

# Teaching Declarative and Procedural Knowledge

KEITH JOHNSON

## 1. Introduction

This paper will consider how best to present *knowledge about* grammar to students, then how to convert this into *knowledge how to* use that grammar. We shall consider these issues in relation to what are in fact two models of learning: Anderson (1982) and Neves and Anderson (1981). These two models are very similar, and will here be regarded as one single model. To an extent, what is here being attempted is a restatement of certain familiar issues in language teaching in terms of this model.

Underlying the approach is a view that language learning is comparable to other types of cognitive skill acquisition, and that foreign language teaching may fruitfully be seen as a type of skill training. An advantage of considering language learning in this light is that it permits one to utilise insights from other sorts of learning. It is in keeping with this approach that we here utilise the Anderson/Neves and Anderson model, which was initially developed to account, not for language learning, but for the learning of geometry.

The model will not here be outlined; only two points about it are central to the argument of this paper:

- (a) It utilises the distinction between *declarative* and *procedural* knowledge (henceforth DK and PK). The distinction is related (though not identical) to others made in applied linguistic discussion: the traditional one is *knowledge about* (DK) and *knowledge how to* (PK). There is also Bialystok's (1982) *knowledge/control* distinction. For general discussion of the declarative/procedural distinction, see Winograd (1975).
- (b) The model uses *productions*. These are rules of an *IF ... THEN* sort, developed to describe procedures. In Figure 1, P1 and P2 exemplify general productions, P3 a specific one.

In the Anderson/Neves and Anderson model the learner begins by referring to general rules like P1 and P2, and relates them to knowledge held in a memory data base. This procedure ('consciously applying the rules') is laborious. With time, the learner develops PK (that is, proceduralises the knowledge), until finally it is held in a form like P3, which can be applied with speed.

Figure 1: Examples of a production

- 
- P1 IF the goal is to form the present perfect of a verb and the person is third singular, THEN form the third singular of *have*.
- P2 IF the goal is to form the present perfect of a verb and the appropriate form of *have* has just been formed, THEN form the past participle of the verb.
- P3 IF the goal is to form the third singular, present perfect of the verb *change*, THEN form *has changed*.
- 

## 2. The importance of procedural and declarative knowledge

It is by now well accepted that there is a role for PK (*knowledge how to*) in all language learning, the teaching of grammar included, and the general case no longer needs rehearsing. Bialystok (1982) argues that the specific role of PK depends on the nature of the language task. Her paper uses different terms and has a slightly different conceptual framework from that used here. Using our, not her, terms: for tasks such as spontaneous conversation where immediate access to knowledge is required, PK is important; it will be less so in tasks like writing where speed of production (and hence rapid access to knowledge) is often not required.

What does perhaps need to be established is that there is an equally important role for DK. There are two sorts of reason for this importance. The first has to do with deployment of knowledge in use. There is the general argument which claims that the student who has a series of specific productions of the P3 type above (but no generalised ones of the P1 or P2 sort) cannot generate present perfect tenses for verbs not met before. This student has learned a series of individual present perfects (*has changed* for *change*, *has worked* for *work*, and so on) but possesses no rule. There is no data base or set of rules for generating the tense, and hence the system cannot easily go beyond data already met. The need for 'generativity' is hence a strong argument in favour of DK. Once again, Bialystok (1982) pinpoints a specific role. While DK may have little part to play in spontaneous conversation, for example, it will be crucial in many writing tasks, where having a DK data base of rules to refer to and manipulate will be an advantage.

The second sort of reason why DK is important concerns learning. There is some evidence from the general skills literature that having *knowledge about* is a useful first step to developing PK *how to*. Fitts and Posner (1967), for example, report that time taken to lead novice pilots to their first solo flight is reduced from ten hours for a control group, to three and a half for an experimental group where some initial explanation was given.

Both DK and PK are, then, important for the language user. How may they be developed? We now consider two sorts of general strategy, referred to as Paths 1 and 2. We suggest that these two paths correspond to the two strategies commonly recognised by many (not only Krashen) under the names of 'learning' and 'acquisition'. We here outline the two paths, and identify for each what needs to be done to meet the established aim of providing both DK and PK.

