

Reading the Discourse of Reading: a window on the meaning
construction processes of third year learners of French
Vol. 1

By Philip Hood, B.A.

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy, May 2000



Dedication

The completion of this thesis was made possible by four groups of people to whom I wish to record my thanks.

Firstly, the very many students who have submitted themselves to the various smaller and larger-scale projects that I have carried out over the last six years. I am especially grateful to Angela, John, Kirsty and Neil for their involvement in the Part One Project, and to the whole of 9K (too numerous to name) for their vital role in the Part Two Study.

Secondly, the many teachers who have so willingly supported my work in classrooms. Special mention of course goes to those teachers at the two schools I used extensively, who were at all times both patient and extremely helpful.

Thirdly, my colleagues at the School of Education, notably those in the Centre for Research into Second and Foreign Language Pedagogy - we have together expended much energy and emotion on the joint pursuit of doctoral degrees across the millennium ! I also wish to record my thanks to Colin Harrison for his advice and to the institution at large, which understands that a project such as this needs space and time.

Lastly, my family, who also understood that need, especially during the last phase of the work, and undertook many distractions (with varying degrees of enjoyment) to allow me to complete the writing.

Thesis contents

1. Introduction and Rationale
2. The Part 1 programme - Methodology and Analysis
3. A Literature Review
4. The Part 2 programme - Methodology
5. The Part 2 programme - Data Analysis - 1
6. The Part 2 programme - Data Analysis - 2
7. The Part 2 programme - Synthesis and Recommendations for Teachers and
Researchers
8. Bibliography
9. Appendices

Abstract

The main study of the thesis seeks to gain insights into the reading behaviour of 27 adolescent learners of French, who were recorded while completing four reading tasks, 2 individually and 2 in small groups. The transcripts are analysed from four perspectives. Firstly (as a baseline for the other perspectives) a quantitative approach examines factors surrounding relative success in the tasks. Secondly a coding system (Pressley and Afflerbach, 1995) is used to analyse strategy use by a selection of subjects in both individual and group contexts. Thirdly, the nature and quality of the talk of four of the groups is explored using Mercer's (1995) three categories of talk and aspects of Almasi 's (1995) work on sociocognitive conflict. Finally the problem-solving discourse of the transcripts is analysed in two ways - using a concordancing program and by investigating the roles in decision-making of key individuals in four different groups through discourse analysis techniques. A discussion chapter seeks to draw together the findings from these four perspectives. The processes used for the study are also discussed through chapters on relevant literature, on methodology and by a report of an earlier preparatory study.

'Too often our focus has been on what students *should be doing*; we must begin by asking what students *are doing*.'

Hosenfeld, 1976, p128,

cited as the conclusion to 'Learner Strategies: State of the Art Article'

by McDonough, 1999, p14.

Chapter One

Introduction and Research Design

1.1 Introduction - the purpose of the thesis

This thesis has its origins in a variety of contexts, which are located in classrooms, in curriculum development as well as in research literature. The academic sources centre principally on the research over the last twenty years into cognitive activity during the reading process, and specifically how to gain access to this activity during the reading of a foreign language. This is such a broad field that it too needs to be located in a more specific research context. Weber (1991, p101) writes (here about second language learning in the United States): ‘Learning the spoken form of the second language is primary, learning to read the language is secondary. Until recently this has been a long-standing notion in American second language instruction ...’. On p102, she adds: ‘The possibility that a learner’s knowledge may be confirmed, elaborated or extended through experience with the written language has not been directly addressed.’ And further, on p108: ‘The assumption seems to be that reading will follow from knowing the structure of the language and knowing how to read in the first language.’ We will be concerned in the thesis with a different learning arena from the one Weber is describing, (the early years of the secondary foreign language learning context in England.) But her words very accurately reflect the role of reading in the foreign language curriculum in the UK. Some reasons for this situation will be discussed in section 1.3.2, but we might also account for this view of the role of reading by noting a research strand which started with Alderson (1984). The theme here is the issue of whether reading problems amongst beginner stage

foreign language (FL) learners are due to issues of language (ie the scope of their FL knowledge) or of reading (ie their reading skills and ability to use strategies in either their first or a subsequent language). Subsequent contributions to this debate, (for example Devine 1988, Donin and Silva 1993, Bernhardt and Kamil 1995, Lee and Schallert 1997) have not concluded that it is the latter, indeed in some cases have suggested that it is definitely the former. The implication of this is that more complex reading should be delayed until the language threshold has been reached. This strand will be explored more fully as part of Chapter Three.

It is this view of reading and learning in a second or foreign language, which the thesis most wishes to address. It has sought to do this by gathering a range of evidence about reading behaviour in the early stages of learning a foreign language, (from an attitudinal questionnaire, different reading task styles, and using as the major instrument two distinct forms of verbal report). It also presents an analysis of that evidence from multiple standpoints. In doing this it is following a judgement made by Hosenfeld (1976, p128) but subsequently cited as still very relevant by McDonough (1999) that ' Too often our focus has been on what students *should be doing*; we must begin by asking what students *are doing*.' The primary intention is therefore not to promote a theory or to find a solution, but to 'discover or uncover propositions' (Maykut and Morehouse, 1994 p12). Using such propositions it will be possible for teachers, course developers and other researchers to form a view of whether reading can and should in future play a larger part in the learning process at an earlier stage and in different ways.

The thesis, then, describes a three-stage investigation, which began with a questionnaire which asked for self-evaluation of foreign language reading skills in Russian and some opinions on foreign language reading. This was administered to a large number of learners in a variety of schools. From the outcomes of this data some questions were formulated. These were focused on a small sample of the questionnaire cohort through small group work involving the completion of three different Russian reading tasks together. This attempted to gain a window on the processes involved in beginning foreign language learners' meaning construction. From the evidence discovered by this stage of the research further questions arose about the respective roles of individual and peer group contexts and of the effect of task on the approach to meaning construction. As a result the literature was fully reviewed and the major Part Two investigation was designed and carried out. This focused on learners of French in order that by using a language with roman script conclusions might be more valid for the majority of language learners in the UK. The study compared the strategies used in small group work with individual reading task approaches by a cohort of 27 learners. The data from this major part of the investigation was used to formulate the final propositions referred to in the previous paragraph.

1.2 Introduction - Research Philosophy

1.2.1 The global - location in a research paradigm

The two major research paradigms (positivism and the phenomenological approach) are often seen in opposition (for example Kuhn 1962, Lincoln and Guba 1985, Strauss and Corbin 1990, Ely 1991, Silverman 1993, Morse 1994). In fact it seems as if they are as

much the expression of a philosophy or conviction about the accurate description or ordering of natural phenomena as approaches to research. In Maykut and Morehouse (1994) the authors present (p12) such a distinction, adapted from Lincoln and Guba (op cit), through a set of postulates, which ask six global questions such as 'How does the world work ?' and 'Are causal linkages possible ?' The set of responses to the six questions define a great divide between the positivist and phenomenological approaches. But other writers qualify this 'divide' differently, eg Mason (1994) and Robson (1993, p19) who differentiates more between the 'laboratory' and the 'real world' than between 'scientific' and 'interpretive' approaches and advocates a greater willingness to cross the divide between quantitative and qualitative approaches if any project so demands. If we look at the literature on research into second or foreign language learning, we also find a blurring of some of the great distinction between paradigms. Whereas Brown (1988) does focus on statistical research almost entirely, he does (ibid p2) concede the value of case study as an approach. Chaudron (1988), Allwright and Bailey (1991), Nunan (1992) and Bailey and Nunan (1996) all take a view that the theme of the research will determine the best methodology and present examples of different projects to exemplify their discussion. Bailey and Nunan, (ibid, p2) also write that, 'It is important to note that global references to "qualitative research" and "qualitative data" can be more productively examined if we separate concerns of *data collection* and *data analysis*.' The approach used in this research is itself a hybrid, as we will now explore.

1.2.2 Fitting and misfitting the paradigm - a general outline

Philosophically this investigation takes a phenomenological starting point. It did not have a question to prove or disprove, but rather sought information about an aspect of foreign language learning. In terms of its observation structure through the textual analysis of transcripts it fits far better Silverman's (op cit, p9, table 1.2) qualitative research category than the quantitative alternative under his heading of methodologies (for example, observation is 'fundamental' and transcripts are used 'to understand how participants organise their talk'). But qualitative purists might argue with some of the organisation of this investigation. It chose to use a survey to gain some initial data and this was quantitatively analysed for broad patterns of opinion. The study then sought information about the way students usually read in a foreign language but to do so constructed a specific context for the observation to take place and specific materials to be used in that context. Thus it was not a naturalistic enquiry in its true sense, although, given that it was using less usual or perhaps even unfamiliar observational methods (think-aloud tasks and group-based interaction for reading), it would not have been natural to do that in the normal classroom setting anyway. Certainly the privacy of a separate room for both individual and group-based tasks appeared to increase the ease of the subjects and make them more 'natural'. But the programme did not proceed from theory. It followed more the principles of grounded theory (Glaser and Strauss, 1967, Strauss and Corbin, 1990) to the extent that, having formed a view on aspects of the theme from the use of one instrument, it sought further clarification and depth from the use of another. This then led to a further design and the data from this was interpreted in the light of the points learned from the earlier stages. While a final 'theory' is not produced, conclusions are drawn, and

so the process has led to propositions rather than proceeded from them. Perhaps most importantly, the study uses throughout an inductive approach to data analysis, which will now be further explored.

1.2.3 Aspects of a quantitative approach

There was an intention throughout to use quantitative methodology as a means of sorting data for further analysis, rather than as an end in itself. The initial questionnaire was designed to discover whether within and across different contexts a large number of respondents shared views on certain issues concerning reading in a foreign language (in this case, Russian). The performance data on all of the reading tasks used and the strategy counting exercises were again carried out to supply a bank of information to give a further perspective to the findings from transcript analysis. Similarly, concordancing was used to produce an easily digestible form of data about the frequency and context of certain key lexical items and discourse elements. But essentially and ultimately the aim is to understand the words rather than to count them.

1.2.4 Aspects of a qualitative approach

The major data set for this research is the words of the subjects. In all, forty five students spoke on tape during the stages of the project. More than thirty of these have had a detailed analysis made of part or all of their output. A further 214 students responded to the initial questionnaire, which involved a small amount of open-ended writing. Thus the voices of the subjects are central, even if they are engaged on more formal rather than more open-ended tasks. The objective was to discover what they do when they read, and

there is no reason to suppose that their actions were any different during the investigation tasks than they would be during normal classroom reading. The tasks here may have been more challenging than the subjects usually met, but this was crucial to the observation. If the focus is meaning construction through strategic reading, then a challenging text and task is best suited to reveal the extent to which this is possible.

The analysis of the data set is then the crucial element of the study. As discussed in section 1.2.3 above, there was some quantitative analysis involved, but the majority of the propositions offered at the end of the thesis depend on close textual readings, examined from a variety of viewpoints. Coulthard (1977), McCarthy (1991), Ericsson and Simon (1993), Schiffrin (1994), Potter and Wetherell (1994) MacWhinney (1995) and ten Have (1999) all give clear overviews of the different approaches possible in either discourse or conversation analysis of such transcripts. In addition to this the work of Barnes (1976, 1992, 1995), Tudge (1990), Goodman and Goodman (1990), Hedegaard (1990), Resnick et al (1991) Edwards and Westgate (1994), Mercer (1995), Almasi (1995), Moll and Whitmore (1998), Light and Littleton (1998) and Forman and Cazden (1998) explores language in classrooms from different perspectives but often via Vygotskian notions of social learning theory. As part of the literature review in Chapter Three and subsequent reviews of the progress of the study in the later chapters, the role of these researchers' work will be more fully explored.

An important aspect of qualitative analysis would be the degree to which the subjects are involved and their voices heard. 'The knower and the known are interactive and inseparable' (Lincoln and Guba, 1985, p37). The structure of the sessions was that individuals were asked to think aloud while reading and completing the tasks, and of

course group members were asserting and defending their thoughts aloud too, as part of the interaction involved in group work. The format of the tasks was that wherever possible subjects were asked to account for conclusions reached, but the researcher remained as separate from the process as possible. In other words for the majority of the observation / recording subjects said only what they chose to say. All of these aspects offer an opportunity to establish an emic voice, where subjects' thoughts and concepts are presented openly and in a fairly unambiguous way, and where the inner voice can be projected. Of course the researcher's etic (distanced and judging, but, we hope, still objective) standpoint does also emerge through the analysis processes. But these include several analysis strategies, so a number of interpretations can be compared with the original text statement of the subject. Readers of this thesis have therefore a series of aspects to choose from. A system of strategy coding and key word concordancing offer a means to compare aspects of strategic processing between individuals and tasks. An examination of the discourse of meaning construction in both the individual and group contexts can enhance teachers' understanding of students' knowledge about language and of their strategic competence. The analysis of interactions within the groups with particular reference to ongoing meaning construction will clarify the role structure and will offer information about the advantages and disadvantages of a group attempting to co-operate on a reading task. Finally, any differences which appear between performances on the different task-types may lead to conclusions about task setting.

1.3 Developing theoretical sensitivity

Strauss and Corbin (1990, pp 41-3) define theoretical sensibility as '... the attribute of having insight, the ability to give meaning to data, the capacity to understand, and capability to separate the pertinent from that which isn't.' They add that, 'It can also be developed further during the research process.' They then name four possible sources of theoretical sensitivity: literature, professional experience, personal experience and the analytic process. This thesis is founded on aspects of all four of these sources, and the remainder of this first chapter describes in outline the place played by each.

1.3.1 The parameters of the research literature

The fields of research literature which could have been considered are huge. Aspects of second language learning, first language reading, psychology, linguistics, educational and research theory are all valid areas. From these three major specific strands were investigated in addition to the research methodology literature (including aspects of discourse and conversation analysis). The first of these centred on the nature of L2 reading and the use or development of reading strategies. The second looked at the think-aloud process as a method of gaining access to information about reading behaviour. The third was the literature on peer interaction, specifically where this involved group-based problem solving. By taking into account writing from different fields and standpoints (which were however all linked in their relevance to the investigation) it was hoped that a broader view of the data could be maintained and that theoretical sensitivity would develop to a greater extent. The research literature is discussed in detail in Chapter Three.

1.3.2 The professional and personal experience

The deepest roots of the thesis stem from foreign language classrooms in secondary schools in England. The researcher's professional (and personal) experience covered over twenty years as both classroom teacher and subsequently teacher educator and researcher, and therefore included access to very many classrooms in very many schools. Here the issue surrounding the nature and role of reading in foreign language learning is influenced by much wider concerns such as those of classroom organisation and general foreign language competence. It is a broad generalisation but, like Weber (op cit), we can say that during the last twenty years these learners have, on the whole, grown in confidence in their use of simple foreign language in dialogue interactions. They have also, in many cases, adapted very well to the stream of comprehensible input (Krashen 1982) generated by sensitive teacher use of target language throughout their typically rather meagre foreign language learning time allocation. But we should also say that the learners for the most part appear not to have adapted well to any stream of comprehensible input through written text, nor indeed have they been offered this for the most part. Teachers and course-book writers have seen reading as a problem area, particularly for the less able, have concocted safe texts which do not spring surprises (and therefore do not spring challenges either). Yet, to return to Weber, she writes, again of learners of English as a second language (ibid, p105): 'Moll and Diaz [1985,1987] view reading as a unified ability to interpret and construct meaning across languages. They remind us that, in Vygotskian terms, instruction must be at the proximal level to be effective and go on to assert – provocatively – that for reading instruction to be proximal

in English, it should not be limited by decoding skills in it, but should rather be aimed at the level of understanding shown in the first language.'

At the point where language competence *has* grown a little, materials in our schools have focused (perhaps too late in the learning process) on all of the techniques which develop reading as a quick fix to find some specified information, and very little on the utilisation of linear reading processes. Reading for learning, that is for the learning of vocabulary, grammar or even 'content', ie new information, has not been seen as a priority. Reading for testing the latest vocabulary set, or the latest grammar teaching point or to demonstrate life-skills such as menu or railway timetable interpretation has been much more the focus. Interest-value or a consideration of the zone of proximal development (Vygotsky, 1978) have rarely been the prime or even a major consideration.

So from the classrooms questions emerged: What do learners think about reading in a foreign language ? Do they regard it as any more difficult or enjoyable than any of the other three skills ? Why do they read ? Do they learn from it ? Can they self-evaluate in this way reliably ? Later this could be developed as: How might this relate to the recommendations of Cohen and Hosenfeld (1981) and others writing subsequently about the development of reading strategies and competence ? What would be a response to the question set by Alderson (op cit) ?

To offer further background to these questions we might first explore briefly, but more generally, the role of talking about strategy use in secondary foreign language learning in England, thereby considering also commonly used teaching styles in those classrooms. This focus on strategy use for language learning, linked to the autonomy movement, is

particularly vibrant in the world of TESOL, rather than in our secondary schools in England. Reading underpins learning, and reading strategies underpin general learning strategies. Even metacognition can be said to exist in a cyclic relationship with reading strategy use. To be mentally organised in one's approach to reading is an important pre-skill for effective reading. But as readers start to use a range of reading strategies to decode and interpret information and gain more from the process, so this forces them to develop new levels of capability in how to self-organise this increased understanding and thereby their ongoing learning. And so they also become metacognitively more sophisticated. Do our foreign language learners as a rule naturally develop such strategies? The answer is almost certainly, that they do not. They can perhaps be 'taught' strategic reading, (Mitchell and Swarbrick 1994, Hood 1996) but there is no way of knowing how far this approach would be internalised outside the classroom or indeed once the focus on it was removed. Can we account for this lack of strategic capability? It is probably true that the preferred teaching model, although described as 'communicative' is almost certainly the presentation, practice, production model, which tends to engender dependence on the teacher rather than independence from her. As Nunan and Lamb write, (1996, p46): 'The three-stage presentation, practice, production procedure was based on the psychological model that viewed learning as a linear process of understanding, internalising and activating knowledge. These three stages represent a gradual movement from high to relative low-structure interactions (though many production tasks give the illusion of student control). Teachers make numerous managerial decisions as they plan each of these stages in a teaching cycle.'

These decisions include the content and procedure ... roles of teacher and students, type of class arrangement for each phase, how students will demonstrate mastery.'

In fact Presentation in this context tends to be an extension of audio-visual methodology, heavily based on visual support and repetition without seeing the written form. The nineteen-eighties represented a time in England and Wales when modern language learning was expanding to be more accessible to learners of lower ability and was also becoming more focused on relevance to learners' needs. This resulted in a very creditable aim on the part of many teachers to ensure success for these learners by limiting the language to which they were exposed to a bare minimum and controlling practice and 'production' very tightly. Such methodology used in mixed-ability groups inevitably affected more able learners too. Learning new language through the reading of texts was almost unknown at this stage.

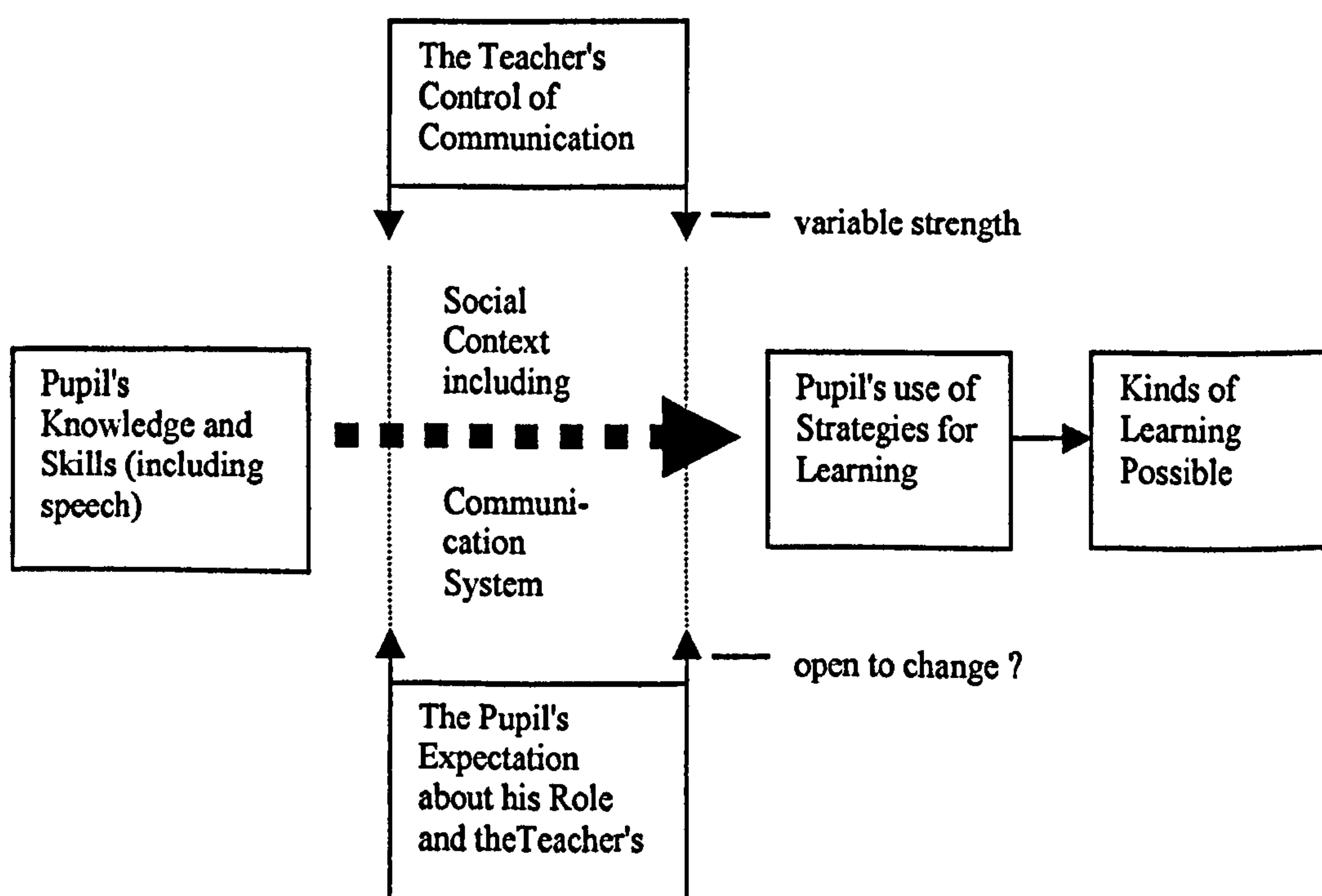
The role of the target language in classrooms in the 1990s has been a further factor in this process. The expectation according to the National Curriculum (NC) (1991, 1995) is that the target language will be used for the majority of the teaching time. Teachers following this instruction have found it impossible to devote time to discussing issues such as how we learn or indeed how we read, at least until the preparations for the General Certificate of Secondary Education (GCSE) national examination at 16 years old. It is as if, through the almost exclusive use of the target language from the beginning of learning at eleven years old, we have created a discourse community which is cut off from the rest of the secondary school curriculum. In other subjects there can be extensive discussion of learning styles if teachers feel this is appropriate to their methodology, but this is much more difficult in MFL. On this subject Mercer (1995) writes: 'We can think of each

teacher as a *discourse guide* and each classroom as a *discourse village*, a small language outpost from which roads lead to larger communities of educated discourse.' Current educational discussion of the role of thinking skills (eg Carvel 2000) underpins this point on a national policy scale.

So the MFL style of a transmission model with extensively controlled peer group interaction (mainly because linguistic levels are not sufficiently developed to allow an alternative) short-circuits what Barnes (1976, p32) calls 'pupils' participation in the shaping of learning'. He represents this as a diagram, reproduced below.

We see that the central area describes how the teacher sets up the classroom context and what the pupils' expectations of this are. This central area is crucial because the learning strategies the pupils use are filtered through it. As Barnes summarises (ibid, p33): 'The communication pattern of any classroom is the outcome of a history of mutual interpretation by teachers and pupils, in each case based upon previous experiences which they bring to the lesson.'

Fig 1 - source: Barnes, D. (1976) From Communication to Curriculum, London, Penguin, (p32)



If teachers' use of the target language limits the variety of working and thinking structures available, pupils will not be able to take an active part in the expression of new meanings, (in this context, the concepts and skills of learning a language, not just of using what has been internalised). The separation of modern language learning from other curriculum areas through the twin elements of a more trivial content (survival in the target language speaking country) and a lack of opportunity to discuss learning structures and styles may play a large part in the lack of popularity of modern languages in the post-compulsory sector. It is important to state here that this is not to advocate an abandoning of current target language use policy but to ask for a refinement of guidance to lessen the guilt felt by teachers when they wish to use English for a valid purpose.

At the GCSE stage, there can be more discussion of, for example, strategic reading. But there is still a history of problems associated with this examination. It began in 1988, and was a 'communicative' examination based loosely on van Ek's Threshold Level (1975).

The syllabuses contained defined content for all levels and rigid restrictions on what could be tested outside that body of language. This mitigated most strongly against any reading of authentic texts which engaged with the material at a level above 'survival'. It worked in favour of key-word spotting and certainly against more linear reading approaches.

So foreign language learning can, as it begins at eleven years old, quickly snuff out the independence, largely based on the reading of instructions and on reading for meaning, which is already well established by the primary school. In this way it often creates very quickly a spoon-feed culture, which then persists throughout the secondary phase.

Foreign language course books continue to underpin the Presentation, Practice, Production methodology. They may supply as an optional extra a bank of mini-readers, but the text quality and the task setting in the early stages of learning are wholly inadequate for the development of more sophisticated reading capabilities. Even if GCSE results do indicate that foreign language competence is growing, this, as we have shown, does not include a sophisticated reading competence and it certainly does not transfer easily into continuing foreign language study.

But do we have any concrete evidence for this analysis ? An illustration of the early teacher-dependence mentioned above can be seen in the following two extracts from small focus group interviews. These were carried out two years apart in different cities with learners of different foreign languages, but with learners of the same age (twelve to thirteen years, in their second year of FL learning).

Extract 1:

T Do you think that if you're really sure about what you like and what you don't like doing that you should be able to choose what you do ? **If you were to come in and I said: "Alright you've got to learn this language, but you choose how you do it" do you think that would be a good idea ?**

S5 Yes

S4 I don't

S1 No, I don't

S5 Well, as long as, I suppose

T D, you tell me why you do and then K can tell me why she doesn't

S5 Well I think you should have a teacher to tell you some things, but put in a bit more fun as well, like more wordsearches because when you're finding them you learn more because you're looking for them. And if you have more games and that you'll have fun and you'll like it more so you'll want to learn more

T Right what do you think, K ?

S4 Well I think if you said: "Right you can choose how you want to do it," somebody would get to something that they really like, like wordsearches all the time, and that's all they'd ever do and they'd never get to learn vocabulary at all

S2 Yes, so we wouldn't learn as much as we would when you tell us what to do

Interview transcript (focus group of 7): (Hemmings, 1996, pB3) (my emboldening)

Clearly here learner choice has been interpreted as a choice between different task-types provided by the teacher. There is no indication of any independent, reading-centred, learning style under consideration. The comment about wordsearches not helping someone to learn vocabulary also hints very strongly at a teacher-centred and teacher-controlled presentation model for vocabulary, as of course, we *could* feasibly learn new words through 'reading' a wordsearch. The image presented, of the teacher 'telling the learners what to do' is not untypical of many other anecdotal instances discussed among teachers and teacher-trainers.

Extract 2: (*teacher in italics, pupil responses, uncoded by individuals, in plain text*)

Would you say that you learn things in Russian lessons in the same way that you learn new things in other lessons ?

I think it's different

How might it be different ?

Because if it's in English and she says "we all go to the park" and then we have to repeat it, it would be
(laughs)

Is it the teacher who tells you new things in other lessons ?

Yes

So the teacher explains things, you're still listening to a teacher like you do in language lessons ?

(Nodding)

Do you ever find out things by yourself ?

Yes

More in other lessons ?

More in other lessons

The same

I think more in other lessons because you can understand the language – it makes it easier

So if they sent you off to read something, you haven't got as many problems in other lessons ?

It's understandable, it's in English

So would that mean that you're - you know what it means by being independent ?

Yeah (chorus)

Would you say you're more independent in other lessons, or the same as in

languages ?

More (chorus)

Is it easier to have the teacher telling you lots of stuff or is it easier to find things out by yourself ?

The teacher telling you

So do you like that better ?

Yeah (chorus)

Interview transcript (group of 7): (Hood, 1998, unpublished data, my emboldening.)

Here there is a realisation that there are other ways of learning, including independently through reading materials, but the preferred model is again that of the teacher having full control and, more importantly, full responsibility.

The professional and personal experience, which includes first-hand experience and the influence of a multitude of discussions with teachers, student-teachers, course writers and examiners provides a 'common sense' background to the study. Silverman (op cit, p5) writes, as part of a critique of scientism: 'Of course social science needs to study how "common sense" works in a way which "common sense" would not and could not follow for itself. In doing so, however, it will inevitably draw upon common-sense knowledge.' This is exactly the role for this aspect of common sense in the study. It would be dangerous to make assumptions based on the professional and personal experience, but

both supply a valuable touchstone for the recontextualisation process (Morse 1994b) at the end of analysis.

1.3.3 The stages of the investigation and the analytic process

The study was not pre-planned in its entirety but grew, stage by stage as evidence from the different elements was used to inform the planning of the later stages of the work. Taking Robson's (1993, p40) 'three traditional research strategies' into account, it became clear very early in design attempts that in addition to the philosophical starting point outlined in section 1.2 there were pragmatic reasons to reject an experimental approach which focused on change in reading capabilities. Measuring such change might in itself not be difficult but ascribing such change to a particular treatment of an experimental group would be dangerous. Given that reading competence is a part of general language competence, any change could be the result of any or all teaching and learning activity over the period in question. But the other two approaches, survey and case study both offered worth while scenarios. Therefore the investigation began with a stage, which looked at attitudes to reading in a foreign language via a questionnaire given to more than 200 learners in their second year of Russian in five different schools. The broad results from this activity suggested that a small scale case study investigation into actual strategy use by learners with only 1-2 years of learning could illuminate some of the questionnaire findings. Also of interest was the notion of reading as a group rather than as an individual activity. The questionnaire was therefore followed by a first investigation (via video recording) of group-based reading tasks. A total group of 18 (the majority of one small class in one of the five schools) from the questionnaire sample were selected and from

these subjects, working in five small groups, the transcript of a pilot group of four students was analysed more thoroughly. A little later the members of the pilot group were also audio-recorded while completing reading tasks individually and being asked to think aloud as part of that process. Further questions were raised by this stage, including that of the specific issue of non-roman scripts in reading, (see Chapter Two for a detailed account of both the first and second stages). Therefore the main investigation was designed to involve further work on the two data-gathering processes (think-aloud and group interaction) with a full class of learners who had completed two years of French rather than Russian. By this stage it was clear that a range of analysis techniques could be applied to the data obtained. As part of that process some quantitative work on areas such as task performance, strategy use and reading approaches would be possible, but a more major strategy would be the analysis of the transcripts from both think-aloud and group interaction activities from two distinct viewpoints. The language recorded would reveal how strategies were being used (or not being used). Of course this could not be guaranteed to be in the sense of a complete 'laying bare' of mental processes, but would reveal strategy use in the crucial area of task completion, ie the comprehension of the chunks of language that the task had selected as important. Furthermore discourse and conversation analysis techniques applied to the protocols would also put the group work on reading into the familiar broader context of Vygotskian social learning theory. The qualitative analysis process is described by Morse (op cit, p24) as 'the cognitive struggle of model or theory construction' which she maintained had not been sufficiently deconstructed. She continues to name four elements of the process: comprehending, synthesizing, theorizing and recontextualizing, and considers each of these specifically

sited within one of four branches of qualitative research. In her analysis of these stages as cognitive processes she demonstrates that each clearly adds to the researcher's understanding of and sensitivity to the core issues contained in the data.

It is clear then, that in addition to the reading and the professional and personal experience discussed earlier, the processes involved in designing and completing the study, and particularly the analytic processes demanded by it, were a further positive means of developing theoretical sensitivity.

In Chapter Two we present the Part One study, which informed both the necessary scope of the literature review and the Part Two study. In Chapter Three we review the relevant literature to provide a more detailed specific context for the procedures of data gathering and data analysis. In the subsequent chapters the main study is described and discussed and propositions for further consideration are outlined. The thesis concludes with a short chapter concerning future directions.

Chapter Two

The Part One Study

2.1 The context

The study began with the questions reported earlier in Chapter One, as follows: What do learners think about reading in a foreign language? Do they regard it as any more difficult or enjoyable than any of the other three skills? Why do they read? Do they learn from it? Can they evaluate their own skills reliably? It had been noticeable during time spent in Russian language classrooms in England that reading was a particular problem. A specific instance which stands for many others occurred when the researcher was observing and supporting a student teacher, and stopped to help a fourteen year old student in the third year of learning doing a pairwork speaking task. The student could not remember an item of vocabulary and asked the researcher for it. The advice given was to look back a couple of pages in the exercise book for the relevant topic vocabulary list. The student replied: "I'm really sorry, but I can't read what I've copied down in my book". An imperfect or partial knowledge of cyrillic script will inevitably hold back learners of Russian from achieving their potential in terms of using well the language they have learned, and will stop them completely from learning more language through reading. Teachers take a lot of time to teach the cyrillic alphabet at the beginning of courses but this knowledge wanes over time without a regular diet of stimulating texts and reading / writing tasks.

With learners of Russian these effects are obviously likely to be more significant. But the problem of recognition of familiar vocabulary can affect some learners of other foreign

languages too, even when the script is roman. A parallel anecdote is one where a 12 year old learner of French asked the researcher for the word *écoutez* in a cartoon story. When the word was pronounced aloud in French the reply was 'Oh that means *listen*.'

We know from the principles of reading theory about the interactive process (for example, Rumelhart 1977, Stanovich 1980, Lesgold and Perfetti 1981, and see the longer review in Chapter Three) that one of the vital aspects involved in reading and meaning construction is a rapid, preferably automatic decoding capability. Obviously this process supplies the basic textual information from which meaning can start to be derived or against which top-down processes such as prediction of meaning can be checked.

Stanovich (op cit) has shown that the use of contextual clues in previous sentences and other top-down skills such as prior subject knowledge can compensate for a lack of bottom-up processing capability. But one result of this could be an inaccurate or even false reading. In addition a reliance on such information does nothing to address the lack of a well developed decoding capability. It is reading speed, itself obtained via automatic recognition, which partly enables developing meaning to emerge through the connections between clauses, sentences and paragraphs. Stanovich's summary section (ibid p64) includes: 'In short, the good reader identifies words automatically and rapidly, whether by direct visual recognition or by phonological recoding..... The result is that more attentional capacity is left over for integrative comprehension processes.' In other words a poor reader will not have the working memory space for a more sophisticated meaning construction. If in the foreign language classroom the completion of a reading task can be

facilitated partly by common sense and world knowledge, students may appear to teachers to be reading (ie constructing meaning from text) better than they are.

2.2 An outline of the process

The Part One data gathering divided into two main sections, a questionnaire given to a large number of 12-13 year old learners of Russian (during their second year of learning) and a pilot study involving recording a small group of students from that cohort while they were reading.

A questionnaire format which included items which asked subjects to rank the four skills and their capabilities in certain language learning areas, and to signal agreement or disagreement with statements about learning had been used before on a project concerning listening while reading (Hood 1994, unpublished data) and by two other researchers (Adwick 1995, Hemmings 1996). The questions outlined in 2.1 and in Chapter One were combined into this format, which was tested on a small group of learners. Only minor adaptations of rubric were found to be necessary. The results, detailed in section 2.3, focused on individual reading skills and the purposes of reading, for example the relative importance in their eyes of text and task, and on the role of collaborative reading. The responses also raised issues about how well learners could monitor their own reading capabilities. All of these were then taken into account when designing the format for the next stage.

The reading behaviour data was gathered by asking five small groups to collaborate on completing three reading tasks in Russian. Of these, one group (referred to as the pilot group) was chosen for detailed analysis. Three tasks were used, each involving a different

means of demonstrating comprehension. The process was audio- and video-recorded, and the transcripts were coded in terms of the purpose of each 'move' within the interaction. This was done not in terms of speech act analysis, but in terms of the progress towards meaning construction and task completion. The coded transcripts could then be analysed to discover the nature of the subjects' reading behaviour, both individually and as a group. Later the four students were asked to do two further tasks individually, in order to trial a comparison process between the individual and group modes of working.

2.3 The questionnaire - methodology and analysis

2.3.1 Description of the questionnaire / Table of results

The questionnaire, which sought to elicit some broad base-line attitudes towards reading in Russian, was given to a total sample of 232 students in their second year of learning, and at the age of 12-13. Of the five schools used one school was selective and single-sex, the others were comprehensive and co-educational. All four comprehensive schools had a policy where Russian was taught to a full range of ability (though not necessarily to all students in the school). The samples were identified by each school as high, low or mixed ability in some way (which could include descriptions such as 'mixed but not including the lowest ability'). As a result of these descriptions, the sample included two sub-groups, which could be separately examined, as well as the whole cohort. These were students who were taught in a high-ability only context (in the selective school or a top set) and those who were always taught in a low ability only context.

