with rules pertaining to a (practically) finite class of utterances. This will almost always be a lexical class. Therefore, scope and reliability can best be illustrated with examples from the lexicon. However, scope and reliability may even be distinguishable when we are dealing with rules pertaining to an infinite class of utterances, i.e. in nonlexical realms (e.g. syntax).
We can classify these four tendencies in terms of scope and reliability, as follows:

1. Large scope & high reliability
2. Small scope & high reliability
3. Large scope & low reliability
4. Small scope & low reliability

Many language teachers believe that it makes sound practice to restrict explicit grammar teaching to rules with both a large scope and a high reliability. In the three remaining cases, learners should not be given the rules. Such rules are believed to bring too little profit considering their costs (the law of "diminishing return"). Thus, in the cases of 2, 3, and 4, learners must be told to simply associate the gender feature to each word individually. On this practice we may base the following hypothesis:

*H4 The advantage of explicit instruction is greater when a rule applies to many cases (large scope) and when it has a high success rate (high reliability) than when it has a small scope and/or a low reliability.

Note that complexity on the one hand and scope and reliability on the other, although to be distinguished conceptually, in the practice of language pedagogy often go hand in hand. For instance, language teachers often decide not to present complex rules if they are small in scope. Furthermore, when teachers wonder whether they should explain a phenomenon either with a more reliable but more complex rule or with a less reliable but more simple rule, they may opt for the latter, sacrificing reliability for simplicity, e.g. presenting a general rule without its exceptions or subrules.

3.4. Rule learning versus item learning

Consider the following four rules of German:
1. The finite verb in declarative main sentences almost always takes second position.
2. When only one constituent is being negated, the negation word *nicht* is placed immediately before it, e.g. *nicht A sondern B* ("not A but B").
3. German nouns ending on *-e* are almost always feminine, as we have already seen; e.g. *Küche* ("kitchen") is feminine.
4. Verb forms in the present tense, indicative, first person singular almost always end on *-e*;
   e.g. *ich lerne* ("I learn English")

All four rules are large in scope and they rank high on reliability. Yet there is a difference between rules 1 and 2 on the one hand and rules 3 and 4 on the other. For the latter two phenomena, there are, in principle, two ways for language production. The language user may either apply the rule, or produce individually stored forms (e.g. *lerne*) or features (e.g. the feature Feminine with *Küche*). Thus, for phenomena 3 and 4, two alternative routes towards language production offer themselves, one based on rule learning and one based on item learning. For phenomena 1 and 2, however, there is only the possibility of rule application.

One could argue that teaching rules 3 and 4 does not foster language acquisition much, since learners may spontaneously learn individual forms or individual features (item learning). For phenomena 1 and 2, however, rule teaching may help learners to apply the principle (which may or may not have originated from previous encounters with a number of individual instances) to an indefinitely large class of instances. This leads to the following hypothesis:

Note that, viewed from a processing point of view, both routes lead to the same end product. They both may be active and both may be initiated automatically (just as the two word recognition routes in reading, one route of direct access and one indirect route via phonological coding).

What is called a "rule" in terms of grammar, may be called an "analogy" in terms of language processing. Furthermore, even the "verb-second" and the "*nicht* + the to be negated constituent" analogies may stem from item learning, i.e. they may be derived from the storage of individual instances previously encountered in the input, e.g. in the form of formulaic utterances (Ellis, 1993; Krashen, 1981, Ch. 7). See also Logan, 1988; Robinson & Ha, 1993; Schmidt, 1992.

One might be tempted to use the label "nonlexical" for rules 1 and 2 and the label "lexical" for rules 3 and 4, and argue that "item learning" offers itself as an alternative to lexical rules only. However, the division...
The advantage of explicit instruction is greater when language production can only be based on rule application, than when it can be based not only on rule application but also on the retrieval of individually stored items ("item learning").

The likelihood of item learning as an alternative to rule learning may be dependent of the frequency with which items occur. Forms occurring frequently in the input may be more likely to be stored and accessed separately than infrequent forms. Since the issue of item vs rule learning mostly pertains to inflectional morphology, a hypothesis concerning the relative merits of item learning might be more relevant for the learning of languages with moderate inflection (French, German, English) than for the learning of languages with high amounts of inflection (Turkish, Finnish, Hungarian). Thus, we can reiterate the above stated complexity hypothesis, while applying it to the domain of inflection:

The advantage of explicit instruction is greater in the case of complex inflectional rules (which are also reliable and large in scope) than in the case of simple inflectional rules (also reliable and large in scope).