### 3. Two paths to declarative and procedural knowledge

#### **Path 1**

Path 1 is the learning path, and it moves from DK to PK. Teaching provides DK; through a process of automisation this knowledge becomes proceduralised, and so automatic that it is eventually indistinguishable in performance from knowledge internalised by that pathway which Krashen (1982) calls 'acquisition'. Hence the model is what Ellis (1990) refers to as an 'interface' one, where (in loose terminology) learning becomes acquisition.

The two most obvious problems which this path poses are how best to present DK to the learner in the first place, and how to help the learner convert it into PK; both these issues will be considered later. But there is a third. A danger with this path is that the DK will disappear once the procedural has been developed; that P1 and 2 will be unavailable for further use once a set of P3s has been developed. Examples of how PK replaces DK abound in the literature: a car route becomes so familiar that you follow it unconsciously, losing the ability to represent it in map form; a much-used phone number sometimes can only be remembered by dialling; changing gear in a car becomes so automatic that you are unable to explain to another person how to do it.

Bialystok and Sharwood Smith (1985) distance themselves from Anderson, because, recognising the importance of DK, they feel that in Anderson's model the declarative is replaced by the procedural. But Anderson certainly recognises that it is important to maintain declarative representation, and indeed a section of Neves and Anderson (1981) is entitled 'Getting the Best of Both Encodings' (the declarative and the procedural). Anderson concedes that DK may disappear, but regards this as detrimental; something requiring remedial action. The problems with Path 1 are, therefore: initial forming of DK; maintaining DK; proceduralising DK.

#### **Path 2**

Path 2 is the acquisition path (in Krashen's sense), and it moves from PK to DK. In the first instance the learner 'directly proceduralises' ('acquires') knowledge. In other words, the learner goes straight to P3, without passing from P1 and P2, and over time develops a number of P3 type constructs. On this path, proceduralising is not an issue – learners do it for themselves in the first place. But Path 2 involves the very great danger that DK will never be achieved. The learner has a set of individual rules; she or he needs to develop the 'cognitive map'. Note that if this is achieved, this model is also an 'interface' one, this time with 'acquisition becoming learning'.

This strategy of acquiring procedural rules directly (without moving from the declarative) is dubbed by Anderson as 'high risk' because productions once formed are, if wrong, exceedingly difficult to change. A comparison with learning a musical instrument: once a wrong fingering has been acquired and is well embedded, it is

exceedingly difficult to eradicate. Hence, presumably, in second language learning the phenomenon called fossilisation (and part of the reason why pidgins, once acquired, fossilise early).

Two further observations on this problem:

- (a) It is difficult not just to 'unproceduralise' the wrong rules (like the wrong fingerings referred to above), but also to declarativise the right ones as well.
- (b) Declarativisation, though difficult, is possible. The L1 situation is one in which learners at school declarativise acquired knowledge as part of the process of learning how to write (among other things). This is achieved partly through tuition, and partly simply in confrontation with tasks that require it. One writes a composition, and in the process recognises, for example, incomplete understanding of a word; so one looks it up.

It is not our aim to arbitrate between the two paths (between 'learning' and 'acquisition'). But perhaps viewing these constructs in the light of a general cognitive learning model may provide a framework in which arbitration becomes more possible than at present. Specifically, we might view the learning acquisition issue as part of the more general debate in cognitive science, referred to earlier (Section 1 of this paper), regarding knowledge in data base versus knowledge in productions.

The remainder of this paper will concentrate on Path 1, and will look in more detail at the problems it poses in relation to the development of DK and PK. During the discussion we shall make reference to the traditional presentation – practice – production (P P P) teaching model, beginning by looking at the presentation stage.

#### **4. Forming declarative knowledge**

We here briefly consider issues related to the question of how language may best be presented to students. In Section 2 of this paper it is argued that there are two types of motivation for providing and maintaining DK, one related to 'deployment in use', and the other to 'learning'. For deployment in use, the learner requires a set of rules that are generalisable enough to apply overall. In the case of the present perfect, for example, the learner has to know how the participle is formed.

