Towards greater objectivity in classroom observation

Graham Howells
Universidade de Brasília
Lynne Meiring
University of Wales

Many teacher trainers have struggled with classroom observation schedules that have tended to consist of grids in which the frequencies of a specified range of teacher-learner interactions are marked. These schedules were originally planned before the miniaturisation of video cameras had made it practical to film lessons relatively unobtrusively. This article argues in favour of a system in which, when classes are filmed, the results may be analysed in conjunction with grid-based evaluation schedules and presented via computer-generated graphs to create an objective and easily interpreted method of reporting to the student teacher the distribution of classroom activities. These activities may then become the subject of discussion with reference to their instructional value.

1. The Increasing Importance of Classroom Observation

The increase in inspection and appraisal in education today means that there is a general need for teachers to have a guarantee of receiving the most objective means of appraisal possible and also to be given a means of preparing themselves for appraisal. Much work has been done in devising observation schedules that describe and analyse classroom discourse and interaction, but there are problems in using such procedures in order to evaluate teacher effectiveness:

I. They have often been devised and are often carried out by an observer who does not regularly teach (and indeed may not have been a classroom teacher for a considerable time). The process may become a "top-down" exercise rather than a sympathetic, collaborative one. (Also, it is very rare for learner feedback to be taken into account concerning the lessons observed, unless the learners are adults.)

II. The schedules have usually been devised according to a tallying system on a printed grid. This method gives a quantitative summation of interactions from which qualitative conclusions must be inferred after analysing total numbers of interactions. Some of these systems are described below.

III. The observation techniques are problematical in that a schedule of easily manageable proportions gives only a partial view of classroom interactions. On the other hand, it is difficult in non-experimental situations to operate systems that pre-specify a wider range of teaching/learning activities to be noted.

IV. There seems to be an assumption (either explicit or implicit) behind research in this area that the analysis of classroom interaction allows teacher trainers to enable students to conform to some ideal teaching style when, as the Pennsylvania Project (Clark, 1969) indicates, we cannot even state with certainty what overall teaching approach *is* the most effective.

V. Observation schedules have often been developed through research in applied linguistics and in teacher training and appraisal, have been made to perform a task that they were not designed for. They do not take into account variables such as: socio-economic environment of the school; learner motivation; quality of teaching materials; the physical environment of the classroom, or the range and quality of its technical equipment.

VI. From a technical point of view there is a problem in that some of the appraisal methods described in the literature were formulated before the widespread use of video-recording. Where video-recording has been used, it has usually merely provided a more methodical way of *describing* interaction, without always taking sufficient advantage of the opportunities offered for *analysis* of the teaching activities.

VII. Observation schedules have a centralising tendency: teacher/learner behaviours have to fit the Procrustean bed of the headings on the report form. Those categories are used to assess the teacher's performance even though, as suggested above, it is impractical to create a usable report form that will encompass all the behaviour variables of normal classroom interaction. It is not therefore surprising that these observation systems pay little heed to variables of individual teaching styles.

2. The Desiderata of Classroom Observation

What is needed is a means of achieving a system of observation that is seen to be fair and objective, easy to implement and which offers the possibility of clear presentation of results to the teacher being observed.

Firstly the whole area of classroom observation must be considered in its methodology and its effectiveness. The recent move towards accountability and teacher appraisal will no doubt have highlighted the need to carry out accurate classroom observation. This observation must look beyond the teacher to what is actually taking place in the classroom. Ellis (1990:68) quotes Mclaughlin's view on this issue claiming that there are dangers in viewing teacher/learner behaviour separately because information is lost about the sequential flow of classroom activities. There are clearly problems with this in that observers have the formidable task of evaluating the teacher, the pupils and their interaction. This assessment must be fair and provide a basis for discussion and therefore must be carefully backed up by fact. It must be far more accurate and beneficial than a "mock bureaucratic ritual", and it must involve "performance review, development review and potential review" for pupils and teacher. Concentrating on any individual focus would fail to appreciate the synergistic wholeness of the lesson. In addition to this there is the problem of subjectivity. Often during concurrent appraisal the teacher is provided with a slightly biased unsubstantiated appraisal. Fanselow (1977) pinpoints in a rather colourful way the problems involved in appraising lessons thus:

