European Journal of Cognitive Psychology

ISSN: 0954-1446 (Print) 1464-0635 (Online) Journal homepage: http://www.tandfonline.com/loi/pecp20

Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge

Nick Ellis

To cite this article: Nick Ellis (1993) Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge, European Journal of Cognitive Psychology, 5:3, 289-318, DOI: 10.1080/09541449308520120

To link to this article: http://dx.doi.org/10.1080/09541449308520120

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 1993, 5 (3) 289–318

Rules and Instances in Foreign Language Learning: Interactions of Explicit and Implicit Knowledge

Nick Ellis

Department of Psychology, University College of North Wales, Bangor, UK

This paper reports a study of implicit and explicit learning of second language (L2) grammatical forms, the "soft"-mutations of Welsh. "Random" learners saw randomly ordered instances. "Rule" learners first learned the rules. "Rule&Instances" learners saw the rules applied to instances. Initial learning, generalisations to new words and constructions, implicit fast performance in a well-formedness RT decision task, and explicit knowledge of the rules were recorded. Analyses of over 71,000 language trials demonstrate: (1) "Random" learners quickly achieve competence on original learning material, but show little implicit learning, performing poorly on well-formedness (or "grammaticality") judgements, and have poor acquisition of explicit knowledge of the underlying rule-structure. (2) "Rule" learners take many trials to learn the rules but this facilitates their understanding of the natural language. However, they often know rules explicitly, yet fail to apply them in practice. Explicit and implicit knowledge are doubly dissociated. (3) Initially, "Rule&Instances" learners learn slowest. However, they alone abstract a working knowledge of soft-mutations. When exposed to new constructions, they generalise and are able both to explicitly formulate the new rules and succeed on implicit well-formedness judgements.

INTRODUCTION

Humans have devised ways of flying and accordingly discuss the mechanics. With swallows it is otherwise. And since swallows know nothing of technique, it never shows, never obtrudes upon their intent, joyous courses, from the day they leave their nest till their brightness falls from the air. (Adams, 1980, p. 197)

Requests for reprints should be addressed to Nick Ellis, Department of Psychology, University College of North Wales, Bangor, Gwynedd LL57 2DG, UK.

I would like to thank Sue Sinclair, Gordon Brown, Carl James, Dianne Berry and Alan Beaton for useful discussions concerning this research project, Annette de Groot and Chris Barry for constructive editorial comments, and my past students, Sandy Britain, Susan Rogowski and Lindsey Hulme, for help in running the experiments.

Some things we just can do, like using our native language (L1), riding a bike, or dreaming. We have little insight into the nature of the processing involved: like swallows fly, we just do them. Other of our abilities depend on our knowing *how* to do them, like speaking pig Latin, multiplication, or cooking from a recipe. Understanding the relationships between implicit and explicit knowledge is a central goal of psychology.

We acquire L1 unconsciously by engaging in natural meaningful communication. From this "evidence", we automatically assimilate complex knowledge of syntactic and morphological structures (Chomsky, 1986). Yet paradoxically, we cannot describe these rules (Seliger, 1979), the discovery of which forms the object of the entire discipline of theoretical linguistics. Some second language (L2) teaching methods are guided by this and renounce explicit grammar-based instruction. These views hold that acquired (implicit) knowledge and learnt (explicit) knowledge are stored separately and that learnt knowledge cannot be converted into acquired knowledge. However, older children and adults can acquire and act upon rules and schemata; they can, for example, be taught grammatical rules for forming a plural. Thus other L2 teaching approaches are heavily rule-based and hold that explicit knowledge is a necessary, or at any rate a desirable, precursor of implicit knowledge.

Radical swings in the history of L2 teaching methodologies reflect this schism (for reviews, see R. Ellis, 1990; Kelly, 1969). Traditional "grammar-translation" methods emphasised study by literacy and translation and had an *explicit* bias with formal explanation of L2 rules and a deductive approach to learning. Come the Second World War, the Behaviourist Zeitgeist in America led to structural approaches and audiolingual methods which outlawed the teaching of metalinguistic rules and which regarded L2 as just another specific domain to be understood by general laws of learning: L2 acquisition involved discrimination and generalisation from structured examples by analogy not analysis, i.e. *implicit*, inductive learning through patterned practice. By the 1960s, critics began to observe that these methods produce fluent but flawed speakers (e.g. "Audiolingual methods have been teaching speech but not language": Donaldson, 1971, p. 123) and *explicit* instruction of grammatical rules was reintroduced in the Cognitive Code Method, "a modified, up-to-date translation theory" (Carroll, 1966, p. 102), which held that perception and awareness of L2 rules precede their use. In the 1970s and 1980s, the pendulum swung back to naturalistic methods (Krashen, 1982; 1985). Krashen's underlying theory, the input hypothesis, is a non-interface position in that it posits that adults can subconsciously acquire languages and they can consciously learn about language. But in this view, learning cannot be converted into acquisition; subconscious acquisition dominates in L2 performance, and conscious learning is used only as a monitor, i.e. as an editor to correct output after it has been initiated by the acquired system. Thus in Krashen's Monitor Theory, *implicit* acquisition of L2 is the essential aim of instructional programmes. Currently, the pendulum is yet again in swing: in the light of analyses of the disappointing abilities of graduates from "grammar-free" foreign language (FL) programmes (Gomes da Torre, 1985), there are new calls for a return to explicit methods (Cox, 1989; James, 1986; Kingman, 1988).

Such swings in educational practice make it clear that there is no simple answer to which of these methods is "best". There is thus a need for empirical investigation into the roles and interactions of implicit and explicit knowledge in the acquisition of foreign and second languages. The present experiment, therefore, investigates *ab initio* students learning an L2. In doing this, for the sake of logging a complete history of language exposure for each learner, the target language was *written* Welsh, which was presented on a computer and the learners' role was predominantly receptive—they were to understand the phrases and show their understanding by translating them into their native tongue. Furthermore, in order to allow sufficient exemplars for abstraction and generalisation, the sample utterances, although meaningful in their own right, were not themselves sequenced in any order which made for overall communicative sense. For these reasons, readers from a more "linguistic" background are likely to find wanting the ecological validity of the present operational definition of "naturalistic learning". While I acknowledge this and caution the reader to remember the same, I believe that there is an important role for this style of investigation which aims to bring together psychological and linguistic theory and application. It is by no means the best or only approach, but it is a useful adjunct to the multiplicity of approaches that must be adopted for a complete understanding of the cognitive processes involved in L2 learning and acquisition (McLaughlin, 1990). Note also that it is only the cognitive processes that concern us here—these studies are blind to the many motivational, instrumental and attitudinal factors that are crucially important in determining learners' progress (Naiman, Fröhlich, Stern, & Todesco, 1978; Krashen, 1982; 1985; Lambert, 1974).

The present study, therefore, compares the effectiveness of explicit (Rule), structured (Rule&Instances) and implicit (Random) programmes of exposure to a complicated rule structure in Welsh, the soft mutation. There are three types of mutation in Welsh—the soft, nasal and aspirate mutations. All of these cause regular changes of certain word-initial consonants and they are triggered by particular grammatical contexts (see Ball, 1988, for descriptions of the mutations and contexts). A wide variety of contexts call for a soft mutation of the initial consonants of nouns, thus t->d, c->g, p->b, m->f, d->dd, etc., and learners have to find out about this quickly or else they cannot even use a dictionary. In phonological terms, different classes of consonant undergo different changes. The voiceless plosives /p, t, k/ are replaced by the voiced plosives /b, d, g/. Similarly, the voiceless alveolar trill /r/ and the voiceless lateral fricative /ɬ/ are replaced by their voiced counterparts /r/ and /l/. The voiced plosives /b, d/ are replaced by the homorganic fricatives /v, ð/. The voiced plosive /g/ is deleted. Finally, the nasal /m/ is replaced by the homorganic fricative /v/.