The initial four questions sought to establish a basic attitude towards reading in Russian and views on issues of personal competence. Four categories were offered for this to

avoid any tendency to take a safe 'middle option'. The fifth item looked at the popularity of reading in comparison with the other three skills. The final six statements (with which respondents had to agree, disagree or signal no certain opinion) considered aspects of the role reading might occupy in their learning. The last item, (C6), was supported by a supplementary question, asking respondents to give a reason for their answer. Finally subjects were asked to complete some very simple statements about reading, their own personality and learning style preferences.

The results of the questionnaire are here presented in tabular form with discussion following the table.

Table 2a: QUESTIONNAIRE INTO ATTITUDES TOWARDS READING IN A FOREIGN LANGUAGE (Integrated version, containing questionnaire text and statistical data)

Key to notation:

W = whole sample: n = 232

H = acknowledged high ability (from selective school + 'top set' of 1 comprehensive school): n = 103

L = acknowledged low ability (as identified from 2 of the 4 comprehensive schools: n = 41)

NB no objective measurement of language attainment was given. Figures are given for the whole sample, and then for those working in a higher or lower ability *context*. The figures in the table are in percentages

In the whole of this questionnaire reading means understanding in your head and possibly telling others what you think. It does NOT mean reading aloud.

A Reading in Russian Lessons

Question / Statement		ability	A lot	It's OK	Not much	Not at all
1	How much do you enjoy reading in Russian ?	W	8.6	64.7	15.9	10.8
		H	9.7	71.8	13.6	4.9
		L	2.4	44	19.5	34.1
2	How much do you usually understand when you read in Russian ?	W	25.9	42.2	26.3	5.6
		H	31.1	53.3	15.5	0
		L	2.4	19.5	53.7	24.4

3	How much of the Russian alphabet do you think you know confidently ?	W	45.3	26.7	23.3	4.7
		H	62.1	24.3	12.6	1
		L	17.1	19.5	46.3	17.1

4	When you read Russian, how far can you usually guess words you don't know ?	W	20.7	44.4	29.7	5.2
		H	16.5	57.3	23.2	2.9
		L	17.1	24.4	44	14.6

B Reading and the other skills

Please write numbers 1-4 next to the four skills below to show which you enjoy most (1), which second (2), which third (3), and which least (4)

LISTENING..... READING..... SPEAKING..... WRITING.....

Percentages putting reading....

	W	H	L
First	9	13.6	4.9
Second	23.3	25.2	17.1
Third	42.3	43.7	31.7
Fourth	25.4	17.5	46.3

C What do you think about reading tasks in Russian ?

In this section there are 6 statements about the reading tasks in Russian. For each one decide what you think and then underline AGREE, DISAGREE or DO NOT KNOW

1. We learn more Russian from speaking than from reading

	AGREE	DISAGREE	DON'T KNOW
W	62.3	19.5	18.2
H	66	17.5	6.5
L	40	32.5	27.5

2. Reading is a personal activity - you can't do it in groups

	AGREE	DISAGREE	DON'T KNOW
W	18.2	70.6	11.2
H	11.6	76.7	11.6
L	32.5	47.5	20

3. I would be happy to read something in Russian even if there were no questions to answer on it

	AGREE	DISAGREE	DON'T KNOW
W	56	15.9	28
H	68	14.6	17.5
L	26.8	36.6	36.6

4. When I have a reading task, I read the questions more than I read the text

	AGREE	DISAGREE	DON'T KNOW
W	40.7	34.6	24.7
H	42.7	43.7	13.6
L	46.5	15	37.5

5. I get a lot out of working with a partner when we do reading tasks

	AGREE	DISAGREE	DON'T KNOW
W	71.9	13	15.1
H	79.6	10.7	9.7
L	60	12.5	27.5

6. Reading in Russian is very different from reading in English

	AGREE	DISAGREE	DON'T KNOW
W	80.6	14.4	5
H	81.7	17.2	9.7
L	68.3	14.6	17.1

Please try to explain your answer to Question 6 in a sentence (empty box given in original)

D Some personal views from you - Please fill in the following gaps: (gaps provided in original)

- Reading Russian is sometimes best done individually because
- Reading Russian is sometimes best done in a group because
- In Russian lessons I volunteer answers
- If I am asked to describe my personality I say I am
- In all my subjects the kind of work I like best is

2.3.2 Discussion of the questionnaire results - whole cohort

Before reviewing these results in detail, some general patterns should be noted about the sample as a whole. In Items 1-4 close to 70% of respondents evaluated themselves on all 4 items within 2 consecutive of the 4 graded categories (ie, the top 2, middle 2 or lower 2). No subject had a response in each of the 4 categories, and few had all four responses in the same category, either of which *might* have indicated a pattern of random selection or lack of thought about the questions and responses. Of the 30% who had a single response within a third category, approximately one-third identified their knowledge of the cyrillic script as being a lower-skill ability. This is entirely in tune with findings about

reading capability and behaviour identified in Section 3.1.

On items 6-11 there was a widely varying response pattern, with no clear tie to ability or school, (but see discussion of the School D top set scores in section 2.3.5). Only 16 subjects in all matched the majority response pattern throughout all six items. This is to be expected when the items were seeking information about personal preferences in learning styles and affective issues such as mode of working, rather than knowledge or skills-related issues.

In analysing the responses, we need to be mindful not just of what they *seem* to tell us, but also of any internal contradictions. It is possible that younger respondents do not always have a developed ability to be accurately self-evaluative, particularly if they are unaccustomed to discussing learning processes.

From item A1, we see that there is a large percentage among all groups except the separately taught lower ability learners which enjoys reading in Russian a lot or a little. When this response is viewed in conjunction with the responses to items B and C1, then this general opinion can be seen more comparatively, and the positive regard perhaps seems less strong. Those items show that a clear majority of the whole sample feels that reading is only third or fourth preference amongst modes of working as delineated by the four skill areas, and that they as students believe they learn more from speaking than from reading.

The responses to questions A2 and A4 clearly show a general confidence about levels of text comprehension and the ability to guess the meanings of new words which would suggest that learners *could* read more independently and even use it to advance their learning at their own pace. But it seems significant that there was among the *whole* sample a greater proportion of respondents who felt very confident about their ability to guess the meanings of new words than there was within the separate high ability sample alone. Such a pattern is unlikely, as more able and successful language learners tend to show greater ability in this area, (Rubin and Thompson, 1982, Devine, 1988a, Grenfell and Harris, 1993/4). Item A3 reflects a better balance, perhaps, in terms of the respective proportions for each 'band'. However, as alluded to earlier in the chapter, work in the field in Russian classrooms over several years has shown that a partial knowledge of the alphabet and a slower automatic decoding rate is a major obstacle to more rapid learning amongst very many learners, and so would suggest otherwise. The proportions feeling confident here are unlikely to be truly representative of the real state of internalised knowledge.

To summarise, then, the findings from items A1-4 demonstrate a general sense of confidence and high self-esteem, which in itself is a positive feature and much to be encouraged. But, it needs also to be viewed with caution. These responses do suggest that the general reliability of younger learners' own ability to evaluate their competence in such an area as reading and learning strategies should not just be assumed. And if we are interested in broadening reading experience and in the development of more autonomous ways of working and learning *through* better reading, then that self-awareness of individual capacity is an extremely important element (Nunan, 1996). The

more aware we can make the learners of the issues surrounding effective reading, the more they will be able to monitor and enhance their own skills.

Items C1-6 reveal a variety of issues, which suggested they might be explored further, (for example the gains to be made through paired or group reading). C1 shows that reading is generally seen as less useful than speaking as a vehicle for learning language. Interestingly this view is not so strongly held by the separately set low ability group. This group have probably been led to believe over a period of time that their lack of reading competence has kept them back in general educational terms. As a result they perhaps see reading as equally important as speaking. But for the majority reading is not especially highly regarded either as a major mode of learning (or as we have seen in terms of being an enjoyable activity). It is not that we would want reading to be seen as *more* valuable than speaking, but we would want it to be appreciated in its dual role as a 'language use' activity and a 'language learning' activity. (Little, Devitt and Singleton, 1989). In the climate referred to in the first chapter, where the teacher often controls quite tightly the range of vocabulary and structures which are presented to learners, it might be helpful to allow learners to meet new language by themselves and within their own spheres of interest. When these learners talk about learning through speaking they may also be referring here to such various elements of the language learning cycle as repetition after the teacher at the start of the unit, pairwork practice tasks and role-play production activities at the end. Listening and speaking often exemplify and sometimes dominate all parts of the cycle, and so they, the learners, perceive these tasks as the most crucial.

Item C2 shows that the majority of these learners like to work with others. The field work suggested that this is genuinely because they can hear differing views and can pool understanding. The lower ability group differed here, again perhaps because they see reading as such a crucial element in their own development that it must be more a personal matter, and perhaps also because they often have far less self esteem in this area.

Items C3 and C4 contain a certain tension when viewed together, and are also the most interesting items when viewed by ability grouping. It is perhaps also significant that these items have the highest 'don't know' count of the six statements in Section C. This may suggest that more detailed discussion of how we learn is in less familiar territory for a number of learners. The high ability group is the most willing to read without a task (ie a more authentic reading purpose), and the low ability group the most unwilling. The high ability group is evenly split on the virtues of focusing on text or task, whereas the low ability group clearly views the task as the more important element. The pilot group proved themselves to be untypical of the whole on this element by being more focused on task than any other single grouping. This is not only expressed in the questionnaire, but is also very evident from the recordings of their interactive work on the reading tasks, where they are very task-led in most instances, (see sections 2.5.1 and 2.5.5 for exemplification).

Item C5 reinforces C2 with whole-sample percentages in favour of collaborative reading very close in both items (70.6% / 71.9% respectively). The high-ability group figures are similarly matched (76.7% / 79.6% respectively). The lower-ability group seem to relate

better to working with a partner (60% in favour) than to working in a group (47.5% in favour), and the phrasing of the item (*'getting a lot out of working with a partner'*) might suggest that the sharing of expertise is an important element here. Item C6 shows broad agreement across all groups, that reading in Russian is different from reading in English, with only a higher number of 'don't knows' in the lower ability group causing a deviation. The students were asked to add a comment on C6. Only 16 subjects failed to respond, but the responses ranged widely. A majority (67.6%) noted differences in the respective alphabets or languages, although of these only just over a third were negative about the differences. Most were neutral in commenting on difference rather than difficulty. A number (13.4%) commented on differences in linguistic structure (such as case endings or word order). A little over 10% felt that reading is the same in any language. Very small numbers mentioned phonetics with a majority realising (rightly) that Russian is a very phonetic language, but still a handful commenting that it does not sound the way it is written. Many of those who commented on the difficulties of the alphabet had earlier indicated such problems in item A3, but many others who did, had not done so. Again, we see a need for an awareness raising of *what* learners know and can do, and *how* they can come to know this better.

The final section of the questionnaire is not analysed in detail, because it revealed very few patterns. The very last question about preferred work style in school revealed a varied pattern of answers, with some mentioning subjects rather than work styles. Many, at least from four of the schools did use words such as *discussion* or *group work* as part of their preferences, although significantly, in the top set group analysed separately as

School D in section 3.3.5, only one subject did so. The remainder mentioned more individual task styles such as listening or writing.

2.3.3 Discussion of the questionnaire results - eager readers

If a small sub-group is viewed (those who identified reading as their *preferred* skill in Section B), then the results show an expected turn in direction towards greater confidence in Section A. Apart from 2 anomalous forms from the lower ability group, the top two categories are selected by the overwhelming majority. The ability to infer the meaning of new words was the least confident of the four categories, but still marked as category 1 or 2 by the majority. In Section C all but one of those working in higher ability sets would be happy to read something without a task (C3), and 80% would read the text than more than the task. A greater proportion of this sub-group than of either the whole sample or the higher ability sample disagreed with statement one while statements two and five scored consistently with that of the whole sample. If the few subjects from mixed or lower ability groupings who put reading first in Section C are also included, the pattern is very similar, especially about the role of text and task. The eager readers then do seem to display some of the qualities which one might tentatively expect: a much higher confidence about the alphabet and about inferring the meanings of unknown words; an interest in text for its own sake; a greater belief that reading might help language learning; and still a willingness to engage with texts in collaboration with others. The notion of the isolated learner reading quietly to her- or himself does not emerge from this data.

2.3.4 Discussion of the questionnaire results - reluctant readers

Equally we can take another sub-set of all those who placed reading *fourth* in their preferences in Section C. This was a small minority from three of the five schools but a significantly higher proportion from two (39% from one and 43% from another). Taken as a whole group these subjects' responses are not radically different in most areas. They may be a little more strongly convinced that we learn through speaking and that the task is more important than the text. They are similarly wedded to the notion of collaborative reading and feel very strongly that reading in Russian is different from reading in English. Within Section C there was a higher proportion of 'don't know' responses than elsewhere and not just within the separately set lower ability groupings. Of course we might find that it was 'don't care' rather than 'don't know', particularly when there was a run of these on the same sheet. But figures in that category ranging between 10% (for C6) and 39% (for C3) were clearly above the equivalent for the whole sample if not for the low ability group alone, where that outcome might be more expected.

2.3.5 Discussion of the questionnaire results - School D

It is worth also considering briefly the response pattern from the school, which included the highest proportion of subjects placing reading as their fourth preferred skill (School D). In all there were 53 respondents, 25 in the low ability set and 28 in the top set. Of these 2 from the lower ability group and just 1 from the top set placed reading first amongst the skills, while 12 from the low ability set and 11 from the top group placed it fourth.

The top set responses are the most significant to examine as here there are definite differences from the pattern set by the whole cohort and the separate higher ability subset. Taking the 28 members of this group as a whole, the figures (in comparison with those of the whole cohort of 232 and the whole top set cohort of 103) are as follows.

Table 2b: Comparison of School D top set scores with those of the whole cohort and the whole top set cohort

A Reading in Russian Lessons

Question / Statement		ability	A lot	It's OK	Not much	Not at all
1	How much do you enjoy reading in Russian ?	W	8.6	64.7	15.9	10.8
		H	9.7	71.8	13.6	4.9
		SchD	3.5	60.7	25	10.7
2	How much do you usually understand when you read in Russian ?	W	25.9	42.2	26.3	5.6
		H	31.1	53.3	15.5	0
		SchD	28.6	35.7	35.7	0
3	How much of the Russian alphabet do you think you know confidently ?	W	45.3	26.7	23.3	4.7
		H	62.1	24.3	12.6	1
		SchD	39.3	28.6	28.6	3.5
4	When you read Russian, how far can you usually guess words you don't know ?	W	20.7	44.4	29.7	5.2
		H	16.5	57.3	23.2	2.9
		SchD	7.1	57.1	32.1	3.5

B Reading and the other skills

Percentages putting reading....

	W	H	SchD
First	9	13.6	3.5
Second	23.3	25.2	21.4
Third	42.3	43.7	35.7
Fourth	25.4	17.5	39.3

C What do you think about reading tasks in Russian ?

1. We learn more Russian from speaking than from reading

	AGREE	DISAGREE	DON'T KNOW
W	62.3	19.5	18.2
H	66	17.5	6.5
SchD	67.8	10.7	21.4

2 Reading is a personal activity - you can't do it in groups

	AGREE	DISAGREE	DON'T KNOW
W	18.2	70.6	11.2
H	11.6	76.7	11.6
SchD	14.2	67.8	17.9

3. I would be happy to read something in Russian even if there were no questions to answer on it

	AGREE	DISAGREE	DON'T KNOW
W	56	15.9	28
H	68	14.6	17.5
SchD	53.6	17.8	28.6

4. When I have a reading task, I read the questions more than I read the text

	AGREE	DISAGREE	DON'T KNOW
W	40.7	34.6	24.7
H	42.7	43.7	13.6
SchD	53.6	17.8	28.6

5. I get a lot out of working with a partner when we do reading tasks

	AGREE	DISAGREE	DON'T KNOW
W	71.9	13	15.1
H	79.6	10.7	9.7
SchD	85.7	7.1	7.1

6. Reading in Russian is very different from reading in English

	AGREE	DISAGREE	DON'T KNOW
W	80.6	14.4	5
H	81.7	17.2	9.7
SchD	100	0	0

It can be seen then that this group, evidently working in a context which is more negative about reading, differ throughout in the proportions feeling confident about their reading capability, and in their view of the value of reading. The responses to Section C show that they are more convinced that we learn more through speaking than reading, more inclined (than other top set students) to think of reading as a personal activity, less happy to read without a task and more task led when they do read. They are however very committed to working with a partner. Finally, everyone in the group considers reading in Russian to be

different from reading in English and the majority commented on the different alphabet as the major difference, and often as a major problem.

These figures are the global statistics for that top set. The group within it who placed reading last in their preferences differ from the whole cohort to an even greater extent on C1, C3 and C4. In Section A 45% of this subgroup opt for the two lower categories on Q1, with 36% doing so on Q2, 27% on Q3 and 45% on Q4. Given that these are top set students these percentages are very high if compared to the whole top-set cohort.

2.3.6 Summary

The questionnaire has therefore produced some varied responses, but responses which are consistent within the subgroups analysed, and responses which make sense when both field experience and reading theory are considered.

It was felt, then, that a valid purpose of further investigation would be to 'open a window' on reading behaviour by subjects involved in the questionnaire.

Propositions for the next part of Part One Study were generated as follows:

- as a result of the above findings the observation / on-line recording of subjects may be an appropriate method of finding out more about real reading behaviour
- Given the emphasis by learners on the positive aspects of collaborative reading, this may best be organised through reading tasks given to small groups
- The focus of the observation needs to be on:
 - How well basic decoding of text operates
 - How far strategic reading takes place
 - How far collaborative reading can add to the processes and product involved

- Using a selection of different task-types may reveal different reading behaviours
- The texts need to be challenging (so as to elicit strategic approaches) but not too difficult
- The observation of some individual reading may offer different or further insights into reading behaviour

2.4 Methodology of the group-interaction tasks

A class group, initially of 23 was chosen from the questionnaire sample¹. The class was given the GAP reading test (McLeod 1965) because it uses a cloze technique and this was felt to focus both on an overall meaning construction process and on the sort of strategic reading which centres on inferring the meaning of specific items. This is clearly an important aspect of foreign language reading when some vocabulary may not be known. Of the class, 18 students, divided into groups of 3 or 4, completed an observation. One of these groups (with a range of GAP scores) was subsequently chosen as the pilot group for intensive analysis.

Three reading tasks were chosen for this stage. (See Appendix A) Each centred on the language topic in focus, (free time activity), but drew in also previously encountered and new vocabulary. In this way all the texts did more than simply consolidate the current topic and the current active vocabulary. The first text was deliberately chosen to be

1 This group which was a mixed ability sample differed little in their percentage profile from the overall mixed ability sample of the rest of the cohort – only in one answer were they radically different (item C4), where they were much more likely to read the questions than the text when doing reading tasks.

familiar - it was taken from the course book in use and was a true/false/impossible to say (TFI) task-type, which the subjects were used to meeting. The second text was adapted from another Russian teaching resource and the accompanying task was to infer the meanings of ten underlined words in the text. The third text was created by the researcher together with a native Russian speaker who taught parallel classes in the school. The task for this text was a Four Penfriends (FP) task, devised by the researcher and used with other classes in German in previous classroom observation activity. This task requires readers to identify the author of a specific letter by matching the content and style with the characteristics of four different penfriends. The students were not allowed dictionaries while they worked on any of the tasks, as it was felt that this would maximise the search for meaning through strategic behaviour.

On a subsequent occasion the pilot group members were asked to complete a think-aloud task individually on two further texts, one a TFI task and one a FP task. (See Appendix B).

2.5 Analysis of the group-interaction tasks

2.5.1 The tasks and the task performance

The Texts were all of equal length (100 words) and produced as follows:

- TFI task - 96 lines of dialogue
- Underlined words task - 179 lines of dialogue
- Four Penfriends task - 104 lines of dialogue

The task performance by the group can be summarised as follows:

A. True false task :

Nine questions out of ten were correct, of which seven definitely appear to be for the right reasons - the one which was incorrect was the item where the subject as well as verb person was different from that in the text (item 4)

TFI statement analysis: All statements were in 3rd person - all relevant text was in the 1st person

Item 1 = parallel sentence - with different place name

Item 2 = parallel sentence - with different number, out of sequence

Item 3 = directly parallel sentence with extra insert in text

Item 4 = parallel sentence form but use of she rather than he changes from consistent approach across most other statements

Item 5 = parallel sentence with different sport

Item 6 = parallel sentence with different number

Item 7 = parallel sentence - statement true - no vocabulary different

Item 8 = parallel sentence although plural verb forms make parallel less easy to spot - the day mentioned in the statement was different from those in text

Item 9 = parallel sentence form across 2 sentences - the item in the statement was different from the 2 items mentioned in text

Item 10 = out of sequence - directly parallel text

B The Underlined words task:

Of the ten words underlined, four were guessed or near guessed. Of these one was already known, two were guessed (close cognates), one closely related meaning was guessed from an identification of the word root.

Thus, six were not guessed, of which

- 1 was not guessed but context produced sensible guess attempt
- 1 was not guessed but a feasible strategy was used to make the wrong guess
- 1 was not guessed but a false friend cognate and world knowledge + knowledge of other part of text were combined to make a generally feasible guess (though not in immediate text context)
- 1 was cognate, but no guess attempted
- 2 had no guess attempted

The rationale for the selection of the ten words to underline was as follows:

Three were cognates, one had a familiar root, five had supportive context, one (an adverb of place) acted as an anaphoric reference linking two sentences.

Four were at beginning of a sentence (in the first two words), two were in middle of a sentence, four were at the end (the last word).

C The Four Penfriends task:

In this task the correct writer was identified and of nine possible reasons five were identified, but mostly at the prompting of the researcher to look for further detail.

2.5.2 Summary of 'moves' across the three tasks

The three transcripts (See Appendix C for full details) were coded on the following basis.

The purpose(s) of each utterance was analysed and assigned a descriptive code as a particular 'move'. These codes were not pre-chosen but developed as the transcripts were read. (See Appendix D for the complete list). The occurrences of the codes were counted for the group within each of the tasks, then for the individuals within each of the tasks and in total. Each table is presented with a specific discussion following it

Table 2c: Whole group moves in the 'True / False / Impossible to Say' task

Function of move made by subjects in peer group interaction - TFI task	Count
Read aloud from the task statement	26
Agree with what has just been said	12
Offer an answer	11
Read aloud from the text	9
Translate	6
Give organisational clarification	7
Seek organisational clarification	6
Correct pronunciation / reading of the text	5
Move on or start to move to another section	4
Disagree with what has just been said	3
Clarify reading / answer	3
Correct a reading of the text	2
Echo what has just been said	2
Not know	2
Defend a reading or an answer	2
Infer meaning through use of a cognate	1
Affirm, admit wrong, withdraw, signal end of task (each)	1

The most common action therefore was to read aloud elements of the statements below the text, which formed the true/false task. These statements contained less than half the number of words than there were in the text itself. Therefore, given that there were only

approximately one-third of the number of instances when the text was read aloud as opposed to the statements, we have strong evidence for this first exercise being very task driven. The instances of translation do tend to be from the text, but even these added to the reading aloud from the text score do not nearly match the total read from the statements. Furthermore the majority of reading aloud from text occurs in the first third of the transcript. Once the initial section of text had been read in sequence, the task statements almost entirely initiated the rest of the discussion.

There is a close balance between agreements, echoes, and withdrawal of wrong statements on the one hand and disagreements, corrections, defences of statements on the other hand. This suggests (as does the recording) that this was a co-operative collaborative exercise, but one in which at least some of the subjects were willing to 'argue their case' and take a lead. This will be further analysed in 3.5.2 when individual move patterns are discussed.

The other major moves are to offer answers to the task, to seek and to offer organisational clarification, ie to find out where the rest of the group is with its reading or where it should look for an answer, or to give information in response to such requests.

Strategic reading in the narrower sense of inferring meaning through a strategy is very sparsely represented here (apart from translation as a meaning constructor), with just one definite use of a cognate noted. Of course there may have been other similar cases which appeared simply as translations (ie using words already known.) when they were in fact inferences from context or of an unfamiliar cognate word.

Table 2d: Whole group moves in the 'Underlined Words' task

Function of move made by subjects in peer group interaction - Underlined Words' task	Count
Read aloud from the text	30
Using a strategy to construct meaning * see list below	17
Translate	11
Agree with what has just been said	10
Defend an opinion / statement	9
Not know / Be unsure	9
Disagree with what has just been said	8
Clarify a meaning, reading or answer	7
Read aloud an underlined item	7
Echo what has just been said	7
Seek clarification	6
Seek meaning of vocabulary	5
Make a suggestion	5
Explain something	4

* list of strategies (apart from translation) used to construct meaning:

4 instances of cognates, 3 each of: using knowledge about language; considering the context of the sentence, 2 of using world knowledge, 1 each of: reading the preceding sentence; continuing to read on; making a guess; skimming the rest of the text; and making an analogy with a familiar item.

Immediately of note in this second task is the greatly increased time and number of moves (approximately 75% more again) needed to complete it, when compared with either Task One or Task Three. It seemed that the pilot group found this task more difficult, and one noticeable feature about the transcript is the different motivations to complete or give up. One subject suggests stopping the task at a point just two thirds through the transcript, yet continues to play a major role to the end, when the others ignore the suggestion. Clearly this may also suggest features of a group dynamic which could result in an individual making more gain more from a group-based task than s/he might as an individual.

The most important difference in the table is the number of strategic moves designed to construct meaning (as opposed to organisational moves or moves such as reading aloud) made by the subjects during this task. If translation were to be included as such a move there would be approximately four times as many as in Task One and Task Three. Clearly this is in itself a strong indication of the ability of certain task types radically to alter reading behaviour.

The reading behaviour is still to an extent task driven, as the focus tends to settle on the underlined words with readings being made either side, rather than any complete reading for global comprehension. Although subjects do suggest reading a preceding or following section as a comprehension strategy, nobody in the group spoke of any need for a fuller reading either in between the underlined words or perhaps as a pre-task activity.

Table 2e: Whole group moves in the 'Four Penfriends' task

Function of move made by subjects in peer group interaction - Four Penfriends' task	Count
Read aloud from the text	16
Clarify reading / pronunciation / tasks	9
Agree with what has just been said	8
Echo what has just been said	8
Translate + continue to give information	7
Suggest a reading or an answer	5
Continue to justify an answer	4
Justify a reading	4
Seek the meaning of vocabulary	4
Seek to end the task	4
Be unsure	3

In this task the focus inevitably starts on the text, as there are neither separate task statements nor the underlined words on which to focus. But the group's essential task

driven character is still shown by the statement (from P2) just 2 lines into the transcript (and after just three sentences have been read aloud) that they can already decide who the author is. At line 15 with still only the first paragraph having been read, P2 attempts to stop the task again. The supplementary questions demonstrate a need to read further, which the other group members see, and so the process continues for a further 90 lines. In the end the subjects supply five of nine possible reasons for the author's identity, but, if compared to Task Two, fail to utilise meaning constructing strategies to any great extent. There is evidence however that this task does draw out some greater strategic behaviour than does Task One. For example there are attempts to guess through identifying cognates (eg line 53, P1 '*Dzhudo sounds like judo*') and a greater awareness of the role of contextual or world knowledge clues. For example when looking at the sentence which means '*I like sport passionately*', P2 suggests (transcript task three, line 39): '*That could be something like really good or really bad*'. Here she appears to have understood that it is adverbial usage, and to have sense through the tone of the text that it is a fairly 'extreme' word. Further on, when considering the text: '*I even wrote a novel about school. It was excellent*' and having originally suggested a translation of '*I'm excellent at reading novels*,' (line 74) p2 suggests: '*I think writing actually would be better. I'm excellent at reading wouldn't sound right, would it, so it makes sense to say excellent at writing.*'

This task at first did initiate a linear reading of text, but again there was no suggestion that to continue that might give any form of global meaning. The table demonstrates that the predominant behaviour was to read aloud and translate, to clarify such readings or what needed to be done, to agree with statements made. The items which were not

mentioned as reasons for the author's identity were not entirely impossible to infer. For example, she plays in both the School and the Petersburg Symphony orchestras. Although orchestra was recognised, the cognates for its name were not, and neither was the familiar root of the word school, which had been recognised in Task Two. The fact that she plays three sports on Saturdays, was only partially appreciated, but perhaps Saturday is a more difficult word to infer (if not known) in this context. She claimed to be a champion (which is cognate) but this was not seen. Finally, and the most difficult, she watches all TV programmes every day on her three televisions. The last fact was recognised but the consequence of it was missed.

2.5.3 Summary of individual differences in 'moves' within and across the three tasks

There are great differences in the amount of involvement by each of the four group members during the three tasks as summarised in the table below

Table 2f: number of 'moves' made by each group member during each of the three tasks

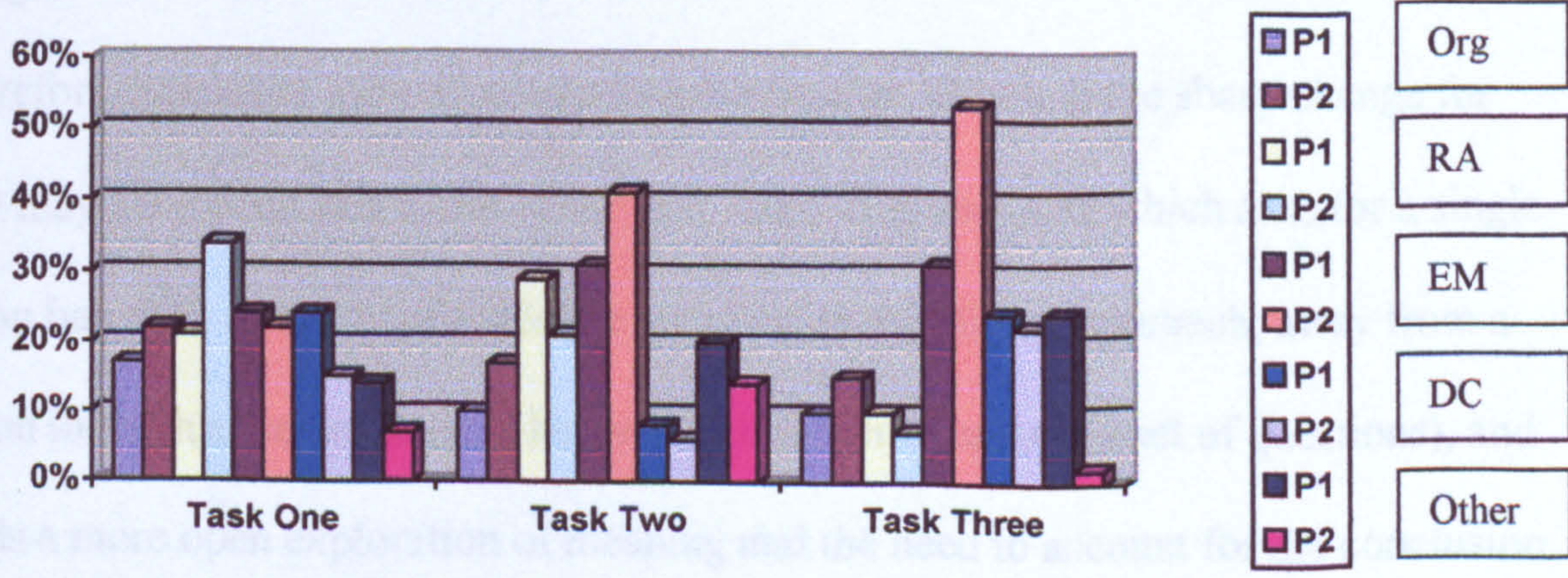
Subject	Task One	Task Two	Task Three
P1	29	52	29
P2	41	70	45
P3	10	18	10
P4	5	10	3

Clearly then the group dynamic was very strongly established throughout the session. P1 and P2 between them had never less than 74.6 % of the interaction, with two of the three tasks involving them in more than 80% of the moves. P2 was clearly dominant even within this partnership, in terms of the amount of times she spoke.

If we analyse the nature of the moves made by P1 and P2 in each of the tasks we can see better what their roles were. To do this the moves in these tasks were classified into four broad categories: those which *organise* self or others; those which simply involve *reading aloud*; those which involve *exploring meaning* (and this would include translation); and those which constitute *drawing conclusions*. A small number (including basic agreements, or saying 'I don't know') would then fall into a general 'other' category.

Table / Figure 2g: Nature of / proportion of moves by task by Subjects P1 and P2

Move type	Subject	Task One	Task Two	Task Three
Organising self or others	P1	17%	10%	10%
	P2	22%	17%	15%
Reading aloud	P1	21%	29%	10%
	P2	34%	21%	7%
Exploring meaning	P1	24%	31%	31%
	P2	22%	41%	53%
Drawing conclusions	P1	24%	8%	24%
	P2	15%	6%	22%
Other	P1	14%	20%	24%
	P2	7%	14%	2%



If we view these results by subject initially, we will see that both P1 and P2 organised progressively less as the session developed and they became accustomed to the group mode of working. For P2 the proportion of time spent reading aloud also declined during that time. Both explored meaning progressively more through the three tasks, and both were able to draw conclusions significantly more in the first and last task than in the more challenging second one.

If we compare the proportions of moves spent on each type of move it is evident that P2 spends a consistently greater proportion of her moves organising than does P1, but P1 tends to read aloud more than P2 (except in Task One). P2 grew significantly more able to explore meaning as the tasks developed and in Tasks 2 and 3 spent more time on this than did P1. Although P1 always spent proportionately more moves engaged in drawing conclusions, P2 came closer to his figure after the first task.

For P1 Tasks One and Three cause an emphasis on exploring meaning and drawing conclusions, whereas Task Two sees him reading aloud and exploring meaning to a far greater extent. P2 is engaged principally on reading aloud and exploring meaning in both tasks One and Two whereas in Three she focuses more on exploring meaning and drawing conclusions.

P2 therefore appears to grow in reading sophistication, although the sharp change for Task 3 may have been as a result of the task itself. A global task which asks for a single decision based on a variety of evidence may lead to a different approach, away from a focus on individual sentences and keywords (in order to answer a set of questions), and towards a more open exploration of meaning and the need to account for the conclusion reached.

Across the three tasks it is P1 and P2 who almost exclusively give the answers. But there is no clear pattern either in which of the two provides the answer or in which is responsible for the reasoning which precedes it. At times P1 appears to have the more secure steer, and in task two (the most difficult of the three) he uses more strategies during his attempts to deduce meaning. He is also the most reliable of the four in his knowledge of the Russian alphabet, and often corrects mispronunciations.

In order to gain a further insight into individual behaviour it was felt appropriate to involve the four in an individual 'think-aloud' reading session. This followed approximately three months after the original group recording.

2.5.4 The think-aloud protocols from the pilot group

On the subject of think-aloud data Ericsson and Simon (1993, p78) write: 'Thinking aloud activity is not entirely alien to everyday life and almost all subjects have probably had some experience of it before they come to the laboratory.' We perhaps need to add the caveats that learners of this age may not be as accustomed to the think-aloud method, and can be more self-conscious about working in unusual ways than adults might be. A small amount of explanation and training was given, but the researcher wished to avoid offering examples of the sort of reading behaviour which was being investigated and was interested to see if approaches which had been used in the group session appeared again in individuals. Therefore it is acknowledged that the protocols may only reveal a part of the cognitive activity involved. Nevertheless the process gives us some data on strategy use, and 'non-use'. P1 and P2 produced protocols which were clear enough in their intentions, if rather sparse in their evidence of reading behaviour. The protocols produced

by P3 and P4 were marked by a tendency to read aloud (often very poorly, and therefore evidently with very little textual understanding) and then to struggle with the tasks and to make very little of them. It seemed that especially for P3 the texts and tasks were in fact too difficult. Her apparent understanding of the three texts used in the group activity may have been enhanced by the discussion about the texts and tasks as they proceeded through them. When isolated with the two think-aloud texts she became extremely nervous and unable to function.

2.5.5 Comparison between one subject's think-aloud protocol and her participation in the group context

Clearly it is valuable to compare the performance of a group member with a corresponding role in the group context. It seems appropriate to take one of the two more engaged members of the pilot group, particularly as the two less engaged members produced very sparse think-aloud protocols (as shown in the preceding section) which consisted mainly of the reading aloud of the text. For this reason P1's and P2's transcripts were examined and given the evidence regarding their respective roles within the group it is P2's transcript (Appendix E) which is discussed in detail below.

There follow two short transcript extracts from the think-aloud session:

The first is an extract from a think-aloud protocol by subject P2, as she attempts a true/false task (with the text and true/false statements in Russian). The statement she is looking at is: *У него проигрыватель (He has a record player)*. The relevant section of text states: *Тoже в универмаге хорошие кассеты и компакт-диски - они много стоят, как майки, но мне очень хочется их посмотреть. (Also in the store are good*

cassettes and compact discs - they cost a lot, like the football shirts, but I really want to have a look at them.)

Number five is (*reads question*) and I'm just looking for I'm trying to work these out by looking at like to see if the words that are in the question and in the actual in the actual paragraph.

But it doesn't all make sense cos I don't know what all of the words mean but So it's (*reads question five*) so I can't yet see that in the paragraph but I'm having a look. I think number five has to be nyet. No actually I think number five's da.

(..... *signifies a significant hesitation*)

The second is an extract from a short interview with her afterwards:

Can you tell me a little bit about doing that, and was it easy or not to do ?

It wasn't, it was quite, I think that one was the easiest, the one about Sasha*, but I didn't understand it all, because some of the words we haven't, I haven't heard of before and things like that.