> When teachers, employers, students or sales people discuss the same lessons, texts, tests, methods and schools of language teaching they often sound like the characters in the Japanese movie, Rashomon – they each give contradictory and equivocal accounts of the same events or items.
> He then goes on to describe an instrument – FOCUS – for observing communication used in specific settings.

The simple conclusion is that meaningful evaluations cannot be made unless an accurate picture can be given of what is actually going on in lessons. A more accurate picture of lessons can be gained by using an observation instrument to create a picture of interaction, which gives a concrete picture of verbal behaviour. Such systems operate on a series of categories which the observer tallies whenever the particular category is discernible in the lesson. These can then be entered onto a matrix to provide a graphic picture of the lesson. This enables teaching/learning patterns to be analysed. Percentages for each category can be calculated and ratios discovered. Most instruments of lesson analysis are essentially adaptations of the original system developed by Flanders.

Instruments should include an analysis of teacher talk and of pupil talk and also some way of monitoring other types of verbal behaviour and the absence of it. Although in themselves they do not demonstrate how meaningful the interaction is, they do provide useful data for this purpose. Many such instruments have been created and trialled since the 1960s, some specifically concerned with language acquisition and others more general and appropriate to a greater or lesser degree to various curriculum areas. These instruments have differing numbers of categories which are ticked as appropriate at time intervals during a lesson, for example every 3 seconds. Appendix 1 shows some language-specific models quoted by Chaudron (1988), based on research by Long (1976). In this selection the number of categories varies from seven to 73 and there is also variety in the specific time period involved for observation.

It would be impractical within the confines of this article to attempt to analyse the many instruments that have been developed to date; it is sufficient to highlight some aspects of them and mention some of the more famous instruments. Perhaps the most well known and certainly the original is the Flanders system, described by Medley and Mitzel (1963) as «the most sophisticated technique for observing classroom climate» [quoted in Allwright (1988:68)].

Flanders employed ten categories, seven were teacher-related and these were divided into direct and indirect ones, then came two student-related and one silent category. «Teacher talk» included the categories of accepting feelings, praising and encouraging, accepting or using ideas of pupils, asking questions, lecturing, giving directions, criticising or justifying authority. The pupil categories were pupil response and pupil initiation. There was then a silent category to cover pauses or any confused communication which could not be understood by the observer.

Moskovitz (1976) created another well-known instrument which is worth mentioning because it took the Flanders categories and, by making additions and adaptations, made an instrument more suitable to the Modern Language Classroom. Under «Pupil talk» were included choral response, reading aloud, responding to the teacher with opinions and feelings, being off the subject, being non-task orientated. The teacher categories were extended to include jokes, repeating student response verbatim, asking cultural questions, discussion, personalising, modelling and several others.

Jarvis' (1968) model is also quite impressive because of the way its 24 categories capture characteristics of the Modern Language classroom, differentiating between drill language and language for real communicative purposes.

A more recent and perhaps less cumbersome instrument is that of Bowers (1980) which drastically reduces the number of categories to seven. Bowers also includes the vital differentiation between language used directly to teach, and language used for normal social and organisational purposes.

There are obviously many problems involved with using these instruments: firstly, using an existing instrument may not enable the researcher to achieve the focus required for a specific purpose and secondly it can take a very long time to become sufficiently acquainted with an instrument to be able to use it effectively, a point

not lost on Moskovitz (1967), who produced a programmed manual for teachers to teach themselves how to use the Flanders' system.