In grammatical terms, there is a wide variety of contexts which cause the soft mutation. Some of these are very specific, for example: (1) after the personal pronoun meaning "his", "son" translates as "mab" but "his son" is "ei fab o"; (2) after the preposition "o" meaning "from", the local town is "Bangor" but one would come "o Fangor". Others are very general; for example, when a feminine singular noun follows the definite article "y", "gwraig"—meaning "housewife"—becomes "y wraig". Thus the "soft mutation of Welsh" is a complicated rule system: the learner needs to know of (1) its existence, (2) the content of the rule system, i.e. the set of letters which mutate and their mutated equivalents, and (3) the contexts which call for this mutation. Like many aspects of grammar, it looks remarkably complicated when described thus, yet native language speakers do it flawlessly and unconsciously.

METHOD

The present experiment compares three regimes of instruction in the task of translating written Welsh into written English in a sequential learning paradigm. In outline, the three styles of instruction were as follows:

The *Random* group was taught by exposure to instances—they saw Welsh words beginning with one of 8 initial letters (5 mutate, 3 do not), two examples of each. They saw these in a random order, sometimes alone, sometimes in a construction that does not trigger a mutation, and sometimes in a construction that does cause a mutation. This group constitutes our broad operational definition of implicit or "naturalistic" learning in that the learners are trying to comprehend meaningful utterances. Of course, this falls far short of Krashen's (1985, p. 4) definition:

> people acquire second languages only if they obtain comprehensible input and if their affective filters are low enough to let the input "in". When the filter is "down" and appropriate comprehensible input is presented (and comprehended), acquisition is inevitable . . . In other words, comprehensible input is the essential ingredient for second-language acquisition. All other

factors thought to encourage second-language acquisition work only when they contribute to comprehensible input and/or a low affective filter.

Although the utterances in this experiment are understandable, the learning situation as a whole is very artificial. On the plus side, however, it could be argued that the content of the phases for the random group in the present experiment is potentially more conducive to abstraction of the underlying rules than truly "naturalistic" exposure since, although they are randomly ordered, the collection of utterances are all pertinent to one particular class of grammatical form, thus allowing abstraction—truly "naturalistic" Welsh conversation, however it might be made comprehensible, contains such a scattering of the variety of forms that it is difficult to conceive of a mechanism which could analyse which instances pertain to which structures (although such a mechanism surely does exist in the native learner of L1: Pinker, 1987).

Operationalising a *Rule* group is somewhat easier. These subjects first learned the vocabulary of 16 words. Next, they were explicitly taught the eight rules of content of the soft mutation to a criterion of complete correctness. The instruction was given by means of a computer and comprised simply the status of initial letters in mutating and non-mutating contexts. At no point did we give learners explicit statements (of the sort used in the Introduction above) about the role of mutations in Welsh, such as: "The first letter of a noun may change into another letter. This is the case in constructions such as . . . , but not in the case of constructions such as . . . , After constructions such as . . . , the following letter changes occur . . . etc." Rather, as described in the Procedure, we simply showed them "protowords" (wildcard words beginning with a given letter) either changing their initial letter in mutating constructions or not in non-mutating constructions and had the subjects learn the corresponding initial letter of the lemma.

Having learned the vocabulary and the rules, they then transferred to the decoding of these items in mutating and non-mutating constructions in the same "naturalistic" exposures as the random group.

The *Rule&Instances* group was first taught each rule and how it applied to two instances of vocabulary (this treatment being influenced by the work on structure of exposure in schema abstraction). Once they had learned these rules and exemplars of their operation, they also then transferred to the decoding of these items in mutating and non-mutating constructions in the same "naturalistic" exposures as the Random group.

Having run these three groups, we realised that the Rule&Instances and Rule learners differed from the Random group in both the structure *and the amount* of exposure to language. It was therefore necessary to run a fourth condition, the *Yoked Random* group, whose amount of exposure was matched (yoked) to the Rule and Rule&Instances groups. These subjects received the same content as the Random group, but they were forced to translate the initial learning material for several more passes than was necessary for them to learn it—thus seeing the material as much as the Rule and Rule&Instances learners. Comparisons of the performance of the Yoked Random group with that of the Rule and Rule&Instances groups informs us about the effects of *method*; comparisons between the Yoked Random group and the Random group informs about effects of *amount* of exposure.

In brief, beginners, assessed for their language learning competence using the Modern Language Aptitude Test (MLAT), were assigned to one of the four matched groups. Tuition was computer-based to allow the recording of a detailed record of how many times the learner has seen any particular word before, in what contexts, for how long, etc. The dependent variables were the number of learning trials to criterion, the learning curve, the types of error, the time spent on each trial and stage of translation, etc. Also assessed was generalisation (1) to new words with similar initial letters and (2) to new mutating and non-mutating constructions. After each stage the subjects were tested for their implicit fast performance of the rules in a reaction time well-formedness decision task, as well as their explicit knowledge of the rules.

Subjects

The subjects, recruited by advertisements on local notice-boards, were paid £2.50 per hour for their involvement. They were allocated to groups on the basis of their scores on the Modern Language Aptitude Test (MLAT: Carroll & Sapon, 1955) to ensure four matched groups. There were 15 subjects in the Rule group, 14 in each of the Rule&Instances and Random groups and 8 in the Yoked Random group. The average (± SD) age was 24.2 ± 9.7 years. The average (± SD) MLAT scores were as follows: Sentence Comprehension 22.1 ± 7.9; Words in Sentences 18.7 ± 6.9; Paired Associate Learning 15.8 ± 5.1; Total 56.7 ± 14.1. The average number of foreign languages spoken was 1.6 ± 1.1. The four groups were well matched on all of these measures, there being far from significant group effects on both one-way ANOVA and Kruskal-Wallis testing.

Procedure

The learning trials were all presented on a Macintosh computer. On each learning trial, a Welsh phrase was shown on the screen and the subject was to type in the appropriate English translation. If the response was null (a simple "return") or in error, then the correct translation was shown and the subject invited to study it; when they had finished study, they pressed "return" and were asked to try the translation again. This sequence was reiterated until the subject was correct on that phrase; thus on any trial, the subject had to respond correctly to proceed. On each trial, the following data were recorded: the response, whether it was correct, the latency of the first letter of each word of the response and of the total response, and, if the initial response was incorrect, the study time on the correct translation and the number of retakes to correct response.

There were a number of different learning phases, each consisting of a number of trials. The stimuli for the phase were presented in random order and the subject went through the complete set of stimuli for that learning phase. They repeated sessions on each phase until they had reached the criterion of a completely correct run with their initial responses to the stimuli, whereupon they moved to the next phase in the sequence. There were also a number of test phases which were the same in every respect except that they were not given feedback if their responses were erroneous and they only participated once in each test phase.

The subjects worked through the sequence of sessions and phases usually for 1 h at a sitting, taking between three and eight sittings in all.

The Rule group had two special initial phases (vocabulary and rule learning phases), and the Rule&Instances group had their particular structured Rule&Instances introduction. Thereafter, the phase sequence was the same for all three groups and included exposure to randomly ordered language, tests of implicit and explicit knowledge, and generalisations to new vocabulary in known constructions and new but analogous constructions. These sequences are as follows.

The Rule Group's Special Introductory Phases

Vocabulary Learning Phase. Here the Rule group learned 16 words of vocabulary, two words each starting with the letter *t, d, c, h, s, p, a, m* (e.g. *trwyn* ->"nose", *tafod* ->"tongue"). The sequential learning procedure was as follows. In each session, the words were presented sequentially in a random order. For each word, the learner had to type in the correct translation. If they did not know it, if they made a mistake, or guessed or simply pressed "return", the correct translation was shown. The subjects could study this for as long as they wished, but they had to type in the correct response before they could proceed with the next trial. A session thus comprised the corpus of words being shown in this fashion. If the learner had made a mistake on any of the words, the session was then repeated. If, however, they had made no errors on the 16 translations, they graduated that phase and moved on to the next. Thus the phase continued until the learner translated all 16 words correctly on their first presentation in a session (the sequential learning criterion).

Rules Learning Phase. Next, the Rule group learned the rules of letter soft-mutation as exemplified in the constructions used in early phases. "Protowords" were made starting with each of the eight letters and followed by three equal signs to denote wildcards (e.g. *t* === which represents any word beginning with *t*). The protowords were shown either (1) alone (e.g. *t* === , where the required response was "t === " to denote that a word on its own beginning with *t* does not mutate), (2) in a non-mutating construction (e.g. *ble mae t* === , which requires the response "where is a t === " to denote that *ble mae* translates as "where is" and that it does not cause a soft mutation for words beginning with the letter *t*) and (3) in a mutating construction (e.g. *ei d* === *o*, which requires the response "his t === " to denote that *ei* . . . *o* translates as "his" and that this construction causes a soft mutation and that the initial consonant *t* soft-mutates to *d*, thus the word in its root form would have begun with the letter *t*). There were thus 24 trials to this phase and these were presented in random order using the same sequential learning procedures as used in the vocabulary learning phase. The phase continued until the learner correctly completed all 24 trials correctly on their first presentation in a session.