The learning motivation relates to the fact that DK acts as the first stage in the process of proceduralisation. The requirement here is that the knowledge representation should be readily translatable into PK: it should be 'proceduralisable'. We have already seen that providing overviews is useful, but there are two qualifications:

- (a) Holding (1965) notes that a little prior teaching is better than a lot, observing that feedback (after rather than before practice) plays an important role in providing assistance for proceduralisation. See Johnson (1986) for discussion.
- (b) Explanation (for a language point in the form of linguistic description of

'how it works') is possibly not the best form of guidance to be given. Certainly elaborateness needs to be avoided. Holding (1965): 'It is possible to disrupt the operator's performance in a quite lasting way by the over-elaboration of instructions'. He goes on to suggest that figurative or emotive description is often better than scientific, and speaks of the 'science of hints' which, he admits, is as yet relatively undeveloped.

An issue of recent applied linguistic interest is the relationship between the initial declarative representation of a rule and the internal representation which a student eventually develops. Some imply that the two sorts of representation should resemble each other. Prabhu (1985) argues that since they cannot, we should not offer an external representation. Pienemann (1985 and elsewhere) argues that the sequence of exposure (to grammatical items) provided for learners should relate to some internal sequence, revealed by research (e.g. morpheme acquisition studies).

Presumably, however, external rules can be internalised and can change form. An anecdotal example: my car cassette player has complex rules regarding how to rewind/fast-forward according to which side is uppermost in the machine. My initial set of (learned) rules came from the instruction booklet, and follow the familiar pattern for such sets of instructions. My final representation, which I now use, is quite different, and involves such personal instructions as 'press the button near the fire extinguisher to rewind side A' and 'to fast-forward side B press the button on my wife's side'. If one accepts Rutherford's (1987: 18) point that 'pedagogic descriptions are *aids* to learning, not the *object* of learning', then there is no necessity that such descriptions should initially attempt to replicate in any way any internal representation.

## 5. Maintaining declarative knowledge

It has been noted in relation to Path 1 that DK may disappear with proceduralisation, and why this disappearance needs to be avoided has been discussed in Section 2 of this paper. This might be done by ensuring that explicit mention of rules regularly occurs, perhaps by consciousness-raising techniques and through exercises that in one way or another 'draw notice' (in Schmidt's 1990 terminology) to rules. What is important is to note the danger, and to recognise that remedial action needs to be taken; the teacher needs regularly to return to rules covered earlier.

## 6. Proceduralising declarative knowledge

### *Automisation*

The process of proceduralisation is one of automisation: making automatic. P3 in Figure 1 is quicker to use than Ps 1 and 2. Automisation is generally recognised as important in skill learning; indeed Schiffrin and Dumais (1981) dub it 'a fundamental component of skill development'. The reason for its importance is that it releases conscious attention for the higher-order activities that require it. To use a

standard example: we automate changing gear in a car so that it no longer takes our conscious attention; we can then concentrate attention on watching and anticipating traffic movement. The importance of automatisation to language has already been touched on above in reference to Bialystok (1982) and in relation to different areas of language use.

A characterisation of automatisation is: 'getting a skill right when minimal attention is available for getting it right'. To interpret this in terms of grammar: a novice learner can usually only get structures right when there is full attention on them; there is, in other words, full *form focus*. But in much natural language use there is full *message focus*. Thus, to restate the characterisation above in terms of the form/message focus distinction (much used in recent applied linguistic discussion), automatisation is: 'getting the grammar right when there is message focus, and no attention is available for form focus'. Note how well this characterisation relates to some contemporary issues in language teaching. For example, Prabhu's (1987) hypothesis is that form is best acquired when focus is on meaning. This may easily be reconceptualised, in the framework used here, as a statement (with which we do not necessarily agree) that 'form is best learned when no attention is available for getting it right'.

### ***How not to automatise: the failed audio-lingual model***

For many, automatisation implies the type of drilling activity associated with audio-lingual (AL) teaching. Why did AL fail, and what lessons are to be learned from its failure?

In AL the view was that practice alone was sufficient for learning to occur. Hence AL practised the present perfect in the language laboratory, then expected the learner to produce it properly in natural communicative (message-focused) situations. The belief was that there would be transfer from class to life, and the actual failure of this strategy to work may be regarded as a failure of transfer of training. In terms of the P P P model, we may say that AL omitted the final P; it expected transfer from the first two Ps to the third, and hence the production stage was seen as unnecessary.

