Part I Classroom research: what it is and why it is important

The two chapters in this first section of the book will provide the framework for interpreting what follows. Chapter 1 explains what we mean by classroom research and traces its history in the language field, in order for the reader to gain a backdrop on the issues which have emerged as lasting concerns. We will then turn, in Chapter 2, to the question which underlies this book: why focus on the classroom? These two chapters, taken together, constitute a general framework for looking at what happens in the language classroom.

A note of caution is in order here regarding the relationship of language classroom research to other fields of science. Classroom research differs from other types of research with which you may be more familiar. Compared with psychology or with medicine, for example, it seldom approaches the relative simplicity and rigorous control of formal experiments. Compared with general educational research, on the other hand, language classroom research adds the extra but necessary complexity of dealing with situations where language is both what is being taught and the means by which it is being taught, and often with situations where more than one language is used in classroom talk. Finally, language classroom research is a little like anthropology in some ways, where researchers try to understand what is going on in particular social or cultural settings. Although language classroom research has only occasionally looked at cultural differences, the individual classroom can be considered (and studied as) a cultural entity.

So language classroom research has its own distinctive tradition. But it is not an entirely novel tradition. The backdrop to be described in Part I is in fact based to a considerable extent on a history of borrowing and adapting from other research traditions. This history has led to a research framework that is still evolving. What it has already achieved is a strong (and still growing) awareness of the tremendous depth and richness of the language classroom as a site for the investigation of language teaching and learning.

1 The development of classroom research

Before we begin to present the findings of classroom research, we need to acknowledge that it has a chequered history. To put the findings in proper perspective, we shall, first of all, loosely define the field and present a brief survey of the different themes that have characterised the last twenty or so years of classroom research on language teaching and learning. Our intent is to show how some themes have retreated into the background in spite of their early promise, and how new ones have come forward to take their place. It is a research field still really in its infancy, still working on major breakthroughs, but at the same time already well established and already making a distinctive contribution to our understanding of classroom language learning and teaching.

1.1 What is classroom research?

Classroom-centred research is just what it says it is – research *centred on* the classroom, as distinct from, for example, research that concentrates on the *inputs to* the classroom (the syllabus, the teaching materials, etc.) or on the *outputs from* the classroom (learner test scores). It does not ignore in any way or try to devalue the importance of such inputs and outputs. Instead classroom research simply tries to investigate what actually happens inside the classroom. At its most narrow, it is in fact research which treats classroom interaction as virtually the only object worthy of investigation.

This conception of classroom-centred research is far too narrow for our purposes here, however. We shall want to include studies that focus on classroom language learning in many different ways, and we shall want to refer to some that do not directly investigate classroom interaction at all.

We take 'classroom-centred research' (or simply 'classroom research' hereafter) as a cover term, then, for a whole range of research studies on classroom language learning and teaching. The obvious unifying factor is that the emphasis is solidly on trying to understand what goes on in the classroom setting. Examples of the types of issues that have been studied so far include how teachers respond to learners' errors, how interaction occurs in classrooms, the type of linguistic input provided in classroom

settings, the feelings of teachers and learners at various points during or after lessons, and so on. Most of these topics will be discussed in some detail in Chapters 5 to 10 of this book.

1.2 How is it done?

Basically research on classroom language learning involves a good many different procedures. As with any other sort of research, the starting point will involve getting as well informed as you possibly can on the issue you wish to investigate. We can call this consulting 'expert opinion', and it will include catching up with the speculative and theoretical literature as well as with earlier research.

Consulting expert opinion in this way does two things. First of all, it tells you what issues have already been thought about and investigated, and should lead you to sort out more precisely whatever you eventually decide to investigate yourself. Second, it should tell you how related questions have been investigated in the past, so that you can make a more informed decision about how to start doing your own study, and about the procedures you will eventually use.