The Rule&Instances Group's Special Introductory Phase

Rule&Instances Learning Phase. This structured instruction was particular to the Rule&Instances group and comprised 64 trials which were arranged into eight blocks, one for each letter. The sequence of eight trials within each block was constant: two exemplars of that letter were shown (e.g. *trwyn* -> "nose"; *tafod* -> "tongue"), next the non-mutating construction with the relevant protoword and then the two exemplars (e.g. *ble mae t* === -> "where is a t === "; *ble mae trwyn* -> "where is a nose"; *ble mae tafod* -> "where is a tongue"), and finally the mutating construction with the relevant protoword and the two exemplars (e.g. *ei d* === *o* -> "his t === "; *ei drwyn o* -> "his nose"; *ei dafod o* -> "his tongue").[1] Although the sequence within each block remained constant, the blocks themselves were randomly ordered within each presentation of the phase. The sequential learning criterion again applied: the phase continued until the learner completed all 64 trials correctly on their first presentation in a session.

[1]In these examples, the stimuli are shown in italic, the symbol -> should be read as "requiring the response", and the correct response is then shown in quotation marks.

Phases Common to All Groups

Thereafter, all groups shared the common sequence of phases as follows.

Natural Learning Phase 1. This comprised a random sequence of 48 trials where the 16 exemplars were shown alone (e.g. *trwyn* -> "nose"), in the non-mutating construction (e.g. *ble mae trwyn* -> "where is a nose") and in the mutating construction (e.g. *ei drwyn o* -> "his nose"). The sequential learning criterion applied.

Rule Test Phase 1. All of the test phases involved just one session (one pass through the stimuli), and no feedback as to correctness of response. As rule test phase 1 was designed to test explicit awareness of the rule-structure underlying the mutations, it was identical to the above rules learning phase except that no feedback was given and the subjects were tested just once.

Well-formedness Test Phase 1. This was designed to elicit fast spontaneous grammaticality judgements. There were 64 randomly ordered trials where each of the 16 exemplar words appeared correctly not-mutating in the non-mutating construction (*ble mae trwyn*), incorrectly mutating in this construction (*ble mae drwyn*), correctly mutating in the mutating construction (*ei drwyn o*) and incorrectly not-mutating in the mutating construction (*ei trwyn o*). In a practice session, the subjects saw grammatically well-formed (e.g. "he had an egg", "he runs down the road") and ill-formed (e.g. "they was ready", "four toy") strings and were asked to press "Y" if the string was good, or "N" otherwise. They were asked to make this judgement as quickly as possible. When they were happy with this procedure, they then moved on to the same judgements for the 64 Welsh stimuli. No feedback was given and there was only one session.

Natural Learning Phase 2. This phase investigated generalisation to new vocabulary. It was identical to learning phase 1 in that the same constructions were used; however, it differed in that there were now two new exemplars for each of the eight initial letters (e.g. *troed* -> "foot", *tocyn* -> "ticket"). There were therefore again 48 trials in each session. The sequential learning criterion applied.

Rule Test Phase 2. Identical to rule phase 1.

Natural Learning Phase 3. This phase investigated generalisation to new constructions. The vocabulary and design was therefore the same as in natural learning phase 2 except that there were two new constructions, one which did not cause mutation (*eich* === *chi* -> "your === ") and one which did (*'sgynnoch chi* === -> "do you have a === "). There were therefore 48 trials per session. The sequential learning criterion applied.

Rule Test Phase 3. Here explicit knowledge of all four of the rules was tested for each of the eight protowords (e.g. *ble mae t* === , *ei d* ===*o*, *eich t* === *chi*, *'sgynnoch chi d* ===). The 32 trials were presented just once, in random order, with no feedback.

Well-formedness Test Phase 2. Well-formedness test phase 1 assessed implicit awareness of grammatical correctness where the subjects had seen the well-formed strings on several prior occasions. In contrast, this phase was designed to tap fast judgements of grammaticality for constructions on which subjects had just been successful in natural learning phase 3, and exemplars where they had been successful in natural learning phase 1, *but* where they had never seen these exemplars in these constructions before. The 64 trials therefore followed the structure of well-formedness test phase 1, using the 16 words of vocabulary from natural learning phase 1, but the new constructions from natural learning phase 3 (*eich* === *chi*, *'sgynnoch chi* ===). The subjects' judgements could therefore not be performed on the basis of their familiarity with these particular strings, and thus the phase is designed to test fast, implicit judgements based on the generalisation and synthesis of different sources of knowledge. No feedback was given and there was only one session.

Rule Learning Phase 1. This phase determined the extent to which learned rules have been retained (in the Rule and Rule&Instances groups), and any potential savings in rule learning consequent on natural exposure in the Natural group. The stimuli and procedure were those of rule test phase 3, explicit knowledge of all of the rules being tapped for each of the eight protowords (e.g. *ble mae t* === , *ei d* === *o*, *eich t* === *chi*, *'sgynnoch chi d* ===). However, feedback and instruction was now given on each error trial, the sequential learning criterion applied, and the subjects thus repeated the session until one completely correct run.

RESULTS

There were over 71,000 trials. This section will report the bare bones of the analyses as briefly as possible. The major dependent variable for the learning phases is the number of errors made in that phase. The more errors, the more the total trials needed to reach the criterion of one totally correct pass, the more sessions and the longer the time taken. A reminder of total learning history is provided by the cumulative trials taken to the end of any particular learning phase. For the rule test phases where there was neither feedback nor repetition, the dependent variables are number of error trials (along with a breakdown by construction) and time taken. For the well-formedness tests, we report the average accuracy and latency of response for each construction presented with both correct and incorrect mutations. For each of these dependent variables, group differences were assessed either by one-way ANOVA or factorial ANOVA with repeated measures across constructions and *post-hoc* Scheffé pairwise comparisons. Thus in Table 1, for example, the column "P of F on one-way ANOVA" concerns tests between the four groups on that measure. As some of the data have skewed distributions, these results are cross-checked against Kruskal-Wallis analyses of the same data, with K if $P < 0.05$ and KK if $P < 0.01$ for the group difference. The columns for Scheffé contrasts contain an asterisk if that pairwise contrast is significant at $P < 0.05$.

Performance on the MLAT (especially the paired associate learning subtest) and the number of foreign languages spoken correlated with scores on the learning and rule test phases. Although the group differences on these phases are even more significant when analysed by ANCOVA with MLAT and number of foreign languages as covariates, we here report the ANOVAs since they are more succinct.

Vocabulary Learning Phase. The Rule group learned the 16 words of vocabulary in 6.6 sessions on average. The mean number of trials was 105.6, of which 38.0 were in error. This took a mean of 18.7 min. They thus needed an average exposure of 6.6 trials per word, but they were correct on the majority of trials, averaging only 2.38 error trials per word. With this procedure, it took something over 1 min of study to learn each word.

Rules Learning Phase. In this phase, the Rule group learned the rules of letter mutation for the protowords alone, in one mutating and in one non-mutating construction. They took on average 8.5 sessions to do this in 38.9 min, with 41.1 errors out of 204.8 trials, the majority of these occurring for the mutating letters in the mutating construction.