3. The Benefits of Classroom observation.

Another issue which needs to be addressed is the usefulness of these instruments in classroom observation. Teacher trainers and those carrying out any form of appraisal would no doubt welcome any system which took the overriding subjectivity from teacher assessment and made it more valid. Dick Allwright gave a paper on this in 1972 entitled «Prescription and Description in the Training of Language Teachers». He outlined the problems faced by supervisors as they carry out prescriptive assessment and arrived at the conclusion that prescription is inappropriate if descriptive techniques are inadequate. A teacher trainer is far better employed encouraging students to examine themselves what actually goes on in lessons.

In any discussion from the perspective of the trainers, it is important to be acquainted with the work of Grittner, State Supervisor of Foreign Language with the Wisconsin Department of Public Instruction in 1969. In *Teaching Foreign Languages* (1969) he concluded, from his own visiting of schools, that teachers have a distorted view of what actually takes place in their lessons, and are astonished when they hear an audio recording of those lessons. Even in what was supposedly an audio lingual lesson, 80% of the classroom work involved teacher talk, most of which was in English, and on the rare occasions when the pupils did speak, their responses were either repetitions of what the teacher had said or responses elicited from written material. All functional communication carried out between teacher and pupils was in English. The teacher's use of the target language was minimal and pupil use nil. In the light of these observations Grittner became interested in what he termed «interaction analysis» as a means of changing teacher behaviour. Subsequently Grittner dropped his role as a supervisor and became involved in encouraging and instigating teacher self observation. He spent

a considerable amount of time working on publishing Nearhoof's adaptation of the Flanders system which stressed the need to distinguish between real and drill use of language.

Moskovitz not only saw very clearly the benefit of using observation systems as a solution to supervisor problems, but also the fundamental benefit in terms of trainee feedback. What she was in fact advocating was a trainee centred approach to training teachers. Teachers can find out many things from the results of this analysis. Moskovitz lists some of these as follows:

1. What percentage of the class time does the teacher talk?
2. What percentage of the class time do the pupils talk?
3. Does the teacher use more direct or indirect influence in the lesson?
4. Is the teacher more direct or indirect in motivating the class?
5. What immediate feedback do pupils receive?
6. To what extent is there pupil participation for extended periods of time?
7. What behaviour is used to elicit pupil response?
8. To what extent can pupils bring in their own ideas?
9. What behaviour does the teacher use extensively in communicating?

These observations clearly illustrate something of the value in terms of personal development. Although evidence of the benefit to teachers is sparse, Moskovitz does refer to some research in this field in her paper, «The effects of training foreign language teachers in interaction analysis». The research referred to was carried out on both in-service and trainee teachers. Fourteen subjects were asked to assess their teaching by using the Flanders system of classroom observation. It was discovered as a result, that teachers used more indirect behaviours in motivating and controlling their classes, more praise, less direct influence and there was more acceptance of pupils' feelings and ideas and less criticism. The students on the other hand, were observed to initiate their own ideas more, give fewer na-

rrow predictable replies, initiate the conversation more and talk for greater lengths of time when they expressed their own ideas. At the end of the training time the student teachers were asked for their evaluation of the system. They were asked initially whether the study of interaction analysis applies to foreign language teaching. They were given a scale of 1-9, with 1 indicating 'not at all' and 9 'a great deal'. The mean score for the group was 7.8 indicating that they felt the Flanders' system had considerable application. They were then asked if they felt the study of interaction should be made a requirement for language teachers, six strongly agreed, eight agreed, no-one chose neutral, disagree or strongly disagree. Two other open-ended questions were also asked. In the first instance they had to complete the phrase, «study of interaction analysis has made me realise...»

Responses included:

> ...that I can plan the way I am going to behave in the classroom, what is going on in the classroom and what is successful under specific conditions; exactly what I am doing in front of the class both good and bad, how my students react to my behaviour, that controlling my behaviour and the behaviour of my students can be done, what behaviours I use, what behaviours I don't use and why.