Doing the actual research is essentially a matter of data collection and analysis. Typically you need some sort of a record of what happened in a particular classroom or classrooms, so that you can analyse the record (your data) and describe the classroom processes in whatever terms interest you. There are, broadly speaking, two different ways of getting such a record.

The first and most obvious way is to develop a data base (the record) by direct observation. You could observe the classroom processes by sitting in the classroom and taking notes. To make your record more focused you may decide in advance what you are going to be looking for, make an observational schedule with the categories you consider appropriate, and then record what happens under those category headings. If, for example, you are interested in seeing whether boys or girls do more of the talking in class, you might have just two categories ('boys' and 'girls') and then make a tally in the appropriate column on your observation sheet whenever a child speaks. For more complicated enquiries, and so that you can have a more complete record anyway, you may prefer to audio-record, or even video-record, what goes on, so that you can go back in detail to what was said, by whom, in what tone of voice, and so on. For that you will probably need to transcribe the interaction from the tape, so that you will end up with a comprehensive written record of the lesson to analyse, as well as the original tape itself.

Direct observation is not always the most appropriate way to gather classroom data. Sometimes it seems too risky because of the likelihood

that being observed will change people's behaviour. Can a teacher really be expected to carry on just as usual with an observer sitting at the back of the room, or with a video-camera there, or even with just a modest audio-cassette machine sitting on the desk? Furthermore, there are many interesting aspects of classroom processes that are not actually observable in any very reliable or measurable way. If we want to know what makes people anxious in class, for example, then just looking at them will not tell us everything we want to know. Sometimes people do look anxious, of course, but a lot of people who in fact feel very anxious may be able to cover it up very convincingly in class. If we want to investigate anxiety, then, some other way of eliciting the data will be necessary.

The second way, the obvious alternative to direct observation, is simply to ask, to give people an opportunity to report for themselves what has happened to them and what they think about it. The traditional way of getting such 'self-report' data is to conduct surveys, usually through interviews or by written questionnaires. Although we must not assume that such survey questions are always answered truthfully, or even carefully, we can try asking people to tell us what has happened to them in class, and treat their replies as our basic data, or at the very least as data to consider alongside whatever we have been able to observe directly. Survey data collection techniques normally involve working out in advance the categories you are going to use for your investigation (but this time for your questions rather than your observations – what you are going to ask about). For this reason, you need to think very carefully about the wording of your questions so that nobody can say you got the answers you did just because of the way you posed the questions.

The problem with observation schedules and with surveys is precisely, however, that someone has to decide in advance what to look for, or what to ask questions about, and that someone is most often a researcher who is not typically directly involved in the classroom being investigated. Such relatively 'closed' techniques may easily miss the insights that could be provided by the participants themselves, the teachers and the learners, beyond anything the researcher has thought to build into the data collection procedures.

To capture these sorts of insights, alternative forms of data collection are needed – some more 'open' form of self-report, for example, where the participants themselves, without specific prompting on specific issues, try to record their experiences and perceptions for the researcher. For example, a learner's diary may reveal aspects of the classroom experience that observation could never have captured, and that no one would have thought of including as questions on a questionnaire. The relationship of competitiveness and anxiety in language classrooms has been explored this way (see Chapter 10 for further details), through the analysis of a number of diaries by different learners in different classes. They were not

asked to write about competitiveness or anxiety, but these proved to be important issues for many of them, given what they did write, unprompted.

The desire not to prejudge the importance of events observed has led some researchers to consider the procedures and philosophies of ethnography as a viable approach to collecting and interpreting classroom data (see van Lier 1988; Watson-Gegeo 1988; Watson-Gegeo and Ulichny 1988). 'Ethnography' has been variously defined, but basically it is a collection of procedures (and an attitude) used predominantly by anthropologists who attempt to document and understand the behaviour of people in cultures. Originally, ethnographers studied exotic cultures, but in recent years, ethnography has been used to study people in many diverse but everyday settings, including restaurants, urban street gangs, schools, etc.