Rule&Instances Learning Phase. The Rule&Instances group took 10.1 sessions to pass this block. This was far longer than the Rule group took on their two separate components (649.1 *vs* 310 trials; 112.6 *vs* 57.6 min). This was in part attributable to the forced lengthy structured procedure, which entailed the subjects in the Rule&Instances group performing correctly on trials where they had already assimilated the correct

TABLE 1
Performance Measures for the Four Groups on the Natural Learning Phases (see Text for Details)

Measure	Rule	Rule& Instances	Random	Yoked Random	P of F on One-way ANOVA	Pairwise Comparisons: Scheffé Rule ≠ R&I	Rule ≠ Random	Rule ≠ Yoked	R&I ≠ Random	R&I ≠ Yoked	Random ≠ Yoked
Natural learning phase 1											
Error trials (n)	17.0	2.1	53.9	130.5	0.0001 *KK*		•	•	•	•	•
Trials (n)	147.2	96.0	284.6	552.0	0.0001 *KK*		•	•	•	•	•
Time (min)	23.0	13.5	56.1	104.4	0.0001 *KK*		•	•	•	•	•
Sessions (n)	3.1	2.0	5.9	11.5	0.0001 *KK*		•	•	•	•	•
Cumulative trials	457.6	754.1	284.6	552.0	0.0002 *KK*	•			•		
Natural learning phase 2 (new vocabulary)											
Error trials (n)	51.0	46.1	53.0	110.3	0.006			•		•	•
Trials (n)	249.6	233.1	267.4	522.0	0.04						
Time (min)	39.4	42.6	43.2	66.1	NS						
Sessions (n)	5.2	4.9	5.6	10.9	0.05						
Cumulative trials	755.2	1026.3	600.0	1122.0	0.005 *KK*				•		•
Natural learning phase 3 (new constructs)											
Error trials (n)	4.7	8.6	13.1	22.0	0.01 *KK*		•	•		•	
Trials (n)	134.4	198.9	216.0	222.0	NS						
Time (min)	20.9	33.9	41.7	30.3	NS						
Sessions (n)	2.8	4.1	4.5	4.6	NS						
Cumulative trials	913.6	1249.1	840.0	1368.0	0.01 *KK*						•

information—the total number of error trials for this group was 75.5, very similar to the 79.1 for the Rule group over their two component phases.

Natural Learning Phase 1. This comprised a random sequence of 48 trials where the 16 exemplars were shown alone (e.g. *trwyn* ->"nose"), in the non-mutating construction (e.g. *ble mae trwyn* ->"where is a nose") and in the mutating construction (e.g. *ei drwyn o* ->"his nose"). This is the first exposure to the language for the Random groups. The performance measures for all four groups on natural learning phase 1 (and the two subsequent natural learning phases) are shown in Table 1.

There were considerable savings for the Rule&Instances and Rule groups from their earlier sessions. Both of these groups passed through this phase in less time, with fewer sessions and far fewer error trials. The Rule&Instances group had a particular advantage over the Rule group making, for example, only 2.1 errors on this phase. The Yoked Random group was forced to take between 11 and 12 sessions on this phase, overlearning the material even though the subjects in this group had met the criterion of correctness, in order that they were matched to the subjects in the Rule&Instances group on exposure (albeit unstructured) to the language strings; by the end of this phase, they had taken 11.5 sessions compared with the subjects in the Rule&Instances group's 12 (10 from their special initial phase and 2 here). Thus the Yoked control group had equivalent *amounts* of exposure to the language strings as the Rule&Instances and Rule groups (NS; Scheffé contrasts on cumulative trials between Yoked control and these groups by the end of this phase), but differed in the lack of structure to this exposure.

There are many instances in the Rule group's data of the dissociation between explicit and implicit knowledge. The subjects in the Rule group had learned the vocabulary in their first phase, they had explicit knowledge of the rules from their second stage, and yet they would often fail to apply this knowledge in this natural learning phase, making on average 17.0 errors.

If, however, we consider the total number of trials to get to this stage of competence, we discover that the Random group appeared to be much faster (285 *vs* 458 and 754 trials for the Rule and Rule&Instances groups, respectively) and more accurate (54 *vs* 96 and 78 cumulative errors for the Rule and Rule&Instances groups, respectively). If one was assessing learning just from the ability to deal with the particular instances that the learner has been exposed to, a criterion oft applied in the classroom concerning the repetition and use of formulaic patterns, then we must advocate the Random, more "naturalist" method as promoting faster acquisition. However, as will be seen in the next two sections, the Random group's fluency which had been acquired quickly by association brings with it little by way of either explicit learning of the underlying structures or generalisation.

Rule Test Phase 1. The performance of the four groups in the first and subsequent rule test phases is described in Table 2. A two-way mixed analysis of variance (4 groups × 2 constructions) on the subjects' mean percent correct on that construction in rule test phase 1 demonstrates significant main effects of group [$F(3,47)=33.25$, $P < 0.001$] and construction [$F(1,47)=26.64$, $P < 0.001$] and a significant group × construction interaction [$F(3,47)=6.68$, $P < 0.005$], whereby the two Random groups were considerably worse on the soft-mutating rule.

Both the Rule and Rule&Instances groups knew the rules quite well, of course—they had been explicitly taught them in earlier sessions. The Random group had not, however, abstracted the rules, making 9.3 errors on the test, the vast majority of which occurred on the mutating construction where the subjects were only 37% correct. As only five of the eight initial letters mutated in this construction (i.e. 37.5% did not), it is clear that the subjects in the Random group had no explicit conception of the structure of the soft-mutation rules after their "natural" exposure. Even though the Yoked control group had as much language exposure as the Rule&Instances and Rule groups, this mere exposure had not allowed the subjects in this group to abstract conscious knowledge of the underlying rule structure; indeed, it was this group which was the most ignorant of explicit rule structure.

Well-formedness Test 1. This was designed to elicit fast judgements of grammatical correctness. There were 64 randomly ordered trials where each of the 16 exemplar words appeared in four conditions: correctly non-mutating in the non-mutating construction (*ble mae trwyn*), incorrectly mutating in this construction (*ble mae drwyn*), correctly mutating in the mutating construction (*ei drwyn o*) and incorrectly non-mutating in the mutating construction (*ei trwyn o*). The performance measures for all four groups on both well-formedness tests are shown in Table 3.

The raw binary (correct/incorrect) data for just the five mutating letters were analysed as a repeated measures MANOVA with the five letters, the two replicates of each letter and the four conditions as within-subjects effects and subjects nested within groups. There was a significant main effect of group [$F(3,46)=8.08$, $P < 0.001$], a significant effect of condition [$F(3,138)=35.9$, $P < 0.001$][2] and a marginally significant group × condi-

[2]This and subsequent MANOVAs were performed using the MANOVA procedure in SPSS-X. The effects concerning within-subjects factors are averaged tests of significance against unique within-cells sums of squares (see Norusis, 1985, ch. 7).

tion interaction [$F(9,138)=1.68$, $P < 0.10$]. On each of these types of judgement, the Rule&Instances group performed best, with very few errors. The Rule group performed next well, at levels close to the Rule&Instances group on the correct forms, but close to chance level on judging that incorrect forms were ungrammatical. This profile of performance was also seen for the Random group, which performed "on the other side of chance" on the incorrect forms, i.e. the subjects in this group were inclined to accept ungrammatical forms as correct. The Yoked Random group performed very much like the Random group.

Overall, there was little difference between the groups on accurately accepting correct responses (both mutated and non-mutated), but the Rule and both Random groups were particularly inaccurate in failing to identify incorrect forms as being ungrammatical, this being especially so for the Random groups. It is only the Rule&Instances group which had an appreciation of incorrect grammar.

In order to ascertain whether any superiority of performance was due to a slower, more careful style of analysis in the Rule&Instances group (e.g. using explicit knowledge to monitor the strings such that the task might be "implicit" for the Random groups but tap laboured "explicit" processing in the others), a similar MANOVA design to that described above was used to analyse the total time taken for each response. There was a significant effect of condition [$F(3,111)=4.14$, $P < 0.01$], but no other significant main effects. Thus the superior performance of the Rule&Instances group, and to a lesser extent the Rule group, does not seem to depend on a different speed or style of processing.

Natural Learning Phase 2. This phase investigated generalisation to new vocabulary. It was identical to natural learning phase 1 in that the same constructions were used, but it differed in that there were now two new vocabulary exemplars for each of the eight initial letters (e.g. *troed* -> "foot", *tocyn* -> "ticket"). In Table 1, it can be seen that there were no significant differences between the performances of the three main groups on any of the measures for the content of this phase. Thus the demonstrated superior implicit and explicit knowledge of the structural patterns of language of the Rule&Instances group did not facilitate their passage through this new phase. This suggests that the major learning that occurs in this phase concerns vocabulary alone, a conclusion supported by the fact that the Random group took nearly as many trials (n = 267) and error trials (n = 53) to pass this phase as they did on its prior equivalent natural learning phase 1 (n = 285 and 54, respectively), and that these measures for the Rule and Rule&Instances groups rose to the same levels.