You know when we did it before, when you were working with the group, and now doing it like this by yourself, is there any difference between reading in the two different ways ?

Yes because when you're with your group, you're all putting ideas together and like you do it together, but whereas like I'm on my own, you have to think of it all yourself and that's umm

How easy was it to say what you were doing while you were doing it ?

It was easier than I thought it would be. It's just normal like you're talking to someone else.

* She had also worked on a further task, a penfriend letter from Sasha

In terms then of her strategic reading, as she was not able to use a dictionary, we might expect a range of strategies used to discover the meanings of unknown words, but in fact

we have only here evidence of word matching between the true/false statement and the text. The word for record player is not 'guessable', but the structure 'he has' is well known and clearly cassettes and compact discs are cognate. We might expect her to look closely at what *is* there and identify words, which then become crucial to answering the question. We might expect her to comment on what she has read and understood in more detail. Even if this use of strategy is not revealed by the think-aloud process, (because she finds it unnatural) the result (ie getting the answer right for the right reason) would indicate that such a behaviour is occurring. The task format (see later section of the article) is here quite crucial, as the emphasis is on true/false task completion and not on text comprehension. She gives an answer, but without a rationale for it.

For the purposes of comparison a third extract comes from the group session where the four subjects are looking at Task Two (as discussed earlier). At the point where they are reading the text states:

Летом я играю в футбол, а зимой в хоккей. Я также плаваю. У нас в городе хороший бассейн. Там можно плавать и зимой и летом. (In the summer I play football, and in the winter ice-hockey. I also swim. We have a good pool in the town. There you can swim both in summer and in winter.) The two underlined words, *in winter* and *there* were the two which the group were asked to infer.

(Italics show where they are reading from the text in Russian, and an asterisk denotes inaudible speech)

37	P2	<i>Let</i>
38	P1	<i>Letom =</i>
39	P2	That's a p =
40	P1	That's an l =
41	Ps	<i>Letom, ya igrayu v futbol, a zimoy =</i>
42	P2	That's winter (1)
43	Ps	*
44	P1	yeah, zimoy is winter
45	P2	<i>Ya *</i>
46	P4	I play [football - football is what I play]
	P2	<i>Tam</i> what's the
47	P1	I'm not sure what tam means * let's see <i>tam =</i>
48	P3	We've done that I play football and in winter I play hockey =
49	P2	we've done it already =
50	P3	I know =
51	P2	Right, <i>ya *</i> we've already had it . We were just wondering out what this tam means (1)
52	P1	Read the sentence after. (1) er <i>tam možno (2) plavat' =</i>
53	P2	<i>i zimoy i letom</i> , so that's like, err, it's winter and spring or is it summer
54	P1	it's summer isn't it ? =
55	P3	summers's em =
56	P2	Spring Spring. So it means (2) hold on I think we've sort of read the question er the thingy wrong. It's <i>ya na (1) gorodye khorosho bassein. Tam mokho. (1)</i> Do you know what that means ?
57	P1	I haven't got a clue (1)
58	Ps	*
59	P2	Right, so then (1) it means something [like he plays
	P1	it's something that he does] in winter and summer
60	P2	yeah, so (2)
61	P1	tam =
62	P2	it could be something what's town, what could tam be, it could be a sport or something
63	Ps	*
64	P2	Could be, could be anything (2) could be swimming ?
65	P1	It's not though
66	P2	I know, that's what I mean it could be anything. (2) Go on

The text was from a book in printed format, but was italicised so that the letters appeared more like handwritten Russian. The group members were all more used to reading conventional handwritten text. Their knowledge of the cyrillic alphabet is still comparatively weak, as some of this material is clearly cognate, but they are still having problems with decoding.

Nevertheless, from this interplay we can begin to gauge better whether in a group setting the overall data may be richer in total, than with a separate individual from the group. In

this short extract P4 speaks once and P3 three times (although both join in with chorused readings or translations), while P1 and especially P2 dominate. Clearly we *do* see much more evidence of different approaches from her in the group context than in her think-aloud protocol. She organises (eg lines 49, 51, 56), explores meaning (eg 42, 53, 56, 59), draws a conclusion (62), and operates in a more sophisticated way, not necessarily because she understands a lot more of this text, but because she gets feedback and extension, principally from P1. This, taken with the contrast already noted in P3's group and individual behaviour, would suggest strongly that peer interaction offers, at least for some learners, an opportunity to activate or even discover strategic reading behaviour. What they see may either be their own revealed strategy use (of which they were not consciously aware) or reading behaviour demonstrated by others. The group discussion thus sets up an opportunity for a zone of proximal development (Vygotsky, 1978, see Chapter 3) to be activated. Of course for this to be effective on a long-term basis there needs to be a regular pattern of working in this way. Thus, both for learning and data-gathering purposes, it is worth pursuing this methodology over a longer period of time.

3.5.5 The apparent role of task

Another aspect of the study which emerged at this stage as potentially significant, is the role of task. Grellet (1981, pp12-13) established a typology for reading activities, consisting of reading techniques, establishing how the writer's aim is conveyed, understanding meaning and finally assessing the text. These categories clearly move from a more teacher-assisted to a more independent mode of working and would therefore demand a growth in strategic competence as a reader progressed through the taxonomy.

In early foreign language reading in schools only a part of the first of these is really attacked, and we have seen evidence of how certain task styles (for example, true/false) can lead to learners being excessively task- rather than text-led. Before Grellet's higher level skills can be considered, even if applied to simple text, there is a need for more coherent and detailed reading. How to achieve this is then itself an important issue for a teacher setting a reading task. The group transcripts do show variation between the three task-types. We saw that a text-centred task such as that of inferring the meaning of certain highlighted words has an effect on the range of resources brought to bear on the text. We saw that a more global or problem-solving reading task (for example, read a letter and decide who wrote it) can lead to more linear processing even if the emphasis is still strongly on completing the task. True/false tasks tend to encourage the practice of key-word spotting revealed in both the think-aloud protocol and the group-based transcript. This heavy focus on *task* rather than text would seem then to be greater with certain types of task, which also happen to be typical of course book and examination use. With more global tasks, it appears that the learner has to read more in order to be sure that s/he is covering enough material to answer the question properly. Making a single global decision about a whole text is therefore perhaps more of a stimulus to linear processing than making a series of individual decisions, such as true-false statements, which can be completed with more selective reading.

2.6 Conclusions from the think-aloud and group-interaction task data

To summarise, the questionnaire and pilot group activity seems to have indicated the following general conclusions about younger foreign language learners:

- many often find it difficult to evaluate accurately aspects of their own foreign language reading competence, such as their ability to draw inferences about the meaning of previously unknown vocabulary items or how well they know an unfamiliar alphabet
- many have very little awareness of the processes which might be involved in strategic reading, although they do appreciate the value of reading with another learner
- few see reading in a foreign language as a prime means of learning language or information
- many tend to see reading as a means to complete a task set by a teacher rather than as an intrinsically useful activity
- the practice of reading in small groups appears to activate greater strategic behaviour, if only because interpretations and conclusions have to be justified; additionally, a desire for either collaboration or competition or both can motivate readers to continue where they might give up as individuals.

2.7 Propositions framing the Part Two study

2.7.1 Previous propositions

To re-cap, the propositions which framed the group interaction stage of the Part One Study were as follows:

- as a result of the questionnaire findings the observation / on-line recording of subjects may be an appropriate method of finding out more about real reading behaviour
- Given the emphasis by learners on the positive aspects of collaborative reading, this may best be organised through reading tasks given to small groups
- The focus of the observation needs to be on:
 - How well basic decoding of text operates
 - How far strategic reading takes place
 - How far collaborative reading can add to the processes and product involved
- Using a selection of different task-types may reveal different reading behaviours
- The texts need to be challenging (so as to elicit strategic approaches) but not too difficult
- The observation of some individual reading may offer different or further insights into reading behaviour

The Part One observation study clearly indicated that observation and on-line study can reveal insights into the reading behaviour of both groups and individuals. There was clear evidence concerning the issues of decoding, the use of strategies and the different contexts of individual and group-based reading. The three task-types chosen brought out

some differences in reading behaviour although the respective roles of task type, task-type familiarity and text difficulty could not be identified with any confidence.

2.7.2 New propositions

As a result of the evidence from this stage further propositions were constructed for a Part Two study. These were as follows:

2.7.2.1 The Context

- Given apparent cyrillic script specific issues, which perhaps exaggerated the difficulties in the recognition of learned vocabulary and the decoding of new items, the foreign language used in the study should be changed from Russian to French or German
- A full mixed-ability class of learners should be used so that any conclusions could have greater validity
- The age of the subjects might better be slightly older to allow for a broader language base to be available

2.7.2.2 The methodology

- There should be an equal focus on the two data gathering methodologies (ie group interaction and individual think aloud protocols) to allow true comparisons to be made

- There should be a focus on only two tasks. The second task-type used in the Part One Study appeared to be anomalous when compared to the other two in both its outcomes and associated motivation
- There should still be as much parity between texts as possible
- In order to analyse effectively meaning construction and other strategy use, further coding and analysis techniques are needed
- In order to analyse effectively the role of group interaction and the role of individual characteristics within it further analysis techniques, for example concordancing and discourse analysis approaches are needed

2.7.2.3 The research questions

- Can observation / online study offer useful information about reading behaviour in the secondary school context ? (ie a confirmation of the conclusion to the Part One study that it can)
- If so, how can we describe the range of behaviours observed ?
- Does the choice of context (individual and group-based reading) cause differences in behaviour for individuals ?
- Does task type affect reading behaviour or performance?
- Can the information generated be used to make proposals for enhanced teaching and learning ?

Chapter Three

A literature review

(Foreign language reading - processes and strategies; the use of think-aloud protocols in gaining access to cognitive activity during reading; monitoring the role of peer interaction in group-based problem solving)

3.1 Introduction

In this literature review chapter three key areas for the study will be considered.

These are:

- three important issues involved in second language reading: a brief overview of important L1 and L2 models of reading; threshold levels of language for the transfer of L1 reading behaviour; and the nature and role of reading strategies
- the role and purpose of self-report (think-aloud protocols) for gaining access to the reading process
- the role in problem-solving activities of peer interaction via small group discussion

Each of these strands has a very broad literature but the chapter will focus on the specific research and writing within the field which is appropriate for a study relating to foreign language reading, and which can shed light on the findings of the study. If we generate transcript-based information about the reading behaviour of a group of learners, we need to be sure we can gain as full a picture as possible of the information it contains. Thus we need to establish the research findings about beginner / intermediate L2 reading behaviour. We need to ask questions about how we find out that information, ie to investigate previous use of our target methodologies: obtaining think-aloud protocols

during individual reading tasks and monitoring small-group collaboration in completing problem-solving tasks. We should further investigate whether such collaboration can enhance learning and by what methods the success or failure of collaborative work can be evaluated.

3.2 Foreign language reading - processes and strategies

3.2.1 Models of reading

Since the late 1960s there have been many attempts to establish models of reading, but these have predominantly stemmed from L1 reading theorists. As we review these models we need to consider whether there are aspects of the reading process which are inherently different in certain L2 settings and which would therefore suggest modification of any given model.

Recent models of reading are loosely categorized into three sets: bottom-up, top-down and interactive. This order also reflects the way in which the theories developed, with bottom-up and top-down models being defined at approximately the same time by different schools of thought. An extreme view of either would constitute a diametrical opposition, but in reality both processes clearly had validity and the resulting switch to the interactive model to an extent combined and therefore supplanted both. But all three models are separately important for L2 reading, as they reflect certain types of reading processes which are very evident in foreign language classrooms.

Three 'bottom-up' models often cited are those of Gough (1972), Laberge and Samuels (1974), (which was later revised) and Carver (1977-8). In all of these, words on the page are decoded in a linear fashion and meaning is constructed from a subsequent processing of the emerging combinations of words into phrases and sentences. In this way comprehension is entirely text-driven, and unfamiliar items have to be identified by whatever means before reading can continue. Carver (op cit) differed in writing about the combination of visual perception (ie words on the page) and audio-perception (ie the articulation of those words either inside the head or aloud) and called this *rauding*. Given that there is a tendency for early L2 learners to read slowly, evidently word-by-word and to read aloud, this model obviously has implications for foreign language teachers and researchers. It is probably true that such readers do not consider anything beyond the words they see in front of them. So much mental 'space' is being taken up in this form of processing that other activity such as the use of metacognitive strategies or prior knowledge is excluded. In addition we must remember that reading aloud in a foreign language does not automatically lead to comprehension if the words read are not yet known. In discussing this issue, Barnett (1989, p18) states: 'Foreign/second language readers frequently read material without understanding it, and the role of internal speech during second language reading is questionable.' But in more general terms about bottom-up models she writes (ibid, p19) 'The text-driven nature of bottom-up models may have more to say about weak second and foreign language readers than has been acknowledged.' Thus we should bear in mind differences between the L1 and L2 contexts, and not reject the potential ability of these models to explain reading behaviour

simply on the grounds that L1 reading theory has moved on. Chapter Seven will review this issue in the light of the findings of the study.

The top-down (psycholinguistic) model originated with Goodman (1968) and Smith (1971). Both emphasise meaning construction as opposed to the decoding process: 'In all this it is meaning which makes the system go.' (Goodman, op cit, p98); 'Reading is less a matter of extracting sound from print than of bringing meaning to print.' (Smith, op cit, p2). According to Goodman (op cit) the reader uses a knowledge of syntax and semantics as well as prior/world knowledge to interpret the text and is not dependent on word-by-word processing. The model defines four elements: predicting, sampling, confirming and correcting. He demonstrates (ibid, pp17-19) that there are three levels of proficiency and that the first two do depend on some decoding and recoding (by which he means sounding words aloud). But the proficient reader samples large sections of text, makes the fit to what has been predicted and gains meaning. As alluded to above, in L1 reading the ability to sound a word (ie decode and recode in Goodman's terms) would lead to understanding since the word would almost certainly already be known to the reader as native speaker (Singer 1981). But in L2 reading simply sounding a word is of course no guarantee to comprehension, and foreign language classrooms would contain many who could read aloud an entire text without understanding it if the text was pitched at too high a level. The process of vocalisation can also seriously interfere with the meaning construction process as many language teachers have found when using a reading aloud activity. Therefore it is at Goodman's proficiency level three that we would find a confluence of effective L1 and L2 reading behaviour. Smith (op cit) similarly stresses the anticipatory nature of reading, the role of existing knowledge and the reader's overall

purpose which also drives the search for meaning. Again the basic textual decoding process is not a major part of his model, as the reader begins from her/his own knowledge base and works down to the printed text. Clearly where readers of L2 texts have particular advance background knowledge the top-down process is of great assistance to them in meaning construction.

Coady (1979) looked specifically at the psycholinguistic perspective on L2 reading, and proposed that readers do move from more text-driven behaviour to more top-down behaviour as their competence grows. But he also concluded that readers shift approaches depending on the text or the reading goals and therefore that not all readers follow the same paths to comprehension or the same developmental pattern.

In summary we could say that elements of such top-down models are entirely appropriate for more sophisticated readers, but clearly a lack of knowledge either of vocabulary or of grammatical and syntactical features of the language could impoverish the anticipatory and confirmation powers of the reader. Either, then, the model would fail to operate or its use could involve too great a dependency on world knowledge. As we have already said there is a tendency in foreign language reading advice given by teachers to encourage intelligent guesswork to compensate for lack of knowledge, and while this is sound advice in some respects it could prolong a stage where reading is always approximate and never precise. Again we shall discuss this issue in Chapter Seven in the light of evidence from the Part Two study.

The interactive model of reading originated with Rumelhart (1977) and features (ibid, p588) a simultaneous processing of a textual input (in what he calls the pattern

synthesizer) from orthographic, lexical, semantic and syntactical knowledge bases. In all cases the sampling of the graphic element will always be just sufficient to make suppositions about developing meaning, ie better readers will use the process faster than poorer readers. Just and Carpenter (1980) proposed a model in which physical features, words, meanings, case roles, clauses, text units, and domains of discourse could all be drawn into play as part of the processing but during a serial model of word recognition and comprehension. Barnett notes (op cit, p29) that this 'may help explain the word for word reading style of some second language readers.' Stanovich (1980) refined the interactive model by suggesting that strength in some areas of processing can compensate for weakness in others, thus to some extent underpinning the teacher advice referred to above.

To summarise, the strength of interactive models is that they allow for any of the processes to occur, as of need, and for feedback loops to exist between these different modes of processing. They do not foreground either bottom-up or top-down processes. According to Samuels and Kamil (1988, p32), '... even if a skilled reader can generate predictions, the amount of time necessary to generate a prediction may be greater than the amount of time the skilled reader needs simply to recognize the words.' We could encapsulate the refined interactive-compensatory process using Stanovich's words (op cit, p63), 'Interactive models ... assume that a pattern is synthesized based on information provided simultaneously from several knowledge sources. The compensatory assumption states that a deficit in any knowledge results in a heavier reliance on other knowledge sources regardless of their level in the processing hierarchy.'

Kintsch and van Dijk (1978) considered comprehension to be far a more important part of the reading process than the recognition of individual words. They ascribed the ability of good readers to achieve comprehension to a combination of main proposition identification, the subsequent drawing together of propositions into the main gist, and a 'new text' generation process. In other words the reader constructed meaning from perception of major units rather than individual lexical items. In this the role of schemata (both general and specific, knowledge-based and linguistic) was also highlighted. Yet they too did not entirely ignore the role of lexical recognition, and in a subsequent work, (van Dijk and Kintsch, 1983, pp23-4), they wrote: 'What is really wrong with poor readers is that they recognize isolated words inaccurately and too slowly, and compensate for their lack of decoding skills with context-dependent guessing or hypothesis testing'.

The interactive model again provides the basis for this summary, as the interconnectedness of text-driven and meaning-driven processing is apparent. And this analysis does hold good for much foreign language reading during the first stages of learning but also among weaker readers long after that time.

Others who wrote about foreign language reading include Bernhardt (1986, p103) whose constructivist model of the L2 reading process postulated that six different features (word recognition, prior knowledge, phonemic/graphemic features, metacognition, syntactic feature recognition and intratextual perceptions) are all present in processing. This data emerged from recall tasks on texts used in German, French and Spanish and although there is no direct explanation of how the elements work together, it is clearly another example of interactive modelling.

More recently Ridgway (1994) defined a model of reading for both L1 and L2 which built on the Rumelhart (op cit) and particularly Stanovich (op cit) models. This showed that different readers might process text in different ways at different times (depending on the nature of the text) and that this process might also be seen as a developmental model of reading. In this model (op cit, pp67-8) Ridgway looks at the processing involved in dealing with a selection of different items encountered in a text. These might range from a well-known content or function word which is automatically processed, to an ambiguous item which needs to go through a selection process, to a less well-known word which requires greater processing, or 'attention' to be applied to it. As with an unknown word this may lead to stage where the word is ignored if it is decided not to be important or it may activate some kind of strategic processing in what is called a 'problem space'. Again, as with Just and Carpenter (op cit) we see a focus on a more linear process, while still one which does not deny simultaneous activity from a variety of knowledge sources.

In addition to this consideration of models of reading, we should also note another factor which contributes to the actual reading process, irrespective of the model. The role of affective elements such as interest in and motivation towards actual text content is especially relevant to L2 reading, which is often harder work than L1 reading would be. Little, Devitt and Singleton (1989) highlight this in their justification of the use of authentic (meaning inherently more interesting and personally significant) texts in language learning, noting (ibid, p5) that 'the greater the personal significance factor ... the "deeper" the processing, and the "deeper" the processing, the higher the chance of

processed material being recalled subsequently.' Clearly the issue of depth of processing could be relevant for any of the three models we have considered, and would contribute to greater understanding of text whether the reader was text-driven, meaning-driven or both. While the texts used in this study were not authentic, in the sense of having been published for native speakers, they were constructed with a native speaker with the intention of making them as authentic as possible for the age-group and within the letter style adopted.

3.2.2 Reading difficulties - a reading problem or a language problem ?

In the brief overview of models of reading given above we have found that attempts to explain the performance of good and poor readers and to define the processes they use while reading suggest the existence of a variety of contributory factors. The issue of whether it is their linguistic knowledge or their capability as readers which most determines the scope of their success or indeed their reading problems has been addressed by several researchers over the last fifteen years, for example Clarke (1980), Alderson (1984), Elley (1984, 1991), Devine (1988a, 1988b), Bossers (1991), Swaffar et al (1991), Bernhardt and Kamil (1995), Lee and Schallert (1997) and Ridgway (1997). Clarke (op cit) wrote about the 'short-circuit hypothesis' where good readers reverted to poor reading habits when confronted with a difficult or confusing L2 task. He concluded that although proficient L1 readers do transfer their skills to L2, limited language proficiency has a great effect on their reading behaviour. As such the good and poor reading behaviours might be present in the same readers at different times. Alderson (op

cit) took up the issue by posing the question: Reading in a foreign language: a reading problem or a language problem ? Although the research he reviewed was inconclusive because it had not been designed to answer that question, he summarised (ibid, p20), 'Considerable support was found for the modified second hypothesis, namely that some sort of threshold or language competence ceiling has to be attained before existing abilities in the first language can begin to transfer.' But Alderson also pointed out that the ceiling might not be a universal and that perhaps it might alter according to the reading proficiency of individuals. He advocated longitudinal studies, which subsequently took place. At the same time, Elley (1984, 1991) firstly diagnosed that a reading problem might be a result of insufficient reading practice and subsequently demonstrated the importance of the activity of reading to improve both reading and language competence and advocated a 'book flood' to address both problems. Devine (1988b) concluded that there is a relationship between success in reading and language proficiency, citing knowledge of syntax as particularly important, and reflecting on the effect on comprehension of reading more slowly as a result of lower linguistic knowledge levels. She also advocated that exposure to language within text is an effective way of learning language and that therefore text should not be held back from beginning learners. Finally she highlighted the importance of the background knowledge learners bring to a text, especially if they were 'meaning' rather than 'sound' (ie text) centred. She found (1988a) that two low-proficiency subjects appeared to differ in the degree to which they transferred strategies from L1 reading, depending on their basic reading approach. Bossers (op cit) reviewed three major studies, all of which pointed to the conclusion that at lower levels of L2 linguistic knowledge the role of L1 reading ability was not

significant, but that it became so when L2 proficiency was higher, thus supporting the threshold hypothesis. There was an indication in one of his reported studies, (Carrell, 1991) that a second rather than foreign language knowledge threshold might be lower because the language is available in the surrounding environments. Swaffar (op cit, p53) took the view that attention needs to be paid, through use of authentic texts, to both language learning and content learning. 'Language knowledge and reading in a second language are complementary but distinctly different abilities.... The two styles [language-based and interactive] do access different learning strategies Students need to practise both styles of reading.' Bernhardt and Kamil (op cit) discussed the claims under the headings of the Linguistic Threshold Hypothesis (LTH) and the Linguistic Interdependence Hypothesis (LIH), (which is related to the work of Cummins, (1979, 1991) and would suggest that literacy concepts transfer across languages). Bernhardt and Kamil noted that the evidence for the LIH is generated mainly from school learners who are in the developmental stages of all language and literacy skills, first and second (ie young bilinguals). After reviewing studies which provide direct evidence of the two hypotheses, they concluded (ibid, p21) 'that L1 literacy is an important contributing factor to L2 reading but that language knowledge appears to be even more substantial.' Their own study also found this conclusion to be justified, and proposed a further statement of the questions along the lines of (ibid, p33): 'How L1 literate does an L2 reader have to be in order to make the L2 work ?' and 'How much L2 knowledge does an L2 reader have to have to make the L1 reading knowledge work?' Lee and Schallert's (op cit) study came to the same conclusions as those of Devine, Carrell and Bernhardt and Kamil. Ridgway (op cit) also concurred with this conclusion, but, like Donin and

Silva (1993), noted the importance of familiarity with topic as an important example of how good reading skills combined with background knowledge could compensate for a certain lack of L2 knowledge. Nevertheless he still found that the lower L2 language competence threshold applied.

In conclusion then it appears that many studies concur that L2 knowledge is a highly significant factor in L2 reading proficiency. Without language you cannot read effectively. But L1 reading ability was also shown to have a significance and for this reason it is important to review the literature also about reading strategies.

3.2.3 The nature of reading strategies

Researchers writing on strategy use in general have often presented pictures of good learners, or in this case good readers, as a way of defining useful strategies and more especially productive strategic behaviour. When discussing these issues we may also be considering unconscious strategy use (mainly cognitive) and a more conscious reviewing and choosing of strategies to attack a problem (mainly metacognitive). There is also a need to note the role of schema theory, eg Anderson and Pearson (1984), which supports aspects of the top-down process by enabling expectations about the text to be generated in advance of and during a reading.

Garner (1987), following Flavell (1981) considers in detail the differences between metacognitive knowledge and experience and strategies. The first is stable and stable information about cognition, the second refers to awarenesses of metacognitive events generated by task processing, while the third results from the first two and describes the

actions taken to problem-solve. Strategies can be, as stated above, both metacognitive and cognitive.

The metacognitive aspect is not especially new as, over fifty years ago, Lewis (1948) noted the importance of thinking and reflection to successful language learning. Carton (1966, 1971) specifically looked at comprehension strategies in examining inferences and summarising three cue-types, intra-lingual, inter-lingual and extra-lingual, (ie using knowledge from within the target language, between languages and from background sources). Rubin (1975), Stern (1975) and Naiman et al (1978) all considered second language learning strategies in broad terms and produced definitions of successful learner characteristics. Hosenfeld, in a series of papers (1977, 1979, 1984 and Cohen and Hosenfeld, 1981) developed ideas specifically about reading strategies and the definition of a 'good reader'. The 1981 article looked at obtaining such pictures from verbal reports and will be reviewed in section 3.3 of this chapter. The 1979 and 1984 pieces provided a case study of two readers and, from this investigation, a two-part and very detailed definition of good reader characteristics. These included those who tended to, (Hosenfeld, 1984, pp233) 'keep the meaning of the passage in mind, read in broad phrases, skip inessential words, guess from context the meaning of unknown words and have a good self-concept as a reader.' In addition they would, (ibid, pp233-4) '(1) identify the grammatical category of words; (2) demonstrate sensitivity to a different word order; (3) examine illustrations; (4) read the title and make inferences from it; (5) use orthographic information; (6) refer to the side gloss; (7) use the glossary as a last resort; (8) look up words correctly; (9) continue if unsuccessful at decoding a word or a phrase; (10) recognize cognates; (11) use their knowledge of the world; (12) follow through with a

proposed solution to a problem; (13) evaluate their guesses.' Clearly this enables both readers and teachers of readers to view a whole range of very concrete possible solutions to addressing reading deficiencies. The following section (3.2.4) will review research on the issue of whether learners can be taught to modify their strategic behaviour.

Other researchers have similarly produced classifications of strategies, for example Knight, Padron and Waxman (1985), Padron and Waxman (1988), Kern (1989). In addition to these perhaps the most different from Hosenfeld was Block (1986), who used two categories, General and Local, which are, respectively, comprehension-gathering and attempts to understand specific words or phrases. She also writes about two modes of reading, extensive (relating to the author's ideas) and reflexive (the reader's personal relationship to the text). This categorisation differs from the very concrete reading process-centred 'list' of Hosenfeld and deals less with a move-by-move description and more with how an integrative global understanding of text meaning and structure is obtained. Sarig (1987) combines elements of both approaches with a four 'move' classification, using different strategic purposes to categorise a large range of strategies. Each 'move' has a purpose concerning decoding, clarification, creating coherence or monitoring. Garner (1987, p50) held the view that strategy use was better seen metacognitively, ie, 'strategies are generally deliberate, planful activities undertaken by active learners, many times to remedy perceived cognitive failure.' She also notes that when strategies become automated they are perhaps better referred to as skills. Importantly, she notes (ibid, pp28-9) that many studies have shown that 'younger and poorer readers have little awareness that they must attempt to make sense of text'. She also writes, citing Garner and Reis, (1981, p571), 'younger children and poorer readers

are unlikely to demonstrate that they notice major blocks to text understanding.' So it is a matter of awareness that strategies exist and could be used which can be as important as the use of the 'right strategies'.

But other authors continued to prefer the 'deconstruction' method of listing potential strategies. O'Malley and Chamot (1990) and Oxford (1990) looked at general learning strategies for second language learning, and in the process produced inventories, some of which are relevant to reading. Oxford (op cit, pp321-4) created a list of strategies useful for reading, classified according to her own system, and illustrated some of these in concrete task descriptions elsewhere in the book. This forms part of the wider Strategy Inventory for Language Learning (SILL), (Oxford 1990) which has been much used by researchers investigating strategic use in a wider context.

3.2.4 Can strategic reading be taught ?

Writings by Hosenfeld (1979, 1981, 1984) and Cohen and Hosenfeld (1981) all addressed the issue of defining programmes which would enhance the strategy use of individuals while reading. These centred (Hosenfeld, 1984, p234) on a three-stage approach. The initial stage provided a reader with an individual strategy inventory through either think-aloud or introspective/retrospective reporting. The second phase was for the reader to compare the strategies used with a successful reader's protocol. The third stage was for the reader to attempt to apply those strategies to a further text. The results of two case-studies (with Cindy and Ricky) are reported, and it appeared that in the short-term they both increased their use of strategies, although to differing degrees. This process acts as a model for many that followed, in that we see a needs analysis followed by a treatment

and then an application. However we cannot be certain from Hosenfeld's description whether the new strategies used were internalised and became part of the readers' permanent behaviour. In other words we do not know whether strategy teaching actually works or whether it remains effective only while it is itself the focus of attention.

Bereiter and Bird (1985) reported on a study which differentiated between two treatments, firstly one of modelling strategies via a think-aloud process and secondly one that added to that modelling process guided practice in using the strategies. The group which received the double treatment made gains in the use of strategies, but neither the first group nor a control group did. Certain strategies seemed more teachable than others. 'It should not be surprising if the most sophisticated reading strategies prove to be the most difficult to teach', (ibid, p150)

Thompson (1987, pp52-4) referred to many studies, mainly involving L1 reading, where the use of a range of task-types to accompany reading were investigated via the affect they had on subsequent recall. The purpose which linked these studies was to involve a specific task-type (which included using flow-charts, titles, embedded headings, pre-reading questions, story-specific schemata, imagery and perspective) which in turn would cause a certain strategy to be used. The many studies referred to had in common that behaviour was affected by these techniques, but again none appeared to report that they had measured the continued presence of such changes at a subsequent date. Thompson concludes (ibid, p54), '[The battle] can only be fully won when learners discover for themselves that certain strategies can enhance their performance, and on the basis of this discovery are willing to continue using these strategies on their own.' However proving a task effect *is* probably a stronger case as we can assume that task-type may maintain an

effect by acting on the reader in the same way on each occasion. This is clearly different from relying on a reader independently to choose to apply the strategy.

Garner (1987) who focuses very strongly on the metacognitive aspects of strategy use, reviewed a series of studies which suggested that strategy training can be effective, but highlighted the need to track use after the training process had been completed. She cites Flavell (1970) who pointed out the problem of children failing to use strategies spontaneously even though they could use them effectively when explicitly directed to do so. Garner also stresses the need to make the strategy training part of the normal classroom activity in order to make it (ibid, p126) 'more personalized and more routinized.'

Barnett (1988) and Kern (1989) both measured the impact on comprehension of strategy training using control and experimental groups. Kern found that strategy use did improve, most by less able but also by average readers, whereas Barnett found little significant difference between groups. Oxford (1990) and Chamot and O'Malley (1994) also produced positive strategy gain results from intervention programmes. But again the longer-term post intervention effects appear not to have been measured.

The use of modelling as a technique (ie teacher and / or peer demonstration of strategies to encourage their use by learners) had been reported on by Palincsar and Brown (1984) in a process involving L1 reading called reciprocal teaching. This was further investigated and with L2 readers by Cotterall (1990a, 1990b, 1993). Her findings are interesting because between the three studies there is some movement in opinion. She demonstrates (1993) the need for any teacher designing a training programme to take into account a host of considerations. She notes the importance of metacognitive awareness as

a prerequisite for any programme and comments too on six further aspects which can influence the success or failure for the group or for individuals within it. These factors range across a broad set of highly important areas, including the role of the teacher, the language of discussion, the nature of the text, the strategies selected for teaching, the dynamic of the group and the understanding of the learners of the programme. Essentially then, although the two articles from 1990 had reported positive effects of such a programme, the deeper the investigations probed and the more Cotterall reflected on the processes, the more their inherent complexity seemed to prevent clear-cut conclusions. After a critique by Rees-Miller (1993) that strategy training had not been assessed empirically, Cohen together with Weaver and Li (Cohen 1998) carried out an extensive programme (the Minnesota SBI experiment) looking at learning strategies in general. This research produced a range of findings demonstrating some clear strategy use gains, and led Cohen to conclude that the notion of integrating strategy training into classrooms was endorsed by the results. Yet the same issue arises as Cohen suggests in his section on further research the need to (ibid, p150-1) 'assess the extent to which the learners transfer their strategy training from this experiment to performance in subsequent language classes.'

Two reviews of strategy research, McDonough (1995) and Gu (1996) found that the success rate of interventions on strategy use was poor, in fact 'fragmentary, unsystematic and narrow in scope (Gu, ibid p22). But McDonough's later (1999) survey of research on learner strategies claimed that the recent evaluations of interventionist programmes showed better empirical designs. In this he cited Cohen, Weaver and Li (Cohen 1998) and other examples such as the Oxford edited collection (1996), Dörnyei (1995) Talbot

(1997) and Nunan (1997). He concluded (McDonough, 1999, p13) that teaching strategies 'is not universally successful, but in certain circumstances and modes success is demonstrable.' But he also added (ibid, p14): 'Clearly research in the evaluation of training programmes has opened up, and one might almost say, come of age in the last few years, but that has merely served to open up further interesting problems for analysis, perhaps by different means.'

It is worth concluding this section by noting the important role of *motivation* on the part of the strategy user or 'trainee'. Both 'skill and will' are needed (Paris, Lipson and Wixson, 1983, p304) or strategic reading simply will not happen, whether training is given or not.

3.2.5 Monitoring strategy use

In his 1998 volume Cohen (ibid pp24-64) discusses at length six different methods of gaining information about strategy use. These are: oral interviews and written questionnaires (as used by Carrell, 1989, Oxford, 1990); observation (as in O'Malley et al, 1985); verbal report (as Cohen and Hosenfeld, 1981, Block, 1986; see also section 2.3); diaries and dialogue journals (as Oxford, Lavine et al, 1996); recollective studies (as Pickett, 1978, Poulisse et al 1987); computer tracking (as Baily, 1996). Clearly, as methods, these can be divided into those where the subject is consciously providing the information and those where an observer can attempt to track on-task behaviour and make conclusions from the evidence seen. Both will contain inherent reliability risks. The subject's own perceptions of strategy use may not reflect the true pattern, for example if s/he overestimates or underestimates the range and frequency of usage. An observer may

see only part of the picture of strategy use which produces certain observable outcomes. Cohen examines each method in turn, noting its potential advantages and disadvantages. It would not be appropriate here to rehearse his entire argument, but it is important to note his overall conclusions and to expand on some of his points by reference to other studies. He highlights very strongly the desirability of using multiple data gathering and interpretation methods, citing an example of a potential investigation, in this case into speaking strategies, where five of the six methods could be used to gain as full a picture as possible. He also notes the desirability of matching the method to the expected outcome of the study. For example, in contrasting structured with unstructured interviews and questionnaires, he points out that the information gained by highly structured questioning will generate data for quantitative analysis, but that it may be superficial (with simple questions) or skewed because of the potential suggestibility of highly detailed questions. Unstructured interview techniques might bring data that was too individualised for later comparison analyses, but on the other hand this might generate effective case-study data. Cohen also discusses the importance of the way in which any investigations are presented, with the full context as important as a detailed rationale and description of both the methodology and analysis. All of these features are vital to the deeper understanding on the part of research readers.

Cohen devotes a majority of the chapter to a more detailed analysis of the contributions of verbal report to knowledge about second language learning processes and strategies.

As this Study has used a form of verbal report (think-aloud protocols) for a major part of its data, the review of relevant literature will now deal with this aspect in a second major section.

3.3 The use of think-aloud protocols in gaining access to cognitive activity during reading

3.3.1 Introduction

In a key article for researchers considering the question of learner strategy use by second language learners, especially readers, Cohen and Hosenfeld (1981) set out a theoretical model for and some evidence towards a justification of using mentalistic data to effect improvements in language learning. At the core of this model were two processes: thinking aloud, where 'the subject just lets the thoughts flow verbally without trying to control, direct or observe them' (p286) and self-observation, which 'can range from largely unanalyzed verbalizations to those that reflect extensive analysis' (p286). They further distinguish within the latter between 'introspection' which occurs immediately or 'within a few seconds', 'immediate retrospection', which is 'a recollection of the experience after the event' and 'delayed retrospection', which may occur 'an hour, a day or even a week after the event.' (p286). After describing the theoretical model, they then (pp289-292) offer a detailed framework for researching mental states in second language learning, including instrumentation, sampling, data collection procedures and data analysis. Although each of these short sections contains a paragraph of discussion, the points raised (including potential pitfalls) are often listed rather than argued to a conclusion and they are rarely referenced to other studies. The authors do refer to the fact that these techniques are viewed with some 'automatic hostility' (ibid p285) by some researchers, but they do not review any concrete reports on this. As evidence for the usefulness of these techniques, Cohen and Hosenfeld reviewed six of their own studies on reading in a foreign language, which had used them. Three of these involved online

protocols to determine strategy use by 'successful' and 'unsuccessful' readers, and investigated whether instruction could influence reader behaviour in this area. The other three tried to ascertain what aspects of text were difficult for nonnative readers. In all cases they report that the think-aloud or self-observation techniques gave them data which allowed them to draw a conclusion to their research question. But the immediate 'leap' into using self-report data as a basis for further inferences about language learning, without having established the validity of the data itself, is a questionable feature, considering other writing of the time. In this review of relevant literature we will concern ourselves both with writing about the nature and value of the data itself and with reports on what it might give us of value for second language learning research, particularly for second language reading issues. Furthermore, although this study is concerned with immediate on-line verbalisations rather than with retrospection, both aspects will be considered in terms of previous research in order to account for that choice of method.