AL's failure can be regarded in terms of remoteness of practice conditions to the real skill required. Holding (1965) makes the point in relation to swimming: 'practising the arm movements of the breast stroke on land, may not be the same thing as using them to swim in co-ordination with the legs'. The conditions of practice are so remote from what actually happens as to be useless; the expected transfer from practice to production does not occur because the practice offered in class is so unlike the production of real life.

In terms of our earlier characterisation, we may say that language laboratory practice is in 'getting it right when all attention is available for getting it right'. This is highly remote from the actual target behaviour ('getting it right when minimal attention is available . . .'). Hence the failure is not surprising – the skills, in class practice and natural production, are simply not the same ones.

What conclusions may be drawn? An important general one would be that the task of automisation is more complex than hitherto thought. The P P \_ model (presentation and practice without production) as conceived of in AL is simply inadequate. It might be argued that a major task of applied linguistics is to consider activation techniques which will ensure transfer from practice to real language use.

More detailed conclusions in relation to the traditional P P P model may be specified. One is that any practice must meet minimum conditions in terms of its relation to the terminal behaviour. In other words, it must not be so remote as to constitute practising a different skill. In terms of Holding's swimming example, one might set a minimum condition that it take place in water. The parallel in language practice terms might be to ensure that certain communicative elements are present, even in the most form-focused practice. One such condition might be to insist on at least a degree of message focus at all stages; another, that there should be an information gap, ensuring at least a degree of information transfer in interactions. A second conclusion is that because the best way of simulating important aspects of communication is through communication, it follows that the final P should be restored; a production stage is necessary.

### ***A form-defocus model of automisation***

One possible strategy for the facilitation of automisation would be gradually to 'deprive' the learner of the attention he/she requires to produce a given structure. We would do this by gradually increasing the burden to be placed on the learner's attention, leaving less attention available for focus on the form being practised. Hence when first introduced to the present perfect (for example), a good deal of attention would be available to the learner to focus on the form. Gradually different conditions would be added to make the task more difficult. More language would be required of the learner (making, as it were, less attention available for each 'piece of language'), and more message focus would be demanded. Faster production could also be required, and one notes in passing that computer-assisted exercises would well manipulate the variable of required speed in production.

This idea is discussed in Johnson (forthcoming) where comparison with an aircraft simulator is made. In that learning situation the novice pilot first practises landing in a clear sky. As she/he improves, 'complications' are added by pressing buttons; one button creates fog conditions, another a faulty engine, a third faulty landing gear. By increasing the complexity of the processing conditions, we correspondingly decrease the 'amount of attention available for getting it right' and hence hopefully assist automisation to take place. Johnson (forthcoming) argues that what is needed in language teaching is a 'present perfect simulator' which can vary processing conditions in the same way. He regards the question of how to achieve this as a central methodological issue of the present time. A number of people (some working within the area of task-based learning) have developed taxonomies of operating conditions relevant to what is outlined here; examples are Brown *et al.*

(1984), Nunan (1989), Candlin (1987) and Skehan (1994) where variables like cognitive complexity, cognitive load and affect, are considered.

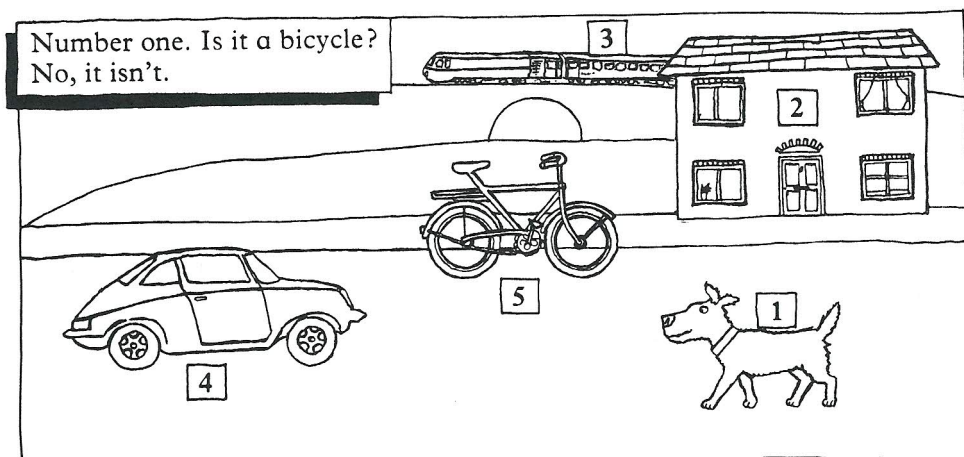
### **Communicative methodology and automatisisation**