The cumulative errors remind us that the superior knowledge of structure in the Rule&Instances group was hard-bought in their many learning trials in the Rule&Instances learning phase.

TABLE 2
Performance Measures for the Four Groups on the Rule Test Phases and on the Final Rule Learning Phase (see Text for Details)

Construct	*Rule*	*Rule& Instances*	*Random*	*Yoked Random*	*Pairwise Comparisons: Scheffé* Rule ≠ R&I	Rule ≠ Random	Rule ≠ Yoked	R&I ≠ Random	R&I ≠ Yoked	Random ≠ Yoked
Rule test phase 1 (% correct by construct)										
ble mae === (no mutation)	93	94	71	61		*	*	*	*	
ei === o (mutation)	93	85	37	41		*	*	*	*	
Rule test phase 2 (% correct by construct)										
ble mae === (no mutation)	89	95	67	78		*		*	*	
ei === o (mutation)	77	86	54	56		*	*	*	*	
Rule test phase 3 (% correct by construct)										
ble mae === (no mutation)	90	92	80	81		*		*	*	
eich === chi (no mutation)	88	85	80	77			*			
ei === o (mutation)	90	88	73	72		*	*	*	*	
'sgynnoch chi === (mutation)	78	87	68	60		*	*	*	*	

Construct	*Rule*	*Rule& Instances*	*Random*	*Yoked Random*	*P of F on One-way ANOVA*	*Pairwise Comparisons: Scheffé*					
						Rule ≠ R&I	*Rule ≠ Random*	*Rule ≠ Yoked*	*R&I ≠ Random*	*R&I ≠ Yoked*	*Random ≠ Yoked*
Rule learning phase 1 (mean no. of errors per session to correct session criterion by construct and group)											
ble mae === (no mutation)	0.49	0.40	1.00	0.96	0.001		•	*	*	•	
ble mae === (no mutation)	0.92	0.95	1.20	1.21	NS						
ei === o (mutation)	0.56	0.72	1.22	1.59	0.01			•		•	
'sgynnoch chi === (mutation)	1.32	1.02	1.66	2.61	0.001			•		•	

TABLE 3
Performance Measures for the Four Groups on Well-formedness Tests 1 and 2

Condition	*Rule*	*Rule & Instances*	*Random*	*Yoked Random*	*P of F on One-way ANOVA*
Well-formedness test 1					
% correct on the words starting with the five mutating letters					
ble mae === (no mutation, correct)	91	95	79	86	0.03
ble mae === (mutating, incorrect)	61	80	43	60	0.01
ei === o (mutating, correct)	83	91	79	78	NS
ei === o (no mutation, incorrect)	47	78	36	39	0.001
Time (sec) on the words starting with the five mutating letters					
ble mae === (no mutation, correct)	3.45	3.10	3.10	2.55	NS
ble mae === (mutating, incorrect)	3.98	3.53	3.40	2.82	NS
ei === o (mutating, correct)	3.92	3.98	3.35	2.50	NS
ei === o (no mutation, incorrect)	4.73	3.85	3.20	2.58	NS
Well-formedness test 2					
% correct on the words starting with the five mutating letters					
eich === chi (no mutation, correct)	82	78	64	64	NS
eich === chi (mutating, incorrect)	56	71	27	53	0.01
'sgynnoch chi === (mutating, correct)	66	72	65	69	NS
'sgynnoch chi === (no mutation, incorrect)	46	70	34	54	0.04
Time (sec) on the words starting with the five mutating letters					
eich === chi (no mutation, correct)	3.82	3.52	2.57	2.80	NS
eich === chi (mutating, incorrect)	4.43	3.42	2.87	2.53	0.01
'sgynnoch chi === (mutating, correct)	4.30	4.25	2.78	2.48	0.01
'sgynnoch chi === (no mutation, incorrect)	4.03	4.28	3.07	2.82	NS

Rule Test Phase 2. This phase was identical in form to rule test phase 1 and the patterns of performance were broadly replicated: "natural" experience with the new exemplars of natural learning phase 2 did not generally affect explicit knowledge of the rules, the Rule&Instances group still performed with very low error rates, and the Random groups were still very ignorant of the rule structure. A two-way mixed analysis of variance (4 groups × 2 constructions) on the subjects' mean percent correct on that construction in rule test phase 2 (Table 2) demonstrates significant main effects of group [$F(3,47)=6.90$, $P < 0.01$] and construction [$F(1,47)=14.36$, $P < 0.001$], but a non-significant group × construction interaction [$F(3,47)=0.49$, NS].

Natural Learning Phase 3. This phase investigated generalisation to new constructions (in contrast to new exemplars in natural learning phase 2). The vocabulary and design were therefore the same as in natural learning phase 2 except that there were two new constructions, one which did not cause mutation (*eich* === *chi* > "your === ") and one which did (*'sgynnoch chi* === > "do you have a === ").

It is clear from Table 1 that, in contrast to natural learning phase 2, there were transfer effects here: the Random groups took more trials to learn these new constructions (of the same rule content as the old) as applied to vocabulary they already knew. It was the Rule group in this instance which was superior, the subjects in this group applying their knowledge to comprehend these new sentences. The Rule&Instances group took somewhat longer, but was significantly superior to the Yoked Random group.

Rule Test Phase 3. Here explicit knowledge of all of the rules was tapped for each of the eight protowords (e.g. *ble mae t* === , *ei d* === *o*, *eich t* === *chi*, *'sgynnoch chi d* ===). The 32 trials were presented just once, in random order, with no feedback.

A two-way mixed analysis of variance (4 groups × 4 constructions) on the subjects' mean percent correct on that construction in rule test phase 3 (Table 2) demonstrates significant main effects of group [$F(3,47)=8.21$, $P < 0.001$] and construction [$F(3,141)=9.44$, $P < 0.001$], but a non-significant group × construction interaction [$F(9,141)=1.27$, NS]. However, it is clear that the Rule group's performance dropped for the new soft-mutation rule (78 *vs* 90% for the old one). In contrast, the subjects in the Rule&Instances group built upon their pre-existing structural knowledge and generalised from it—it is they who, untutored, abstracted their superior explicit knowledge of the content of the new soft-mutating construction (*'sgynnoch chi* ===) by transfer from that for the prior one (*ei* === *o*), performing at 87% correct. The Random groups were still significantly more ignorant, especially regarding this structural content of the soft-mutation rules.

We are now in a position to compare the explicit knowledge of all four groups over time and experience. A three-way ANOVA on the two constructions (*ei* === *o*, and *ble mae* ===) which have appeared in all three rule test phases (Table 2) across groups demonstrates significant main effects of group [$F(3,47)=20.46$, $P < 0.001$], phase [$F(2,235)=11.16$, $P < 0.001$] and construction [$F(1,235)=39.73$, $P < 0.001$]. The group × phase interaction was significant [$F(6,235)=6.58$, $P < 0.001$], demonstrating that whereas the performance of the Rule&Instances and Rule groups remained fairly constant over time, that of the two Random groups steadily increased over rule test phases 1–3. The group × construction interaction was also significant [$F(3,235)=4.01$, $P < 0.01$], with the two Random groups being particularly poor on the mutating construction. The construction × phase interaction [$F(2,235)=3.50$, $P < 0.05$] demonstrates that performance increased over the three rule test phases for the mutating construction. The three-way interaction was insignificant [$F(6,235)=1.73$, NS]. These data are illustrated in Fig. 1, where it can be seen that the Random groups slowly assimilated knowledge of the soft-mutation structural content from experience.