The second open-ended statement was, "I believe the most important things the foreign language teacher can gain from knowing the Flanders system are"; and the responses were as follows:

> An understanding of how to elicit student responses, original ideas and their feelings, how to get the students to contribute in class without fear, how to develop and use behaviours which accept encourage and praise the student, the importance of having some tangible checklist to analyse behaviour in class, the need for more and varied encouragement so as not to sound trite with merely "bien", the basics of teaching no matter what the subject, how to react to student responses, varying teaching techniques.

The conclusions were that the system led to more positive attitudes toward teaching, more positive attitudes by pupils, the use of more indirect teaching patterns and more expression of the pupils' own ideas in the foreign language. This is backed up by in-service teachers who recorded similar results. They claimed increased classroom sensitivity, more indirect behaviours, improved interaction, increased understanding of pupil perceptions, greater variety and trying to get more participation. According to the Moskovitz research the only really negative responses encountered were concerning the relationship with the co-operating teacher or supervisor, but this is a relationship which needs nurturing and which can be fraught with difficulties whether an instrument is used or not, in fact the element of producing concrete evidence should ease the relationship rather than create further difficulties.

4. Practical Applications of Classroom Observation

At this point we hope to have established two points: that accurate classroom observation leads to helpful feedback, and that accurate classroom observation is achieved with difficulty, by means of cumbersome instruments. If we are to use classroom observation as a daily technique in our training of teachers, we must find a way to make the existing instruments more "user-friendly". The solution we have arrived at and which indicates interesting directions for the future, is the combination of the video camera with the types of instruments described earlier.

Video recording is regularly done in teacher training, but the tendency is to use the film as a basis for discussion rather than objective analysis of the type that the observation instruments described earlier have tried to achieve. The advantage of using the video camera in conjunction with those instruments is that the film allows the observer to overcome the main problem of observation schedules — the time and focus element. It is very easy to say: «Record interactions every three seconds and check them off on a chart»; in

practice this is neither easy nor accurate. First, there is the problem of focus: if the observer is logging teacher/learner interactions, where do simultaneous off-task learner/learner interactions fit in? Second, there is the problem of observer fatigue: a 40-minute lesson has 2,400 seconds, which means 800 observations. This is a very laborious process in which the observer is being used as a recorder — something which modern technology has now rendered irrelevant. Making a video-recording of a lesson affords the observer the opportunity to check off interactions more accurately because the tape may be stopped and re-run, so the observer is able to rest and to go back over a particular point. The focus problem is partially solved by filming also; careful positioning of the camera can permit the inclusion of a significant proportion of the class, so interactions other than those involving the teacher may also be logged.

The technical problems of filming deserve an article of their own. Although the modern video recorder is a very effective instrument, the varied lighting conditions in classrooms mean that good results can only be guaranteed with a larger camera, which is an intrusive instrument of which the learners are always aware, especially when the observer or a technician is operating it. An alternative which can work if the facilities are available is to set up two cameras on tripods, one at each end of the room and to leave them recording. The results may be reviewed independently or edited together. Another important advantage of the film is that different aspects of the lesson may be observed in different viewings. One viewing could concentrate on verbal interactions, while another session with the same recording could analyse body language, distribution of questions, board work and the timing of elements of the lesson. It is one thing to suggest that a teacher adopt a sunnier appearance, it is another for that teacher to see his/her own tense, nervous expression.

The use of two cameras goes some way towards solving another problem: sound quality. If a teacher is expounding or asking ques-

tions to individuals, the voices are usually recorded with reasonable clarity. In the modern classroom however, where pair and group work is encouraged, much of the interaction is lost on the recording of a single camera. If a really accurate representation of classroom interaction is required, it is possible to place cassette recorders on tables where groups are working. It is *not* a good idea for the camera operator to 'go walkabout' in an attempt to sample pair or group work.