3.3.2 Self-report as data

The Cohen and Hosenfeld article represented an attempt to personalise for second language learning an issue which had been much discussed during the late 1970s (and continued to be so in the 1980s) in more general terms concerning memory, metacognition and first language reading. In four articles (which appeared either in *Psychological Review* or *Child Development*), Nisbett and Wilson (1977), White (1980), and Cavanaugh and Perlmutter (1982) all raised doubts about self-report when used to try to gain a view on the cognitive processes involved in problem-solving, although Ericsson and Simon (1980) proposed alternative views. (Ericsson and Simon's work is now

regarded as definitive and will also be reviewed more fully later in this section on the basis of their 1993 volume.).

A discussion of these issues should give a sounder basis from which to analyse further and more language-learning specific writing, which followed this particular flurry of activity between 1977 and 1982.

Nisbett and Wilson maintained that verbal reports (in this case retrospective rather than concurrent) revealed products not processes, and they argued for three major conclusions, which set out much of the 'debate area' (Nisbett and Wilson, 1977, p233). Firstly, about general reliability: 'The accuracy of subjective reports is so poor as to suggest that any introspective access that may exist is not sufficient to produce generally correct or reliable reports'. Secondly, on why there might be confusion about the origins of the self-report: 'When reporting on the effects of stimuli, people may not interrogate a memory of the cognitive processes that operated on the stimuli; instead they may base their reports on implicit, a priori theories about the causal connection between stimulus and response.' And thirdly, taking these two points further: 'Subjective reports about higher mental processes are sometimes correct, but even the instances of correct report are not due to direct introspective awareness. Instead they are due to the incidentally correct employment of a priori causal theories.' The authors reviewed a host of previous reports across a range of relevant areas including problem-solving processes. They claimed that in general there was little or no correlation between verbal report and behaviour, and so contested that the reports accurately reflected the actions of the subjects. They accepted that these processes might have inherent in them reasons for this lack of correlation and

so designed some brief experiments specifically to investigate people's ability to report accurately on the effects of stimuli on their responses. The results of these were that (ibid pp242-3): 'Subjects, as it turned out, were virtually never accurate in their reports. If the stimulus component had a significant effect on responses, subjects typically reported that it was non-influential; if the stimulus component had no significant effect, subjects typically reported that it had been influential.' The authors then went on to maintain a large role for a priori causal theory, showing that subjects (participants in a genuine evaluative process making judgements based on stimuli) and observers (subjects asked only to relate the relevant stimulus factors and judgements as a theoretical exercise) could come to significantly similar conclusions. They highlighted why verbal reports might be wrong, specifying circumstances such as removal in time, mechanics of judgement factors (such as serial order effects, contrast effects), context factors, non-events, non-verbal behaviour, and any discrepancy between the magnitude of cause and effect. They even attempted to account for this 'unawareness of unawareness', citing mainly that we confuse content with process and are thus unable to be sure how we solved a problem, or how our attitudes to a situation developed. This is further confused by the fact that feedback on the accuracy of verbal reports is so sparse.

Following Nisbett and Wilson, it would seem unwise then to place any emphasis on the value of self-report data. The way in which they argued their case (though not necessarily their view on self-report data) was challenged by White, (1980), who maintained that they had not justified their claim that conscious awareness was limited to the *products* of a cognitive process rather than the *process* itself. He went on to assert that there (ibid, pp105-6): '..are no criteria by which a mental event could be located under

one heading and discriminated from the other.' Therefore we do not know whether this confusion exists. He also claimed that using a retrospective report caused memory of the process to be lost, so supporting what Nisbett and Wilson had said, but accounting for it in this different way: 'It is a strain on cognitive capacity to keep in mind the route by which the solution was achieved after the achievement of that solution.' (ibid, p106). He developed still further the reasons for inaccurate reports (ibid, p107): 'In practice we tend to confuse, for instance, consciousness of the "doing" of some action with the attention we pay to the perceptual feedback of the action. We are more likely to say that we did something consciously when we have paid close attention to it, as in the case of some difficult and non-programmed sequence of behaviour, or one that must be acutely sensitive to changing circumstances..... Once we have started speaking, we are no longer in a position to know whether, or in what sense, our utterance was "done consciously", or whether its contents were conscious before they were uttered.' In other words the consciousness of performing the self-report may influence the subject's view of the process or perhaps even the process itself. So he does not dispute the accuracy of retrospective reports, but differs from Nisbett and Wilson about why this may be so.

Cavanaugh and Perlmutter (1982, p17) expanded on this when commenting on the use of post-experiment interviews, which sought to clarify cognitive processes: 'Most important, demand characteristics, such as perceived expectations to report effects of variables included in the experiment, even if none was experienced, play a major role in the subjects' responses. In addition, due to their retrospective basis, it is impossible to determine if reports obtained during post-experimental interviews are true reflections of what subjects did during the experiment or if they are post hoc rationalisations or

hypotheses designed to account for their behaviour.' They also raise the problem of conflict between the original task which is the subject of the monitoring and the monitoring process itself (ibid p19): 'The most serious problem with think-aloud methods is the possibility that the mediating processes responsible for translating processing into verbal reports interfere with carrying out the task. The result is incomplete verbal report, both in terms of quantity (some knowledge is never translated into verbal reports) and quality (some verbal reports do not accurately reflect underlying knowledge).' They conclude that verbal reports alone will not give sufficient or sufficiently reliable data and recommend that (ibid p22): '..... multiple assessments of memory knowledge (ie the method of converging operations) must become standard.'

There is then a very strong view, expressed in different ways, but nevertheless seeming to find agreement, that we would be unwise to regard self-report data as a true reflection of cognitive processes. Its main danger would seem to be that it might sometimes reflect these processes, but equally it might not. This apparent randomness makes conclusions very difficult to draw. But there was another case put at that same time.

In the longest review (of the four articles) of this subject matter and in some contrast to many of the above expressed views, Ericsson and Simon (1980) addressed the issue of verbal reports being used as data in the following ways. They insist that the methodology for collecting data be properly defined and point out that, at that time, any type of introspection was regarded as the same as any other by critics of the method, whereas it was easily possible to delineate different approaches which would also have greater or lesser validity. They examine in great detail different types of verbalisation procedure, including an examination of the relative roles of the two 'tasks' (ie the set task, which is

being monitored and the verbalisation task – both White and Cavanaugh and Perlmutter had of course commented on this difficulty). The two tasks can be distinct or interrelated, concurrent or subsequent. Clearly, these factors will affect the verbal report, as it may be subordinate to and dependent on the ongoing cognitive process (and so more likely to reveal that process) or it may be concurrent and separate (and so less likely to do so). These differences must be held in mind when the data is examined. Ericsson and Simon then distinguish between three levels of verbalisation: (ibid p219): ‘When information is reproduced in the form in which it was acquired from the central processor, we will speak of *direct* or *Level 1* verbalization. When one or more mediating processes occurs between attention to the information and its delivery, we will speak of *encoded Level 2* or *Level 3* verbalization.’ These last two categories are then expanded to draw out a difference between an intermediate recoding into verbal code (Level 2) and an intermediate scanning / filtering or an intermediate inference or generative process (both of these Level 3). The authors discuss forms of probing, which might be used to elicit the data and give both a general and detailed specification for a processing model. Already then it is evident that by breaking down the ‘problem’ areas of elicitation technique, respective role of task and verbalisation, and type of verbalisation process, and by locating this discussion in a model of information processing, accepted by psychologists of the time, the authors begin to be much more specific about how we might judge the accuracy of verbal report data. They show in detail for example that probing by the researcher is a highly complex issue, and they conclude that asking subjects to focus on a specific action rather than a generalised one, and by asking for a general verbalisation rather than directing a concrete (possibly leading) question, we are more likely to receive a

representative picture of the processes. They accept that such data will always be incomplete, but maintain that what is given will be accurate if these guidelines are followed. They then review a whole series of studies using the theoretical model they have established and conclude, (ibid p235): ‘... these studies consistently support our model’s prediction that producing verbal reports of information directly available in propositional form does not change the course and structure of the cognitive processes. However, instructions that require subjects to recode information in order to report it may affect these processes.’ And further: ‘With increase in experience with a task, the same process may move from cognitively controlled to automatic status, so that what is available for verbalization to the novice may be unavailable to the expert.’ They offer three reasons for the incompleteness of verbal reports (ibid, p236): ‘(a) The information is not heeded, hence not stored in short term memory (STM), hence not accessible for verbal reporting. (b) Not all the information available in STM at the time of the report is actually reported. (c) Not all of the information previously available in STM has been retained in long term memory (LTM), or is retrievable from LTM.’

The authors finish by an extensive review of the evidence presented by Nisbett and Wilson (op cit). Their conclusions are that there *are* defined conditions under which one can assume that verbal reports are valid, (but that these were not the ones spoken of by Nisbett and Wilson), and that under other experimental situations and with other procedures: ‘veridical reports could hardly be expected.’ (ibid, p247). In other words introspection of this sort is a perfectly reliable methodology as long as it is carried out in a rigorous manner, taking into account the most likely pitfalls.

It is perhaps worth jumping forward at this point to their later volume, (Ericsson and

Simon, 1993), where they review their previous work, and maintain broadly similar conclusions, especially about the key issues of experimenter prompting and the relative strengths and weaknesses of concurrent and retrospective verbalisations. They do add that very short tasks can be reliably reported upon retrospectively, but that previous caveats about time delay otherwise apply. They discuss the think-aloud process specifically as follows: 'Thinking aloud and talking aloud can be elicited almost instantaneously by the appropriate instruction from virtually all human adults.' (ibid p224). They add that 'most people's normal mode of thinking is silent', but note that 'even a brief reminder from the experimenter will start them speaking again.' (ibid p256). This has one drawback, however. 'When subjects are verbalizing concurrently, they will verbalize the information as it is heeded. However, when subjects have been reminded to think aloud they will verbalize the information in STM. Almost any model of the cognitive processes would predict that goals would be kept in STM longer than intermediate results ... Thus our model would predict a relatively high frequency of reports of goals in response to reminders.' (p257)

With regard specifically to thinking aloud while reading they explain what may actually occur and why. One problem may be that in the protocols produced little more than the text is vocalized. Ericsson and Simon's explanation for this is either that 'the internal representation (and hence the vocalization) simply mimics the text' or that 'during reading with comprehension information must be accessed in LTM to generate a coherent representation of the text's meaning. A protocol that simply follows the text omits this kind of information.' (p254). They also cite the work of Waern (1979) who found significant differences between protocols produced when readers were asked to complete

a task compared with those produced when they were asked simply to read. This raises the possibility that task-design may have an effect on the reading process itself, that we may be able to activate strategy use by asking for certain approaches to a text.)

A further recent work on thinking aloud was produced by van Someren, Barnard and Sandberg (1994). They consider cognitive processes involved in problem-solving and also stress the importance of defining closely how the information is accessed. They share the other writers' mistrust of the reliability of retrospective techniques and are wary of the influence which certain types of prompting might cause. They sum up the alternatives to concurrent think-aloud in a useful table:

Table 3.1 van Someren, Barnard and Sandberg, 1994, pp24-5

The table below summarizes the techniques with respect to the method (apart from think aloud):

	Disturbance	Memory errors	Interpretation
Retrospection	No	Yes	Yes
Introspection	No	Little	Yes
Prompting	Yes	No	Little
Dialogues	N/A	No	No
Structured techniques	Yes(?)	Yes(?)	Yes(?)

In the case of the structured techniques the risk of distorted data depends on details of the structuring and timing of the questions.

The reasons for invalid data, according to the table above (ibid, all p24):

'The disadvantage caused by disturbance ... lies not so much in the hampering of concentration, but in the possibility that thought processes take directions different from those they would have taken had the subjects been left on his or her own.'

'Memory errors can produce both incomplete and false reports.'

'If the subject is asked to interpret his cognitive process or if he is required to do so by a structured technique that does not fit the content of the process, this is a source of distortion and invalidity of the data with respect to the cognitive process.'

So, of the methods discussed which are: observation of problem-solving behaviour; structured techniques; verbal reports via retrospection or introspection; prompting; dialogue observation and think-aloud, it is clear that the authors consider the last two to be more reliable when all factors are considered. About each of these respectively they write, (ibid, p23) '... dialogues have the advantage that they can be recorded under more natural circumstances than a think-aloud session.' 'This kind of instruction seems to work quite well with (young) children to whom the think-aloud instruction is not very clear.'

'Another possibility to enhance talkativeness is to get people to collaborate on a task.'

'When people collaborate they will sometimes have differing opinions. Thus they are forced to give arguments to clarify steps of the thinking processes.' 'People will not verbalise all their thoughts in a dialogue situation.'

(ibid, pp25-6) 'Thinking aloud is a method which in principle does not lead to much disturbance of the thought process. The subject solves a problem while the talking is executed almost automatically. The data so gathered are very direct, there is no delay.'

The subject does not give an interpretation of his or her thoughts, nor is he or she required to bring them into a predefined form as in structured techniques.' Clearly then, the two methods adopted for this study would seem to have validity as long as the structures for obtaining the data are based on the principles outlined by the researchers cited.

3.3.3 Thinking aloud and L2 reading

Since the 1980 article by Ericsson and Simon many researchers have used a think-aloud procedure to obtain data about the reading process.

Brown (1980, p455) raised the problem of memory capacity when a *concurrent* protocol is used: 'Consider skilled readers... they can proceed merrily on automatic pilot, until a *triggering event* alerts them to a comprehension failure. ... The difference in time and effort between the normal rapid *automatic pilot* state and the laborious activity in the *debugging state* is the difference between the subconscious and conscious level.' And further, 'The debugging activities themselves occupy the lion's portion of our limited capacity processor and the smooth flow of reading abruptly stops.' So it may be that the most interesting aspect of cognitive processing while reading, the way in which we deal with a problem, is difficult to monitor because of the conflict in demands on working memory of the two tasks, solving the problem and reporting the process.

Seliger (1983) explored this potential difference between using self-reporting for certain types of problem-solving and for introspection on the language learning process: (p188) 'Linguistic processes are clearly more abstract than visual processes used to move blocks around in the typical problem-solving type of experiment. In asking subjects to describe ongoing or previously experienced linguistic processing, we are asking the subject to describe the very means of description itself. That is we are expecting language to function as both process and product at the same time.' He continued: '... memory load on [language learners] is even greater than on monolinguals....Therefore the likelihood

of much reliable data having anything to do with actual linguistic processing is quite small.' Of course, Ericsson and Simon's (1993) view on this would be that it would depend on whether subjects were simply reporting the process as it is or being asked to recode or interpret. But there may be a rationale, taking both Seliger's and Brown's views into account here, for considering, with second language reading, the effects of the difficulty of the given text. The more there are unknown lexical or structural items, the more the focus would be on deciphering meaning, and the less processing space there could be for reporting what that problem-solving process actually consists of. Seliger also questions the reliability of self-report as a conscious activity (ibid p188): '... the learner's report might very well be the result of examining his own second language production and surmising how he *might* have reached this product, rather than a description of what went on at the time of producing the second language utterance. In fact learner / subjects with nothing to report might report something in order to satisfy the experimenter, and the experimenter would not be able to test for the validity of the report.' This view seems to be in agreement with the Nisbett and Wilson and White attitudes to the problem, but again could be confounded by taking into account the Ericsson and Simon position on asking for concurrent verbalisation rather than pre-primed probes. If the subject is not really aware of what the researcher is looking for s/he cannot deliver it.

Hosenfeld (1984) recommended that one should 'begin each interview with a practice session' (ibid, p232). This is in tune with Ericsson and Simon's (1993) recommendations if the practice is merely intended to establish the habit of talking while reading. It might raise a further issue about the training of subjects, whether to read in a certain way or to

do the think-aloud process: if the wrong sort of training is given, does this influence the reading or reporting behaviour of the subjects so that the protocol ceases to be a true representation of their internal mental processes, but merely a reflection of perceived teacher or researcher expectation? Hosenfeld stresses the need to get as complete a report on their behaviour as possible and gives criteria for how she will decide if they are to introspect or retrospect. (depending on whether or not their reading normally involves a translation process). She also notes the need to phrase questions very carefully: (ibid p232) 'Questions should be worded so that they do not impose directions upon students' thought processes or self-report.'

In the same volume, Harri-Augstein and Thomas (1984) review their conversational paradigm which is aimed at (p253): 'enabling L2 readers to arrive at personal descriptions of their reading process, so that they can reflect upon and develop their competence.' This process involves the reader in commenting on how s/he 'maps meaning onto the words on a page,' and also uses 'personally relevant criteria for assessing comprehension and how learners invent, review and change meaning.'

They conclude the value of this process as follows (ibid p264): 'When students embark on conversational investigations of their own reading they are seldom aware of the cognitive and affective processes which underlie their reading behaviour..... By guiding learners into contact with their own processes, they become more aware of their skills and attitudes towards reading and so bring these under review.'

This again raises the question as to whether teachers and researchers are guiding subjects into working in a certain way during the experiment. We would need to know whether a) this represented their normal cognitive process during reading or at least b) whether there

was subsequently permanence in any resultant learned behaviour. If neither applied, we would have learned nothing from the protocols about 'normal' cognitive processes, but merely about influential processes during experimentation.

An important next stage in the debate was offered by Olson, Duffy and Mack (1984) who looked specifically at the role of 'thinking-out-loud' (TOL) as a method for studying real time L1 comprehension processes. They give a general rationale for using TOL, list and discuss a variety of different types of the process, review a series of studies, look at other applications of the TOL task and summarise the advantages and disadvantages of the method. In so doing they pay due attention to both the Nisbett and Wilson and Ericsson and Simon positions, and elaborate (ibid, p254): 'Furthermore, TOL [thinking-out-loud] data should not be taken as *direct* reflections of thought processes, but rather as data which are correlated with underlying thought processes. TOL data provide a sample of what's on the subject's mind during the task. But they will not necessarily reveal the strategies, knowledge sources or representations actually used. These theoretical constructs must be inferred from the TOL data.' This reflects the fact that in Ericsson and Simon's (1993) volume they devote an entire chapter (pp169-220) on 'Inferences from Verbal Data'. Clearly it is helpful to isolate from an analysis of the reading process exactly which aspects might best be revealed from the use of think-aloud protocols. According to Olson, Duffy and Mack (ibid, p255): '... the TOL task is best used to study the higher level processes in reading: the inferences, predictions, schema elaborations, and other complex cognitions that occur as part of skilled reading.' They continue on p257: 'These data should reveal the kind of *strategies* used by readers in accomplishing

these tasks, the kinds of *knowledge sources* employed, and the kinds of *representations* constructed.' And crucially: 'While memory measures like recall have provided useful information about the knowledge sources and representations used in text comprehension, they tell us very little about the strategies employed or about the sentence-by-sentence interactions among the knowledge sources and representations.'

Like Ericsson and Simon, Olson Duffy and Mack analysed in quite some detail a selection of TOL task types, which included sentence by sentence talking with either general or focused instruction probes, selective talking (at controlled probe points) or after the fact talking (which the authors agree is problematic).

Thus, the more open-ended instruction to talk at any time about whatever is going through the subject's mind was not featured in this survey. Obviously, the results in the above TOL tasks used by the authors reflected what subjects were asked to focus on, ie certain aspects of what they were thinking about while reading were investigated, rather than the question – what are they thinking about ? Their summary tables (ibid, pp266 and 268) show that the proportions of TOL productions were directly dependent on three focus areas, with a minimum of 73% of the productions concerning these three of the eight items represented in the table. Even though this is what the authors *wanted* to research, it again demonstrates the need for care as regards how much subjects will be influenced by instructions. The more researchers specify, the less they can be sure that the information they receive is reflective of natural and normal processes. The authors list (ibid pp284-5) a summary of advantages and limitations of using TOL data. The limitations outweigh the advantages, but only in the same areas as had already been indicated by Ericsson and Simon and the other writers. The need for sensitivity to

instructional variables, and to text type, the difficulty of TOL data analysis, the big capability differences with regard to carrying out TOL by different subjects, and the potential of the TOL task to influence comprehension processes are all familiar caveats.

Afflerbach and Johnston (1984), in a review of L1 self-report methodology, underlined the importance of several areas. Two of these were training for self-report and heeding the influence of task (for example text summary demands more processing space than comprehension questions). They also noted the need to choose subjects carefully (will self-report reveal whether they are more '*verbal*' or better readers ?) and to use other data gathering methods such as eye movement data to validate self-report. Finally the analysis process was very complex and needed to reflect verbal features (eg intonation, tone, pauses) in written transcripts as well as to evolve a classification system for observed behaviour. For the relevance of this last point for this study, see Chapter Four, and specifically the references to the Pressley and Afflerbach (1995) coding system.

Garner (1987) reviewed research regarding metacognition and L1 reading and demonstrated that links between good reading and metacognitive awareness were not just age-related, as younger *and* poorer readers seemed at a disadvantage when compared with good readers. Often strategies were either not recognised or identified wrongly as positive or negative by poorer readers. Better readers were also more able to detect errors or inconsistencies in a text and a reliance on intra- rather than intersentence meaning construction was more typical of poorer readers. She addressed various methodologies for gathering verbal report data, including thinking aloud, where she drew from: Hayes

and Flower (1980) and Hayes-Roth and Hayes-Roth (1979) to illustrate how protocols can be divided into sections for analysis and some aspects of that analysis. Garner showed that accessibility, inadvertent cuing and verbal facility can all be problems with think-aloud as a process, but the other problems mentioned above would probably not be. She summarised (ibid, p78): 'We will ... have three applications for the think-aloud procedure: in research, to find differences among groups of learners in (verbalised) strategy use; in diagnosis, to find deficiencies among particular learners in strategy use; and in instruction, to model appropriate strategy use for reading or a host of other complex cognitive tasks.'

Cohen restated and reviewed his thoughts about the positive use of mentalistic data, (Cohen, 1987) by discussing in more detail the issues surrounding the practice. He now divided the categories of learner-report data into three groups. These were self-report (the learners' descriptions of what they do, characterised by generalised statements about learning behaviour), self-observation (the introspective or retrospective inspection of specific language behaviour) and self-revelation (a think-aloud stream of consciousness, unedited and unanalyzed).

The dilemma was whether learners engaged in the last of these categories are thinking aloud and describing the process or observing (by which he meant starting to analyse the process). He cited 6 important factors including the number of participants in the process, the research context, the recency of the event for the subject, the mode of elicitation and response, the formality of the elicitation and the degree of external intervention.

Cohen commented that the data available was generally about language use rather than

language learning, but noted that the language which accompanies cognitive tasks such as solving anagrams or mathematical or logic puzzles is often accessible. Therefore applied linguists should also be able to access reliable verbal reports.

Cohen admitted that Seliger (1983) made good points with regard to L2 learners verbalising, (Cohen, 1987 p36): 'Memory of mental events is problematic for the learner, and could lead to faulty reporting, but it then is the challenge for the researcher to tap this information while it is still available'

In the same volume, Abraham and Vann (1987) carried out both interviews and think-alouds with a group of ESL learners. They reported on two of these subjects in the chapter, one successful and one less successful learner. The think-aloud protocols were obtained via four tasks on: verb tense usage, article usage, cloze and composition. Clearly here the nature of the task determined the range of strategies which would be most used. The authors took their strategy typology from Rubin (1987) and included cognitive and communication strategies.

Each of these categories was then further broken down as regards the interview protocols which also formed part of the investigation, but the think-aloud protocols were categorised only under those general headings as the authors felt that further breaking down would have been unreliable.

They found that reader behaviour as well as ability made a difference; both the time taken and the number of strategies used differed greatly between subjects. The exact procedures followed are not described in detail so we are unsure whether any probes were used or how task was set up.

Roehler and Duffy (1991) wrote about the instructional actions which could be performed by L1 language arts teachers and recommended that teachers could model think-aloud processes. They cite two long examples from other studies, but the practice raises the familiar objection that this is a form of training rather than a method of eliciting normal behaviour and that, as such, there is no built in monitoring of its long-term success.

Trabasso and Suh (1993) opted in their study of whether global causal inferences are made during L1 comprehension to allow their subjects to think aloud freely, having considered and rejected the notion of regular probes. 'We wanted them to be as free as possible regarding how they processed the text. In particular we wished them to be free to infer goal plans and to use them to interpret actions if they wished to do so without being told or asked.' (ibid, p12). The authors found that subjects' protocols revealed that they maintained, retrieved, elaborated or explained information while thinking aloud. These functions broke into nine categories of inferences and mental operations. The findings were triangulated by using other methods such as recall summarising and reading time measurement. They sum up their view of the think-aloud process as follows (ibid, p31): 'We believe that talking aloud *reflects* what information the subjects have accessed and what operations they have performed during comprehension. We do not wish to claim that what becomes conscious and what is talked about in communication is identical with how and what the subjects thought about during silent reading. We do claim, however, that the text information and the reader's knowledge and motivation jointly determine, constrain, or lead to what is thought about and what is talked about consciously.' This

judgement has great relevance for a study of thinking aloud while reading where some transcripts are very sparse, as it foregrounds the importance of basic understanding of and interest in the text as well as the capability of using specific strategies.

Kern (1994) investigated L2 readers' use of translation as a comprehension device via a think-aloud process. He found that the data was able to demonstrate a decline in the use of translation as reading became more confident over time and that different types of translation could be associated with greater or less success in the task. Word-by-word translation which did not integrate meaning led to less successful performance than that which (ibid, p455) 'facilitated the synthesis of meaning, presumably by increasing the functional capacity of working memory.' Although his study uses think-aloud to investigate another area, we can also utilise his findings to predict what might appear or not appear in a think-aloud protocol offered by a reader at a particular level of attainment.

Pressley and Afflerbach (1995) carried out a major research review of investigations which had used the think-aloud process in L1 reading, and produced an extensive coding (see discussion in Chapter Four) of the actions revealed by the protocols. They refer to their book (ibid, p139) as a 'cornerstone' for further research rather than as a summary of all there is to know. But they also use their data to define 'constructively responsive reading' (ibid, chapter 4, especially p98 onwards). This they regard as a model which encompasses previously proposed theories of text processing (ibid, p117). Thus verbal protocols are here seen not just as a window on the reading process but as a key piece of evidence for theory formulation.

Crain-Thoreson et al (1997) looked at two models of think-aloud procedure accompanying L1 texts. They compared marked texts (ie those that have typically a red dot placed at certain points to encourage verbalisation) and unmarked texts. In so doing they wished to evaluate the conflicting views of Olshavsky (1976-7) and Afflerbach and Johnston (1984) who respectively preferred the 'red-dot' method because unmarked passages were unlikely to elicit few comments, and rejected it because it might elicit superfluous or place-holding comments. The authors concluded that marking produced (Crain-Thoreson et al, 1997, p587) 'the more veridical picture of text processing.' They felt that the process encouraged those who could not understand to report the fact, and those that could to report *how* they were understanding the text. Nevertheless, as previously stated, this method does tend to dictate an approach to reading, and bearing this in mind as well as the view of Afflerbach and Johnston the current study opted not to mark the texts.

Cohen (1998) outlined six different methodologies for the discovery of information about reading and learning processes, and recommended multiple approaches wherever possible. He devoted as much space to verbal report as to all of the other five together. In doing so he maintained his definition of the three forms of verbal report from the chapter of 1988, and provided (ibid, pp49-62) an extensive exploration of the need to structure the method very carefully. This perhaps best summarises the change over the two decades from the 1981 Cohen and Hosenfeld article which failed to address many of the important issues which were just then becoming evident. Cohen sets out seven 'issues in verbal

report methodology' and makes nine major suggestions 'towards more robust verbal report methods and more complete write ups'. Thus he covers all of the areas highlighted so far in this research review: immediacy; respondents' roles; probing; guidance; the effects on task performance; the nature of the subjects and of the materials; methods of analysis; reliability. The purpose of this section is to (ibid, p62) 'provide greater systematicity both in the collection of such data and in the reporting of such studies through the research literature.

In conclusion, the literature on think-aloud reporting to accompany reading more than validates it as a method provided that the rigour demanded by Cohen (1998), echoing Ericsson and Simon (1980, 1984, 1993) is present at all stages of the process. While we know that as regards on-going meaning construction, including concurrent strategy use, it will only ever reveal what the subjects are prepared to tell us, it does reveal at least that. (Without a think-aloud process we might know something of what readers are doing, through eye movement data for example, but we will not be sure that our inferences are correct, nor will we know why the movements are as they are. With a think-aloud process the subject may reveal overtly or covertly the rationale behind her or his reading. When a text is accompanied by a task, a think-aloud protocol will at some point express a rationale for decisions made in response to questions or instructions to decide, for example the truth of statements accompanying the text. As we have seen a lack of verbal reporting may simply indicate that a reader is not 'verbal', and successful task responses will then demonstrate that there is a need for further study, perhaps from a different methodological standpoint. But a lack of responses to the task as well as a lack of verbal

reporting will then reveal more about the reader as a reader and this in turn will shed light on the lack of verbal data offered. Chapter Four will take up the methodological implications for the Part Two Study.

3.4 Monitoring the role of peer interaction in group-based problem solving

3.4.1 Introduction

In a positivist paradigm cognition is seen as an entirely individual concern, but current researchers working from a constructivist approach have demonstrated that this is not the complete story. Resnick (1991, p2) writes: 'Cognitions *about* social phenomena have long been of concern to social psychologists. But constructivism forces students of many social phenomena to treat social processes *as* cognition, leading them to analyze the ways in which people jointly construct knowledge under particular conditions of social purpose and interaction.'

In order to reflect on issues which are of close concern to this study, this section will survey literature in three connected areas. Firstly, the major theories on relevant sociocultural issues by Vygotsky, and Bakhtin will be briefly reviewed. This provides the framework for our further investigation. Here we will also consider the application of theory: evidence that collaborative learning has been successful in small group format. Secondly we will consider the work on the nature of positive group talk by Barnes et al, Phillips, Mercer et al, Almasi and Brewster. As part of this we will consider issues of group constitution, including gender. Finally, we will examine the perspective on dialogic interaction which can be brought by a more linguistically-based analysis, based on approaches drawn from discourse and conversation analysis. But it is important to

recognise that in some respects the second and third of these strands are intertwined. It is for the sake of increased clarity that they are separated here.

3.4.2 Socially shared cognition and collaborative learning

Vygotsky's writings (eg, 1978, 1981, 1986) present two very important constructs which lay behind the rationale of the study. The general genetic law of cultural development states (Vygotsky, 1981, p163) 'Any function in the child's cultural development appears twice, or in two planes. First it appears on the social plane, and then on the psychological plane. First it appears between people as an interpsychological category, and then within the child as an intrapsychological category.' Furthermore the emphasis he gave to the role of the mediating tools, eg language, was such that they were shaping the action which they framed. Although taken literally these viewpoints would indicate little or no development emanating from *within* an individual, Wertsch and Tulviste (1998, pp26-7) note that 'such action always involves an inherent tension between the mediational means and the individual or individuals using them in unique, concrete instances.' Thus the language used is inextricably linked with the specific context, the specific intra- and interpsychological processes taking place and therefore both shapes and is shaped by the individuals concerned. It is possible then, within the field covered by this study, that for some participants at least the experience of group work and the learning gained from group work may be qualitatively different from the parallel experience and learning associated with the individual tasks.

The second aspect is that of the zone of proximal development. Vygotsky (1978, p86) explains this as follows: '[The zone of proximal development] is defined as the distance

between a child's actual developmental level as determined by independent problem-solving and the higher level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers.' He felt that when assessing a child's developmental level it should not just be the level at which the child functioned independently that should be recorded, but the potential supported level s/he could reach.

Both aspects of Vygotsky's theory are combined in his notion, discussed in Wertsch (1991), that two planes of development, the natural and the cultural coincide and mingle and eventually become a single line of personality. Thus guided learning, for example within the scaffolding system defined by Wood, Bruner and Ross (1976), plays a part in this line of development. Social interaction, including that with more capable peers or adults is a major source of cognitive development. Clearly this has relevance for a study which involves both individual and collaborative tasks, not in a developmental sense, where cognitive change is measured, but as part of an analytical process which examines the difference between the processes and products arising from interaction with parallel tasks in the two contexts. As part of this we may see that for some participants the benefits of working with others gives added value to their own individual capability.

Bakhtin has an additional relevance as he focused more on the interactions at utterance level, but not in isolation: 'Speech is always cast in the form of an utterance belonging to a particular speaking subject, and outside this form it cannot exist.' (Bakhtin, 1986, p71), and further, (ibid, p84) 'Any utterance is a link in the chain of speech communication.' He used the term dialogicality, defined by Maybin (1994, p132): 'Every utterance is always

also a response to some previous utterance and every utterance also anticipates and takes into account its own possible responses.' Bakhtin himself (1986, p68) defines this process as: 'Any understanding is imbued with response and necessarily elicits it in one form or another: the listener becomes the speaker.' He links with Vygotsky in his view of the power of the mediating tool of language, as summarised by Wertsch (1981, p95): 'In Bakhtin's view, speakers always use social languages in producing unique utterances, and these social languages shape what their individual voices can say.'

Clearly in our analysis of the talk generated in the study we need to be mindful of the presence of social speech genres, perhaps defined partly by the presence of the background institutional context. Where relevant, and indeed possible, we may attempt to discern where individual actions and responses are controlled by the context rather than by their interaction with the task.

Evidence that demonstrates the positive effects of small groups engaged in collaborative tasks on learning comes from many sources. Larger scale surveys were carried out by two researchers each on two parallel occasions (Sharan 1980, 1990, Slavin 1980, 1990) and all four surveys testified to the benefits of cooperative learning programmes which were well-planned, well-structured and closely monitored. In addition Johnson and Johnson (1989) reviewed extensively the issue of cooperation versus competition and again showed the benefits of structured collaboration. The importance of the *nature* of those collaborative interactions was demonstrated by several researchers, eg Perret-Clermont (1980), Forman (1981), Glachan and Light (1982). All found that discrepant views within groups contributed to cognitive growth as a result of the articulation and resolution of

positive conflict. It is important for this study to establish that research has found collaboration to be an effective learning tool as this validates the purposes of any comparisons made here between the individual and group modes. This is especially relevant since the tasks used in this study were genuinely collaborative, ie demanded a group response, as opposed to the type of collaborative context where individuals can cooperate where they wish as part of producing individual responses. But since this is not a longitudinal study where the long-term success of groups is to be measured, there is less value of a more extensive review of this area. The study looks at how discussions between participants extend or inhibit their individual and collaborative understanding of the texts and tasks, and as such it is the nature of the talk in those groups and the nature of the individual protocols which are paramount as data.

3.4.3 The nature of talk

3.4.3.1 Introduction

As a result of the general theoretical view of social interaction outlined above, studies were made of the talk produced in classrooms either under teacher guidance or simply between peers. As stated earlier, the separation of the literature around the global nature of group talk from that which examines more specifically discourse elements in the talk is not entirely clear cut. It does, however, allow a degree of separate focus on the same material which will facilitate analysis. In terms of over-arching principles it is important to note the definitions made by Kress (1985, pp142-3) of the terms genre, discourse and texts. 'At any point in the history of a social group there exists a repertoire of linguistic textual forms ... These are the genres which determine the textual form in which a

specific discourse finds its expression.' And on discourse, (ibid, p139) 'the institution will produce a set of statements which largely define, describe, delimit and circumscribe what it is possible and impossible to say...'. 'Texts then are constructs produced in the relatively constrained interrelation of discourse,' (ibid, p144). Kress also notes that although (at that time) the formal structures of non-literary genres were not much researched, those of classroom lessons (Sinclair and Coulthard 1975) and pupil group discussions (Barnes and Todd 1977) had been addressed. We will review findings regarding the nature of group talk shortly, but it is worth perhaps noting here some issues regarding genre and discourse in the setting of this study. Extensive discussion of foreign language reading tasks was not a normal activity for these participants, and indeed discussion about learning processes was also less common than it might have been in other curricular areas owing to the target language issue. While there would be clear expectations surrounding expression in the whole-class context and especially in interaction with the teacher, there would be fewer accepted 'rules' around group interaction. Erickson (1982) wrote about the two interrelated factors of academic task structure (ATS) (a patterned set of constraints provided by the logic of sequencing in the subject matter content of the lesson), and social participation structure (SPS) (a patterned set of constraints on the allocation of interactional rights and obligations of various members of the interacting group). His description of lessons in action showed that a gifted teacher can create a positive lesson structure by improvising her or his control of these two factors in order to maximise the benefits to the learners of the processes involved in the task structure. Clearly this attribute of experienced and sensitive teachers may or may not (and to varying degrees) be managed by learners within a small group. Additionally, part of the

learner expectation about tasks such as those used in the study would be that they were product- rather than process-oriented. This too would have a shaping effect on the discourse that emerged.