Well-formedness Test 2. Well-formedness test 1 assessed implicit awareness of grammatical correctness where the subjects had seen the well-formed strings on several prior occasions. In contrast, this phase was designed to tap fast judgements of well-formedness for constructions on which subjects had just been successful in natural learning phase 3 and exemplars where they had been successful in natural learning phase 1, but

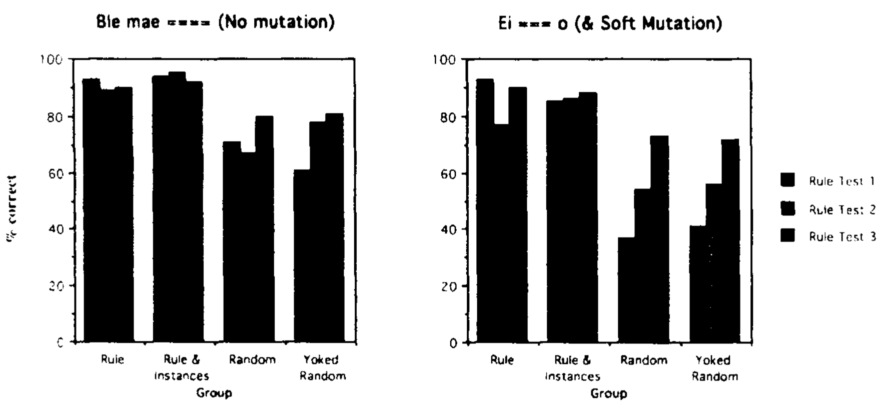

FIG. 1 Explicit knowledge of the mutation rule system of the four experimental groups as a function of content and amount of language experience.

where they had never seen these exemplars in these constructions before. The subjects' judgements could therefore not be performed on the basis of their familiarity with these particular strings, and thus the phase was designed to test fast, implicit judgements based on the generalisation and synthesis of different sources of knowledge. The data for well-formedness tests 1 and 2 are shown together in Table 3 and Fig. 2 for purposes of

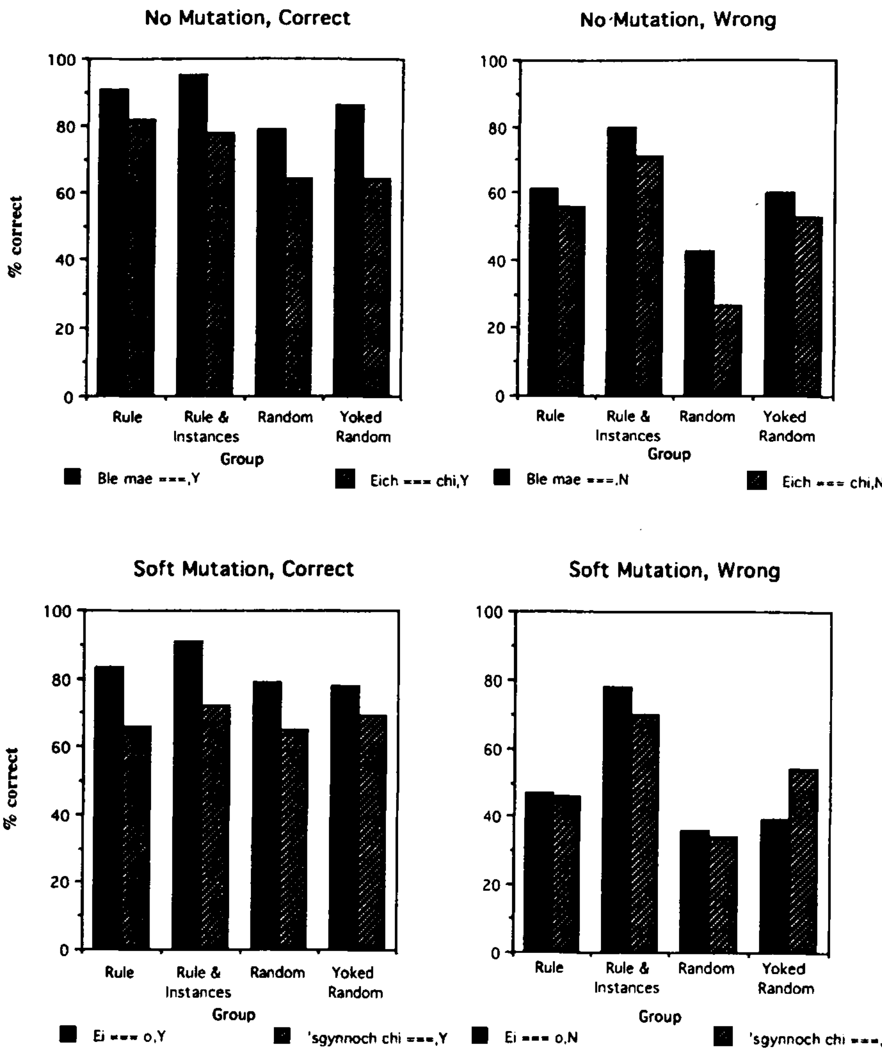

FIG. 2 Performance of the four groups on well-formedness tests 1 and 2. These involved fast, "implicit" grammaticality judgements for old (■) and novel (▨) utterances which were correctly and incorrectly mutated.

comparison. All of the groups now performed at lower levels on accepting the correct strings in contrast to well-formedness test 1, this demonstrating that some of their accuracy in the prior grammaticality test was attributable to their correctly recognising that they had seen this particular string before.

The raw binary (correct/incorrect) data in well-formedness test 2 for the five mutating letters were analysed like those from well-formedness test 1 as a repeated measures MANOVA with the five letters, the two exemplars of each letter and the four conditions as within-subjects effects and subjects nested within groups. There was a significant group effect [$F(3,45)=2.90$, $P < 0.05$], a significant effect of condition [$F(3,135)=16.84$, $P < 0.001$] and a significant group × condition interaction [$F(9,135)=2.38$, $P < 0.02$]. There was little difference between the groups on accurately accepting correct responses (both mutated and non-mutated), but both the Rule and Random groups were particularly inaccurate in failing to identify incorrect forms as being ungrammatical, this being especially so for the Random groups—the pattern from well-formedness test 1 repeating itself. The Random group performed badly on all of the conditions, suggesting very little implicit acquisition of structure; the Yoked Random group, with somewhat more language exposure, appeared somewhat (but not significantly) better. However, both Random groups were particularly poor, as were the Rule group, at identifying ungrammatical strings as being incorrect. The Rule&Instances group alone had this ability.

A similar repeated measures MANOVA on the latency data demonstrated no significant effects or interactions, although one-way ANOVAs within condition do suggest that the Rule group was somewhat slower than the two Random groups where letters were soft-mutated. However, the significantly superior performance of the Rule&Instances group was not associated with slower responses.

Rule Learning Phase 1. This phase determined the extent to which learned rules had been retained (in the Rule and Rule&Instances groups), and any potential savings in rule learning consequent on natural exposure in the Random groups.

It can be seen from the bottom of Table 1 that the Rule&Instances group had greater savings, learning the rules faster than all the other groups. The Random groups were, as with all prior tests of explicit rule knowledge, particularly poor; whatever untutored accommodation was observed over rule test phases 1, 2 and 3 still left much explicit learning to be done. The Rule&Instances group had abstracted the knowledge relevant to the new constructions automatically, generalising from the exemplars of natural learning phase 3 and the structures they had acquired for the old constructions during the Rule&Instances learning phase and natural learning phase 1.

None of the groups had explicit instruction on the two "new" rules (*eich* === *chi* and *'sgynnoch chi* ===). Yet the Rule&Instances group and, to a lesser extent, the Rule group, were able to automatically generalise and extract this knowledge (as shown both by their superior performance in rule test phase 3 and in their lower learning errors on these constructions here). In contrast, the Random groups, who had more exposure to the relevant language evidence (natural learning phase 3), have abstracted little (rule test phase 3) and had to learn the new rules afresh in rule learning phase 1.

The Rule&Instances group retained their old knowledge *and* abstracted and transferred this structure to the new constructions automatically from their "natural" experience of relevant exemplars during natural learning phase 3.

DISCUSSION

To summarise the findings of this experiment, three major styles of instructional exposure produced three very different patterns of performance in the learners.