It is possible to reinterpret what occurs in communicative methodology in terms of automatisisation. Two examples follow. Figure 2 exemplifies a drill intended for young

Figure 2: Conventional drill

#### ***Fred's picture***

Look for one minute.



(from Johnson, 1982, Workbook, p. 11)

learners. The original version did not include the words 'Look for one minute', and was a straightforward (and exceedingly boring!) drill done around the class (Teacher: *Number one. Is it a bicycle?* Pupil: *No, it isn't*, etc.). The instruction 'Look for one minute' turns the exercise into a memory game. Pupils now have a minute to memorise objects and their corresponding numbers. Then they put away the picture, and the teacher's questions (as above) become more meaningful, essentially having the underlying form: *Can you remember whether number one was a bicycle?* The reason for adding the memory element was motivational, to make the exercise more interesting. But the effect has been to achieve 'form defocus'. That is, a memory element has been introduced which deflects the attention away from focus on form and onto the task of remembering. Such an exercise may be seen as a step towards automatisisation.

The second example (Figure 3) is a conventional information gap exercise. Student A conceals the answers and asks the questions; Student B looks at the

Figure 3: Conventional information gap exercise

## 2. Pairwork

*Student A*

You'd like to go out, but you've got a sore throat, and you'd better stay at home. Ask your partner for details of radio programmes, and fill in the table. Ask:

What time. . . . ?

What's on at. . . . ?

Which station. . . . ?

<i>Time</i>	<i>Programme</i>	<i>Station</i>
_____	Late Night Pop	Radio 1
9.00	The World This Weekend	Radio ____
10.00	_____	Radio 3
_____	Schumann's 3rd Symphony	Radio 4
10.15	Interview with Colin Davis	Radio ____
9.00	_____	Radio 2

*Student B*

Your partner is not very well, and has to stay at home for the evening. He will ask you for information on radio programmes. Tell him, using the programme guide below.

**Radio 1**

- 7.30– 9.00 **The Top Thirty Show**  
 9.00–11.00 **Jimmy Savile's Request Show**  
 11.00–12.00 **Late Night Pop** The best of the new L.P.s

**Radio 2**

- 7.00– 9.00 **Your Hundred Best Tunes**  
 9.00–10.00 **Waltz Times** A programme of waltz music  
 10.00–12.00 **Sounds of Jazz**

**Radio 3**

- 7.00– 7.30 **News**  
 7.30– 9.00 **The Evening Concert** (Beethoven's 3rd Symphony, and 2nd Piano Concerto)  
 9.00–10.00 **Bartok's String Quartets** (This week: Nos 3 and 4)  
 10.00–10.15 **News**  
 10.15–12.00 **Music Now** An Interview with Colin Davis

**Radio 4**

- 7.00– 8.30 **Play of the Week** ('The Robot' by Andrew Jameson)  
 8.30– 9.00 **Letter from Japan** Some thoughts on contemporary Japanese life  
 9.00–10.00 **The World This Weekend**  
 10.00–11.00 **The Drama of T.S. Eliot** (Talk by Paul Sinclair)  
 11.00–12.00 **Schumann's Symphony No 3**

(from Johnson and Morrow, 1979: 91–2)

answers but not the questions. Such exercises are often justified in terms of a degree of message focus. It is argued that since the questioner does not have the answers sought, all her/his attention will be turned towards the message (*what* the interlocutor will say). But it may be claimed that such exercises are in fact no more than pseudo-communication; students are not *really* interested in radio programmes and are merely playing what they know to be the language teacher's game, of practising English while pretending to do something else. There is no real message focus, just disguised form focus.