Finally there is the issue of peer relationships and role expectations that exists within any class of learners, and especially amongst the adolescent age group. This is impossible to quantify or qualify, but we may suppose that on a single occasion with a comparatively formal setting, it might play somewhat less of an important role than would emerge in a long-term study. Nevertheless, very relevant to this issue is Kress' point made later in the chapter (op cit) that texts which arise from interaction also demonstrate the power struggle between the participants and can show dominance or collaboration, depending on the nature of the interaction. This factor clearly has the potential in the study to influence the group talk very strongly. What we will be able to observe and describe is exactly that behaviour; what we will not be able to do is account for it.

Therefore in this review and later in the analysis we shall be considering firstly the characteristics of the discourse within the genre of school-based group talk, and then more specific areas concerning relationships between the participants as shown by the texts which they produce.

3.4.3.2 Classifications of group talk

Barnes described certain positive elements of talk within small groups (Barnes 1976, Barnes and Todd 1977) as 'thinking aloud' and 'exploratory talk'. Both are defined variously as 'a means for controlling thinking' and 'groping towards a meaning' (ibid, p28). The importance of this was that language was being used not just for

communication but as a tool for learning. Both the speaker and the audience can be part of this learning, and as Barnes characterised exploratory talk as 'marked ... by hypothetical expressions' (Barnes 1976, p28) it is easy to see how a less didactic discourse might enable dialogue, discussion and the shaping of ideas to occur. Bennett and Cass (1988) found that certain types of group constitution (either homogenous groups of high attainers or mixed groups where a high attainer was placed with two low attainers) produced better results than others. But they also repeated earlier warnings (Bennett 1985) about assumptions of automatic benefits from group talk, ie that some types of grouping for tasks and hence some types of talk did *not* enhance learning. Barnes subsequently further refined his description of positive styles of group talk, developed the contrasting categories of presentational and exploratory talk (Barnes, 1992) and revisited content and interaction frames (Barnes and Todd 1995). In comparing presentational and exploratory talk he adds to his previous definition of exploratory talk by noting (ibid p126) that 'learners are unlikely to embark on it unless they feel relatively at ease, free from the danger of being aggressively contradicted or made fun of.' In contrast presentational talk tends to occur in situations which 'discourage exploration: they persuade the speaker to focus on "getting it right", that is, on appropriate speech and the expected information.' (ibid p126). Wells (1992, p289) develops this theme in a different way. 'Thus while it is true that one function of a text is to enable the listener to reconstruct the speaker's meaning as accurately as possible, there is a second and equally important function, which is to provide the occasion for the generation of new meaning as the listener makes sense of what the speaker says by responding It is in this second "dialogic" function that a text acts as what Lotman (1988) calls a thinking device.' He

goes on to note that under such dialogic conditions pupils can provide the teacher with evidence of their levels of skill and understanding. The notion of the 'dialogic' model is also discussed by Maybin (1994, p132) where she describes talk as 'not a transparent conduit through which knowledge is passed, but an integral part of how understanding is collaboratively accomplished.'

An alternative model is that of Phillips (1985), who listed five modes of talk which he found in the talk of 10-12 year olds. These modes were: hypothetical; experiential; argumentational; operational; and expository. Of these the operational and argumentational were the most commonly encountered. Expository talk reflected more the style of teachers in classrooms and was found least. Of the remaining two he writes (ibid, pp76-77) 'Both oblige group members to review the conversation itself; to treat the text as a shared field and to treat remarks made at any point as remaining present for contemplation during an extended period of time.' Clearly these would equate with the notion of exploratory talk. Phillips is very definite on a further important issue (ibid, p74). '... children negotiate together to arrive at a way of talking and are therefore not constrained to produce any one style more than another.' This view clearly reflects that different learning styles and abilities will shape individuals' approach to groupwork but it also needs to be considered in the light of Kress' view on genre and discourse as well as Brewster's views on the role of task.

Mercer (1995, pp104-7) described three categories of talk between peers in classrooms, (disputational, cumulative and exploratory) and three possible levels of analysis,

(linguistic, psychological, cultural). He developed the description of each talk category as follows:

(ibid, p104) 'Disputational talk is characterised by disagreement and individualised decision-making. There are few attempts to pool resources, or to offer constructive criticism of suggestions.' (ibid, p105) ' information is flaunted rather than shared, differences in opinion are stressed rather than resolved.'

(ibid, p104) '[In] cumulative talk ... speakers build positively but uncritically on what the other has said. Partners use talk to construct a "common knowledge" by accumulation. Cumulative discourse is characterised by repetitions, confirmations and elaborations.'

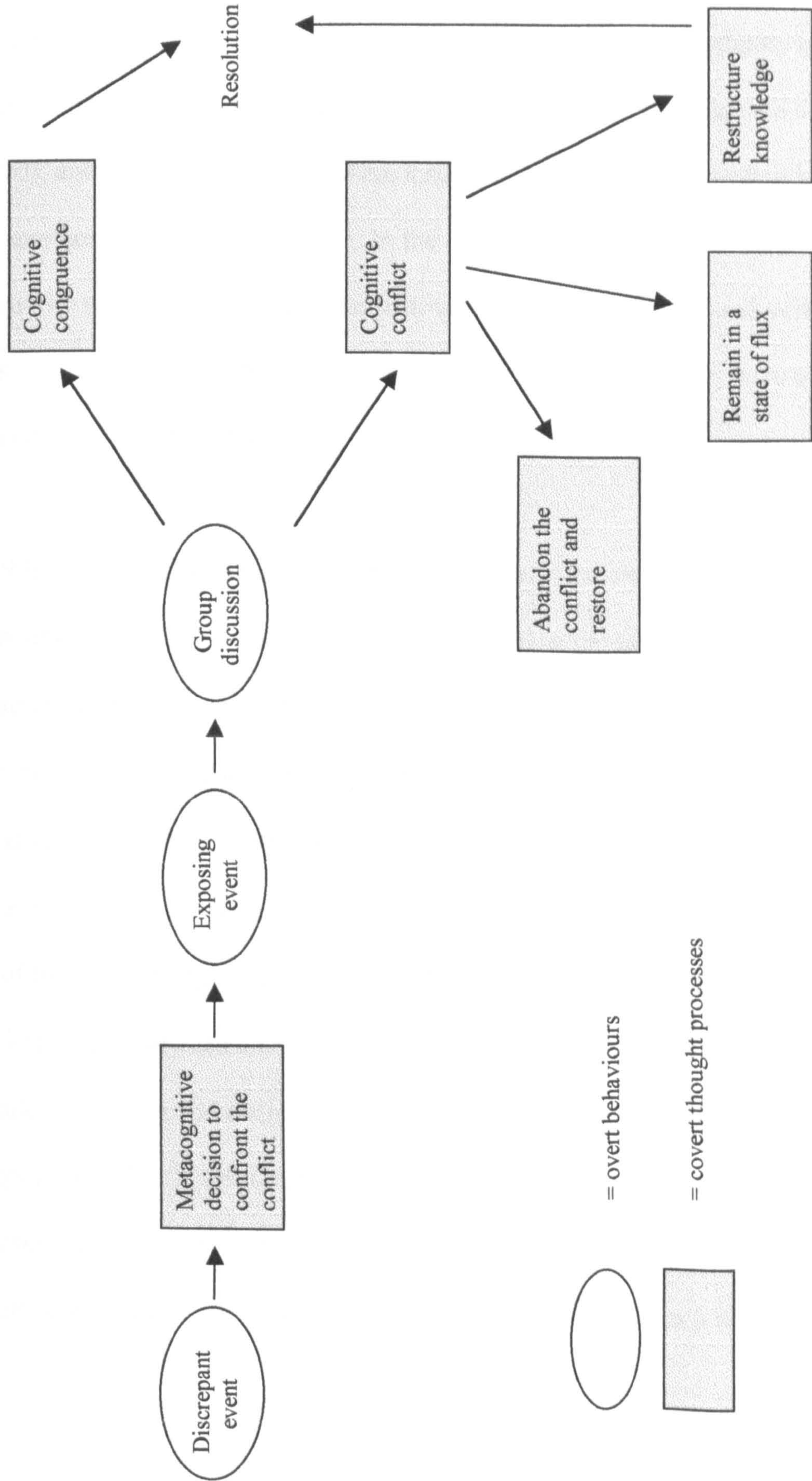
(ibid, p104) '[In] exploratory talk ... partners engage critically but constructively with each others' ideas Statements and suggestions are offered for joint consideration. ... Challenges are justified and alternative hypotheses are offered.'

He concludes: (ibid, p104) 'In exploratory talk *knowledge is made more publicly accountable and reasoning is more visible in the talk*. Progress then emerges from the eventual joint agreement reached.' (original italics)

The possible levels of analysis noted are linguistic, psychological and cultural, which deal, respectively, with the talk as: text (through an analysis of speech acts); thought and action, (including the visibility of reasoning); evidence of a discourse community (in this case with an analysis of the 'educated' discourse in use).

Almasi (1995) uses the term 'sociocognitive conflict' to describe (p317) 'the conflicts that emerge ... as readers encounter alternate interpretations or discourse that forces them to

Fig 2 A heuristic model of sociocognitive conceptual change in reading. Source: Almasi, J. (1995) 'The nature of fourth graders' sociocognitive conflicts in peer-led and teacher-led discussions of literature' *Reading Research Quarterly*, 30, 3 pp314-351



reconsider and update their own interpretations of text.' In fig. 2 she enlarges on this process, showing the possible courses for the group and individuals within it. Almasi (ibid, p341) found in her investigation that there were three types of sociocognitive conflict: 'conflicts within self, (where the subject was aware of some uncertainty in interpretation); conflicts with others (where differing interpretations are realised as they are articulated); and conflicts with text (where a subject advances a firm reading but is then told by another that it is not accurate).' In the peer-led groups she used, individuals engaged primarily with conflicts of the first sort, which, when recognised and verbalised, then led to a situation where '... through the interpretive community [they] were able to experience conceptual change through the resolution of their conflicts.'

Brewster (1999), in her exploration of gendered talk, recasts the classification of talk-types as four strategies (along 2 axes) and four styles (within the fields between the axes and thus describing the interaction of two strategies). This model contrasts, on the vertical axis, co-operating in and controlling the task, and, on the horizontal axis, impeding and facilitating task completion. It locates the four styles, compliant, collaborative, persuasive and coercive between these two axes. Each axis can therefore have either of two contrasting styles, eg task control can be either coercive or persuasive. Equally the interaction between any two axes can be through a style. Thus a coercive style links task controlling and completion impeding strategies and is equivalent to Mercer's disputative talk. A compliant style links task co-operation and completion impeding strategies and mirrors cumulative talk. But Brewster differs from Mercer in her division of exploratory talk into persuasive and collaborative styles, each with its own

characteristics of topic management and turn-taking and its own functions. This was partly due to gender roles. Key differences centred on the degree of skilled listening involved, the number and length of turns, the use of *I* and *we* in discussion and the use of disagreement.

Brewster found some specific gender-related differences during a jigsaw task with boys more likely to challenge (persuasive style) and girls more likely to collaborate. But along with other researchers (eg Jenkins and Cheshire 1991, Wareing 1994) she found that individual relationships as much as the sex of the participants determined the style of interaction. On the other hand Smith (1985) had catalogued a great deal of research that showed that men tended to interrupt much more often than women and have longer turns in paired mixed- and single-sex interaction. Webb and Kenderski (1985) found that this was true of males in high achieving groups, where they obtained more help with problems than they were prepared to give, and thus did better in the tasks. These studies would affirm Brewster's view of the gender link to the two types of exploratory talk. In this study the sample is not large enough to permit firm observations about gender, but the division of exploratory talk into two sections may be helpful for our analysis of the talk as a whole, and information about the sexes of the participants who utilise such talk types can at least be noted.

3.4.4. Linguistic / Discourse analysis

In addition to a more global judgement about the nature of talk within the groups, based on sequences of moves which demonstrate progress (or a lack of progress) towards meaning construction or decisions about the texts, we can gain further insights into the

group-based process by analysing individual moves by individual participants more closely.

Discourse analysis techniques provide a method for investigating such phenomena, but there is a need to draw from a very broad literature the most appropriate approaches. Coulthard (1977) and Schiffrin (1994) both give an overview of up to six different approaches to discourse analysis and the manner in which these approaches are appropriate to varying research traditions. Schiffrin also demonstrates that in certain situations these approaches can be used beyond their conventional arenas, noting for example that (ibid, p381), '... although speech act theory and Gricean pragmatics are two approaches to discourse that do not begin from the analysis of utterances, the insights that they offer for the analysis of a single utterance can be extended to multiple utterances.' Similarly, Blum-Kulka (1997) demonstrates a possible link between a speech-act based approach from pragmatics (encompassing Grice's work (1975) on maxims which systematise the process of inferencing and Brown and Levinson's (1987) work on politeness) and an ethnography of communication approach, eg Gumpertz (1982). She notes that this could enrich a detailed analysis by highlighting further the role of context. She also reports that Levinson (1983) finds that elements of conversation analysis can enrich a pragmatics based approach.

This study sought to discover specific aspects of the group interaction and therefore, like Blum-Kulka, found itself drawing from more than one discourse analysis practice. In terms of approaches to transcript analysis there appears to be a series of links between the work of Mercer (1995), of Barnes and Todd (1995), of Sinclair and Coulthard (1975), of

Levinson (1983) and of Schegloff (1991) or Pomerantz and Fehr (1997). The differences tend to involve the respective roles of general context and individual utterances. Each link moves the next writers further along a path from a more global towards a more microsociolinguistic approach. Similarly with each move the focus is more on the language of each utterance rather than on either discourse rules or on the context in which it is situated. The context remains important, but for example in conversation analysis it tends to emerge from the analysis of utterances and is justified by a microgenetic analysis of that language.

Thus in summary each pair of our writers has an overlap, but those at each end of the chain are rather different from one another. A brief summary of their approaches now follows.

Mercer, as stated earlier, was concerned with concrete examples of global descriptions of talk-types. For example in Mercer et al (1999) the researchers characterised exploratory talk mainly through the sense of the children's discussions, but also partly by monitoring the use of longer utterances. They also counted occurrences of the three key terms, *because*, *Do you agree ?*, and *I think*. Barnes and Todd (1995, p79) described five different aspects of the social and cognitive functions of conversation (on two levels, one of discourse moves and the logical processes associated with them, and the other of social skills, cognitive strategies and reflexivity). These categories were broken down into constituent elements and collaboration was defined through four types of moves (initiating, eliciting, extending and qualifying) plus additional elements of task management. The work is very strongly based on the dialogic model of conversation described by Bakhtin (1981), which considers all talk to be linked to what has preceded it

and what will follow it. And in addition this is also strongly commensurate with Mercer's three types of talk, so the links are strong between these researchers. The Barnes and Todd system perhaps most importantly allows the combination of 'content' and 'interaction frames' so that each piece of talk expresses a speaker's ideas and simultaneously signals her/his view of the relationships between the individuals involved in the talk.

The work of Sinclair and Coulthard (1975) in identifying typical classroom talk focused mainly on teacher-pupil interaction. But the authors also supplied a basic set of descriptors for discourse elements of a lesson which might be applicable to group talk as well, especially if in their group talk the learners are consciously or unconsciously imitating the style of language used by teachers. For example, the initiation, response, evaluation (IRE) pattern, consisting of teacher moves of questioning (first) and evaluating (third) with a pupil response in between, (eg Mehan, 1979), is well-known to modern language learners through patterns of teacher questioning in the whole class context, (eg Edwards and Westgate, 1994, p150). It would be interesting to discover whether any such pattern is present in unguided group talk. Certainly Barnes and Todd (1995) found in their research few examples of what they referred to (ibid, p161) as 'chairperson's moves' and stated further (ibid, p161) that 'where they did occur they often achieved premature closure...'. Clearly, if one participant is liable to use such a pattern it would be an indication of that person's own perception of role within the group. It might also indicate that others in the group shared that perception if s/he were allowed to function in that manner to a large extent. It could also demonstrate how far the group members saw that pattern of discourse as the norm.

But the IRE structure is only one aspect of the description of classroom discourse promoted by Sinclair and Coulthard (1975), and in using the definitions of acts, moves, exchanges and transactions we can examine the patterns of discussion prevalent in different groups. For example is one participant responsible for the majority of the framing moves ? Do most transactions involve all participants ? Do exchanges involve similar patterns of moves and acts ? The purpose of such analysis is to attempt to gauge how many of the participants are benefiting from the group context. Is it generally helpful and supportive of learning or does it only really offer extension to those socially equipped to input and extract from the process ?

Through conversation analysis techniques (Pomerantz and Fehr op cit, ten Have op cit) eg, sequence selection, the characterisation of actions, the analysis of timing and turn-taking we can gain a further view of the experience that subjects might have had during group work..

Finally, the work of social science 'philosophers' such as Billig (1987) is relevant for one aspect of the analysis. Within any group set-up dependent on discussion for an outcome, there will be both conscious and subconscious awareness of roles. 'In going about our everyday business, we need to possess the appropriate scripts for the scenes in which we find ourselves at any given moment.' (ibid, p13). This will involve at least superficial collaboration, but what can result from this is (ibid, p16) 'the deliberate suppression of argument', as the script may predominate over actual personal preferences. But if we use also a 'game' as well as a 'theatre' metaphor, we additionally need to take into account such phenomena as the general rules of the game and the existence of competition as well as collaboration. But (ibid p23) 'an argument, like a game, depends on a wide area of

agreement, in order for the disagreement to be aired.' All learners in school are playing a game and following a script to an extent. But a new situation such as an investigation of this sort may cause its participants to follow the usual rules or break them, or a combination of both. We need to consider the various implications of this when we try to generalise beyond the individual or group and the point in time of the investigation.

Thus Schiffrin's views expressed throughout Section 3 of her 1994 book, that it is healthy to locate and utilise both the common ground and the contrasting elements of the various approaches to discourse underpins the standpoint taken by this study. That is that we wish to describe as fully as possible the nature of the reading and the talk which occurred both in the individual and group sessions. Chapter Four looks at the methodology of that analysis in more detail.

Chapter Four

The Part Two Programme - Methodology

NB The processes described in sections 4.1 - 4.3 are summarised in the analysis flowchart in fig 3 on p.145

4.1 The methodology of the data gathering process

For the Part Two study it was felt more appropriate to investigate the reading behaviour of learners of a Western European foreign language. The use of Russian, and therefore of cyrillic script, in the Part One study brought an extra dimension, in that it added to the difficulty of recognition of familiar vocabulary items for a significant number of learners. The advantages of using French were that the language uses roman script, has a larger number of cognate items than Russian, and is learned by a greater number of school students in the UK than any other foreign language. Although there is no intention to generalise from a single class of learners, it was considered likely that the findings might be recognised as immediately relevant by a greater number of teachers, learners and researchers if French was the medium chosen.

It was also felt very important to expand on the role of the individual reading programme in the Part Two study in order that a more systematic comparison could be made between the two modes of operation. It was seen as important to maintain two of the original three task-types and to ensure some comparability of readability between all of the texts used (see section 4.1.3). The original second task-type (inferring the meaning from context of underlined words) was not used again, partly because it had already demonstrated its

particular qualities very fully in the Part One study, and partly because it was a completely unfamiliar task-type to most learners. In addition, there was an intention in the Part Two study to allow dictionary use in the form of a 'live' dictionary (the researcher), because this represents a more natural mode of reading in a foreign language. It was also felt that denying dictionary use completely might have demotivated some participants in the Part One study. The second task-type would not have been appropriate in that case, as it centred on the need to infer meaning without recourse to a dictionary.

4.1.1 The Subjects

The class used for the study was a Year Nine mixed-ability French group from a comprehensive school situated on the outskirts of Nottingham.

The school is described in a 1996 report (<http://www.ofsted.gov.uk/>) as 'fully comprehensive and [representing] a full range of attainment and socio-economic backgrounds' (there were then 12% of the school population entitled to free school meals). It states further that: 'The overall attainment of pupils on entry is skewed very slightly towards average and below average attainment.' There are currently (1999) approximately 1300 students on roll and of these some 8.7% are identified as having special needs at stages 1-5. 54% of the students in Year Eleven achieved 5 A*-C GCSE grades in 1999.

The same OFSTED report described attainment in the modern languages department as being 'in line or better than that expected nationally for their ages.' Additionally it comments: 'Pupils show that they have developed the receptive skills of listening and

reading and show good levels of understanding.' It also states that: 'Pupils are very willing to work in pairs and small groups.'

The class consisted of 30 students, one of whom was absent for the entire study. Of the 29 who took part, one attended only the pre-testing and training session. A further student missed the think-aloud session but took part in the group session. A further three students missed the group session. There are therefore 27 individual think-aloud protocols available and group interaction transcripts for 25 students. This gives a total of 24 students who completed all parts of the study.

The students had completed two years of French. The ability of the participants ranged from very able to below average ability, but no students in the class were statemented for special needs. Macmillan reading scores for this group from entry to the school two years previously ranged from 9.3 to 14+ with two subjects having no score available. The GAP tests (McLeod 1965) administered (see 4.1.2 below) produced reading age scores of 9 years 11 months and above on the Blue variant of the test and 10 years 11 months and above on the Red variant.

4.1.2 Pre-testing and Training

The preparation process for the study consisted of a lesson run by the researcher with the class teacher present in the week preceding the beginning of the Study. The full script outline of this lesson can be found in Appendix F. A summary only appears here. The lesson included three major elements: the administration of the GAP Test; training for the think-aloud process and an outline summary of the study. The GAP test was administered according to the instructions in its manual (Unwin 1970). In common with the

recommendations of Ericsson and Simon (1993) and van Someren, Barnard and Sandberg (1994) the think-aloud training began with simple mathematical calculations. The researcher used a procedure of modelling thinking aloud to a task and then inviting the class to try the same process with a parallel task. Feedback was then taken from as many students who wished to speak about the processes which had been involved. There were two initial stages in English with a progression from a simple calculation to a problem which required a reading and a subsequent calculation. The class members were then asked to do a short and simple reading task in French and to articulate (in English) the processes they used to decide on an answer. Finally they were shown a more complex but still short text in French with two true/false statements attached. Again feedback was taken on their reading processes. The researcher then outlined the scope and purpose of the study.

4.1.3 The Reading Tasks

Following the findings of the Part 1 Study it was decided to use two task-types in the Part 2 Study. The 'true-false-impossible to say' (TFI) task type was chosen because it is a common task used by foreign language teachers and course writers in the UK. The three-option variant (ie not just true-false) was selected as it made simple guessing a little less likely. By using some items which would require inferences to choose and account for the correct response, the task satisfied the cognitive level associated with problem-solving. The Four Penfriends (FP) task type was chosen because in the Part 1 Study it had seemed to motivate discussion and to cause a change in reading behaviour. It too

constituted a problem-solving task, as there was a need to consider evidence and make a decision.

Four new texts and tasks (two of each type) were created and checked for accuracy by a native speaker and teacher of French. Two centred broadly on the language topic of family and two on the language topic of free time activity. According to the American Council for the Teaching of Foreign Languages (ACTFL) Reading Proficiency Guidelines (1986) a mid-intermediate reader should be: 'Able to read consistently with increased understanding simple connected texts dealing with a variety of basic and social needs. Such texts are still linguistically noncomplex and have a clear underlying internal structure. They impart basic information about which the reader has to make minimal suppositions and to which the reader brings personal interest and/or knowledge.

Examples may include short, straightforward descriptions of persons, places, and things written for a wide audience.' The National Curriculum for England and Wales (1995) has the following descriptor for Level Four of its attainment target 3 (Reading and Responding): 'Pupils show understanding of short stories and factual texts, printed or clearly handwritten. They identify and note main points and some details. In their independent reading, in addition to using a bilingual dictionary or glossary, they are beginning to use context to deduce the meaning of unfamiliar language.'

The texts created matched both sets of criteria while still representing a challenge to the participants, (see below). They were also analysed for readability using the Flesch Reading Ease and Flesch-Kincaid grade level indicators (Flesch 1962). Although these instruments are designed for texts in English and may not produce reliable scores for French texts, it was possible through this process to establish an equality *between* the

texts not just in terms of total word score and words per sentence count, but in addition by establishing closely parallel scores in the two measures indicated above (see Appendix G). As stated above, the texts were deliberately pitched at a level which would challenge the subjects as this would be most likely to activate as many of the strategies available as possible, (Afflerbach 1990, Ericsson and Simon, 1993). Therefore some past tense verb forms were included and there was a greater use of conjunctions and adverbs than would often be the case in texts encountered by learners at this level. It was especially important in the group tasks to create materials which would not be automatically independently accessible to all. The context created was intended to be appropriate for the emergence for the group of a Vygotskian zone of proximal development (Vygotsky 1978), where the more able peers might enable the less able subjects to work at a higher level. For this reason the two texts which were felt (subjectively) to be the more challenging were chosen for the group interaction tasks.

The TFI task statements were deliberately varied so that a mixture of skills and strategies was demanded. Some items were more dependent on prior knowledge, some on cognates, some on directly parallel forms between task and text and some on a need to gain a more general understanding and make a judgement or an inference. A fuller discussion of this can be found in the analysis sections 5.2.2.1 and 5.2.3.1.

4.1.4 The Organisation of the Study

Following a study by Hockaday (1984) the subjects were initially grouped, predominantly combining two friendship pairs in a mixed-sex group where possible and creating mainly friendship-based all-female triads where numbers and the sex balance in

the class dictated this. One mixed sex triad of three able L1 readers was formed also on the basis that they were a friendship group in the normal classroom. This also allowed a test of whether such a group would emerge as more able in the FL reading context, as had been found in a History study by Bennett and Cass (1988). Otherwise the groups were of mixed ability (based on the GAP reading scores) as far as possible. In the event because of absences the groupings emerged as follows: 2 groups of 4 students (2 male 2 female) 3 triads (each consisting of 2 female and 1 male) 2 triads (all female) and 1 dyad (both female)

Of the 27 subjects who undertook the individual think-aloud tasks 14 completed them before the group tasks (4 days apart) and 13 after the group tasks (5 days apart). The programme started 6 days after the training session.

The think-aloud tasks were scheduled for 1 subject every 20 minutes during normal lesson times. These tasks were audio-recorded. The tapes used were C46 and no subject overran the 23-minute side. The Group tasks were scheduled for 30 minutes each during a French lesson and another lesson which either immediately preceded or followed it.

These tasks were both audio- and video-recorded.

While it is possible that some information about the content of the texts and tasks might have been communicated between subjects there is only very little evidence of this to be found in the transcripts. This issue was particularly relevant in the think-aloud protocols where on each of the two occasions a whole day was needed to complete the task with all subjects, and where therefore collaboration would have been more possible.

In both individual and group task sessions the subjects were told that the researcher would play the role of 'dictionary' if they decided they could not work out the meaning of

any word in any other way. This was done to facilitate the speed of working and to ensure that the researcher knew which words were needed.

The red dot method, (Olshavsky, 1976-7, Crain-Thoreson, Lippman and McLendon-Magnuson, 1997), where a dot was placed in the text at the end of each sentence to remind the reader to verbalise their thoughts, was not used. It was felt that this method would tend to dictate a linear approach to reading which had been shown in the Part One study to be alien to many of the subjects' normal reading behaviour. Part of the material under investigation was the approach to reading adopted for each of the tasks so as much freedom as possible was preferred. The researcher sometimes prompted the subjects to say what they were thinking / doing, because this is recommended as a method of maintaining verbalisation (Ericsson and Simon, 1993 pp256-7) but generally left them to speak as and when they wished. This was partly because some subjects opted to read silently rather than aloud and it was not always possible to know whether they were processing 'automatically' or considering meaning or interpretation. In addition, some exhibited signs of anxiety during the programme. An over-zealous pattern of reminders to speak might well have increased that anxiety and produced false results. It was as far as possible the subjects' *natural* reading behaviour that the investigation was seeking to discover. But this approach of course resulted in a number of the think-aloud protocols being comparatively sparse. In general terms, this might be a sign that the text was very easy, there was no comprehension issue, and as a result the text itself was read aloud rather than any cognitive processes being articulated, (Ericsson, 1988, Kintsch 1988). This tends then to make of the reading of such a level of text a bottom-up perception rather than a problem-solving process. Equally, however, the sparseness of the protocol

might indicate that the text was very difficult and that, although this could result in greater strategy use (Kintsch, 1988), the strategies available still did not bring any success in meaning construction. This would indicate that it was a language rather than a reading problem (Alderson, 1984). Additionally, it could also be that the subject possessed a lack of strategic awareness about reading, and so had no direction when material was even a little difficult to comprehend. The problem is not unknown in the field - cited by Ericsson and Simon (1993, p252) are Hayes and Simon (1974) and Bree (1969) who found that 'subjects reading text ... sometimes give rather scanty and uninformative thinking-aloud protocols'. The task completion statistics and the full transcripts would be able to offer an idea of why silence might occur, as comprehension would be certain to be revealed by the subjects' attainment in the tasks and rationale for their choices. The time when probes *were* used regularly was if subjects failed to offer a rationale for decisions on true-false statements or on the identity of the penfriend. The think-aloud literature (for example Ericsson and Simon, 1993, van Someren, Barnard and Sandberg, 1994, Pressley and Afflerbach 1995) stresses that probes should not ask subjects to account for their actions, but merely to report them. However, task completion was seen as a separate process from reading and understanding and such probes offered an important back-up method of discovering some elements of reading behaviour and strategy use which their thinking aloud did not reveal. The researcher did not respond to questions about the correctness of answers, but adopted an encouraging mode throughout so that each subject felt that her or his particular reading approach was valued.

4.2 The methodology of task performance analysis

4.2.1 Introduction

A certain amount of quantitative data could be obtained from listening to the recordings without full transcription. This was a useful first analysis base, which gave an overview of the performances of the whole group. Tables reporting on this first analysis of issues shown below in 4.2.2 to 4.2.4 can be found in Chapter 5, Section 5.2.

4.2.2 Individual think-aloud tasks

The success indicators on each task were of course the degree to which the subjects obtained the correct answers.

This was measured in the TFI task by a score out of eight. Importantly however, this was taken not just on the subject's response (ie saying that a statement was true or false or that it was impossible to say), but on the correct response being supported by a correct reason from the text.

On the FP task from a global total of five major sections, fourteen possible individual reasons were isolated which could have been given for the decision on who wrote the letter, and the number of these mentioned was recorded.

4.2.3 Group interaction tasks

The success indicators on each task were of course the same, ie the number of correct responses.

In the TFI task it was also a score out of eight, again dependent on the correct response being supported by a correct reason from the text.

On the FP task there were four major textual elements which could have been given for the decision, and the number of these mentioned was recorded. Equally, other potential authors could be excluded by a view, supported from text, that their particular characteristics were not present in the letter.

4.2.4 Other factors

The task performance data from both tasks can additionally be considered in relation to other items of statistical evidence: the number and type of vocabulary items requested by the subject(s) and the reading approach(es) to each task.

4.3 The methodology of transcript analysis

The performance data revealed a very broad variation among the participants in both individual and group contexts, and, for some individuals, *between* the individual and group contexts. This naturally raised questions about the reasons behind both variant and similar performances. How were successful and unsuccessful individual and group readings different, and were they successful or unsuccessful in the same ways ?

A series of methods of investigation was then implemented, summarised at the end of section 4.3 as fig 3. Each viewpoint contributed to a better understanding of the questions posed above. No individual mode of analysis gave a definitive answer, but the sum of the different findings was used to create a summary which is presented in Chapter Seven.

4.3.1 Presenting the transcripts

The transcripts were presented in the following ways:

All transcripts are arranged according to the principles of vertical running text (Edwards, 1993). Individual think-aloud protocols follow a pause-based / intonation unit method (Chafe 1993) where utterances are separated into chunks of language by pauses. Each chunk then occupies a separate line in the transcript, as does each separate turn taken by the subject and researcher. The lines are numbered consecutively throughout the task and additionally coded by the subject's allocated study number or an R for researcher. A pause longer than 1 second is recorded additionally on a separate line.

Group transcripts are separated into numbered and coded lines essentially by turn-taking. Thus pauses between utterances by the same subject are integrated into the line where the pause occurs, and therefore some utterances take up several lines of the page while remaining coded as a single line of the protocol.

For strategy coding purposes the transcripts were tabulated with separate columns for discourse and coding notes.

4.3.2 Analysis of strategy use

4.3.2.1 General reading strategies list

It appeared that there were three main options available for defining the major instrument to be used for strategy coding of selected transcripts. Firstly, it would be possible to build on the somewhat ad hoc methodology of the Part One Study where the transcripts were coded as they were analysed with codes being generated to describe the behaviours

observed. If this approach were used it would clearly be better to generate a list of reading strategies from a range of sources and to match these to what was found in the transcripts, thus to create a new instrument for this research. After trials (see sections 4.3.2.2 and 4.3.2.3) this option was rejected as it was felt that such an instrument should be extensively pre-tested to ensure reliability. A second option would have been to use one or more of the recognised instruments used in the field of learning strategies in second language learning, for example the *Strategy Inventory for Language Learning* (Oxford, 1990). This had the drawback that reading strategies were only a part of the whole and that the method of use of this instrument was primarily via a questionnaire. The third option, which emerged as the favoured method, was to take a coding system which was based on self report data (ie a relevant methodology) from the reading process. The system developed by Pressley and Afflerbach (1995) for L1 reading seemed very appropriate for this purpose being described by Cohen (1998 p35) as 'an excellent compendium of ideas for L2 researchers' and by McDonough (1999 p5) as having 'enormous implications for the theory of reading in a second language'. Its use would also give the facility to compare near-beginner foreign language reading with first language reading in terms of categories of strategies used and not used.

The full coding system which is edited from the very long original in Pressley and Afflerbach appears in Appendix H. The coding is separated into three or four broad sections (with Monitoring and Activation of cognitive processes both being part of the same section in the original), as in Table 4a.

Table 4a: Key to strategy coding in the transcripts

The code descriptions on the right are taken from Pressley and Afflerbach's work. The coding abbreviations on the left were created for the purposes of this study to allow efficient coding of the transcripts

MC	=	meaning construction, ie identifying and learning text content
MC/BR	=	meaning construction before reading
MC/DR	=	meaning construction during reading
MC/DRcim	=	MC/DR conscious inference making
MC/DRintg	=	MC/DR integrating different parts of text
MC/DRintp	=	MC/DR interpreting
MC/AR	=	meaning construction after reading
Mon	=	monitoring
MonTC	=	monitoring text characteristics
MonPT	=	monitoring processing of text
MonProb	=	monitoring problems
MonPost	=	post-reading monitoring / decisions to process additionally
Act	=	(monitoring and the stimulation of cognitive processing) activation of processes/processing
ActTask	=	activation of processes to accommodate text/task demands
ActDiffwp	=	activation of processing due to awareness of difficulties at the word / phrase level
ActDiffglob	=	activation of processing due to awareness of difficulties beyond the word / phrase level
E	=	evaluating
Emind	=	consistent evaluative mindsets
EStyle	=	evaluating the style of the text
ECont	=	evaluating the content of the text

The codes adopted on the left side of the table were entered into a separate column on the transcripts. This enabled them to be sorted by both code and subject to gain an awareness of the frequency of use of the various categories. The more precise identification of

exactly which strategy was in use was entered into a parallel column which was also capable of being sorted for analysis.

4.3.2.2 Predicted strategies specifically needed for the tasks

But before deciding to adopt the Pressley and Afflerbach coding system, it had seemed appropriate to take two initial steps towards coding. Firstly it was of course possible to take the specific texts and tasks used in the study, and to analyse what knowledge and strategies would be needed to make the decisions demanded by either task. It was also possible (using an extensive field knowledge of adolescent foreign language reading behaviour) to predict likely approaches to successful completion by the subjects. The full text of these predictions can be found in Appendix I. In addition to providing baseline of realistic expectations this stage in the process was also intended to supply an analysis of which strategies would be more likely to be in use or not. Clearly specific tasks demand certain approaches and there would be little point in commenting that a specific reading strategy was under-used if it were unlikely to be demanded by the specific tasks under investigation. Alongside this a set of further strategies which related to metacognitive and social actions (Wenden and Rubin, 1987, Oxford, 1990) was outlined to cater for strategic activity beyond simply processing. Thus the initial coding was in sections entitled Metacognitive, Cognitive and Social strategies.

The second step was to use a grounded approach. An initial set of transcripts, three think-aloud and three group transcripts were coded using the baseline set of criteria from Step 1 above and was expanded according to what appeared in the transcripts. These codings are shown in Appendix Ji. When a full list of strategies was eventually produced these were

cross-referenced to the Pressley and Afflerbach (P&A) system. On the very rare occasions where an action did not seem to fit at all with the system this was noted on the transcript. Clearly this might have affected the metacognitive and social categories since these do not appear as such in the P&A system. Almost all metacognitive actions were classified under the monitoring section. The social categories were only applicable to the group transcripts which would make direct comparisons more difficult. In the event again there were very few social strategies which could not be classified according to the P&A system, which increased the comparability of the two types of transcript. The cross-referencing process was extremely useful in two ways. Firstly it revealed different perspectives on meaning construction as the P&A system offered a greater number of elements, thus allowing more refinement in the initial coding. Secondly it allowed the L2 analogies of the P&A L1 notation to be clarified from real examples of strategy use.