Ethnography has gained considerable support as an approach to classroom research in education in general (Wilson 1977; Erickson 1981; Green and Wallat 1981), as well as in studies of language teaching and learning. For example, van Lier (1988) makes a convincing case for ethnography as potentially the most useful means to study classroom phenomena. He stresses a point which is relevant here, namely, to understand what happens in classrooms, researchers must try to get at the meaning given to these events by the participants themselves.

Thus, the main tools of the classroom research trade are observations (including recordings), surveys, and other forms of self-report. These are tools that set classroom research apart from research in the so-called 'hard' sciences and which draw from social science and ethnography. There is a further tool that classroom researchers often utilise: test data. Of course, the use of test results is in no way special to classroom research. Although there is a lot that can be studied without anyone worrying about whether or not a particular teaching point has been successfully taught (that is, learned), in the long run we must all be interested in the notion of effective learning, and we want to know whether the particular issues we study (sex differences in participation, or classroom competitiveness, or whatever) are important in relation to effective learning. So classroom researchers, like most others, need from time to time to collect test data (and/or other sorts of performance data – spontaneous conversations, essays, and so on) along with information gathered from their observations, their surveys, or students' and teachers' self-reports.

1.3 Where did classroom research come from?

Classroom research is certainly not unique to language teaching and did not even originate among language-oriented researchers. In fact, it took

language teaching some time to even begin to catch up with the rest of the educational world in this respect.

Modern classroom research began in the 1950s, among teacher trainers, in response to the problems involved in helping student teachers in subject-matter classes by giving them feedback about their performance in class during their teaching practice. The trainers realised that they needed to investigate what constituted effective teaching, and then find a way of incorporating their findings into effective teacher training. As we shall see, the first of these issues – what constitutes effective teaching – has proven to be so complex and so fascinating a research problem that teacher training itself has slipped gradually into the background.

1.4 How has it developed?

Just as teacher training provided the earliest concerns, which centred on the attempt to determine what constitutes effective teaching, so teacher training also provided the basic tools of classroom observation – the 'observation schedules' themselves. Researchers such as Flanders (1970) had used direct observation to study teaching, and had developed observation schedules – essentially lists of the categories of teacher and learner behaviour thought to be most closely related to successful teaching. (See Appendix A for a copy of Flanders' system, sometimes called 'FIAC', for 'Flanders' Interaction Analysis Categories'.) These observation sheets were used to help teacher trainees to see just how well (or how badly) their teaching behaviour matched the patterns that research (or consensus) had suggested would be most effective.

Unfortunately the early, almost euphoric, confidence in the findings of such research did not survive many years of scrutiny by an increasing number of researchers who found classroom behaviour altogether too complex to be reduced to a few categories. It seemed that applications in teacher training were therefore premature, and that a major effort should first be put into trying to unravel the enormous complexities of classroom behaviour.

One probable reason why the language teaching profession came late to classroom research was that, just when teachers of other subjects were losing confidence in their methods, language teachers were enjoying a period of unprecedented confidence in theirs, with audiolingualism very widely used and intellectually accepted, especially in the United States. The audiolingual method was related to highly developed theories in linguistics and behaviourist psychology, and had apparently proved its practical worth by its success in military language training programmes during World War II. However, in 1959 the confidence in audiolingualism as a theory was seriously undermined, by Chomsky's (1959) all-out

attack on Skinner's (1957) behaviourist views on language and language learning. And in practical terms audiolingualism was further damaged by reports that learners were not only bored by the repetitious drills that occupied so much of their time, but were not really learning any more than they ever had. Nevertheless there was no immediate loss of confidence in the concept of method itself, since an alternative method to audiolingualism was quickly found. In fact, the alternative drew from cognitive psychology (hence the new method's rather odd name – 'cognitive code') and used arguments from that field to bring together some ideas that were genuinely new and some from the already well-established but otherwise outmoded grammar–translation method.