This criticism does indeed have a degree of justification. A much more honest way of justifying such exercises might be that they offer, not message focus, but 'form defocus'. In this conceptualisation, what the information gap achieves is a complication to the task, a complexity of operating condition which absorbs an amount of the learner's attention, thus preventing a narrow focus on the form of what she or he is saying. As such, one might imagine this to be an aid to automatisa-

This reconceptualisation would, if accepted, rather change the face of communicative methodology. The standard justification for message focus in communicative methodology is that this is what happens in real communication (being message focused), and communicative methodology is in the business of simulating real communication. The altered justification is that we are helping the student towards a situation of total form defocus, which is another way of looking at what occurs in real communication.

## References

- Anderson, J.R. (ed.) 1981, *Cognitive Skills and their Acquisition*, Lawrence Erlbaum Associates, Hillsdale.
- Anderson, J.R. 1982, 'Acquisition of cognitive skills', *Psychological Review*, vol. 89, no. 4.
- Bialystok, E. 1982, 'On the relationship between knowing and using linguistic forms', *Applied Linguistics*, vol. 3.
- Bialystok, E. and Sharwood Smith, M. 1985, 'Interlanguage is not a state of mind: an evaluation of the construct for second-language acquisition', *Applied Linguistics*, vol. 6, no. 2.
- Brown, G., Anderson, A., Shillcock, R. and Yule, G. 1984, *Teaching Talk*, CUP, Cambridge.
- Candlin, C.N. 1987, 'Towards task-based language learning', in Candlin, C.N. and Murphy, D. (eds.), *Language Learning Tasks*, Prentice Hall, Englewood Cliffs, NJ.
- Ellis, R. 1990, *Instructed Second Language Acquisition*, Blackwell, Oxford.
- Fitts, P.M. and Posner, M.I. 1967, *Human Performance*, Brooks Cole, Belmont, Calif.
- Holding, D.H. 1965, *Principles of Training*, Pergamon Press, Oxford.
- Johnson, K. 1982, *Now for English 1*, Thomas Nelson, Walton-on-Thames.
- Johnson, K. 1986, 'Language teaching as skill training', Centre for Applied Language Studies, University of Reading, colloquium paper.
- Johnson, K. forthcoming, *Language Teaching and Skill Learning*, Blackwell, Oxford.
- Johnson, K. and Morrow, K. 1979, *Approaches*, CUP, Cambridge.
- Krashen, S.D. 1982, *Principles and Practice in Second Language Acquisition*, Pergamon Press, Oxford.
- Neves, D.M. and Anderson, J.R. 1981, 'Knowledge compilation: mechanisms for the automatization of cognitive skills', in Anderson (ed.), *Cognitive Skills and their Acquisition*.
- Nunan, D. 1989, *Designing Tasks for the Communicative Classroom*, CUP, Cambridge.
- Pienemann, M. 1985, 'Learnability and syllabus construction', in Hyltenstam, K. and Pienemann (eds.) 1985, *Modelling and Assessing Second Language Acquisition*, Multilingual Matters, Clevedon.

- Prabhu, N.S. 1985, 'Guided and unguided grammar construction', paper delivered at BAAL seminar on grammar teaching, Bath, 1985.
- Prabhu, N.S. 1987, *Second Language Pedagogy*, OUP, Oxford.
- Rutherford, W.E. 1987, *Second Language Grammar: Learning and Teaching*, Longman, Harlow.
- Schiffrin, R.M. and Dumais, S.T. 1981, 'The development of automatism', in Anderson (ed.), *Cognitive Skills and their Acquisition*.
- Schmidt, R.W. 1990, 'Consciousness in second language learning', *Applied Linguistics*, vol. 11, no. 2.
- Skehan, P. 1994, 'Second language acquisition strategies, interlanguage development and task-based learning', this volume.
- Winograd, T. 1975, 'Frame representations and the declarative-procedural controversy', in Bobrow, D.G. and Collins, A.M. (eds.) 1975, *Representation and Understanding: Studies in Cognitive Science*, Academic Press, New York.