There are two motivations for this hypothesis. The first reason is that grammar instruction may be more helpful to clarify complex, less salient structures than to clarify simple, more salient structures (section 3.2). The second reason is that, if there are fewer forms in an inflectional paradigm, each one of them is more likely to occur more often and is therefore more likely to be learned individually, than if there are more forms per paradigm. In the latter case, each form is likely to occur less frequently than all forms together, and is therefore less likely to be learned as an individual item.

### 3.5. Reception versus production

There is ample evidence that L2 learners often keep making errors in spontaneous speech long after the grammar rule in question has been presented and practised in written exercises. In fact, it is on the basis of such evidence that proponents of the noninterface position argue that explicit instruction in L2 pedagogy is of little importance and should perhaps be abandoned altogether. It is surprising, however, that the debate of whether explicit instruction is useful or not, focuses exclusively on E-knowledge's impact on speaking and writing and almost never on its impact on reading and listening. One might argue, however, that E-knowledge (and hence explicit instruction) affects language comprehension during reading and listening positively, in that it helps L2 learners to discern the meaning of the input. Thus, although we might acknowledge that explicit instruction has limited effect on language production, we may hypothesize that explicit instruction has considerable effect on language comprehension.

### 3.6. Semantic versus formal redundancy

Some language forms are semantically redundant (e.g. the -s suffix in third person singular verb forms in English) while others are not (e.g. the -s plural noun suffix in English). This may be of more importance to language comprehension than to language production, since knowledge of semantically redundant phenomena may not be required for language comprehension. Hence, we frame the following hypothesis:

As for L2 comprehension, the advantage of explicit instruction is greater in the case of grammatical features with semantic implications than in the case of purely formal (semantically redundant) features.\(^\text{13}\)

---

between lexical and nonlexical rules may differ from one grammatical theory to another. It would therefore be wise to distinguish between constructs of a theory of grammar (lexical vs nonlexical) and constructs of a theory of learning (rule learning vs item learning), although the constructs may be related across disciplines.

\(^{13}\) Our hypothesis 8 is paralleled by hypothesis 1 in VanPatten's contribution to this issue. The difference is that VanPatten has formulated his hypothesis from the perspective of learners' spontaneous input processing.
In language production on the other hand, learners may have more difficulty to apply purely formal rules (e.g. adding third person singular -s to English present tense verb forms) than to apply combined formal-semantic rules (e.g. adding a plural -s to count nouns in English). Terrell (1991) states that explicit grammar instruction helps to segment the otherwise overwhelming input and to establish meaning-form links, especially for morphology that is neither salient nor semantically essential. This leads us to the following hypothesis: *H9 As for L2 production, the advantage of explicit instruction is greater in the case of purely formal (semantically redundant) features than in the case of grammatical features with semantic implications.

4. Moderating factors

4.1. The conveyance of grammatical information

In all hypotheses mentioned so far, the phrase "explicit instruction" was used. Anyone who wants to test these hypotheses empirically, however, must become more specific as to how to provide the explicit instruction. In this section, we restrict ourselves to mentioning some options:
- With how much grammatical jargon will the rule in question be stated? Can the formulation of an explicit rule, with the use of grammatical terms, be avoided?
- Will the rule first be given together with the examples, or will the examples be given first and will the learners be invited to discover the regularity illustrated in them?
- Will grammar teaching be restricted to the presentation of the rules, i.e. without form-focused practice, or will grammar teaching consist of rule presentation plus form-focused practice?
- After the rule has been stated, will the learners first be involved in comprehension tasks (what VanPatten, this issue, calls "processing instruction", see also VanPatten & Cadierno, 1993a, 1993b; Ellis, 1993b), and will production tasks be delayed until much later? Or will the learners be requested to apply the rule immediately in rule-focused production tasks (what VanPatten calls "traditional instruction")?

4.2. Fluency in language production

When conducting empirical research on the influence of grammar instruction on L2 acquisition, we must measure performance of L2 learners on a variety of tasks. Three independent factors offer themselves as candidates for the measurement of L2 performance:
(1) Task: e.g. comprehension, production, or grammaticality judgment tasks.
(2) Time of testing: during or after instruction and practice.
(3) Task condition: subjects can be instructed to take as much time as they want (monitoring grammatical accuracy), or to perform the task as fast as they can. By manipulating task conditions, the researcher may be able to ascertain subjects' fluency.\(^\text{14}\)

The third factor is especially relevant since we want to test hypotheses concerning the acquisition of I-knowledge. I-knowledge is a theoretical construct, not directly accessible by means of language tests. However, when a L2 learner can produce a target structure correctly only in a slow, and not in a fast pace, it is not likely that (s)he can be said to have I-knowledge of it. Fluency is the behavioral correlate of I-knowledge (assumption 6; see also Schmidt's recommendation 7 in this issue).