4.3.2.3 Initial coding of transcripts (using self-generated strategy categories)

The sample set of transcripts were coded using a lay-out which created a table from the original transcript and allowed a coding or codings to be inserted next to each line. This was completed in a draft format which was handwritten. The Group transcripts naturally revealed a wider range of strategies because they were longer and involved three or four people as opposed to one.

For the group transcripts three were coded, in order of task success (ie A1 then B1 then B3) with each adding successively to the possible list of categories. The total occurrences for each category was then counted and for the TFI task the three groups were compared (see Appendix Jii)

Owing to the concerns mentioned above about the lack of any certain reliability in a self-generated coding system, the P&A system was then used to recode the Group B3 transcript. This appeared to offer a wider range of elements in the broad cognitive / meaning construction field, thus allowing further refinement of the categories available for analysis and perhaps therefore greater potential for insight into the processes involved. The table structure was then revised and the transcripts recoded (see next section for the selection process) to allow a category drawn from the P&A system and a further descriptor of the precise strategy used.

4.3.2.4 Selection of transcripts for analysis

Seven individual transcripts have been extensively analysed for their strategy use. These were chosen to exemplify several different issues, which are listed and explained briefly below

a) Individuals:

- The strategies used by an 'average' member of the class - Subject 7
- The strategic behaviour of the best individual performer compared with that of the worst - Subjects 14 and 27
- The strategies used by subjects who performed differently across the two tasks - Subject 3 (high L1 reading score) poor TFI performance, but the best FP score), and Subject 25 (low L1 reading score) good TFI performance but poor FP score
- The strategic behaviour of a subject with very high dictionary use - Subject 18 (who performed significantly better than his L1 reading scores might predict)

- Differences in strategy use between the individual and group contexts - Subject 10

b) Groups:

Four groups (A1, A4, B1, B3) were selected for strategy use analysis for the following reasons:

- Their respective performance in the two tasks as groups
- Their respective approaches to the tasks

The comparisons made centre on the performance of the groups in the two tasks and the strategic evidence behind the different degrees of success. This includes the total range of strategy use by individual subjects or the whole group and the reading approaches used and their effect on strategy use.

4.3.3 The analysis of group discourse

4.3.3.1 Introduction

As well as the analysis of specific strategy use by individuals and groups it is important to focus on the overall discourse of the transcripts. The purpose of this layer of analysis is to investigate the value of the group context for enhancing learner capability to make sense of text and to create an understanding which goes beyond that which could be managed by the individuals working alone. In doing this we must take into account several factors. These are, firstly, the varied contexts that are created when a teacher or researcher creates groups for a task, secondly, the varied purposes in addition to task completion that individuals within those groups may pursue and, finally, the varied power of individuals to enforce their purposes on their group.

Thus this approach can shed further light on reasons behind the relative success of some individuals or groups and on the reasons for the nature of the discourse which exists in group contexts. It allows the exploration of the role of talk types, and it can offer information on the interplay of ideas, the nature of turn-taking and the role-structure within groups. The section is divided into three broad approaches, the analysis of group talk type, the use of concordancing software to search for specific use of discourse and further discourse analysis techniques to analyse especially roles and decision-making. This multi-directional approach is underpinned by Potter and Wetherell (1994b, p48), who enumerate three features of discourse analysis as being 'particularly pertinent for its research practice'. These are (ibid, p48) the concern with 'talk and texts as social practices' (considering both content and form), a concern with 'action, construction and variability' (which examines the purposes behind choice of discourse), and thirdly a concern with the 'rhetorical or argumentative organisation of talk or texts' (how alternative versions might compete).

Groups A1 A4 B1 B3 were again analysed in order that fuller descriptions of their discourse could be assembled.

4.3.3.2 The nature of group talk

The discourse of the four groups was analysed for indications of exploratory, cumulative or disputational talk (Mercer 1995, Brewster 1999), and of the three types of sociocognitive conflict found by Almasi (1995). (See Chapter 3.4 for a full discussion of the issue of peer collaboration and Chapter 6.2 for the relevant data analysis). This was specifically investigated by examining the examples of decisions taken by the groups (ie

decisions about meaning, judgements on the TFI statements or on the author of the penfriend letter). In this way the nature of the talk used to determine the completion of tasks was matched to the descriptors given by Mercer, Brewster and Almasi, and further evidence was sought from the rest of the transcripts. This purely qualitative methodology was preferred to a further coding exercise at this point, as there is a greater number of problematic issues associated with coding the general interaction of a group discourse than with coding the specific strategic moves made by individuals within the group. This issue is discussed in detail by Mercer and Wegerif (1999), who review several studies which did use coding to describe the nature of group talk, and note with caution that some fail to present the original data, and do not discuss sufficiently the 'prior interpretative analysis that generated the codes' (ibid, p82). They also cite (ibid p82) Draper and Anderson (1991) who enumerate four problems coders encounter when dealing with language in use. These include the ambiguity of many utterances and the multiplicity of interpretation of many utterances. They also mention the need to see a chain of utterances as a whole rather than as a series of codable chunks and the incidence of meanings changing during talk. Potter and Wetherell (1994b) write about the issue of quantification with similar caution, noting the separate issues of reliability and validity. They also point out that readers of papers concerning discourse analysis will need to make *their own* discourse analysis of the discourse chosen to present the content of the article.

4.3.3.3 Concordancing as a discourse analysis tool

Through the investigation of strategy use and, in the case of groups, the nature of the group talk we can begin to see the interplay of ideas within and between individuals. The use of concordancing software to analyse the use of certain discourse markers can add to that picture by highlighting the frequency and the context for use of key terminology associated with thinking, negotiating or arguing skills. In terms of the individual transcripts it can be used to investigate the instances of ideas exploration which are articulated by the subjects, while in group talk it can offer insights also into the interaction of the ideas of individuals involved.

A list of discourse markers for problem-solving tasks was predicted and further items were added which would be relevant for group interaction, (see Appendix K). The transcripts were searched for these items but also via the analysis of a general frequency list for any other possible markers.

The program MonoConc (Athelstan) was used to achieve this. The resulting files were then edited, for example to remove utterances by the researcher and were analysed to establish whether individuals or particularly groups differ in the quality of their discourse.

4.3.3.4 Analysis of roles within groups, turn-taking and decision-making

The discourse and conversation analysis techniques described in Chapter 3, Section 3.4.4 were used in the following ways.

Following Mercer and Wegerif's (1999) view on the dangers inherent in using a quantitative approach to discourse coding (see above, section 4.3.3.2) it was decided to

utilise the selected analysis tools in a more qualitative (both descriptive and evaluative) way. Schiffrin (op cit, p360) differentiates the study of structure (a unit of language that is larger than a sentence) from function (the use of language for social, expressive and referential purposes). She concludes (ibid, p361) that neither radical structural nor radical functional analyses are appropriate, but that a combination of both will balance the strengths and weaknesses of each. Barnes and Todd's (1995, p79) analytical scheme on two levels, (the first of discourse moves and the logical processes associated with them, and the second of social skills, cognitive strategies and reflexivity) provided in its social domain cell descriptors regarding progress through the task, competition and conflict and supportive behaviour. Such descriptors allow one aspect of the shape of a section of transcript to be evaluated. But to make more sense about the boundaries of related utterances and of the individual utterances themselves we need to draw from Sinclair and Coulthard's (1975) work on acts, moves, exchanges and transactions. To be sure that we are using as much evidence as possible to ensure any inferences we make about intended purposes we need to consider Grice's (1975) maxims and Brown and Levinson's (1987) concept of 'politeness'. Specifically looking at a section of transcript with a view to determining when, by whom and how turn-taking occurs reflects a concern of conversation analysts such as Pomerantz and Fehr (1997).

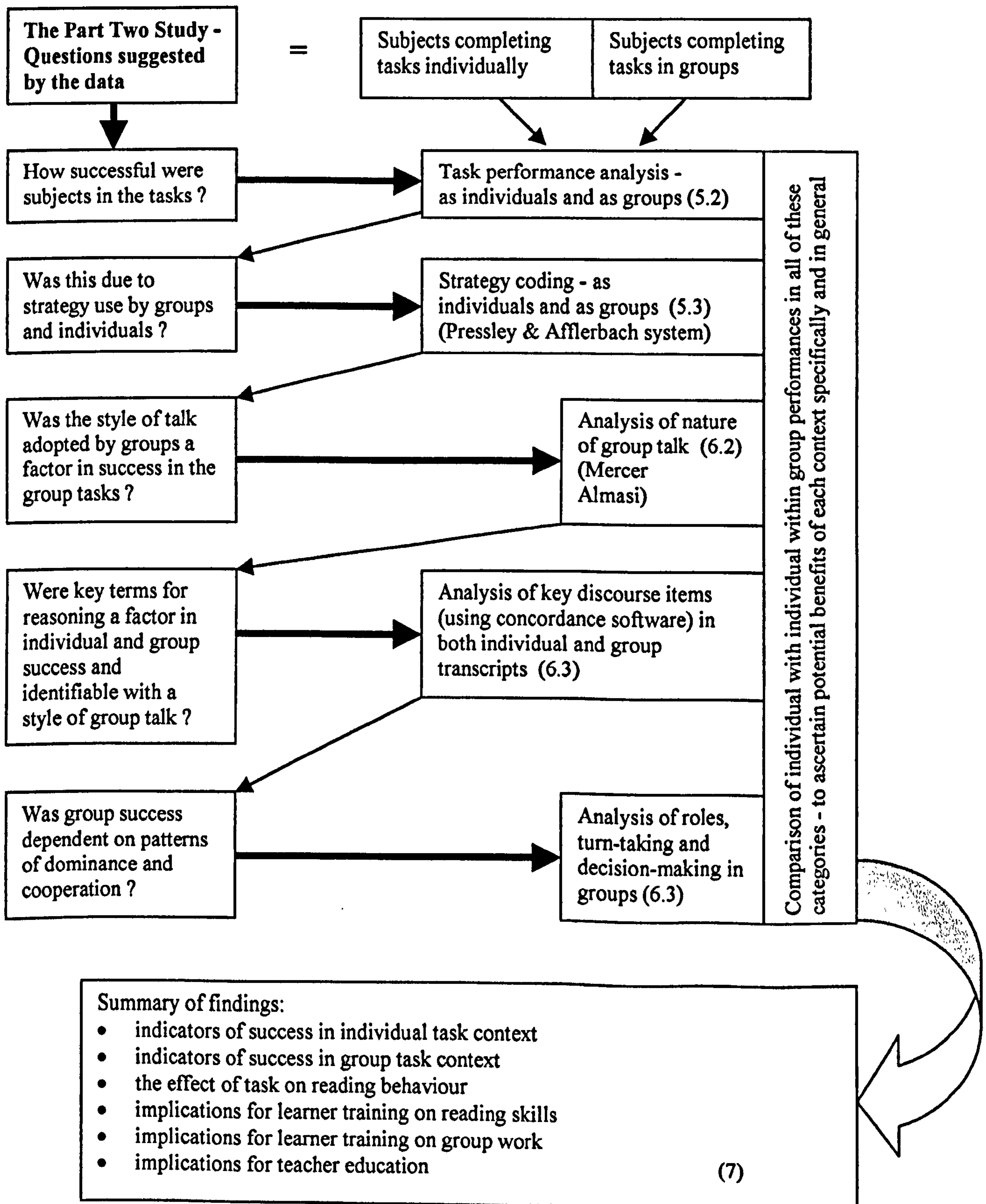
We should therefore select material and view it in turn from a functional and a structural viewpoint as we follow a logical investigation of the discourse contained in it. If we do this with a view to establishing facts about both the content and the context, we have a process which encompasses the approaches of both interactional sociolinguists (eg Goffman 1974) and conversation analysts. Thus Schiffrin's advice is heeded.

As a backdrop we will remember the points made by Billig (1987) about the conscious and subconscious awareness of role, using the metaphors of 'game' as well as a 'theatre'. Thus the ethos both of competition and of collaboration will shape the discourse, and this might change between participants in a group as the circumstances surrounding task completion alter and develop.

The intention then is to reveal specifically how the nature of the turn-taking and responsibility for decision-making within groups framed the success of this context on this occasion to offer learning opportunities to the various participants.

The four focus groups were again used for this analysis and the roles specifically of Subjects 7, 10, 16 and 3, and 19,26 and 28 were reviewed.

fig 3: Transcript analysis flowchart



4.4 The follow-up questionnaire

Six weeks after the programme the students in the class completed a brief follow-up questionnaire (Appendix L). This set out to discover the subjects' opinions on the issue of group-based versus individual reading contexts. It was felt to be important to establish whether the very positive views towards collaborative reading which had been expressed in the original Part One questionnaire would be repeated by a group which actually had a recent experience of the two modes of working with directly comparable tasks. To gain as rounded a view as possible in a single sheet questionnaire the preferences were elicited from the differing angles of enjoyment, achievement and the more neutral and 'objective' viewpoints of advantage and disadvantage summaries. It was felt that because there was some repetition between the categories, the responses would, overall, encompass a more considered view, as subjects revisited for the later more objective probes the attitudes they had signalled to the initial, more personal probes.

The questionnaire responses were analysed quantitatively by producing a summary of the preference figures and qualitatively through the use of a frequency table and concordancing on the most used items. Where there were discrepant or minority responses there was comparison between the response and the relevant transcripts to try to establish whether there was a clear rationale behind the response, eg an evidently less positive experience of the group context. Equally if a subject claimed to have performed better in one context this was investigated to see if the self-evaluation was accurate.

Chapter Five

The Part Two Programme - Summary of Data Analysis:

Task performance and Strategy Use

5.1 General Introduction

This chapter deals with the data as analysed according to the methodology outlined in Chapter Four for task performance and use of strategies. It presents results in various formats and summarises each element as it presents it. A more general discussion of the findings is located in Chapter Seven, which attempts to draw together the conclusions which can be made from individual data analysis sets.

The chapter consists of the following sections:

- 5.2 Task performance analysis
- 5.3 Strategy use analysis

5.2 Task performance analysis

5.2.1 Introduction

The first bank of data to be available for analysis was a global overview of the task performance by individuals and groups. This showed whether the class-members as a whole were generally successful in their reading of the four texts and their task completion. The data is primarily quantitative. It does not at this stage relate to the discourse of the transcripts, but to measurable performances such as the number of correct scores in the tasks, the number of vocabulary items requested by individuals and

the approach to the reading of each text by both individuals and groups. It is an important initial overview and contributes to the findings by allowing questions to be posed about the reasons for poor (or good) performance, the unevenness of performance by some individuals and any links between performance, approach and strategy use by the subjects. It is in three broad sections, (5.2.2 - 5.2.4) the first two of which focus respectively on individual think-aloud data and group interaction data. The third section looks briefly at comparisons between individuals' task performances and their reading scores. Sections 5.2.2. and 5.2.3 consist of three discussions: of the TFI ('true-false-impossible to say') task, then of the FP ('four penfriends') task, and finally of some comparisons of the approaches to the readings and of the nature and number of vocabulary item meanings requested over the two tasks.

5.2.2 The individual think aloud data

5.2.2.1 The 'True-False-Impossible to say' task

The summary of performance on this task is as follows:

From Table 5a it can be seen that 9 of the 27 subjects scored half marks or above. This demonstrates that the text was more challenging than their usual reading texts. There would seem to be a task-order effect, since a significantly greater proportion of A group members (6 of 13) scored half marks or above than B group members (3 of 14). The A group members scored slightly less than the B group members on the GAP tests (see section 5.2.4) but they worked as groups first and were therefore 'primed' to a certain extent for task style and level before they came to the individual tasks. Clearly, although

this may not have affected the very top performers (see section 5.2.4.2) the general effect was very marked. In addition, this would also lead us to expect a reverse effect on the group tasks, (see section 5.2.3).

Table 5a: Think-Aloud 'True-False-Impossible to say' scores achieved by learners, arranged in rank order (n=27)

Subject	Group	correct answer correct reason	correct answer incorrect reason	incorrect answer	no answer	correct answer for possibly correct reason*
14	A4	8	0	0	0	0
4	A3	6	0	2	0	0
1	B2	5	1	2	0	0
8	B3	5	0	3	0	0
18	A2	4	1	3	0	0
5	A3	4	0	4	0	0
9	B3	4	0	4	0	0
25	A2	4	0	4	0	0
26	A1	4	0	4	0	0
17	A1	3	5	0	0	0
7	B3	3	3	0	2	0
15	B1	3	2	0	3	0
10	A4	3	2	2	1	0
16	B1	3	2	2	1	0
28	A1	3	2	3	0	0
22	A2	3	1	4	0	0
23	A3	3	1	4	0	0
20	B2	3	0	3	2	0
13	B4	2	1	2	2	1 *
2	B1	2	3	3	0	0
12	B4	2	2	3	1	0
21	B2	2	2	4	0	0
6	B2	2	1	4	1	0
11	A4	2	0	3	3	0
3	B1	1	3	2	2	0
29	B4	1	2	2	3	0
27	A2	0	0	5	3	0
19	A1					
24						

* the reason given by Subject 13 for item 2 ('it doesn't say anything about ...') is not incorrect but does not give a textual reason for the response

An important implication for TFI tasks in general can be seen if we analyse simply the accuracy of the responses rather than the accuracy of the *reasons* given for the responses

(ie columns 3 and 4 added together). In this case we could believe that a further 12 met the half-marks standard. Since this represents an almost exact reversal of the true achievement with either one third or two thirds succeeding at this level depending on the accuracy of the response monitoring, teachers and researchers should be encouraged to design tasks so that monitoring is in-built.

Table 5b: Range of scores for each Think-Aloud True-False item, rank order

Item	correct answer correct reason	correct answer incorrect reason	incorrect answer	no answer	correct answer for possibly correct reason
2	25	0	1	0	1
6	15	2	8	2	0
5	14	4	7	2	0
8	10	1	12	4	0
1	7	13	7	0	0
3	7	6	8	6	0
7	4	5	13	5	0
4	3	3	16	5	0

Table 5b demonstrates that there was a very broad range of response accuracy to this task. One item (2) was found to be significantly easier because it depended on world knowledge (of pop culture) and language which was familiar and cognate. The next most accessible items were 6 and 5. These items had in common that they linked simple sentence task statements with accessible and broadly parallel text statements. Since a major strategy used by foreign language learners while reading is to look for such parallel readings in task and text, it is not surprising that such items were found to be easier. With item five this is especially interesting since 78% and 63% of subjects asked respectively for the key words *boucherie* and *rayon* in the task statement and 37% asked for the key word *boulangerie* in the text. This demonstrates two things: that the strategy of looking for parallel forms provides a stimulus to identify a key word for understanding, and that the majority of subjects were task led since more requested the task statement vocabulary

than the item from the text. If task completion rather than text comprehension is the goal, it will not matter what the final (and different) element in the jigsaw means because readers already know that *because* it is different the answer is false ! Item 8 was correctly answered by 37% of subjects, and this item, although linguistically accessible, also demanded that readers remembered an anaphoric reference (that *le soir* here referred to *dimanche soir*) or that they knew to return to the beginning of the text for further evidence. Four other items (1,3,4 and 7) were correctly answered only by 26% or fewer of the cohort. These items had in common less accessible task statements each with some unfamiliar vocabulary and an absence of the more directly parallel forms found in items 5,6 and 8.

These findings provide a strong case for more detailed investigation of strategy use and reasoning in the transcript analysis

5.2.2.2 The Four Penfriends task

The summary of performance on this task is as follows:

Tables 5c and 5d show how many times each of the possible 'outrageous lie' items were mentioned by subjects. In Table 5c these are given in rank order, and it is clear that there are just two 'popular' items (ie mentioned by around 50% or more of the subjects). These items seem to have two factors in common. They are both centred on cognates (*problèmes, parlement européen, président de la France*), they both stand in a sentence which clearly starts with the family member and name. In other words syntactically and lexically they are the most accessible.

Of the 12 subjects who mentioned between 1 and 3 items, all mentioned the fact that the father was the president of France, and a further 9 also mentioned the mother's problems with the European parliament. It is possible that information such as this might perhaps have been 'passed down the line' by subjects. Nevertheless all of those who located these items had to read to the end of the text as they appear in the final paragraph, and it can be seen from the transcripts that the items are actually located before being mentioned. The items concerning the writer's brother were the next most mentioned (see Table 5d) with again a probable reason being the cognates *foot*, *nationale*, *célèbre* and *télévision* being highlighted together with the reappearance of the word *équipe* which was featured in the TFI task. The least accessible paragraph emerged as the first, which concerned the older sister. Only 8 subjects requested the word *chante* and no-one appeared to know it already. Since recognition of this word would to an extent unlock the meaning of the rest of the paragraph it was crucial that it was seen as a key word. For a discussion of the amount and types of vocabulary which were requested see section 5.2.2.3. The family members' first names were part of the 'lie structure' as they referred to four well-known French-speaking personalities plus Charlotte (Church). No subject realised the significance of the first names chosen taken together with the descriptions and although in the case of Edith Cresson and Jacques Chirac this is perhaps not surprising, the failure to spot the links at least to Céline Dion and David Ginola is interesting.

Table 5c: Number of mentions of each item as rationale for choice of penfriend author of letter, ranked

Item / Reason	No.
Father is French president	22
Mother in European Parliament	13
Brother international footballer	9
Is famous / on television	8
She is popular in Europe	7
Now plays in London	5
Sister is beautiful	3
Only 12 yrs old	2
Has recorded film title song	1 + ?
Younger sister is singer	1
Sisters are famous	1
Sings in French and English	0
Has made records	0
Names of family members	0

Table 5d: Number of mentions of each family member / paragraph as rationale for choice of penfriend author of letter, (in text order).

Paragraph / Family member	number
1. Older Sister	5 + ?
2. Brother	22
3. Younger sister	11
4. Names	0
5. Parents	35

Table 5e shows the number of items mentioned by each subject in rank order. As can be seen 5 subjects actually failed to choose the correct author of the letter, not realising that it contained any outrageous lies, with most of these preferring option 3, the writer of 'ordinary letters'. But the large number who mentioned only a few items reveals again that there was only a partial understanding of the text by the majority of subjects. This is not surprising as the task type was unfamiliar and it is possible that the reading approach that needed to be adopted (see next section) was also abnormal. On this task just 4 of the top ten scoring subjects were from A groups. This may be because the majority of the most

successful readers (from the L1 reading tests) were in B groups, since Group B3 deliberately contained three of them. The FP task, when compared with the TFI task, demands more linear reading, the creation of macrostructures and a greater integration of textual elements. A closer analysis of the transcripts of two subjects (subjects 3 and 25) who scored very differently between the two tasks transcripts is offered in section 5.3.

Table 5e: Number of items mentioned as rationale for choice of penfriend author, ranked

Learner No.	Items mentioned
3	8
2	6
14	5
22	4 + ?
4	4
8	4
9	4
12	4
13	4
18	4
7	3
15	3
17	3
26	3
1	2
10	2
20	2
21	2
28	2
11	1
27	1
29	1
5	0
6	0
16	0
23	0
25	0
19	
24	

5.2.2.3 Comparison factors between the individual TFI and FP tasks

It is worth also comparing the relative performance across both tasks by individual subjects, (Tables 5a and 5e). Five of the top ten performers on the FP task appear in the top nine of the TFI task. In this sense there is an evenness in reading performance across the two tasks. But individuals do also differ. For example the top performer on the FP task was rated joint 25th on the TFI task . A further three subjects (2,12,13) who appear in the top ten in Table 5e scored only 2 of 8 correct answers in the TFI task. Conversely two of the top nine in Table 5a failed to identify the author of the letter in the FP task. This is not indicative of any pattern in itself but makes transcript analysis more purposeful if such anomalies are subsequently investigated.

The reading approaches used by subjects across the two tasks need also to be compared. Table 5f shows that a total of 9 reading approaches were identified of which seven were used at some point by subjects. The clearest contrast occurs with the first and last named approaches (Read whole text silently and Read TFI task statements or pen-friend descriptors individually and seek answers from text). This highlights a fundamental difference in reading approaches, ie whether a reader reads initially in order to complete a task or for more general text comprehension. Fifteen subjects, including three of the six top readers who completed the individual tasks, used a task-led approach on the TFI task, eleven of these exclusively. Only one used the same approach to the FP task and this in combination with another approach. The number of subjects who read the whole text silently rose from 4 to 10 between the two tasks and there was also a slight increase in the number who translated the whole text into English

Table 5f: Approaching the reading of the texts, individual think-aloud tasks

NB: for *each task type* the left-hand column denotes the total number of subjects who used this approach.

The right-hand column denotes the number of learners who used this approach exclusively (ie *not* in combination with another approach)

Approaches used	On 'true-false-impossible to say' task		On four penfriends task	
Read whole text silently	4	3	10	10
Read whole text aloud	2	2	2	2
Read whole text in French, translating sections (ie words, phrases, sentences) at a time	0	0	0	0
Translate whole text into English	3	2	5	5
Read sections of text silently	6	2	5	4
Read sections of text aloud	0	0	0	0
Read sections of text in French, translating sections (ie words, phrases, sentences) at a time	2	1	2	1
Translate sections of text into English	1	0	5	2
Read TFI task statements or pen-friend descriptors individually and seek answers from text	15	11	1	0

Another important area for comparison was the number of words requested by subjects in and across the two tasks. Table 5g allows various comparisons to be made. We can look at the task performance and vocabulary requested for individuals and compare between subjects for each task type. We can also compare the same factors across task types.

The mean number of items requested was 11.26 for the TFI task and 5.48 for the FP task.

The actual number of items requested ranged from 4 to 39 in the TFI task and from 0 to 27 in the FP task. But the great majority of the subjects fell within a much smaller range

(of 4-14 items in TFI and between 0 and 9 in FP). Three subjects (16,18,26) requested significantly more 'dictionary' help during the TFI task (24, 25, and 39 items respectively), and of these both 18 and 26 were also more demanding during the FP task (18 and 27 items respectively). The three most successful subjects on the TFI task asked for 13, 9 and 11 items respectively (ie very close to the mean) while in the FP task the three best asked for 9,3 and 4 items respectively. It does not appear therefore that there is any correlation between seeking dictionary help and being highly successful in the task. Knowledge of vocabulary does not always equate to meaning construction (Oakhill and Cain 1997) and so not all of the TFI tasks were unlocked in this way. Clearly, though, individual performances will have been enhanced to some extent by gaining speedy access to unknown lexis.

The nature of the words requested is also noteworthy. The most requested items in the TFI task principally appeared in the task statements and were mainly nouns, adjectives or adverbs. *Fume, sentent and commencé* were requested by a comparatively large number, but other verbs featured much less in this respect. In the FP task *chante* was the only verb to be requested by as many of 30% of the participants even though nobody actually appeared to know it. Singleton (1997, p217) cites Rodgers (1969), Phillips (1981) and Ellis and Beaton (1993, 1995) as demonstrating that learners find verbs more difficult to remember. Here we can go further and surmise that many of the participants in the study found it difficult even to engage with the verbs in the text, preferring to focus on the noun-content words

It is interesting to note that for all but 3 of the subjects the number of words requested fell between the TFI and the FP tasks. The most likely explanation for this is that in the first

task it was necessary to focus on and comprehend a body of language (the task statements) in order to carry out the instruction. (Furthermore, these are additional to the text word count). The scanning process which was then instigated led to the text being searched for appropriate chunks of text and although perhaps less of the text was read overall, more was deciphered in a formulaic way, which would involve clarifying vocabulary. In contrast, the FP task demanded a more linear and perhaps fuller reading of the text, but the emphasis was more on skimming for some kind of gist understanding and scanning for anything which might very obviously indicate an author. There was no specific guidance as to which words needed to be understood - the reader determined how much comprehension s/he wanted to gain. It might be that an instruction to give at least a specific number of reasons for choice of author would cause greater focus.

5.2.3 The group interaction data

The performance data from the group tasks is presented below in table 5h. The seven categories selected for comparison (Approach to TFI text and task, Success in the TFI group task, Mean score and range on *individual* TFI task, Approach to FP text and task, Success in the FP task, Success in the *individual* FP task, Structure / Ethos of the groupwork) allow various comparisons to be made. The approaches to each task can be compared across all eight groups, as can the approaches between the two tasks made by each group. Additionally it is possible to examine any differences between the approaches made to the tasks by individuals with those adopted by the groups. The performance scores for each task completed as a group can be compared with the

equivalent scores from the individual tasks. Finally the A group performances can be compared with the B group performances since each cluster attempted the individual and group tasks in a different order. Again, as with the individual performance data, issues for closer investigation from the transcripts can then be identified.

5.2.3.1 The TFI task

As for the A Groups in the individual TFI tasks there was a noticeable gain in performance by the B groups who had already completed their individual TFI tasks before coming to the groups. Whereas there was little if any gain between the mean individual performances and the group score for the A groups there was dramatic improvement on mean individual scores by all four B groups. This demonstrates that students can quickly become accustomed to a more difficult level of both text and task, especially if a group-based structure is used at some point to multiply possible approaches. The five 'highest achievers' over both individual tasks (see section 5.2.5.2) were spread between four groups, one each in A2, A3 and A4 and two in B3, with four other high scorers on the individual TFI task alone coming from groups A1, A2, A3 and B2. In the B groups it did tend to be the higher scoring individuals who performed some important role within the discussion (although this does not mean they had the highest number of turns), but in the A groups this was not the case at all. The most graphic example of this was Group A4 where the best overall individual performer (Subject 14) played in certain respects a very subdued role in the group. (See sections 5.3 and also 6.2 and 6.3 for more analysis of this case).

Table 5h Group-based tasks compared:

1.) A groups

	Group A1 (2M 2F)	Group A2 (1M 2F)	Group A3 (2F)	Group A4 (3F)
Approach to text and task (TFI)	Read / translate text, then turn to task	Task-led, very little detailed translation	Task-led, self-sufficient, least words requested	Read / translate text, then turn to task, then repeat task responses
Success in the task*	$\frac{1-3}{8}$	$\frac{3-4}{8}$	$\frac{5}{8}$	$\frac{3-6}{8}$ #
Mean score on TA individual TFI task	3.33 / 8 (range 3-4)	2.67 / 8 NB 2 subjects scored 4, 1 scored 0 (& had very little input to group)	5 / 8 (range 4-6)	4 / 8 (range 2-8, with 2 subjects scoring 2 and 1 subject 8)
Approach to text and task (FP)	Translate first para, make a guess, and then continue to translate to verify	Silent reading, then translation of random elements	Some linear translation before decision made	Limited reading / translation before decision made. Subject 14 encourages closer reading to verify
Success in the task	Correct author chosen No concrete textual reasons 'trying to explain something in a weird way'	Incorrect author chosen Gist of letter summarised to give rationale for Sasha 3	Subject 5 chose correct author (prices). Subject 4 disagreed and eventually Subject 5 changed view	Incorrect author chosen Some uncertainty between Sashas 1 and 3
Success in the individual FP task	All identified author 3,3 and 2 reasons given	18/27 identified correct author and gave 4 and 1 reason	4 identified correct author and gave 4 reasons	All identified author. 1,2 and 5 reasons given
Structure / Ethos	Two members (f) work extensively together or individually (seated together) with two other members (m) (seated opposite each other) contributing far fewer turns. Tone co-operative.	Members seated separately. Subjects 18 and 25 contribute equally, with Subject 27 saying very little throughout. Tone co-operative	A dyad, which worked closely, looking frequently at task sheets together. Dialogue often clipped and reasoning less articulated because understanding was good.	Two of three members working closely together, (due to seating pattern chosen) but all three co-operating throughout

* Two items were intended to be answered as 'impossible to say', but were interpreted as false by three groups (Statement 8) and by two groups (Statement 5) This response was perhaps not completely incorrect (as would a True response have been). The answers are therefore noted as part of the range of correct responses given. In each case the argument utilised demonstrated that the *thought processes* were not erroneous.

Group A4 gave their answers twice. On the first occasion their answer to statement 2 was correct, on the second it was incorrect

Table 5h Group-based tasks compared:

2.) B groups

	Group B1 (2M 2F)	Group B2 (1M 2F)	Group B3 (1M 2F)	Group B4 (3F)
Approach to text and task (TFI)	Start by reading / translating text, but very soon become task-led	Start by reading text silently (2 mins approx) then task-led	Read aloud / translate majority of text before task is started	Task-led
Success in the task*	$\frac{5-7}{8}$	$\frac{5-6}{8}$	$\frac{7-8 \#}{8}$	$\frac{7}{8}$
Mean score on TA individual TFI task	2.25 / 8 (range 1-3)	3.33 / 8 (range 2-5)	4 / 8 (range 3-5)	2 / 8 (range 1-3)
Approach to text and task (FP)	Reading approached differently by different subjects. 1st suggestion as to author made early and then discussion accompanies reading	Start by reading text silently (2 mins approx) then begin task with initial decision on author	Reading aloud / translation interrupted by Subject 7's wish to make an early decision. Translation less linear thereafter	Some silent reading followed by some translation
Success in the task	Incorrect author chosen Rationale given for Sashas 1 & 3	Incorrect author chosen Rationale given for Sasha 3 (with mention of Sasha 1)	Incorrect author chosen Rationale given for Sasha 3 by excluding other 3 authors. Hurried by Subject 7, but the other 2 agree	Incorrect author chosen Rationale given for Sasha 3 by excluding other 3 authors
Success in the individual FP task	All identified author 3,6,8 reasons given	1 / 20 identified author 2 reasons given	All identified author 3,4,4 reasons given	All identified author 4,4,1 reasons given
Structure / Ethos	Two (m) members seated opposite each other 'lose battle' with 2 (f) members seated side by side over text / task approach. Little overt co-operation	Three members seated separately (m in centre). Subject 1 had leading organising role and spoke most. Subject 20 (m) spoke little but to good purpose. Tone co-operative	Three members worked very co-operatively for TFI task, but a gender separation occurred for FP task with Subject 7 working more alone. Subject 8 was prime organiser	Three members sometimes worked separately (though in chorus) Subject 29 (in centre) became more involved as session progressed. Subjects 12 and 13 had equal share and showed some competition

* Two items were intended to be answered as 'impossible to say', but were interpreted as false by B1 (Statement 8) and by B1 and B2 (Statement 5) This response was perhaps not completely incorrect (as would a True response have been). The answers are therefore noted as part of the range of correct responses given. In each case the argument utilised demonstrated that the *thought processes* were not erroneous.

One item (6) correctly identified by Subject 8, but it is unclear whether this was accepted by the other two subjects who appeared to disagree

5.2.3.2 The FP task

The Four Penfriends task for the groups was clearly found to be more difficult than was the individual version. This was true whether the individual tasks preceded or succeeded the group tasks. Although the language was of a comparable standard, the sentence structure was perhaps more challenging because of the characteristics of the author. Sasha Four is the penfriend who writes in puzzles. In the text there are three major sections where she defines something rather than naming it (her preference for ice-hockey as a sport, her favourite composer and her dilemma over which CDs to buy.). One group, ironically the first to tackle the task, seized on the difficulty of the language as a clue and without being able to define precisely why, decided correctly that it was Sasha Four. One of the pair to which Group A3 was reduced owing to absences, also correctly identified the author, giving the CD section as her reason, but her partner disagreed and eventually persuaded her to alter her view. All the other six groups chose the wrong penfriend, although they all found the key passages very difficult. Here the content involved in this version of the task-type, although notionally the same as in the individual stage, proved to be different enough to become apparently impossibly difficult. This is shown by the fact that 11 of the 13 subjects involved in the B-groups had correctly identified the penfriend in the individual mode, when the task-type was entirely new to them.

5.2.3.3 Comparison factors between the individual TFI and FP tasks

As with the individual tasks there was a tendency to be task-led on the TFI task and to be more linear in the approach to the FP task. Three of the eight groups read and translated

the whole TFI text before starting the task with one other group taking 2-3 minutes to read it individually and silently before turning to the task. Three groups were task-led from the outset with group B1 beginning by translating and then turning to a task-led approach under pressure from one of the members. (See section 5.4.2 for an extensive analysis). This compares very closely with the approaches to the TFI task on the part of individuals. For the most part the same individuals adopted this approach in both modes of working, although it is noticeable that Subjects 8 and 9 who were task-led for the TFI task in their individual protocols were keen to read and translate the text first in their group. Subject 8 (the main organiser) had opted to translate the FP text in her individual session, and it may be that this approach had appealed to her as a result. Or it may be that they realised as the most successfully collaborative group that joint meaning construction was of benefit and somehow more democratic. The third member of the group (Subject 7) was a full participant in this process for the TFI task, but seemingly lost patience with it in the FP task, and very quickly suggested a possible penfriend. Whether further translation would have unlocked the correct author for the group cannot be known.

The amount of vocabulary assistance sought again varied between groups as it had with individuals. Groups A1, A2, A4 and B4 asked for between 12 and 20 items per task whereas the other four groups asked for less than 10 per task. (B3 asked for only 9 in total). Again this may show that the B group members were more accustomed to constructing meaning without asking for words because they knew the task format already. There were comparatively few examples of direct vocabulary enquiry between group members, although where material was being translated it was naturally the

subjects who knew the vocabulary or could guess it who took the lead at any one point. Thus the others were in effect getting any vocabulary queries answered without having to ask for them. Group A4 was perhaps the most interesting in this respect because Subject 14 who was the most successful student overall in the individual tasks did not take a very noticeable lead in the group. However all vocabulary queries to the researcher were directed through her, perhaps because the other two members waited to see if she knew first. In two other groups it was almost exclusively one member who asked for the words, in the others it was more open. The individuals who asked for most vocabulary in the individual sessions again featured here with Subject 26 in Group A1, Subject 18 in Group A2 and Subject 13 in Group B4 asking for the majority of words in each case. The exception was Subject 16 in group B1 who played the smallest role in the group but then asked for 24 words in her own TFI task.