At that time (the early 1960s) language teacher training, naturally enough, revolved around the issue of which of the major competing methods one should prescribe to beginning teachers. Had Chomsky really succeeded in refuting Skinner and the whole theoretical apparatus of audiolingualism? Had audiolingualism really failed in the classroom, or did we just need a much more imaginative approach to drill writing? The underlying assumption was still that what mattered was finding the right method. In effect, what happened in the classroom, and therefore what was learned, was thought to be completely determined by the choice of method.

Unfortunately for the methodologists, however, the major experiments conducted in the 1960s to decide between the main competing methods proved miserably inconclusive. Scherer and Wertheimer (1964), working at university level in Colorado, compared what was at the time the modern audiolingual method with traditional grammar–translation teaching and found no significant differences overall, for a two-year trial period. Soon afterwards came the massive Pennsylvania Project, where, this time at secondary school level, audiolingualism was now being compared with traditional teaching informed by the latest 'cognitive' ideas. After a total of four years of teaching, this experiment (Smith 1970) also failed to demonstrate the expected superiority of either of the methods being compared, although the 'traditional' method yielded better results on some tests of reading skills.

1.5 What has happened to the early concerns?

In the long run, the ultimate loser in this inconclusive battle was neither audiolingualism nor its cognitive rival but the very notion of making global methodological prescriptions. It no longer made sense to imagine that any one method would ever prove in some absolute way to be superior to its competitors. No one method could therefore be

prescribed, like some patent medicine, with absolute confidence in its global, universal, effectiveness.

Grittner (1968:7) summed up the situation at the time by suggesting: '... perhaps we should ask for a cease-fire while we search for a more productive means of investigation'. Some researchers decided to move a step down in the Approach, Method, Technique hierarchy described by Anthony (1963) and do small-scale research at the level of technique instead of large-scale research at the level of method. (In Anthony's terms, the 'approach' was the philosophy of language and learning that provided the theoretical underpinnings of language teaching. The 'method' was a systematic collection of activities and procedures which were derived logically from the approach. The 'techniques' were the various activities implemented during a lesson, which stemmed, in turn, directly from the method chosen.)

Researchers on the Gothenburg English Teaching Method Project (which is often referred to by its Swedish abbreviation, GUME) compromised by working at the level of technique but still in the hope of establishing the validity of such comparisons. (See Lindblad 1969.) They were in fact hoping to test the usefulness of grammatical explanations framed according to Chomsky's 1957 version of Transformational Generative Grammar. In essence, this meant trying to test an 'approach' (Chomsky's theory) by experimentation with 'techniques' (in this case the provision of certain types of grammatical explanation as compared with no explanation at all). The results were again inconclusive, at least until the researchers switched from children to adults, who did seem to learn better from explanations and practice than from practice alone. Even these positive findings, however, did not amount to verification of the absolute validity of any global methodological prescription. The mere fact that they had obtained different results with children and with adults damaged any truly global claims. And, in any case, the relatively small-scale nature of the GUME project – in terms of the number of lessons involved, the number of teaching points covered, the fact that the teaching was not live but on audiotape, and so on – would suggest that we ought to be very cautious about drawing global conclusions from this research.

Meanwhile, in the United States, Politzer (1970) had already conducted and reported upon a seminal study in which he videotaped a number of secondary school French classes. He recorded the frequency with which certain techniques (mostly different types of structural pattern practice) were used, and related the frequencies to learner achievement in the different classes. His results were complex and make very interesting reading. However, they add up to strong evidence that small-scale research at the level of technique was by no means ready to support a prescriptive approach to teacher training. In fact, Politzer

retreated from prescriptivism himself when he articulated his 'principle of classroom economics': he noted that the value of any technique depends in part on the relative value of other techniques that could have been used in place of the one actually selected by the teacher.