The most important benefit of this technique has to be the opportunity for the trainee teacher to be presented with a written analysis of his/her teaching together with a visual record against which the analysis may be compared. The traditional lesson evaluation involves relationship in which one party gives to another, one «Rashomon» version of the lesson, against which there can be no appeal because the only record is that of the observer's tallying. If both parties are viewing the filmed record from which the lesson analysis has been made, first of all the relationship has to become more collaborative, secondly both parties are viewing the same objective record, even if they may not agree on their interpretations of that record.

5. The Application of the Method

The method used in the present study was to take a tally of activities every four seconds rather than the three seconds which is quoted in earlier studies. It was found that the shorter interval led to questionable accuracy, even with the advantage of being able to stop the tape, and that the longer interval was a more realistic division. The headings on the tally sheet were the following:

Whole Lesson Activities
Total mother tongue
Total target language
Confusion

Silence

(Within the "confusion" and "silence" classifications were included reading and writing activities.)

Breakdown of Speech Activities
Teacher real communicative purposes L2
Teacher real communicative purposes L1
Teacher lectures drill L2
Teacher lectures drill L1
Teacher questions L2
Teacher questions L1
Teacher responds drill L2
Teacher responds drill L1
Pupils respond drill L2
Pupils respond drill L1
Pupils volunteer information drill L2
Pupils volunteer information drill L1

These categories concentrate on the oral/aural elements of language teaching, which is justified in that the National Curriculum has directed our teaching more in this direction. Where learners are engaged in pair or group work, the observer has to decide whether to film a representative group or to take the camera off its tripod and circulate around the class. Although pair and group work tends to be in the reading/writing areas, it is important to record how much learner/learner interaction takes place in L1 and L2.

Six lessons were observed to test the methodology described in this article. The total durations of the activities were added up and expressed in percentage form (see Appendix One). The pie-charts were generated using the Microsoft Excel package.

A possible area of error was in the accuracy of tallying the activities, but a test was carried out in which two observers compared the results of observing the same lesson and no significant difference was found.

6. Results of the Experiment

The trainee teachers were surprised and disappointed at the amount of mother tongue they were found to have used in their lessons. Their in-college training had emphasised the importance of L2 communication and the very nature of modern teaching materials imposes an oral/aural emphasis on the taught language. Nevertheless the analysis showed a bias towards using the mother tongue and subsequent discussion with the observer meant that the trainees could see the mathematical evidence and, by watching the filmed record, see in what circumstances they were using L1 and make notes on strategies to solve the problem in future.

7. Conclusions

The conclusions that we may arrive at as a result of this experiment are:

• there are practical problems in making a comprehensive filmed record of a lesson;

• the tally sheet used in this experiment recorded only oral/aural activities;

• the film allowed easier and possibly more accurate tallying than live observation;

• the mathematical analysis of activities was advantageous to the trainees in that it gave them an instant, visual breakdown of their work;

• the experiment supported the findings of Grittner in that trainees were surprised at the difference between their perceptions of the lesson and the objective record;

• despite the emphasis on L2 in communicative teaching methodology, the teachers and learners observed in this experiment relied heavily on L1 in classroom interaction.

We feel that the combination of filming and physical recording of activities illustrated in this experiment has shown itself to be va-

lid in terms of analysing the spoken content of language lessons. Future developments of this methodology should include:

• a more scientific method of observing pair and group activities and learner/learner interactions;

• an objective methodology for evaluating reading and writing activities;

• opportunities for learner feedback to be included in lesson evaluation.

WORKS CITED

Allwright, R.L. 1972. Prescription and Description in the Training of Language Teachers. In Quistgard *et al,* (eds.)

— 1988. *Observation in the Language Classroom.* Harlow, Longman.

Allwright, R.L. & Bailey, M. 1991. *Focus on the Language Classroom.* Cambridge, Cambridge University Press.

Bowers, R. 1980. Verbal Behaviour in the Language Teaching Classroom. University of Reading, unpublished Ph.D. thesis.