Again the amount of vocabulary requested did not correlate with success in the tasks - the most successful group in the TFI task was B3 who requested the smallest number of words.

5.2.4 Comparisons between reading scores and task performance

5.2.4.1 The reliability of the reading test used

Table 5ji summarises reading test and task performance data about all 29 subjects. The reading scores provide one comparison point with task performance, but should not be regarded as in any way definitive about potential performance. There were apparently significant anomalies with the GAP test, and particularly with the Blue version. (For this

Table 5ji: School R. data-set for Reading in French tasks, October 8-20 1999

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
97 McM	12.3	13.3	14+	14.0	10.9		12.9	14+	12.9	9.3	11.6	12.3	13.3	12.9	11.6	9.6	11.6	12.9	11.6	10.9	12.9	12.6	9.3	9.3	12.9		12.3	9.3	
99 GAP	R34	B28	R38	B30	R31	B20	R32	B37	R37	B29	R34	B27	R36	B37	R36	B29	R27	B29	R40	B30	R35	B28	R27	B26	R30	B22	R35	B22	R32
Group	B2	B1	B1	A3	A3	B2	B3	B3	B3	A4	A4	B4	B4	A4	B1	B1	A1	A2	A1	B2	B2	A2	A3		A2	A1	A2	A1	B4
Th. Aloud	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Group	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gender	F	M	M	F	F	F	M	F	F	F	F	F	F	F	F	F	M	M	M	M	M	M	M	M	F	F	F	F	F

TFI individual task

Items ✓	5	2	1	6	4	2	3	5	4	2+	2	2	8	3	3	4	4	3	3	2	2	3	3	4	4	0	3	1
Voc req	11	11	10	9	6	9	12	9	8	4	9	7	13	8	24	10	25	6	6	9	8	6	13	39	14	4	7	
Voc clar	2	2	0	2	0	1	3	1	2	1	2	1	3	2	3	2	3	1	1	1	0	0	3	5	1	1	0	
Voc giv	0	0	0	0	0	0	0	0	2	0	1	0	0	0	1	0	0	0	0	0	0	0	2	0	1	0	0	

FP individual task

Items noted	2	6	8	4	0	0	3	4	4	2	1	4	4	5	3	0	3	4	2	2	2	4+	0	3	1	2	1	
Voc req	3	3	9	0	1	0	3	8	5	1	4	9	16	4	6	5	2	18	2	3	0	?	0	0	27	5	7	4
Voc clar	0	5	0	0	0	1	0	3	0	0	1	2	4	1	0	2	0	5	0	0	0	0	0	1	5	0	3	0
Voc giv	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Table 5jii: School R. data-set for Reading in French tasks, October 8-20 1999

In top 12 RS'97	✓	✓	✓	✓			✓	✓	✓				✓	✓				✓	✓		✓	✓			✓				
In top 13 GAP	✓		✓				✓	✓	✓	✓	✓		✓	✓	✓			✓	✓	✓					✓				✓
In top 9 of TFI	✓			✓				✓	✓				✓	✓											✓				
In top 10 of FP	✓	✓	✓	✓				✓	✓			✓	✓	✓								✓							
97 McM	12.3	13.3	14+	14.0	10.9		12.9	14+	12.9	9.3	11.6	12.3	13.3	12.9	11.6	9.6	11.6	12.9	11.6	10.9	12.9	12.6	9.3	9.3	12.9		12.3	9.3	
99 GAP	R34	B28	R38	B30	R31	B20	R32	B37	R37	B29	R34	B27	R36	B37	R36	B29	R27	B29	R40	B30	R35	B28	R27	B26	R30	B22	R35	B22	R32

reason the 1997 reading scores have remained in the table as a further indicator.) For example two subjects with 1997 reading ages of 12.9 and 12.3 apparently emerged from the GAP test (two years later) with ages of 10.2. Even on the Red test one subject with a 1997 reading age of 12.6 came out with one of 10.11.

In table 5jii we see data about the 'good readers' compared. The top twelve scores on the 1997 tests include all subjects with a reading age at that point of 12.6+. (The equivalent range from the GAP test includes the top thirteen subjects.) Of these 12 subjects, 7 appear in the top 13 of the GAP test ratings, with 4 of the missing 5 having taken the Blue test version. The biggest anomaly here is clearly Subject 4 who with a 1997 score of 14+ failed to make the top thirteen of the GAP test, yet was one of five subjects to score highly on both individual tasks.

5.2.4.2 Task performance and reading scores compared

Five subjects (4F, 1M) scored within the top third of the class on both individual tasks. Three of these were from A groups and two from B groups so at this level the task order effect seems to be less marked. Of the five, three appear in the top band for the two reading tests, and one (Subject 4) has been identified as a GAP test anomaly. The fifth subject (Subject 18 and the only male among the five) had reading scores of 11.6 on both the 1997 test and the GAP test. He was the second highest user of the researcher as a live dictionary, and the only one of the five to use this facility at such a level. As such he is clearly a candidate for more detailed transcript analysis.

5.3 Strategy Use analysis

5.3.1 Introduction

As shown in Chapter Four, the coding procedure adopted after trial versions was that originated by Pressley and Afflerbach (1995), referred to in this thesis as the 'P&A system', and presented in Appendix H. This system was used to code the categories of strategy shown by L1 readers while engaging in self-report exercises, and as such represents broadly a more sophisticated and less bottom-up reading process than would be the case with adolescent foreign language learners. Nevertheless, as we will see the subjects in this study did utilise elements of two of the three major sections (identifying and learning text content and monitoring) and within these sections they also used elements from a majority of the sub-categories. The least utilised strategies were those referring to more global meaning construction and the consideration of any sub-meanings of the text. Prior or world knowledge driven interpretation was also generally missing from the processes observed. In many ways this is not surprising, since language level dictated that the texts would be on non-contentious familiar topics, and here, as we have seen it was those of family and free-time pursuits that were chosen. The group transcripts provide more examples of such meaning-driven processing, as on occasions subjects used world experience to make sense of more difficult text. Examples of this are: 'split families often argue' (Group A2) or in terms of reasoning about what is likely to appear in normal penfriend letters: 'You wouldn't say what they like, you'd just say you had brothers and sisters.' (Group B1). But these are clearly strategies used to compensate for uncertainty in basic decoding, rather than as part of a deeper level interpretative process.

After transcription and coding several sets of potential comparative analyses were identified, and these are dealt with in Section 5.3.2. Items of interest within the individual protocols were the performance of an 'average' subject, 'best' and 'worst' performance compared, subjects whose performance across the two task-types was very different, and the performance of a subject who made very extensive use of the live dictionary.

Additionally a subject was chosen whose performances alone and within the group were radically different. For the group tasks four groups were identified whose performance ranged from low-scoring to high-scoring, and whose group ethos and approach to the tasks varied. By analysing the strategy use of the individuals and groups it is possible to determine the importance of varied strategy use, the significance of certain strategy types and to investigate whether strategy coding alone can account for these differences.

5.3.2 The think-aloud protocols

5.3.2.1 The strategies used by an 'average' member of the class

An investigation of the individual transcripts of Subject 7

Subject 7 appears in the top half of the class in both of the reading scores (joint 6th with 5 others on the 1997 scores and joint 12th with three others on the GAP tests). His scores in the two tasks were 3/8 on the TFI task and 3 reasons mentioned for the correct author on the FP task. (This placed him joint 10th with seven others on the TFI task and joint 11th with three others on the FP task). He was in fact the only subject to appear on both reading score top lists who failed to appear on the comparative top lists for the two tasks,

but the rankings above show that this was not in fact an especially anomalous performance.

The aim of this section then is to review the role of the strategy coding for a representative subject and to discover how much information about the reading processes can be gleaned from this method of analysis. Subject 7's fully coded transcript appears as Appendix M.

Subject 7 used six of the P&A sub-sections and a total of 23 different strategy categories for his TFI task and six sub-sections, 11 categories for his FP task.

In the TFI task, as for all subjects, the most common section used was meaning construction during reading (MC/DR), while after this he used strategies from the sections: activation of processing due to difficulties at word/phrase and global levels next most often. (ActDiffwp and ActDiffglob). He also used conscious inference making strategies (MC/DRcim) extensively. Thus he was drawing principally from two of the major sections Identifying and Learning Text Content and Monitoring.

In the FP task two of the six sections were from Monitoring, with the other four from Identifying and Learning Text Content.

It is possible to tabulate this information for further investigation, although we should remember always that it is the language of the transcript which remains of paramount importance and not simply a decontextualised survey of strategy use.

For Subject 7 the strategy patterns are as follows

Table 5ki: Subject 7 TFI task strategy use
Identifying and Learning Text Content

		Monitoring		
MC/BR		MonPT		
Constructing a goal	2	Achieving a goal		7
MC/DR		Overall meaning of part of text is comprehended		1
Adjusting an interpretation	1	MonTC		
Explicitly looking for related ideas ??	9	Text content not relevant to goal		1
Jumping back to reconsider prev. read information	1	ActDiffglob		
Linear reading ??	7	Reading slowly and carefully ??		2
Maintaining an hypothesis	3	Shifting to a different part of text		1
Paraphrasing text	29	States failure to understand		3
Pausing to reflect	10	Suspending judgement		2
Reading aloud	10	Task not complete, continuing search		1
Restating text	5	ActDiffwp		
Tentative interpretation	7	Using dictionary		11
MC/DRcim				
Contextual clue - cognate	3			
Drawing a conclusion	6			

NB: ?? = not overtly demonstrated but we infer this to be his strategy at that point

Table 5kii: Subject 7 FP task strategy use
Identifying and Learning Text Content

		Monitoring		
MC/DR		MonPT		
Linear reading	2	Achieving a goal		3
Maintaining an hypothesis	3	ActDiffwp		
Paraphrasing text	2	Using dictionary		3
Reading aloud	2			
Restating text	1			
Tentative interpretation	2			
MC/DRcim				
Drawing a conclusion	3			
MC/DRintg				
Looking elsewhere in text for related information	1			
MC/AR				
Rereading parts of text following reflection	4			

These tables allow us to see that Subject 7 read and translated substantially less text aloud in the FP task compared with the TFI task. In fact, in the second task he made an early

decision about the author of the letter and was then able to skim silently to find two further reasons to back up his choice. Only when pressed to look for opinions within the text did he find a further reason which was technically incorrect, and he also made 2 of his 3 dictionary enquiries at this point. From table 5ki we can see that he made significantly more dictionary enquiries during the TFI task and spent longer involved in considering evidence. Strategies such as *pausing to reflect* and *explicitly looking for related ideas* (which we surmise rather than know was part of his activity), plus *restating text, making tentative interpretations* and *adjusting an interpretation* all feature either more often or exclusively in the TFI transcript. This is almost certainly a feature of a heavily structured task which prompts the reader to look for specific evidence, often in specific areas of text (if the TFI statements mirror the text progression, for example.) We can see also that he overtly uses contextual clues such as cognates in the TFI task, and may do so also in the FP task, but because we have less formalised and visible meaning construction we cannot be sure *how* he understands the material. This tabulation exercise therefore can give a general overview of activity, can characterise an approach to a task and can offer very real contrasts between two or more task types. If we wish to go deeper into the processes involved we need to consider extracts of transcript along with the strategy coding applied to them.

The opening section gives an overview not just of Subject 7's reading behaviour when engaged in processing connected text, but also serves as a good example of how many of the subjects in this Study behaved.

2.	7 Em	MC/DR	Paraphrasing text
3.	7 The weekend		
4.	7 Em at the weekend.		
5.	7 activities.	MC/Drcim ??	Contextual clue - cognate ??
6.	7 Em Samedi Saurday	MC/DR	Paraphrasing text
7.	R Mhm.		
8.	7 Em	MC/DR	Paraphrasing text
9.	7 Disco		
10.	7 avec		
11.	7 with my friends.		
12.	7 Em I play sport	MC/DR	Paraphrasing text
13.	7 on		
14.	7 Sunday?	MC/DR	Tentative interpretation
15.	7 em	MC/DR	Paraphrasing text
16.	7 disco		
17.	7 em that means	MC/DR	Pausing to reflect
18.	7 em		
19.	7 oh it's got something to do with music.	MC/DR	Tentative interpretation
20.	7 Em mais mes artistes préférés sont Robbie Williams et Lauryn Hill.	MC/DR	Reading aloud
21.	7 Er. My favourite artists are Robbie Williams and Lauryn Hill.	MC/DR	Paraphrasing text
22.	7 Em.	MC/DR ??	Linear reading ??
23.	7 Em.		
24.	7 Entrainement, what does that mean?	ActDiffwp	Using dictionary
25.	R Entrainement means training.		
26.	7 Em I'm training for rugby	MC/DR	Paraphrasing text
27.	7 at		
28.	7 no em	MC/DR	Adjusting an interpretation
29.	7 carry on carry on	ActDiffglob	Suspending judgement
30.	7 Amies	MC/DR	Reading aloud
31.	7 Amies fumes.		

32.	7 Friends.	MC/DR	Paraphrasing text
33.	7 No.		
34.	7 Yeah. Friends.		
35.	7 Rugby again.	MC/DR	Paraphrasing text
36.	7 Cigarettes.	MC/DR	Paraphrasing text
37.	7 What's sentent mean?	ActDiffwp	Using dictionary
38.	R Er smell.		
39.	7 Smell.	MC/DR	Paraphrasing text
40.	7 Ah. I get it now.	MonPT	Overall meaning of part of text is comprehended
41.	7 At the rugby club, there is a smell of cigarettes.		

The code *paraphrasing text* is used to denote translation, but clearly this can vary in its scope from a single word to a complete sentence (line 21). The structure of the transcript means that intonational units are presented, and the pauses between lines are often brief (commonly less than a second). Where the reader indicates through a rising intonation, or the insertion of a qualifier such as 'something' or 'I think' the paraphrase is classified as a *tentative interpretation*. An action, for which there is strong evidence, but no definite confirmation, such as the use of a cognate on line 5, is included but with a question mark following. A silence is naturally a coding problem. Where the reader does not move any further on in the text, it is classed as *pausing to reflect* (which can include the notion of re-reading). Where a reader is clearly moving on and processing for the first time it is classed simply as *linear reading*. Hesitations, principally signalled as em, um or er, are classified in the same way, eg lines 17-18 are pausing to reflect while 22-23 are linear reading. A direct question, asking for vocabulary is classed as *using a dictionary*. If it was felt that the reader was asking her/himself the researcher delayed offering the translation. Lines 28/29 show far less commonly used coding categories, *adjusting an interpretation* and *suspending judgement*, as do lines 40-41, *the overall meaning of part of text is comprehended*.

Perhaps more importantly we can see this extract as a representation of the flow of meaning construction with reading aloud, paraphrasing, using cognate clues, making tentative interpretations, and asking for single item assistance all involved, often in that sequence. Sometimes (as in this example) this process is supplemented by recapitulation statements at certain points. Sometimes (with the less successful readers, many of these elements are omitted).

In the transcript sections relating to task completion Subject 7 voices overtly the process which most other subjects seem to be using, ie looking for matching vocabulary and structures between task statements and text. He comments that he is: '*scanning the page to see if there are any other words to connect with it.*' This is borne out by his rationale for the decision that task statement five is false: '*because it doesn't say she goes to the butcher's*' rather than identifying that she does work in the baker's. When decisions are made they commonly arise from a sequence which involves further text paraphrase, a silence probably spent looking for matching lexis, perhaps some tentative interpretation, a comment that the statement is true, false or unclear and the rationale for that decision.

The sequence below exemplifies this process.

144.	7 (7)	MC/DR ?? + MC/DR ??	Linear reading ?? + Explicitly looking for related ideas ??
145.	7 Sunday.	MC/DR	Paraphrasing text
146.	7 Number six I think is false	MonPT	Achieving a goal
147.	7 because it says em	MC/DRcim	Drawing a conclusion
148.	7 Elle aime regarder les matchs de rugby le dimanche.		
149.	R Mhm.		

150.	7	That means she watches a match of rugby	MC/DR	Maintaining an hypothesis
151.	R	Mhm.		
152.	7	on Sunday, but in the text it says she plays rugby.		
153.	R	Okay.		

If further justification is offered after the first inference has been drawn (as in lines 150-2) the coding *maintaining an hypothesis* is used.

There are also examples of less successful meaning construction processes, both in Subject 7's transcript and in most of the others. This sequence is taken from the FP task at a point when he has been asked to find some opinions to agree/disagree with (as stated earlier, as a pretext for encouraging further processing of the text)

41.	7	(11)	MC/AR	Rereading parts of text following reflection
42.	7	What does belle mean? B E L L E?	ActDiffwp	Using dictionary
43.	R	Beautiful or pretty.		
44.	7	(19)	MC/AR	Rereading parts of text following reflection
45.	7	Em.	MC/DR	Paraphrasing text
46.	7	My brother		
47.	7	No my sister		
48.	7	is very beautiful.		
49.	7	Nat – naturellement ?	ActDiffwp	Using dictionary
50.	R	Er naturellement can mean like naturally, but it can also mean of course.		
51.	7	French of course. Em.	MC/DR	Paraphrasing text
52.	7	Même	MC/DR	Reading aloud
53.	7	anglais aussi. Elle a fait plusieurs disques et		
54.	7	I think that's someone. That means like her sister	MC/DR	Tentative interpretation
55.	R	Mhm.		
56.	7	is like a very popular film star.		
57.	7	[Or something.		

Here he is re-reading and attempting to get a sense of the first paragraph of the FP text.

This was the most difficult paragraph and produced many other transcript extracts like this. He does not ask for verbs as his key words, a common aspect of the 'dictionary'

patterns across the study, despite the fact that these items are often key to meaning. Here there are three, and *chante* (8 times requested), *fait* (2) and *enregistré* (4) were essential for the major ideas in the paragraph, yet were not clarified either by Subject 7 or by more than a few of the subjects (see figures in brackets). He makes a global reading which has seized on the very last phrase of the paragraph: *d'un film très populaire*, but has misinterpreted it. In fact he was probably led in that direction by a desire to maintain the hypothesis that the author was Sasha Two.

Thus Subject 7 provides an appropriate exemplar for the way in which subjects approached the tasks. Many were task-led in the TFI task from the beginning, but almost all read the text to some degree aloud either in French or English for the FP task. The three extracts above, showing successful meaning construction, decision making with regard to the tasks set and unsuccessful meaning construction, demonstrate the majority of high-frequency strategic use.

5.3.2.2 Best and Worst performers compared: a comparison of the individual transcripts of Subjects 14 and 27

Quantitative evidence:

These two subjects achieved, respectively, the best and worst performance on the TFI (True/False/Impossible to say) task, with scores of 8/8 by Subject 14 and 0/8 by Subject 27. (This subject furthermore did not achieve either any correct answers which were unsupported or wrongly rationalised. All of her responses were actually wrong or simply omitted.) She was also the only one of the twenty seven subjects not to get the second

question right.) A similar difference appeared in their performance on the FP (Four Penfriends) task, where Subject 14 noted five reasons for her choice of letter writer (which ranked her third of the 27) and Subject 27 noted just one reason.

Their approach to the two tasks was, however, broadly similar, although the way each used that approach differed. In the TFI task both read sections of text silently, but Subject 14 also translated sections into English aloud. Subject 27 read text and task statements silently and then responded to the statements afterwards. Although she appeared to be working out her responses before she gave any, in fact she needed clarification still on the section where she summarised her answers. Subject 14 used 81 lines of transcript to clarify the text, whereas Subject 27 used 42 to cover text and task statements silently. On the FP task both read the whole text silently, and then gave a view on identity of the author of the letter. They also showed a similar pattern in the number of words requested, clarified or given with Subject 14 asking for 17 words in the TFI task and 5 in the FP task, while Subject 27 asked for 16 and 5 respectively. (The mean for each was 13.22 and 6.74 respectively.)

But the ratio of speech to silence on the two tapes is an area of significant difference. Both had a comparatively high level of silence recorded, as both opted to read silently for part of both tasks. Subject 27's transcript runs for 182 lines and Subject 14's for 257 lines. Subject 27 has 15 minutes and 42 seconds of silence recorded on the tape, whereas Subject 14's silent total is 8 minutes and 58 seconds. Clearly then there is a large part of Subject 27's activity which we are unaware of, but which does not seem to include

reading with understanding, (as the responses were all incorrect). It does not either include asking for problematic vocabulary at a rate which we might expect if comprehension is not good.

The use of the concordance programme is described in detail in Chapter Six, section 6.3.2.1, but we can use an extract from its findings to illustrate further quantitative evidence about these two subjects. The number of instances of certain key terms for reasoning shows a big difference between them. For example Subject 14 used the word 'so' fifteen times in her transcript as opposed to a single use by Subject 27. 'Because' was used thirteen times by Subject 14 and twice by Subject 27. The word 'like' occurred eight times with Subject 14 and only once with Subject 27.

It is therefore important to look in more detail at the transcript to find possible reasons for the great difference in performance between these two subjects.

Qualitative evidence:

If we compare the opening sections of each transcript we find an initial difference in how the approach (ie silent reading of whole text) is actually utilised. Subject 27 reads with almost no request for assistance over a period of 95 seconds and then turns to the task. Subject 14 overviews more briefly (46 seconds) and then sets about organising her meaning construction of the text. She does not turn to the task until much later. Below are two sequences which show approximately the first 30 lines of each transcript.

*(NB * The category 'Affective' and the coding note: 'resistance to probes' does not appear in the P&A system, but is used here to draw out an important element of Subject 27's performance)*

Subject 27

Turns /speakers / transcript	PA section	Coding notes
1) R Okay. Start when you're ready.		
2) 27 ((Reads silently (32)))	MC/DR ??	Linear reading ??
3) 27 What does that sentent mean?	ActDiffwp	Uses dictionary
4) R Erm. Smell.		
5) 27 Oh right.	MC/DR	Pausing to reflect
6) (63)	MC/DR ??	Linear reading ??
7) 27 Does the first question mean	MC/DR	Tentative interpretation
8) 27 does she go to the		
9) 27 supermarket on Sundays? Um		
10)R Erm.		
11)R The second word?		
12)R Do you know ?		
13) (3)		
14)27 Oh no.	MonProb	Unfamiliar terms in text
15)27 No.		
16)R Travaille means works.		
17)27 Oh right.	MC/DR	Pausing to reflect
18)27 Okay.	MC/DR	Pausing to reflect
19) (108)	MC/DR ??	Looking for information relating to specific goals ??
20)R Are you still looking at number one?		
21)27 No.	Affective*	Resistance to probes*
22) (13)	MC/DR ??	Looking for information relating to specific goals ??
23)27 What does the	ActDiffwp	Uses dictionary
24)27 third word on the fourth one mean?		
25)R Er. Commencé means started.		
26)27 Oh.	MC/DR	Pausing to reflect
27) (65)	MC/DR ??	Looking for information relating to specific goals ??
28)27 What does the last word mean on the fourth one?	ActDiffwp	Uses dictionary
29)R Erm. Dernière means last so		
30)R you put it with something else like la semaine dernière means last week.		
31)27 Oh.	MC/DR	Pausing to reflect
32) (109)	MC/DR ??	Looking for information relating to specific goals ??

Subject 14

Turns / speakers / transcript	PA section	Coding notes
1) 14 Alright.	MC/DR	Pausing to reflect
2) 14 I'm just gonna read it first.	MonPT	Own behaviour
3) R Okay.		
4) 14 ((Reads silently (46)))	MC/BR ?	Overviewing text ?
5) 14 Right. So this first paragraph.	MonPT	Own behaviour
6) 14 Em. At the weekend I am very active.	MC/DR	Paraphrasing text
7) 14 Lundi Mardi Mercredi Jeudi Vendredi Samedi	ActDiffwp	Use context clues - draw on set of PK
8) 14 so that's on Saturday.		
9) 14 Em	MC/DR	Pausing to reflect
10) (2)		
11) 14 Go to a disco with my friends.	MC/DR	Paraphrasing text
12) 14 I do sport on Sunday.	MC/DR	Paraphrasing text
13) 14 At the disco I listen to discs.	MC/DR	Paraphrasing text
14) 14 Plus plus plusieurs. Is that something like all different kind of thing?	ActDiffwp	Candidate meaning evaluated in context
15) R Yeah. That's right. Yeah.		
16) 14 All different styles. Em.	MC/DR	Paraphrasing text
17) 14 But my favourite artists	MC/DR	Paraphrasing text
18) 14 is Robbie Williams and Lauryn Hill.	MC/DR	Paraphrasing text
19) 14 I dance	MC/DR	Paraphrasing text
20) 14 We dance	ActDiffwp	Self correction - structural clue #
21) 14 em	MC/DR	Pausing to reflect
22) (17)	MC/DR ??	Linear reading ??
23) 14 What does bon mean?	ActDiffwp	Use a dictionary
24) R Er. Bon just means good.		
25) 14 Er.	MC/DR	Pausing to reflect
26) 14 It's good	MC/DR	Restating text
27) 14 entertainment?	MC/DRcim	Contextual clue - cognate
28) 14 And something rugby.	ActDiffwp	Keep reading

The coding note: self-correction, structural clue does not appear in the P&A system, as that was designed to deal with L1 readers

From the beginning Subject 27 is attempting to construct meaning and settle on her responses with minimal recourse to assistance. The researcher's interventions are also at a minimal level for affective reasons, (that she was clearly nervous and might have been struggling with the text). This is demonstrated by the response to the question on line 20, where a single 'No' reveals that she does not wish to expand on her current actions. Her

difficulty with the material is shown by the long silence at line 27 during which she is clearly looking at item four, but not managing to understand either the statement or the relevant text. The one word she asks for during the initial reading (line 3) is an unfamiliar word to all of the subjects, but it is difficult to be sure why she chooses that word in the text rather than any other. She has used 7 different categories of strategies in 3 of the P&A sub-sections during this section.

If we contrast this approach with that of Subject 14, we may not see a direct extreme (as could be provided for example by the protocols of subjects 18 and 26 who requested the most vocabulary assistance), but the manner in which she uses assistance is notable. Her initial reading is briefer, and she then organises herself aloud (line 5) so that we can be certain she has a particular game plan. (Subject 27 may also have a strategy but does not reveal it.) Her approach is to attempt to solve comprehension problems by herself and ask for help in the last instance. So we see on line 7 that she recites the days of the week to decide which one she is meeting in the text. On line 14 she makes a contextual guess at an unfamiliar word. On line 20 she corrects a translation she has just made. On line 27 she makes an (incorrect) assumption about a near cognate. On line 28 she reveals that she will look at a whole group of words to make sense of an unknown item and if need be keep reading beyond that section (*'and something rugby'*). Only (in this extract) at line 23 does she ask for a word. It may be that Subject 14 is better at French, knows many more words and can therefore make better overall comprehension of the text than Subject 27, but even if this is the case, the contrast in techniques used is a separate factor from relative competence. In this time Subject 14 has used 12 different categories of strategies in 5 of the sub-sections identified, significantly more than Subject 27.

This contrast is graphically demonstrated by the approach each made towards resolving statement three: *A la disco elle boit et fume beaucoup*. The text source for this statement reads: *Une de mes amies fume. Elle m'invite tout le temps fumer aussi, mais je dis toujours, 'non', parce que je joue au rugby*. The expected strategies (see Appendix I) to cope with a difficult task statement might be:

- to scan for the presence of a form of *boire* and/or *fumer*
- if any vocabulary is not known then to infer meaning from the known items / cognates such as *invite / non / joue au rugby*
- to identify *fumer* as a word to look up if not known

A likely 'real' process (stemming from 'common sense'-based field experience) might be to ask for meanings of *boit* and *fume*, then to scan for either and find *fume* twice in extract above. The meaning of *invite* can be inferred as a cognate. The familiar words *non* and *je joue au rugby* can be identified and so the answer False can be given. If there is still uncertainty, the meanings of *dis* and *parce que* could be checked. In fact, in reality, even this process was under-used, as many subjects turned to the following sentence in the text which read: *Et les cigarettes sentent vraiment mauvais*. Having established the meanings of the last three words they inferred that the item was false because they now knew her opinion of smoking.

Below are the approaches used to this item by, first, Subject 27 and then Subject 14.

3.	27	What does that sentent mean?	ActDiffwp	Uses dictionary
4.	R	Em. Smell.		
5.	27	Oh right.	MC/DR	Pausing to reflect
59.	27	Third one.	MonPT	End of a unit of meaning
60.	(9)		MC/DR ??	Looking for information relating to specific goals ??
61.	27	Miss that one out. What does the fume?	ActTask + ActDiffglob ie (ActDiffwp)	Decision to skip material + Formulates a question on material ie (Uses dictionary)
62.	R	Fume means smokes.		
63.	(33)		MC/DR ??	Looking for information relating to specific goals ??
64.	27	What does the, the B one that one?	ActDiffwp	Uses dictionary
65.	R	Boit means drinks.		
66.	27	What does the last one mean?	ActDiffwp	Uses dictionary
67.	R	Beaucoup means a lot.		
68.	27	Oh.	MC/DR	Pausing to reflect
69.	(8)		MC/DR ??	Looking for information relating to specific goals ??
70.	27	I don't think it says anything.	ActDiffglob	Pausing to scan for source of difficulty
71.	27	Fume. ((whispered))	MC/DR	Explicitly looking for related words
72.	(27)		MC/DR ??	Looking for information relating to specific goals ??
73.	27	Em.	MC/DR	Pausing to reflect

The two extracts come from an initial reading of the text, and from a section where she is working through the TFI statements and trying to make decisions. At line 71 she has perhaps spotted the word *fume* in the text, but then possibly decides she cannot make use of it. Her quite unusual (as regards the rest of the subject cohort) behaviour of not echoing the meanings of words given and not piecing sections together aloud, mean that we cannot have a definite view on her own global understanding of the text or the

statements. So the expected strategic procedures outlined above do not appear to have been kicked into action, beyond the stage of requesting a few items of vocabulary. Four times we assume she is looking for information relating to specific goals, but cannot be sure. She asks for three words but does not seem to combine them into a coherent meaning. She probably does scan for the source of the difficulty, she reflects and formulates one question. The conclusion is that there is nothing relating to that task statement in the text.

In contrast with this we have the longer extracts (itself of significance) from Subject 14.

31.	14 What's that last sentence?	Monprob	Failure to understand
32.	14 Et les cigarettes sentent vraiment mauvais? Or something?	MC/DR MonTC	Reading aloud linguistic charac.
33.	R Right. Which words are you		
34.	R not sure about?		
	In lines 35-39 below, subject is	MC/DR	Looking for acquiring key words
35.	14 Erm. Sentent.	ActDiffwp	Use a dictionary
36.	R Er. Smell.		
37.	14 Vraiment	ActDiffwp	Use a dictionary
38.	R Is really.		
39.	14 And the last one mauvais.	ActDiffwp	Use a dictionary
40.	R Mauvais means bad.		
41.	14 So	MC/DR	Tentative interpretation of text meaning
42.	14 is it		
43.	14 It might be I don't smoke		
44.	14 because it smells really bad, or		
45.	14 it smokes of cig- it smells of cigarettes?		
46.	14 At the rugby place or something.		

107.	14 What does beaucoup mean?	ActDiffwp	Use a dictionary
108.	R Beaucoup means a lot.		
109.	14 What does fume mean?	ActDiffwp	Use a dictionary
110.	R Smokes.		
111.	14 At the disco she smokes a lot.	MC/DR	Paraphrasing text

112.	(12)	MC/DR	Explicitly looking for related ideas
113.	14 No.	MonPT	A goal is achieved
114.	14 I think there it says	MC/DRcim	Drawing a conclusion
115.	14 em friends		
116.	14 kind of like ask her, kind of told her		
117.	14 invited her to smoke		
118.	14 and she said no		
119.	14 because she plays rugby		
120.	14 and the cigarettes smokes		
121.	14 em smell		
122.	14 em.	MC/DR	Pausing to reflect

In the first extract Subject 14 is clarifying a section of text as she reads through it in sequence. She has found a block of words that she does not understand and the researcher (as was standard practice in this investigation) offers a word at a time (in case she is able to infer a second or third without asking). She then pauses to construct meaning in a very tangible manner: she frames questions to herself, uses the words *might* and *or something* to demonstrate that she is seeking options and considering different possibilities. Her phrasing of this tentative meaning shows that she has not understood the text in detail (they are at the disco, not the 'rugby place', for example). At this point she has not considered the task statements, so the suggestion that '*it might be I don't smoke*' is part of the getting of global meaning rather than a task response. In fact she does not ask the meaning of fume here but later does so when looking at the relevant task statement. So her transfer from 'cigarettes smelling bad' to 'I don't smoke' is a meaning inference to be stored in memory. This process is crucial to gaining an understanding which will inform the task subsequently, and to which reference could be made later. (There is a further example of this in Subject 14's group-based TFI task (Group A4, TFI task, lines 88 and

184) where she firstly clarifies a word as the group read through the text and later remembers and uses that evidence to make a decision about an item.)

When she comes to Statement 3 she needs to clarify the meaning of *fume* and *beaucoup*, and she then re-states / translates the task statement. She can then pause to consider the evidence and give a decision (that the statement is false). Her reasons then appear as a paraphrase of that evidence, about the friend's invitation to smoke, her negative reply, the role of the rugby playing in this. She has demonstrated an understanding of the overall meaning of that quite difficult section of text, but, like Subject 27, at the beginning did not know the two key words, *fume* and *sentent*.

Perhaps it is because Subject 14 knows more general vocabulary that she can fill in the gaps between key words more successfully than can Subject 27, but the evidence of her concentration on meaning construction is radically different than that suggested by 27's silences. Subject 14 uses just two more strategies than Subject 27, but the qualitative difference is greater. She uses the dictionary facility more, looks to acquire key words from the text from a sentence she has identified as important (rather than from the task statement). The words asked thus are combined into an idea through a tentative interpretation of meaning which she holds for later. She also reads and translates aloud, and appears thus to be searching for related ideas with more purpose. By simply concentrating on holding meaning and generating hypotheses she appears to be behaving very differently from Subject 27.

On the Four Penfriends task both subjects read silently with very little recourse to vocabulary checks, and both make a decision quite quickly. The A group subjects had completed their group-based tasks before their think-aloud session, so the format was

already familiar to them. The difference between the two task-types is that when making a global decision a reader can check what s/he *does* know to see whether such a decision is possible, whereas in a TFI task particular vocabulary needs to be understood.

Inevitably the process is quicker, particularly on a second occasion, because if even a small piece of evidence is clear, then a decision can be postulated initially at least. With an accompanying drive to complete the task, this is more likely.

So Subject 27 finds a single piece of evidence near the end of the text, makes and reports her decision, but then when pressed for more information, cannot find any.

19.	27 Em, I think it's the second one.	MonPT + MC/DRcim	A goal is achieved + Drawing a conclusion
20.	27 Cos em	MC/DR	Tentative interpretation
21.	27 she says that em		
22.	(16)	MC/DR ??	Looking for information relating to specific goals ??
23.	27 Erm. Does that bit there mean her dad's the president of France?	MC/DR	Maintaining a hypothesis
24.	R Mhm.		
25.	27 That wouldn't *	MC/DR	Maintaining a hypothesis
26.	27 *		
27.	(12)	MC/DR ??	Looking for information relating to specific goals ??
28.	R What did you find the letter was mainly about?		
29.	27 Em. About her family and	MC/DR	Tentative interpretation
30.	27 what they do.		
31.	R Mhm.		
32.	27 What they're like.	MC/DR	Tentative interpretation
33.	(5)	MC/DR ??	Pausing to reflect ??
34.	R So was there anything else that was like outrageous?		
35.	R *		
36.	27 Erm.	MC/DR	Pausing to reflect
37.	(3)		
38.	27 Not sure.	ActDiffglob	States failure to understand

Subject 14 has used the same approach, but can offer more evidence from the beginning:

22.	14 I think	MonPT + MC/Drcim	A goal is achieved + Drawing a conclusion
23.	14 that it's probably Sasha Two.		
24.	R Mhm.		
25.	14 Because	MC/DR	Maintaining an hypothesis
26.	14 she says that		
27.	14 well it might be true but		
28.	14 her brother		
29.	14 what's it?		
30.	14 plays for the national team and he's on		
31.	14 television or something.		
32.	R Mhm.		
33.	14 And	MC/DR	Maintaining an hypothesis
34.	(3)		
35.	14 her mum or something		
36.	14 something to do with parliament.		
37.	14 and her dad was the president of France.	MC/DR	Maintaining an hypothesis
38.	R Mhm.		
39.	14 And then	MC/DR	Pausing to reflect
40.	14 I didn't know if there is a president of France, but	MonTC	Relationship between own background knowledge and text content
41.	14 em.	MC/DR	Pausing to reflect
42.	R There is. ((laughs))		
43.	R Are there any		
44.	R are there any other reasons there.		
45.	R Obviously you've given three. Are there any others you could mention?		
46.	14 Em.	MC/DR	Pausing to reflect
47.	(12)	MC/AR	Re-reading
48.	14 What does	ActDiffwp	Use a dictionary
49.	14 em		
50.	14 cadette chante mean?		
51.	R Em.		
52.	R Cadette means younger.		
53.	14 Mm. So what does chante mean?	ActDiffwp	Use a dictionary
54.	R Er. Chante is sing sings.		
55.	14 So that's my younger sister sings, is it? Em	MC/DR	Tentative interpretation
56.	(10)	MC/DR	Pausing to reflect

57.	14 Does that say something about she's popular in Europe?	MC/DR	Tentative interpretation
58.	14 like a singer so that might be	+ MC/DR	+ Maintaining an hypothesis
59.	14 a bit of a lie as well.		
60.	14 Em.	MC/DR	Pausing to reflect

Again Subject 14 pieces together some relevant information from asking a couple of vocabulary items (lines 48-50). She has already used a newly learned word from task one (*l'équipe*) and two cognates to generate a view based on material about the author's brother, and to this she has added information about the parents, again based heavily on cognates. When asked to look for more evidence she turns to a paragraph which she did not understand as well and clarifies the items, *cadette* and *chante*. The knowledge about the younger sister singing seems to open up the significance of the phrase: *elle est déjà très populaire dans toute l'Europe*, and so she can add that fact to those she has already identified. Interestingly, she has taken nothing from the fairly difficult first paragraph. This would suggest that although more confident in her vocabulary, she is not so far ahead of the others in the group that she can make sense of text which challenges the rest. She has managed to locate five facts, about all but one member of the family so has completed the task more than adequately. With an undefined number of items, a halt can be called more quickly than with, say the 8 TFI items.