Random Learners

The subjects in the Random groups at early stages of exposure (natural learning phase 1) seemed to be working by simple association learning and to learn these associations quickly. If assessed solely in terms of ability to deal with the particular instances that the learner has been exposed to, a criterion oft applied in the classroom concerning the repetition and use of formulaic speech, then one would advocate the random, more "naturalistic" method as promoting faster acquisition—*people can learn quickly and remember rather well the meanings of phrases which they have met and comprehended before*. But they do not unpack the meaning—they seem to be dealing with them as prefabricated patterns rather than as signs combined according to grammatical constraints. Unfortunately, although they can understand the language that they have seen before (natural learning phase 1), they have neither abstracted explicit knowledge of the rule system (rule test phase 1), nor are they proficient on the implicit well-formedness judgements where the stimuli were incorrect (well-formedness test 1)—there has been no assimilation of structure at either level. The "natural" exposure for the Random subjects appears initially to have led them to approach language functionally for its meaning, with no focus on structure and no knowledge thereabouts at either the implicit or explicit level.

However, as they are exposed to more and more language "evidence", so there does seem to be a slow accretion of explicit knowledge as indexed by the steady increase in performance over rule test phases 1–3 in Fig. 1. We cannot be sure in this experiment that the observed development of explicit metalinguistic knowledge would necessarily have taken place had the rule test phases themselves not made the problem of structure salient (Schmidt, 1990), and thus we make no general conclusions from this finding. It is clear from linguistic analyses of language acquisition that implicit knowledge is often consciously impenetrable (e.g. Seliger, 1979), yet at the same time learners do develop some spontaneous metacognitions about language (Bialystok, 1988; 1991; Durkin, 1989; Karmillof-Smith, 1986; O'Malley & Chamot, 1990), the distinction likely resting on the salience and complexity of the language structures involved.

In well-formedness test 1, the Random group demonstrated fast implicit working knowledge of the correct soft-mutation [79% identification of correct soft-mutation after ei === o *vs* only 37% correct (and note the ⅜ non-mutating letters—see above) explicit production of these rules in rule test phase 1]. Thus, at least at this early stage, explicit and implicit knowledge is dissociated in the Random learners.

That the Random learners can operate above chance on the well-formedness tests for correct stimuli which they have seen before *and* which they have not seen before ("novel utterances" in Fig. 2) suggests that *there is some implicit learning in these subjects*. However, although they are often correct in identifying "good language", they are also prone to "allow through" ungrammatical forms (the wrong utterances of Fig. 2)—the Random subjects do not know when language is ungrammatical. Further evidence for implicit learning comes from this latter error tending to decline with increased language exposure, the non-significant trend whereby the Yoked Random group, who have had more exposure to language, perform better in this respect (right-hand graphs of Fig. 2). Whatever implicit learning there may be in these groups, it is certainly laboriously slow and it cannot compare with the superior levels that can be encouraged by structured exposure to relevant exemplars.

Rule Learners

The subjects in the Rule group had a solid explicit knowledge of the rules which they had been taught (a solid *learning* in Krashen's terms), but a somewhat poorer application of that knowledge (poorer *acquisition*) in identifying incorrect language forms in the well-formedness tests. Their focus on form allowed them some transfer to new analogue structures (the new soft-mutation construction seen in natural learning phase 3 and tested in rule test phase 3 where their explicit knowledge of this rule was better than that of the Random groups but somewhat worse than that of the Rule&Instances group). *There does, therefore, appear to be some useful transfer from explicit to implicit knowledge in these aspects of L2 learning.* However, this is by no means fast or automatic: in the early stages (e.g. natural learning phase 1), there were many instances of their knowing the rule explicitly and yet failing to apply it in practice, i.e. they had the knowledge of the rules and the vocabulary sufficient to their understanding the utterances of natural learning phase 1 but still made many errors. At this early stage of learning, at least, there is thus a clear dissociation between explicit and implicit knowledge (cf. Krashen, 1982) and this is the obverse pattern to the Random groups who could operate implicitly without explicit awareness—in other words, explicit and implicit knowledge can be *doubly* dissociated.

The later transfer from explicit to implicit knowledge in these Rule learners confirms the applied linguistics findings (R. Ellis, 1990; Long, 1983) and the psychological ones concerning the learning of artificial languages where explicit instruction is clearly beneficial (Danks & Gans, 1975; Reber, Kassin, Lewis & Cantor, 1980). They do not support a strong "non-interface" position (Krashen, 1982) unless it is argued that (1) this whole task was so artificial that it resembles L2 acquisition not in the least and subjects adopt an "explicit" style throughout, or (2) that Rule subjects deal with the well-formedness tests in a different way from the Random subjects, using their explicit knowledge to monitor the strings (however, the failure to find significant differences between groups in response latencies in well-formedness test 1 do not support this argument). What transfer is seen in these learners is compatible with both the automaticity views of the transfer of explicit to implicit knowledge (Anderson, 1983; Bialystok, 1979; McLaughlin, 1985; McLeod & McLaughlin, 1986) and the "grammatical consciousness raising" (Sharwood-Smith, 1981) and "acquisition facilitators" (Schmidt, 1990) theories whereby explicit knowledge focuses the learner's attention on criterial attributes of the language strings, thus making their inductive hypothesis testing more efficient and allowing the accommodation of relevant schema for the structural patterns of language.

Rule&Instances Learners

The subjects in the Rule&Instances group alone could understand the language and had both explicit and implicit appreciations of its structure. The early instructional procedure of the subjects in this group taught them the rules explicitly (as was so for the Rule group) and thus it is little surprise that they performed well on the explicit rule tests (rule test phase 1). However, in contrast with the Rule group, exposure to the rule along with two exemplars of its application allowed a binding of explicit knowledge to instances so that in the Rule&Instances group there was a new form of implicit learning which produced an organisational structure which allowed generalisation under rapid performance conditions and also permitted these individuals alone to identify incorrect language (well-formedness tests 1 and 2). Furthermore, these subjects were more able to generalise from their grammatical schema for one soft-mutating construction and automatically transfer this system to an analogic construction both implicitly (Fig. 2: 'sgynnoch chi) and explicitly (Table 2, rule test phase 3: 'sgynnoch chi).

Structured instruction allows for generalisation at both the explicit and implicit levels. It was the Rule&Instances learners who abstracted a functional schema for soft-mutations since, when these subjects were next exposed to new constructions (some mutating, some not), it is they who generalised and were able to both explicitly formulate the new rules and to perform well on the appropriate well-formedness judgements. This group alone knew when novel phrases are ungrammatical. In recommending this method of instruction, we are not advocating the slavish repetition of drills of rules and examples. But the results do suggest the advantage of the inclusion of some explicit instruction within any teaching programme, and it is clear that the abstraction of functional schemata for grammatical structures is facilitated by the illustration of pedagogical rules with exemplars.

These results mirror those of Gick and Holyoak (1980; 1983) on analogical problem solving, suggesting that it is the *blend* of abstraction and structured exemplars that is important. However, it is likely that at least some of the effect is due to the conducive structuring of the exemplars alone (Foss, 1968; Palermo & Howe, 1970; Palermo & Parrish, 1971). And this structuring effect is likely to have effects at both the explicit and implicit levels. For the former, helpful structuring of examples is going to make the criterial features more salient (Berry & Broadbent, 1984; 1988; Reber, 1976; Reber et al., 1980; Schmidt, 1990). As for the latter, we still do not properly understand the effects of the blocking, frequency and juxtaposition of types of exemplar on implicit learning, although there are relevant observations in the concept/prototype formation literature (e.g. Homa & Vosburgh, 1976; Peterson, Meagher, Chait, & Gille, 1973; Posner & Keele, 1968: who demonstrated that variable and dissimilar training examples caused slower initial learning but broader and more flexible transfer from the resultant prototype). More recent promise comes from connectionist work on implicit grammar learning: it has been claimed that connectionist (parallel distributed processing) networks, using simple learning algorithms, can exhibit the behaviour previously assumed to be characteristic of rule-governed systems, even though the connectionist nets do not contain explicit rules. Parallel distributed processing modelling is a medium for investigating implicit learning in humans, that which takes place "naturally, simply and without conscious operations".