B.B.C. *A Child's View.* «Horizon» documentary. Date unknown.

Chaudron, C. 1988. *Second Language Classrooms.* Cambridge. Cambridge University Press.

Clark, J.L.D. 1969. The Pennsylvania Project and the «Audio-Lingual vs Traditional» Question. *Modern Language Journal* 53: 388-396.

D.E.S. 1988. *Education Reform Act.* London, HMSO.

D.E.S. 1988. *Statutory Orders and Non-Satutory Guidance for Modern Foreign Languages.* London, HMSO.

Ellis, R. 1990. *Instructed Second Language Acquisition.* Oxford, Blackwell.

Fanselow, J. 1977. Beyond Rashomon. *TESOL Quarterly,* 11, 1, 17-39.

Flanders, N. 1970. *Analysing Teacher Behaviour.* Reading, Mass. Addison-Wesley

Geddes, M. & Sturtridge, G. 1982. *Video in the Language Classroom.*London, Heinemann.

Good, T.L. & Brophy, J.E. 1984. Looking in Classrooms (3rd Ed.) New York, Harper and. Row

Grittner, F.M. 1969. *Teaching Foreign Languages.* London, Harper & Row.

Hammersley, M. 1985. *Ethnography Principles and Practice.* London, Tavistock.

Hill, B. 1982. *Making the Most of Video.* London, CILT.

Jarvis, G.A. 1968. A Behavioural Observation System for Classroom Foreign Language Skill Acquisition Activities. *Modern Language Journal* 52: 335-341.

Klapper, J. 1991. The Role of the Video Camera in Communicative Language Teaching and Evaluation. *Language Learning Journal,* No.4, Sept., 1991.

Krashen, S. 1981. *Second Language Acquisition and Learning.* Oxford, Pergamon Press.

— 1983. *The Natural Approach.* Oxford, Pergamon Press.

Lonergan, J. 1990. *Making the Most of the Video Camera.* London, CILT.

Malamah-Thomas, A. 1987. *Classroom Interaction.* Oxford, OUP.

Marsh, C. 1989. Some Observations on the Use of Video in the Teaching of Modern Languages. *British Journal of Language Teaching,* 27, 1, Spring, 1989.

McKenzie, P. 1989. Teacher Appraisal. Victoria, *Journal of Australian Education,* Review No. 28.

McMahon, A. 1992. Teacher Appraisal. *Journal of Teacher Development,* 1, 1.

Medley, D. & **Mitzel, H.** 1963. Measuring Classroom Behaviour by Systematic Observation In Gage, *Handbook of Research on Teaching.* 247-238. Chicago, Rand McNally.

Moskowitz, G. 1967. The Flint System. In Boyer, *Anthology of Classroom Systems* (Philadelphia Section) 15, 1-15.

— 1968, The Effect of Training Language Teachers in Interaction Analysis. *Foreign Languages Annual, 1:218-235.*

— 1976 «*The Classroom Interaction of Outstanding Language Teachers.*» *Foreign Language Annals 9/2.*

Nearhoof, O. 1969. Teacher Pupil Interaction in the Modern Language Classroom. Unpublished paper cited by Grittner (op. cit.).

Nunan, 1992. Research Methods in Language Learning. Cambridge, Cambridge University Press.

Pearson, J. 1990. Putting Pupils in the Picture Using the Video Camera. *Language Learning Journal,* 2, 71-77.

Sinclair, J. & **Coulthard, R.M.** 1975. *Towards an Analysis of Discourse.* Oxford, OUP.

Ullman, R. & **Geva, E.** 1984. Approaches to Observation in Second Language Classes. In Brumfit, C. J. (ed.) *Language Issues and Education Policies* (ELT Documents 119) Oxford. Pergamon.

Wragg, E.C. 1987. *Teacher Appraisal.* Basingstoke, Macmillan Education.