\(^{14}\) Hulstijn (1989c) has tried to demonstrate the crucial importance of the distinction between Task and Task Conditions in the design of empirical L2 learning investigations.
4.3 Learner characteristics

Every language teacher has experienced that some learners appreciate grammar rules, others detest them. It is not clear whether learners who appreciate grammar, actually profit from it more than learners who don't. Indeed, the literature on the role of learner characteristics such as cognitive style and learner strategies is inconclusive (Abrahams, 1985; Skehan, 1989). To the empirical researcher, two obvious options present themselves:

(1) No additional hypotheses on the role of learner variables such as cognitive style are formulated. However, if learner variables can be simply and reliably measured, they can be treated as control variables. For instance, subjects may be assigned to experimental and control groups according to a principle of matching them on the basis of their scores on a previously administered cognitive style or learning style test.

(2) When an additional hypothesis on the role of a learner variable such as cognitive style is formulated (e.g. that reflective L2 learners benefit more from explicit grammar instruction than impulsive learners), this variable must be incorporated into the design as an independent factor and treated accordingly (i.e. analyses must be conducted for reflective and impulsive subjects separately).

5. Laboratory versus natural settings: a twin approach

In the previous sections we have shown that the empirical study of the influence of grammar instruction on I-knowledge is a complicated affair. One of the most notorious methodological problems is to adequately manipulate all those independent variables which one wants to manipulate while keeping all other potentially interfering variables constant. This is almost impossible in "normal" classrooms with real L2 learners (Harley, this issue). It comes as no surprise, therefore, that the outcomes of such studies often form the object of considerable disagreement. This was recently illustrated by the controversy between Lightbown & Pienemann (1993) and Krashen (1993). Although research in "real" classrooms should bring the final proof because of the alleged validity of the classroom setting, a researcher may want to exert more control than is possible in a normal classroom setting.

There are various ways to bring instruction and learning under control of the researcher. The first way is to control the language to be learned. This can be done by teaching an artificial or partly artificial language (e.g. Issidorides, 1988; Hulstijn, 1989b; Yang & Givón, 1993; DeKeyser, 1994, this issue). The advantage of this method is that the researcher can be certain that no subject in the experiment can have knowledge of the target structures and that performance on tests must stem from learning during the experiment.\(^\text{15}\)

The second way is to control the instruction proper (i.e. the explanation of grammar rules). One way of doing this is to replace the live explanation of the teacher by a prerecorded explanation. Using a computer-controlled learning setting is an obvious way of controlling instruction (Doughty, 1991).

Finally, the researcher may control input in a quantitative way by specifying in advance how much instruction and practice subjects will receive. Here again, a computer-controlled

\(^{15}\) Note that the nine hypotheses advanced in section 3 must ideally be tested with meaningful language materials, not with meaningless, purely formal symbols. Meaningless materials are used in Miniature Artificial Language experiments (e.g. Brain, 1971; McAndrews, Pat & Moscovitch, 1985; Nation & McLaughlin, 1986; the experiments conducted by Reber and associates, reviewed in Reber, 1989). What subjects are to induce in these experiments, are abstract letter patterns generated by a Markov type Finite State Grammar. Such experiments must be distinguished from studies which aim to test hypotheses based on the interface idea, as suggested here. To this extent, we agree with VanPatten (this issue).
learning setting presents itself as an obvious choice (e.g. Chapelle & Jamieson, 1986; Yang & Givón, 1993). The final proof should always be given in a "normal" language class, with "real" teachers and "real" L2 learners, learning a "real" L2 in the framework of a "real" language course, including "natural" exchanges between teacher and learners and among learners themselves (see Van Lier, this issue). However, for methodological reasons it may be recommendable to also conduct more laboratory-type of experiments in which the researcher can exert control of the targets to be learned, the instruction and the execution of receptive and productive tasks, both quantitatively and qualitatively. Natural learning experiments may score high on validity but are likely to score low on reliability. (Semi)artificial learning experiments may score high on reliability but are likely to score low on validity. That is why, in our own research, we adopt a so-called twin approach, combining an artificial or semi-artificial experiment with a "natural" experiment (Hulstijn, 1989a, 1989b, 1992, 1993).

6. An example

In this section we give a brief description of a study currently underway at our department. It is an example of a twin study which aims to test two of the hypotheses based on the position that E-knowledge facilitates the acquisition of I-knowledge under certain conditions. Subjects are university students, who follow an experimental self-study course in an artificial language, based on Esperanto (first experiment), or in a natural language, viz. Spanish (second experiment). For Spanish, four target structures have been selected. Esperanto was supplemented (and accordingly altered) with the same four structures. In this way, the results of the two experiments will be optimally comparable. Individual language learning takes place in a laboratory setting, during 10 sessions of 2 hours each, by means of a computer assisted learning program. This enables us to have a major control over the L2 input and students' activities.