Politzer concluded that 'the very high complexity of the teaching process makes it very difficult to talk in absolute terms about "good" and "bad" teaching devices' (1970:43). It was becoming increasingly clear that language teaching is vastly more complicated than that. Even now, over two decades after Politzer's study was first published, teacher trainers are not in a position to tell their trainees just which techniques to use and which ones not to use – not in a position, that is, that can be justified by solidly established research results.

Having already retreated from a focus on 'method' to one on 'technique' it therefore seemed necessary to retreat at least one step further back into the unknown. In fact two moves were involved. First, it meant retreating from prescription altogether in favour of a descriptive approach, and second, it meant retreating from 'techniques' to 'processes'.

These two moves, taken together, meant trying to find ways of describing classroom processes to find out what actually happens in language classes. We could no longer assume that all that happens is that a particular method, or a particular set of techniques, is simply implemented. Instead, we began to feel that something below the level of technique (something more interactive and less obviously pedagogic) takes place, and that this interaction (that is, whatever actually happens in the classroom) is likely to provide a fruitful topic for investigation.

In the pursuit of these two retreats, first from *pre*scription to *de*scription, and then to the description of classroom processes, two somewhat different viewpoints have emerged in the last decade. Some researchers (those with more of a sociological outlook on education, in fact) have tended to look at the language lesson as a 'socially-constructed event', which all the people present produce through their interactive work. Put more simply, such researchers have stopped looking at teaching as if everything of importance came from the teacher, and have instead started looking at the way people interact in the classroom to collectively produce the learning opportunities that arise there. For example, one of the present authors (Dick Allwright) has gone from looking at how teachers correct errors (see Chapters 5 and 6) to investigating how teachers and learners together determine each learner's level of participation in classroom activities (see Chapters 7 and 8). He has now moved on to investigating how learners' contributions to classroom interaction affect the syllabus the teacher is trying to teach, and therefore what learners can get from lessons.

One viewpoint, then, has stressed the importance of social interaction.

The other viewpoint is the one adopted by more directly language-oriented researchers who have chosen to look at the classroom as a setting to study how language might be acquired from the input provided by the teacher's talk (see section 8.1). Language-oriented investigations have also examined the language produced by learners in classroom settings.

As we shall see in the remainder of the book, these two viewpoints (the sociological orientation and the linguistic orientation) can be viewed as complementary, rather than as being in competition with each other. For example, the 'sociological' interest in the way learner behaviour affects the syllabus can easily be related to the 'linguistic' interest in the nature of the language the learners are exposed to in the classroom. In fact the two viewpoints *must* come together if we are not to miss what might be valuable insights into classroom language learning.

1.6 What has happened to the basic tools?

In the move from teacher training to more fundamental research, typified both in the move from prescription to description and in the moves from method to technique to process, what has happened to the basic tools of classroom research? You will recall that they were originally borrowed from general educational research, and consisted mostly of techniques for using observation schedules (lists of categories) for the classification of teacher behaviour. (Relatively little attention was paid to *learner* behaviour at the time, given the focus on *teacher* training.) These observation schedules had to be modified to make them appropriate to the obvious complexities of language teaching, where language is the content as well as the medium of instruction, where more than one language may be used, and where, as in pronunciation practice, all the learners may need to have a chance to try to produce the same answer to exactly the same question, sometimes simultaneously.

The starting point for many people was Flanders' pioneering work on 'Interaction Analysis' (1970). Flanders used this term for his ten-category observation schedule, which is reprinted in Appendix A. He designed it for general educational purposes, to be relevant to a variety of lessons, rather than for any subject in particular. In his work he combined a politically powerful idea with a very practical simplicity. The powerful idea was that teaching was more or less effective depending on how 'directly' or 'indirectly' teachers influenced learner behaviour. Flanders' was a pro-democratic, anti-authoritarian position that therefore looked for a positive relationship between a 'democratic' teaching style (that is, using indirect rather than direct influence) and learner achievement. He built this general idea into his ten categories, so that some would be used