For example, Rumelhart and McClelland (1986) claimed that a simple learning model reproduced, to a remarkable degree, the characteristics of young children learning the morphology of the past tense in English—the model generated the so-called U-shaped learning curve for irregular forms, it exhibited a tendency to over-generalise and, in the model as in children, different past-tense forms for the same word could co-exist at the same time. Yet there was no "rule": "it is possible to imagine that the system simply stores a set of rote-associations between base and past-tense forms with novel responses generated by 'on-line' generalisations from the stored exemplars" (Rumelhart & McClelland, 1986, p. 267). Marchman and Plunkett (1989) and Plunkett and Marchman (1989) have extended this work using miniature artificial languages and varying factors such as the size of the "irregular" classes and the frequency of the exceptional items within those classes. When the size of an irregular class was small, increasing the relative frequency of the items within that irregular class facilitated learning of that sub-regularity without seriously impairing performance on the large "regular" class. This was found to contrast with the case where there were many, highly frequent "irregular" items, in which case learning of the dominant regularity was significantly impaired. The connectionist enterprise can thus inform us about just what aspects of grammar (for example, the extraction of grammatical class information; Elman, 1990) could be acquired by implicit association learning in both L1 (Lachter & Bever, 1988) and L2 (Gasser, 1990).

Only when we have a handle on implicit learning, when we discover just what associations and language structures can be learned implicitly by this class of systems, will we be able to address the much less tractable problem of the mechanisms of interaction between explicit and implicit knowledge, i.e. between conscious and unconscious cognition. Meanwhile, we are now prosecuting other implicit/explicit interaction research which better approximates real-world language acquisition in that it addresses spoken language, production and generalisation to truly naturalistic conversational situations.

Of course, the end-point of L2 acquisition—if the learners, their motivation, tutors and conversation partners, environment, and instrumental factors, etc., are all optimal—is to be as proficient in L2 as in L1. So proficient, so accurate, so fluent, so automatic, so implicit, that there is rarely recourse to explicit, conscious thought about the medium of the message. The question addressed here is how best to promote that journey, and it appears that explicit instruction and carefully structured exemplars speed at least initial progress.

The four main conclusions to be drawn from this study are:

1. People can learn quickly and remember rather well the meanings of phrases which they have met and comprehended before.
2. There can be implicit learning even with random exposure.
3. Explicit instruction which explains the structure and content of the rules can facilitate acquisition.
4. It is better still to conjoin abstraction and instances by demonstrating the rules in operation with a number of illustrative exemplars of their application.

Manuscript received 5 November 1992
Revised manuscript received 15 January 1993

REFERENCES

Adams, R. (1980). *The girl in a swing*. Harmondsworth: Penguin.

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

Ball, M.J. (Ed.) (1988). *The use of Welsh*. Clevedon: Multilingual Matters Ltd.

Berry, D.C., & Broadbent, D.E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology, 36A*, 209–231.

Berry, D.C., & Broadbent, D.E. (1988). Interactive tasks and their implicit–explicit distinction. *British Journal of Psychology, 79*, 251–272.

Bialystok, E. (1979). An analytic view of second language competence: A model and some evidence. *The Modern Language Journal, LXIII*, 257–262.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology, 24*, 560–567.

Bialystok, E. (1991). Metalinguistic awareness: The development of children's representations of language. In C. Pratt & A. Garton (Eds), *The nature and use of representation in children*. London: John Wiley.

Carroll, J. (1966). The contributions of psychological theory and educational research to the teaching of foreign languages. In A. Valdman (Ed.), *Trends in language teaching*. New York: McGraw-Hill.

Carroll, J.B., & Sapon, S.M. (1955). *Modern Language Aptitude Test*. New York: The Psychological Corporation.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. New York: Praeger.

Cox, B. (1989). *English for ages 5–16 (The Cox Report)*. London: HMSO.

Danks, J.H., & Gans, D.L. (1975). Acquisition and utilisation of a rule structure. *Journal of Experimental Psychology: Human Learning and Memory, 104*, 201–208.

Donaldson, W. (1971). Code-cognition approaches to language learning. In R. Lugton & C. Heinle (Eds), *Towards a cognitive approach to second-language acquisition*. Philadelphia, PA: Center for Curriculum Development.

Durkin, K. (1989). Implicit memory and language acquisition. In S. Lewandosky, J.C. Dunn, & K. Kirsner (Eds), *Implicit memory: Theoretical issues*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Basil Blackwell.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science, 14*, 179–211.

Esper, E.A. (1925). A technique for the experimental investigation of associative interference in artificial linguistic material. *Language Monographs, 1* (cited in Foss, 1968).

Foss, D.J. (1968). An analysis of learning in a miniature linguistic system. *Journal of Experimental Psychology, 76*, 450–459.

Gasser, M. (1990). Connectionism and universal of second language acquisition. *SSLA, 12*, 179–199.

Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306–355.

Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38.

Gomes da Torre, M. (1985). *Uma Análise de Erros*. PhD thesis, Faculty of Letters, University of Oporto.

Homa, D., & Vosburgh, R. (1976). Category breadth and the abstraction of prototypical information. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 322–330.

James, C. (1986). Five uses for grammar in foreign language teaching. *Revista da Faculdade de Letras Do Porto, Línguas e Literaturas, II Série, III*, 95–113.

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

Kelly, L.G. (1969). *Twenty five centuries of language teaching*. Rowley, MA: Newbury House.

Kingman, J.F.C. (1988). *Report of the Committee of Enquiry into the Teaching of English Language*. London: HMSO.

James, C. (1986). Five uses for grammar in foreign language teaching. *Revista da Faculdade de Letras Do Porto, Línguas e Literaturas, II Série, III*, 95–113.

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

Kelly, L.G. (1969). *Twenty five centuries of language teaching*. Rowley, MA: Newbury House.

Kingman (1988). *Report of the Committee of Enquiry into the Teaching of English Language*. London: HMSO.

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

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

Lachter, J., & Bever, T.G. (1988). The relation between linguistic structure and associative theories of language learning—A constructive critique of some connectionist learning models. *Cognition, 28*, 195–248.

Lambert, W.E. (1974). Culture and language as factors in learning and education. In F.E. Aboud & R.D. Meade (Eds), *Cultural factors in learning and education*. Bellingham, WA: Fifth Western Washington Symposium on Learning.

Long, M. (1983). Does second language instruction make a difference? A review of the research. *TESOL Quarterly, 17*, 359–382.

Marchman, V., & Plunkett, K. (1989). Token frequency and phonological predictability in a pattern association network: Implications for child language acquisition. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*, pp. 179–187. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

McLaughlin, B. (1985). *Second language acquisition in childhood, Vol. 2: School-age children*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

McLaughlin, B. (1990). "Conscious" versus "unconscious" learning. *TESOL Quarterly, 24*, 617–634.

McLeod, B., & McLaughlin, B. (1986). Restructuring or automaticity? Reading in a second language. *Language Learning*, *36*, 109–123.

Naiman, N., Fröhlich, M., Stern, H.H., & Todesco, A. (1978). *The good language learner: A report*. Toronto: Ontario Institute for Studies in Education.

Norusis, M.J. (1985). *SPSS-X advanced statistics guide*. Chicago, IL: McGraw-Hill.

O'Malley, J.M., & Chamot, A.U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Palermo, D.S., & Howe, H.E. (1970). An experimental analogy to the learning of past tense inflection rules. *Journal of Verbal Learning and Verbal Behaviour, 9,* 410–416.

Palermo, D.S., & Parrish, M. (1971). Rule acquisition as a function of number and frequency of exemplar presentation. *Journal of Verbal Learning and Verbal Behaviour, 10,* 44–51.

Peterson, M.J., Meagher, R.B. Jr, Chait, H., & Gillie, S. (1973). The abstraction and generalization of dot patterns. *Cognitive Psychology*, *4*, 378–398.

Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Plunkett, K., & Marchman, V. (1989). *Pattern association in a back propagation network: Implications for child language acquisition*. Technical Report 8902. La Jolla, CA: University of California, San Diego, Center for Research in Language.

Posner, M.I., & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.

Reber, A.S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 88–94.

Reber, A.S., Kassin, S.M., Lewis, S., & Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 492–502.

Rumelhart, D.E., & McClelland, J.L. (1986). On learning the past tense of English verbs. In J.L. McClelland, D.E. Rumelhart, & the PDP Research Group (Eds), *Parallel distributed processing, Vol. 2: Psychological and biological models*. Cambridge, MA: MIT Press.

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

Seliger, H. (1979). On the nature and function of language rules in language teaching. *TESOL Quarterly*, *13*, 359–369.

Sharwood-Smith, M. (1981). Consciousness raising and the second-language learner. *Applied Linguistics*, *2*, 159–168.