Of the two subject groups in both experimental studies, one does and the other does not receive explicit information (EI) about grammatical features. However, the explicit-implicit distinction only exists in explanation and feedback; when practising, both groups have the same opportunity of focusing their attention on both form and meaning. The treatment contains, apart from output-processing activities, input-processing activities as have been proposed by VanPatten (this issue; see also Ellis, 1993b). In both the Esperanto-based and the Spanish language course, presentation and activities are directed at reaching communicative goals. Subjects' proficiency in the four target structures is tested during and after the course, in various comprehension and production tasks under both monitored and unmonitored conditions.

There is a further possibility that should be mentioned. It consists of the presentation of L2 input with some artificial targets to "real" L2 learners. Although this option is methodologically very attractive, it may not be acceptable for moral reasons. In one experiment, reported in Hulstijn (1993), it was morally acceptable to adopt this approach, however. In this study, which lasted only a half hour for each subject, subjects read a L2 text containing 16 pseudo words without knowing this, and answered comprehension questions. Subjects, who were paid for their out-of-class voluntary participation, were debriefed afterwards. Since the tests session contained some "natural" tasks as well, from which subjects profited (in that they found them very instructive), none of the subjects complained about having been misled during the reading task.

This research project, entitled "The Influence of Grammar Instruction on Foreign Language Acquisition", is supported by the Foundation For Language, Speech and Logic (under grant 300-173-028), which is funded by the Netherlands Organization for Scientific Research (NWO). The research is conducted by Rick de Graaff and supervised by Jan Hulstijn (May 1993-May 1997).

Subjects are pretested on verbal aptitude and learning style, and assigned to the experimental groups on the basis of a matching principle (cf. section 4.3. option 1).
In connection with the dimensions presented in section 3, we distinguish between one simple and one complex lexical morphological rule, and one simple and one complex non-lexical syntactic rule. The distinction between lexical and non-lexical features is motivated by our aim to investigate the interface position in potential coalition with item learning as well as in isolation. The distinction between simple and complex features is based on the declarative cognitive notions as explained in 3.2.

The following four target structures of Spanish have been selected:
1. Plural inflection of the noun as the simple lexical rule.
2. Verb inflection for the imperative as the complex lexical rule.
3. Position of the negation particle as the simple non-lexical rule.
4. Position of the object and the use of the object clitic as the complex non-lexical rule.

In conclusion, the aim of the study is to investigate how and under what circumstances, test performance is influenced positively by providing explicit information on structural features in comparison with the non-provision of such information. On the basis of the general research question concerning the interface hypothesis, the analyses are geared in particular towards the testing of the following hypotheses:
* The positive effect of EI will be more substantial in the case of the acquisition of complex rules than in the case of the acquisition of simple rules (cf. H3 in section 3.2).
* The positive effect of EI will be more substantial in the case of the non-lexical, syntactic rules than in the case of the lexical, morphological rules (cf. H5 section 3.4).

To accommodate the moderating factor of fluency, the following additional hypothesis is tested:
* A fluent, correct command of the four grammatical structures will be attained sooner by the group receiving EI than by the group not receiving EI, not only in the Plus Monitoring condition, but also in the Minus Monitoring condition of the proficiency tests (cf. section 4.2).

Conclusions

One of the most nagging questions which has haunted researchers and practitioners alike for a long time, is the question of whether grammar instruction aids L2 acquisition or not. The aim of this contribution was to show that it is highly unlikely that this question can ever be answered with an unqualified "yes" or "no". With the aid of the distinctions proposed by Schmidt (this issue) between explicit knowledge, explicit learning and explicit instruction, and working in the framework of an interface position between explicit and implicit knowledge, we have tried to demonstrate that it is possible to formulate a fair number of hypotheses, specifying under what conditions explicit knowledge may facilitate the acquisition of implicit knowledge. Our second aim was to demonstrate that it is possible to go beyond the stage of claiming (on the basis of acquisition theories or pedagogical experiences and intuitions) that explicit knowledge may or may not help acquisition. Many hypotheses can and should be tested empirically. The preceding sections lay out a vast research program. It is our hunch that the results of empirical studies, conducted in laboratories as well as in "real" classrooms, will replace the unqualified interface and noninterface positions, which have hitherto dominated the debate, by qualified positions. These conditions depend on factors such as linguistic domain, complexity, scope and reliability, and semantic or formal redundancy of the target structures. Furthermore, they depend on factors such as rule versus item learning, reception versus production. And finally, they depend on moderating factors such as type of instruction, fluency of performance, and learner characteristics.

References