5.3.2.3 Differences between performances on the two tasks

Subjects 3 and 25 were selected for close strategy use analysis for two reasons. Both represented a reading score category (3 a higher reading score for the class, 25 a lower reading score for the class on both the 1997 and GAP scores), and both performed

unevenly across the two individual tasks, Subject 3 well on FP and poorly on TFI, while 25 performed well on TFI and poorly on FP.

Across all subjects there was a clear tendency to use a greater number of strategies to complete the TFI task than for the FP task, for reasons already discussed above. Both of these subjects used more of the P&A sub-sections and of individual categories on the first task than on the second. Both attempted at different points to integrate different parts of the text, both read aloud at some points, translated at some points, paused to reflect and made tentative interpretations. Subject 25 adjusted ideas on the basis of new information as it arose, scanned for related ideas, found evidence to maintain hypotheses she had made and used the dictionary facility more often. She also overtly held ideas in working memory. Subject 3 read aloud more and translated in shorter units of meaning than Subject 25. He also stated his failure to understand several times, decided to continue without having understood, and although he made more tentative interpretations these could be far *more* tentative than those of Subject 25. For example, he used the word 'something' 18 times during the task but this word occurs only 3 times in the Subject 25 transcript. Where he has acknowledged difficulties in constructing global meaning (here sentences rather than words), Subject 3 tends to move on without solving them. Subject 25 makes better use of the dictionary facility perhaps by being better able to identify the key words for understanding. Subject 3 did begin to translate the whole text, but was not succeeding in very detailed or coherent meaning construction, and uniquely within the study the researcher guided him towards the TFI statements after the first paragraph. Although this was not in tune with the approach used with any other subject it was done

to aid meaning construction, and could not have impaired that process since the subject was clearly having comprehension problems up to that point.

TFI statement 3 again demonstrates these contrasts. Although both subjects failed to get the correct answer the processes used to approach it are different. Subject 3 has difficulties with that section of text before he starts on the task. The transcript splits here into very short intonation units, often of just one word as he shows clearly that he lacks confidence in his understanding of what he is reading. He uses the two words he has asked for prior to this section (*aussi* and *mais*), but as link words they do not really offer clues as to the core meaning. He notes the word cigarettes, which will be important for the task, but in the following extract seems not to have recalled this.

37.3	Um	MC/DR	Pausing to reflect
38.3	Er		
39.3	er my friends	MC/DR	Paraphrasing text
40.3	um	MC/DR	Pausing to reflect
41.3	she	MC/DR	Paraphrasing text
42.3	invites		
43.3	us I think or		
44.3	to	ActDiffglob + ActDiffglob	Suspends judgement + carefully analyses information presented so far
45.3	the		
46.3	also again		
47.3	um but		
48.3	um I		
49.3	I'm not sure what that means toujours	MonProb	Unfamiliar terms
50.3	are parce que something about the rugby I think they go and watch it or something	MC/DR	Tentative interpretation
51.3	and then something about cigarettes or something	MC/DR	Tentative interpretation

He then returns to the section when trying to make sense of the statement.

77. 3	What's the um	ActDiffwp	Using dictionary
78. 3	fume word ?		
79. R	Fumer means to smoke		
80. 3	Right	MC/DR	Pausing to reflect
81. 3	I'm not sure about the answer to that one. I don't quite understand it	ActDiffglob	States failure to understand

There is apparently a need here to elucidate the statement itself. Having asked for the word *fume* he might have returned to the sentence containing the item *cigarettes* and begun to look there, but comes to a full stop and moves on. This raises the issue of task difficulty and the fact that in a TFI task the statements are sometimes linguistically more challenging than the target text. Here the syntax is very straightforward : *A la disco elle boit et fume beaucoup*, but we know that neither verb was commonly known to this group of subjects, and few knew *beaucoup* either. This combined with a common confusion around the preposition *à*, which is often remembered as *to* more than *at* and it could seem a daunting sentence. Nevertheless other subjects did break it down, ask for meanings and go on to offer an answer. Subject 25 works in this way. She has not read the text in advance so considers statement and text together. She establishes the core meaning of the statement by asking for the key items *fume* and *beaucoup* and, we assume, starts to scan for related information back in the text. She quickly finds cigarettes. (We do not know whether she has spotted the two occurrences of the word *fumer* before that) She starts to make a connection between the cigarettes smelling and Chantal not smoking, but does not see this through, and concludes that the statement is *pas clair* rather than *false*.

Nevertheless the basis for dealing with unknown items, moving through tentative

interpretations to concrete decision-making about meaning is laid out clearly here in a way that we did not see with Subject 3's treatment of the same section.

47.	25	What does fume mean?	ActDiffwp	Using dictionary
48.	R	Er. Smokes.		
49.	25	She smokes.	MC/DR	Restating text
50.	25	Does it mean she smokes or	MC/DR	Tentative interpretation
51.	25	*		
52.	R	Er. Elle fume means she smokes. Yeah.		
53.	25	Yeah.	MC/DR	Pausing to reflect
54.	25	Cos it says elle boit et fume	MC/DR	Reading aloud
55.	25	So she smokes	MC/DRintg	Holding ideas in working memory
56.	25	What does beaucoup mean?	ActDiffwp	Using dictionary
57.	R	A lot.		
58.	25	Oh.	MC/DR	Pausing to reflect
59.	25	So at the disco she smokes a lot.	MC/DRintg	Holding ideas in working memory
60.	25	(22)	MC/DR ?? + MC/DR ??	Linear reading ?? + Scanning for related ideas ??
61.	25	What does em	ActDiffwp	Using dictionary
62.	25	les cigarettes sentent vraiment mauvais?		
63.	R	Er. Sentent means they smell.		
64.	25	Cos it says sentent	MC/DR	Restating text
65.	R	Mm.		
66.	25	So she doesn't like smoking cigarettes cos they smell or	MC/DR	Tentative interpretation
67.	25	is it cos	MC/DR	Pausing to reflect
68.	25	I think that one is	MonPT	Achieving a goal
69.	25	pas		
70.	25	pas clair.		

We can reinforce this point about a failure to identify which might be key items from a comparison of two extracts from Subject 3's transcript, one from the TFI task and one from the FP task, on which he was the most successful subject in the study. Firstly we see him working on the text extract concerning Chantal's job in the bakery department of the

supermarket, and then on his attempt to respond to the task statement which reads that she works in the butchery section.

97. 3	I travel	MC/DR	Paraphrasing text
98. 3	I think that's pastries or patisseries or something	MC/DR	Tentative interpretation
99. 3	cakes and gateaux is cakes		
100. 3	and that's with	MC/DR	Tentative interpretation
101. 3	le pain is bread I think		
102. 3	um then pain again and chocolate sont supers	MC/DR	Reading aloud
103. 3	So I think she likes chocolate or something	MC/DR	Tentative interpretation

121. 3	And then it says she travels dans le rayon boucherie	MC/DR +	Paraphrasing text +
		MC/DR	Reading aloud
122. 3	she travels	MC/DR	Paraphrasing text
123. 3	ray- what's rayon ?	ActDiffwp	Using dictionary
124. R	Rayon means, like, department		
125. 3	Oh right	MC/DR	Pausing to reflect
126. (6)		MC/Drintg ??	Looking for related information ??
127. 3	Um	MC/DR	Pausing to reflect
128. 3	I'm not too sure on that one	ActDiffglob +	Suspends judgement +
		ActDiffglob	Shift to a different part of text

The need to ask for the item boucherie seems paramount, but we can only assume that the meaning construction of the previous section some 20 lines previously was not retained in working memory, perhaps because of the comparatively slow processing speed. So the significance was not realised. Similarly the assumption that *travaille* means *travel* was not questioned, even though it was blocking a sensible reading. (In fact only eight subjects did ask for the word, with three more clarifying what it meant.) The text may have been just too difficult for him to allow both processing and strategic behaviour to take place adequately. Yet we could compare the approach here with that used by Subject

3 in the next task when dealing with the second paragraph concerning the author's brother.

37. 3	And then she's talking about her brother- she says my brother Annick has	MC/DR	Paraphrasing text
38. 3	his hair is quite long	MC/DR	Paraphrasing text
39. 3	he is very sporty	MC/DR	Paraphrasing text
40. 3	He plays football	MC/DR	Paraphrasing text
41. 3	for		
42. 3	the national team		
43. 3	((laughs))		
44. 3	I think that's what it says	MC/DR	Tentative interpretation
45. 3	Um	MC/DR	Pausing to reflect
46. 3	um		
47. 3	What does mais mean, I've forgotten	ActDiffwp	Using dictionary
48. R	But		
49. 3	But, but	MC/DR	Restating text
50. 3	he	MC/DR	Tentative interpretation
51. 3	something he		
52. 3	he plays		
53. 3	something about, I think it's about Londres, what's Londres ?	ActDiffwp	Using dictionary
54. R	London		
55. 3	I thought it was. So it's something about he plays in London or something	MC/DR	Tentative interpretation
56. 3	He is	MC/DR	Paraphrasing text
57. 3	also		
58. 3	célèbre pour travail	MC/DR	Reading aloud
59. 3	à la télévision something about television he likes watching it. What is devenu ?	ActDiffwp	Using dictionary
60. R	Devenu means he's become		
61. 3	He has become also	MC/DR + MC/DR	Paraphrasing text + Tentative interpretation
62. 3	a celebrity on television I think that's what it might mean		

Here he is able to build on the presence of cognates to recognise the less obvious key words, and he also is more prepared to make inferences and summarise units of meaning

before building on them. His use of the word *something* in this extract is much more positive, because he is using it to create a coherent meaning of the rest of a phrase, and therefore as a strategy *towards* understanding rather than as an indication that the meaning has not been grasped. Three times in these 25 lines he stops to summarise, and thereby creates a block of meaning to extend his more global understanding. The process of stating meaning so overtly acts as a memory spur to allow the building that he did not manage in the previous task. Clearly then Subject 3 is able to fulfil the expectations we might have from knowing his reading scores in some contexts at least.

Subject 25, on the other hand manages less well without the individual task statements to prompt her readings of the text. She reads the text silently for 43 seconds and immediately makes a decision (incorrectly) about the author of the letter. She offers a partial translation of the same section featured in the extract from Subject 3 above, as follows:

10.	25	Her hair is long.	MC/DR + MC/DR	Paraphrasing text + Maintaining an hypothesis
11.	25	Very long and		
12.	25	she is very sporty.		
13.	25	And she likes national um team.		
14.	25	And she plays for the national football team, for a national football team.		

She has not realise that this concerns the author's brother, nor does she question the fact that the person plays for the national football team. A little later when she is prompted to look for opinions (a third part of the task designed specifically to make subjects look further at the text) she returns to that section:

43.	25	But er	MC/DR	Tentative interpretation
44.	25	it says something about		
45.	25	London.		
46.	25	it says mais maintenant il à joue à Londres.	MC/DR	Reading aloud
47.	25	does that mean she	MC/DR	Tentative interpretation
48.	25	plays for something?		
49.	R	Yeah.		
50.	R	Londres actually means London.		
51.	25	Yeah. So she plays for Lon	MC/DR	Tentative interpretation
52.	25	she plays against London or?		
53.	R	In London.		
54.	25	She plays for England or against England?	MC/DR	Tentative interpretation
55.	R	Er. It just says in London. But [em		
56.	25	Yeah.]		
57.	R	Yeah.		

She still sees this paragraph as about the author rather than about her brother, and is again not concerned about the apparently unlikely activity being described. It is as if she is so concerned with establishing the smaller units of meaning here that she does not stand back to consider the global sense.

We have seen examples where Subject 25 manages to construct more global meanings than does Subject 3. In these examples he seemed to have difficulty moving from a word level to a sentence level. Subject 25 on the other hand made those links successfully in some instances. But Subject 3 in the FP task demonstrates an ability to move from a sentence level to the more global level of whole text, and to perceive a macrostructure which underpins the major thrust and meaning of the text. The latter ability would be the sign of the better reader overall, and in this sense the two performances are perhaps comparatively more in tune with their reading scores than might first have appeared.

5.3.2.4 Making extensive use of the live dictionary

Subject 18 was ranked joint 16th on both reading tests, but managed to be placed joint 5th with either four or five others on both of the reading tasks in the study. He was amongst the top three users of the dictionary facility and warrants an analysis on the basis of how he managed successful meaning construction and task completion.

On a purely quantitative count Subject 18 appears to be an extremely wide-range strategy user. On the TFI task he used items from eight of the P&A sub-sections, 25 categories in all, and on the FP task he used six sub-sections and 18 categories. Clearly many of these appeared only three times or fewer, but this exceeds the strategy use of any other subject in the study. At 249 + 167 lines he also had the longest aggregate transcript, yet in the TFI task he was task-led from the beginning.

The first sequence shows Subject 18 using the dictionary facility and general meaning construction strategies to get the correct answer to a relatively straightforward item, statement 5, *Elle travaille dans le rayon boucherie*. It is noticeable that this takes a long time to achieve, but that he assembles meaning very carefully.

126.	18 She works in	MC/DR	Paraphrasing text
127.	18 what does	ActDiffwp	Uses dictionary
128.	18 is that rayon a name?		
129.	R Rayon means department.		
130.	18 And what does that boucherie?	ActDiffwp	Uses dictionary
131.	18 Boucherie is the butcher's.		
132.	18 Oh she works in the butcher's department or something	MC/DR	Adjusting initial ideas based on new information
133.	R That's it.		
134.	18 So	MC/DR	Pausing to reflect

135.	18 (10)	MC/DR ?	Looking for information relating to specific goals ?
136.	18 Is that maintenant. Is is that job?	MC/DRcim + ActDiffwp	Inferring meaning - context + Uses dictionary
137.	R Erm. Maintenant means now.		
138.	18 What is that? Forgotten what that mais means.	ActDiffwp	Uses dictionary
139.	R Er. But.		
140.	18 But now she works in the butcher department.	MC/DRintg	Holding ideas in working memory
141.	18 Is the wo-	ActDiffwp	Uses dictionary
142.	18 Wh- what's that one just after that rayon bou something in the actual paragraph?		
143.	R Er.		
144.	18 That just there.	ActDiffwp	Uses dictionary
145.	R Er.		
146.	R Boulangerie ?		
147.	18 Yeah.	ActDiffwp	Uses dictionary
148.	R Boulangerie means the bakery.		
149.	18 Oh right. She works in	MC/DR	Adjusting initial ideas based on new information
150.	18 oh so this one's false. It says she works in the	MonPT + MC/Drcim + MC/DRcim	Achieving a goal + Drawing a conclusion + Inferring meaning - cognate
151.	18 bakery and is that pâtisserie the pas- pastry.		
152.	18 So she cooks gateaux.	MC/DR	Paraphrasing text
153.	18 Yeah.	MC/DR	Pausing to reflect
154.	18 So no.	MC/DR	Maintaining an hypothesis
155.	18 The fifth one's false		

In total he asks for five words, the first two to be sure about the meaning of the task statement, the next three to check items in the text, until he finds the key word he is looking for. We could argue that a more strategic reader would establish *rayon boucherie* and the immediately find *rayon boulangerie*, thus short-circuiting the reading process, and eventually he does this. The two other items he requests show that he is processing in a linear fashion, ie trying to comprehend whole sentences rather than key words. This

assertion is underpinned by the fact that he continues to read after the word *boulangerie*, even though he knows already that the statement is false. Other transcripts show that simply asking for words does not automatically lead to comprehension, but here we see a combination of approaches which enables him to gain a more global understanding as well as to respond to specific prompts. Before he reaches his conclusion he has considered the new information and adjusted his working premise (lines 132, 149), and has restated information as an aid to keeping it in working memory (line 140). This continual focus on growing meaning is perhaps the key to translating intensive dictionary use into coherent meaning construction. It also leads him to remember ideas from earlier in the text. This is perhaps shown also by his response to statement eight which says that she goes to the disco on Sunday as well. In his reasoning here he remembers that the beginning of the text contains the statement that she goes to the disco on Saturdays .

244.	18	Is that	MC/DR	Tentative interpretation
245.	18	Is number eight on Sunday she goes to the disco?		
246.	18	Then that means number eight must be wrong.	MonPT	Achieving a goal
247.	18	False.		
248.	18	Cos it says here that she goes to the disco on Saturday.	MC/DRcim	Drawing a conclusion

It is true that he does not also appear to remember that on Sunday she plays rugby, but in giving a correct reason he is fulfilling the task.

Of course there are examples of a failure to gain from the dictionary use. Item 3 again shows that readers can lose track of their understanding when they have to clarify too many unknown words. The need to read, process, gain understanding and make a

decision about comparative statements quickly or risk losing the sense is very apparent from this extract.

40.	18	So is the third one does she like going to the disco a lot?	MC/DR	Tentative interpretation
41.	R	Em.		
42.	18	Would it	MonProb	Unfamiliar terms
43.	18	would it help if you tell me what boit et fume?	+ ActDiffwp	+ Uses dictionary
44.	R	Okay. So boit means drinks.		
45.	R	Fume means smokes.		
46.	18	Oh. Does she	MC/DR	Tentative interpretation
47.	18	does she drink and		
48.	18	at the disco does she drink and smoke?		
49.	18	(12)	MC/DR	Looking for information relating to specific goals
50.	18	On the last paragraph what do those last three lines (sic) mean? The sentent vraiment mauvais?	ActDiffwp	Uses dictionary
51.	R	Erm.		
52.	R	The		
53.	R	sentent means smell.		
54.	R	Vraiment just means really		
55.	R	and mauvais means bad.		
56.	18	(10)	MC/DR	Looking for information relating to specific goals
57.	18	Is that m-	ActDiffwp	Uses dictionary
58.	18	what does that m'invite?		
59.	R	Er. Sort of to invite or to ask someone to do something.		
60.	18	(6)	MC/DR	Looking for information relating to specific goals
61.	18	So	MC/DR	Tentative interpretation
62.	18	I reckon that's false.	MonPT	Achieving a goal
63.	18	third one	MC/DR	Maintaining a hypothesis
64.	18	false.		
65.	18	I reckon the third one's	MC/DR	Adjusting initial ideas
66.	18	tr-		
67.	18	(6)	MC/DR	Pausing to reflect

68.	18 I reckon the third one's	MC/DR	Concluding previous hypothesis invalid
69.	18 true.		
70.	18 Cause it says une de mes amies m'invite	MC/DR + ActTask	Tentative interpretation + Attempt to pinpoint confusions
71.	18 Does that mean like she likes smoking		
72.	18 or drinking?		
73.	R Em. So which		
74.	R which word?		
75.	18 Em	MC/DR	Pausing to reflect
76.	R are you checking up on now?		
77.	18 I don't know.	ActDiffglob	Suspends judgement
78.	18 I'm gonna go for the third one I mean.	MC/DR	Concluding previous hypothesis is valid
79.	18 Yeah. I reckon that the third one's true.		

Here Subject 18 clarifies six single items and two broad sentence interpretations. He establishes the meaning of the statement efficiently and then probably spends time looking for related information. He identifies three key words which others have used to rationalise their response and then reaches the correct conclusion (although without a rationale). Significantly, perhaps, he has not restated the information about the smell of the cigarettes, and has not therefore internalised its significance. He could have given this as evidence for his decision that the statement is false, but not only does he fail to do so, but he then changes his decision without a reason. He subsequently appears to reach the point where he does not want to consider any further, refuses an offer to clarify more vocabulary and so settles on a conclusion without any evidence to support it.

In the FP task Subject 18 opts to attempt a running translation of much of the text. He makes a tentative decision about the identity of the author just over half way through but states himself that he must carry on, He confirms that decision at line 156 and the recaps the four reasons he has found over the final ten lines. Again he asks for a considerable number of words, but chooses quite important items and appears to have a continuing

grasp of the gist of each paragraph as he proceeds. The use of tentative interpretations and questions relating to global meaning again assist this process. He formulates such enquiries more often than anyone else, but they are frequently phrased as requests for confirmation rather than simple dictionary enquiries, for example:

25.	18 Says	MC/DR	Paraphrasing text
26.	18 says she is one of her sisters sisters is very beautiful.		
27.	18 is that what	ActDiffwp	Using dictionary
28.	18 is that what très is very?		
29.	R Mm. Yeah.		

and:

33.	18 is Aînée a name?	ActDiffwp	Using dictionary
34.	R Er. Aînée means older.		
35.	18 Oh. That's it. Her older sister is very beautiful	MC/DR	Adjusting initial ideas based on new information

We also see in that extract again the tendency to restate after new information has been gained, to close off a piece of text and 'file' it but to maintain it in working memory to aid the decoding of the next chunk.

In conclusion then the dictionary facility is not by itself a guarantee of success, although clearly it allows learners to rise above a basic level of performance. (We saw that Subject 27 hardly used the facility at all, for example.) The very top scorers appeared to be good readers who used the live dictionary very judiciously, but Subject 18 has shown that a tenacious approach to discovering unknown items in the text does pay off as long as it is accompanied by reflection, restatement and overt reasoning around the task. The other high dictionary users also managed an average performance across the two tasks and were successful or unsuccessful in terms of individual items in just the same way as Subject 18.

5.3.2.5 Anomaly between individual and group performance

Subject 10 had taken the leading role in the group session despite having a significantly lower 1997 reading score and 1999 GAP score than either of the other two group members. It might therefore have been expected that she would build on that role and experience success in the individual tasks. But her performance did not match this expectation in either task, with 2 correct answers in the TFI task and 2 items noted in the FP task. Her transcript is shorter than average for the TFI task (at 103 lines) and the shortest of all on the FP task (at only 29 lines). Nevertheless she used an average number of P&A sub-sections and categories (5 and 12 for the TFI task and 6 and 8 for the FP task, respectively). In the group TFI task, she personally used 6 sub-sections and 15 categories and her turn count at 109 was greater than in her individual session for the parallel task. Although this gives an indication of some level of difference between the two performances, it does not really tell us about the quality of those differences. Thus, the results of this method of analysis also demonstrates the limits of quantitative strategy counting without deeper analysis of exactly which strategies are in use and more importantly, how they were used.

In the group context, because the participants decided to read silently, translate the text as far as they could and consider the TFI statements twice, there is a very full record of the various articulated processes which took place. In the individual context Subject 10 opted to read the text silently before beginning the task. As a result the most common strategy used by her in the group context (paraphrasing text) does not appear at all in the individual TFI task protocol. But perhaps even more significantly the second most used

category within the group (tentative interpretation) does not appear either. Since the two of these categories combine to form nearly 35% of her turns within the group, we now begin to see something of the qualitative difference. Clearly meaning construction is at the heart of comprehension but we see almost no evidence of how this is achieved in the individual context. Along with this is the strategy of maintaining an hypothesis which she uses a dozen times in the group context but only six times in the individual session. In both, the number of times she achieves a goal (ie decides about a statement) and draws a conclusion are, as expected, nearly the same, since the two strategies go hand in hand on the TFI task. But there are actually fewer examples of these strategies in the individual than in the group context despite her having to share the task with two peers.

The above comparisons are significant but any precise rationale for differences in strategy use between the two contexts is problematic since others are involved in the process in the one and not the other. The actions of the group participants *may* cause certain strategies to be used, but we cannot be certain of this. Nevertheless, the following two extracts demonstrate something of the different processes Subject 10 herself appears to use on the different occasions.

Extract A

70.	10	Number six is false cos she says she doesn't like rugby.	MonPT + MC/DRcim +	Achieving a goal + Drawing a conclusion + Explicitly looking for related idea
71.	10	I think		
72.	10	somewhere.		
73.	R	Mm.		
74.	10	Yeah.	MC/DR	Maintaining an hypothesis
75.	10	She doesn't like rugby.		

Extract B

146.	10	Maria	MC/DR	Paraphrasing text
147.	14	And Guy	MC/DR	Paraphrasing text
148.	10	and Guy, I think that's live with their mum (4)	MC/DR	Tentative interpretation
149.	14	Where's it say that ?	ActDiffglob	Formulating a question
150.	10	* ((points))	MonProb + MC/DR	Clarifying with peer + Maintaining an hypothesis
151.	14	*		
152.	10	It said there's a lot of other people who live there as well. Yeah look ((points))	MC/DRcim	Drawing a conclusion
153.	14	Yeah	MC/DR	Maintaining an hypothesis
154.	10	Yeah, chez nous il y a - Oh no they live with their dad, look, chez nous	MC/DR + MC/DR	Maintaining an hypothesis + Adjusting an interpretation
155.	14	Yeah	MC/DR	Maintaining an hypothesis
156.	10	il y a mon père ((points))	MC/DR + MC/DR	Reading aloud + Maintaining an hypothesis
157.	10	So that's false (3)	MonPT	Achieving a goal

In both of these extracts Subject 10 has actually misunderstood either text or TFI statement and the response she has decided on is correct but not for the right reasons. The difference, however, is that in Extract A she, individually, does not locate any text to support her argument, and rather than look more methodically for it, she relies on an imperfect memory of meaning to justify her response. In Extract B she finds the exact piece of text, and perhaps because of the question from Subject 14 talks through her conclusion much more fully.

This can be seen even more clearly in Extracts C and D where it appears that her perception of the presence of competition within the group (which seems far greater than that of either of the other two participants) drives her on to engage with text and meaning

construction to a far greater extent than appears to be the case in her individual transcript. The individual task (Extract D) has her reading silently for 44 seconds while she considers evidence, yet in that time she does not discover the parallel structure we have mentioned before between *rayon boucherie / rayon boulangerie*. She does note the contrast between the butchers and the pâtisserie, but in such an imprecise way that we cannot credit the response as demonstrating comprehension. This is partly because she has apparently inferred that *travaille* means travel and is reading the text throughout as if the subject is going to rather than working at the supermarket. Other subjects did clarify this item when readings became difficult, but she does not.

Extract C on the other hand has her trying to get to a far closer reading. She speaks more than the other two participants (eight times as opposed to five times and twice, respectively, with a further occasion when 14 and 11 speak together). She states and restates her translation of one section, moves on to the next sentence and then when the other participants beat her to the one after that, she projects a tentative interpretation which tries to answer their uncertain reading. At no point here do they ask for vocabulary help, (and in this sense there is a point of comparison between the two extracts) but the group does generate a variety of possible readings, some close, some quite imprecise. More specifically, Subject 10 herself most of all seems driven to establish meaning. The fact that the extract itself begins and ends with her making a tentative interpretation is a significant factor (especially since she did not do this at all in the individual transcript). This, rather than simple paraphrasing, indicates the important distinction between the process of creating microstructures and simple decoding.

Extract C

102.	10	Is it they set out their knives and forks or something on the table ?	MC/DR	Tentative interpretation
103.	11	Um on the table	MC/DR	Paraphrasing text
104.	14	Small Maria, aime means like	MC/DR	Paraphrasing text
105.	10	Like (4)	MC/DR	Restating text
106.	10	Like setting the table, putting the	MC/DR	Paraphrasing text
107.	14	Yeah setting the	MC/DR	Restating text
108.	10	Putting the knives and forks [etc on the table	MC/DR	Restating text + paraphrasing text
109.	11	Etc on the table]	MC/DR	Restating text + paraphrasing text
110.	14	Table. Natalie	MC/DR	Restating text + paraphrasing text
111.	10	She likes a green salad	MC/DR	Paraphrasing text
112.	14	And I like (5)	MC/DR	Paraphrasing text
113.	10	Yeah (3)	MC/DR	Maintaining an hypothesis
114.	14	Does that mean in the garage ?	MC/DR	Tentative interpretation
115.	10	In the garage	MC/DR	Restating text
116.	14/11	Is that eat food in the garage ? ((laugh))	MC/DR	Tentative interpretation
117.	10	Hang on, does it mean like he takes it downstairs and eats it on his own ?	MC/DR	Tentative interpretation

Extract D

54.	10	What's	ActDiffwp	Using dictionary
55.	10	boucherie?		
56.	R	Er, means like the butcher's.		
57.	10	(44)	MC/DR ??	Looking for information relating to specific goals ??
58.	10	And that number	MonPT	Achieving a goal
59.	10	five's false.		
60.	R	Okay why		
61.	10	Because	MC/DRcim	Drawing a conclusion
62.	10	she says she goes		
63.	10	she went to		
64.	10	somewhere where they sell chocolate.		
65.	R	Mmhm		
66.	10	I think.	MC/DR	Pausing to reflect

Subject 10 then clearly behaves rather differently between the two contexts and further analysis of her group role in Chapter Six will reinforce the most likely reasons for this disparity. Her habit of paraphrasing text aloud, restating text while she continues to analyse it and making tentative interpretations seems to bring her a great deal more success than the silent approach used in the individual tasks. Given that the group context came first, it is surprising that she did not imitate this process more during that session. A strong competitive instinct seems to be a very real possible explanation for this, and this would underline the potential for group based tasks for some learners.

5.3.3 The group interaction protocols

The four groups chosen for comparison demonstrate some contrasting strategy use which in broad terms does comply with their task performance. Thus Group B3 who were the most successful group used more strategies than any others. (Their coded TFI task transcript is included as Appendix N) Group B1 used fewer overall but this was mainly because they were task led for the majority of the time in the TFI task and spent less time on the FP task than the other groups. A4 who were more successful than A1 used more strategies. The coding remained virtually identical as for the individual transcripts, except that certain types of interaction needed codes which did not appear in the P&A system, as that was developed from individual protocols. The categories *clarifying with peer* and *correcting peer* were added.

To compare in a meaningful way, it seems best to take a parallel section from each of the coded transcripts and to investigate the strategy used made by each group to cope with

that particular demand. Again, this is easier to demonstrate by using sections from the TFI task transcripts, as it possible to isolate both meaning construction and decision-making around a particular section of text and the relevant TFI statement. The section chosen here is the text from paragraph 2, which relates how two of the step-siblings are related to Guy through his parents' second marriages. The text here reads: *'Mon frère le plus jeune (Luc) a deux ans. Il est le fils de ma mère et de son nouveau mari. Mon père a une fille de quatre ans (qui s'appelle Maria) avec sa nouvelle femme.'* The relevant task statement (No. 3) reads: *'Maria et Guy ont la même mère.'*

An initial comparison of the time taken to work on the text and task statement shows immediately the variation in approaches. Group B3 took 54 lines to discuss the text and 27 lines to decide on the statement. Group B1 considered the text only as part of the task and did this in 23 lines. A4 spent 29 lines on the text and then 12 lines on their main consideration of the task statement, plus another two when they later recapped. A1 took 31 lines to work on the text and then just 5 to decide on the statement. A1's response was wrong, A4's was correct but the reasoning does not indicate that they understood why. B1 seem to have understood enough to give the right answer for the right reason, but this is not absolutely clearly articulated. B3 were completely correct in both their text analysis and their task statement response. The length of time taken by B3 to decide this is deceptive as they had problems interpreting the task statement. Once they had understood it they clarified their response very rapidly.

The major strategy used by all four groups was paraphrasing text. In a group context the need to articulate emerging meaning is paramount or collaboration does not become possible. But it is perhaps the degree to which this process is shared and what happens

when there are problems which differentiates the four groups. In Extract A Subjects 2 and 3 in Group B1 make tentative interpretations about both text and statement, but at no point is the whole meaning of the text drawn out. They disagree about some essential translation (at lines 58-61), but do not really resolve this, check two words, and when clear about the task statement quickly make a decision. Subject 15 is partially involved with this process, but in a very different mode from much of her input elsewhere (see Section 6.2). Subject 16 is silent throughout. The overall sense we are left with is that partial understanding has led to a correct task response, but that this may not be sufficiently reassuring about their reading processes. Extract A shows this section in full:

Extract A - Group B1

54. 15	What's the, what's the next question ? (1)	MonPT	End of a unit of meaning
55. 3	Who's Maria ? (3) My brother	ActDiffglob +	Formulating a question +
		MC/DR	Paraphrasing text
56. 2	Is called Luke	MC/DR	Paraphrasing text
57. 3	And he's (-) two	MC/DR	Paraphrasing text
58. 2	What's, it says it says that his dad is called	MC/DR	Tentative interpretation
59. 15	No, it's	MonProb	Correcting peer
60. 2	No that's his mother	MC/DR	Concluding previous hypothesis invalid
61. 3	That's his brother Luke (-), mon frère (4)	MC/DR	Concluding previous hypothesis invalid
62. 2	Maria is [his dad's new partner (2)	MC/DR	Paraphrasing text
63. 3	He is] he is (2)	MC/DR	Paraphrasing text
64. 15	What does that mean ?	MonProb	Clarifying with peer
65. 2	What's même ?	ActDiffwp	Using a dictionary
66. R	Er, where it says elle-même, that means she herself (4)		
67. 3	What's nouvelle ?	ActDiffwp	Using a dictionary
68. R	It means new		
69. 3	I think that says she's got a new, a new baby called Maria, because it says Maria et Guy	MC/DR	Tentative interpretation
70. R	Ah, when you asked for même, did you mean même in Question Three or même in the text where it says elle-même ?		

71. 2	Même in Question Three		
72. R	Right, même in Question Three, sorry, that means the same. That word's a bit strange because it has more than one meaning, but in item three, Maria et Guy ont la même mère, même there means the same		
73. 3	I think it means, did Maria [and Guy have the same mum	MC/DR	Tentative interpretation
74. 15	Did Maria and Guy have the same mum]	MC/DR	Tentative interpretation
75. 3	Which is false because it says he is his dad, I think	MonPT + MC/DRcim	Achieving a goal + Drawing a conclusion
76. 2	Yeah, that's his new sister, Maria	MC/DR	Maintaining an hypothesis

The strategy use is quite repetitive, with only little amounts of questioning of each other or of self to try to draw out greater layers of meaning. Subjects 2 and 3 do both try to draw out a general sense of a section (at lines 62 and 69), but it is this more macro- level which is absent. If we compare this to the much greater variation of coding which emerges from the interactions of group B3, we can see this in more concrete terms.

Extract B shows the section where B3 are establishing the textual meaning, and that is followed by their decision-making on Statement 3 as Extract C.

Extract B - group B3 - meaning construction:

49. 8	... Um, my brother (6)	MC/DR	Paraphrasing text
50. 7	Er	MC/DR	Pause to reflect
51. 9	Well we	MC/DR	Pause to reflect
52. 7	Plus jeune, Luc	MC/DR	Reading aloud
53. 9	What does that ((points to sheet, shows 8))	MonProb	Unfamiliar terms in text
54. 7	Deux ans, two years	MC/DR	Paraphrasing text
55. 9	What does, what does that mean ? ((points to sheet, shows 8))	MonProb	Unfamiliar terms in text
56. 8	Yeah, what does jeune mean ? ((shows 7))	MonProb	Unfamiliar terms in text
57. 7	Jeune, I don't know	ActDiffglob	States failure to understand
58. 9	Yeah, what does, what does that mean ? ((shows R))	ActDiffwp	Use a dictionary
59. R	It means young		

60. 9	Oh	MC/DR	Pause to reflect
61. 8	Oh right, so	MC/DR	Tentative interpretation
62. 9	So that means younger	MC/DR	Using patterns to adjust interpretation *
63. 7	A young	MC/DR	Paraphrasing text
64. 9	Younger	MonProb	Correcting a peer
65. 7	Younger	MC/DR	Restating text
66. 8	My younger brother is two	MC/DR	Paraphrasing text
67. 9	Yeah. Um he is (4) Do we know fil ?	MC/DR + MC/DR + MonProb	Concluding previous hypothesis is valid + paraphrasing text + Unfamiliar terms in text
68. 8	((shakes head))	ActDiffglob	States failure to understand
69. 9	Er what's fil ?	ActDiffwp	Use a dictionary
70. R	Fils. Er means son, er you know son, son-daughter son, not sun, the sun		
71. 7	Mon père a un fils	MC/DR	Reading aloud
72. 8	He is [the son of	MC/DR	Paraphrasing text
73. 9	The son]	MC/DR	Paraphrasing text
74. 8	my mum	MC/DR	Paraphrasing text
75. 9	And	MC/DR	Paraphrasing text
76. 8	And (3) No not sure about	ActDiffglob	States failure to understand
77. 9	Well she said she had two half brothers so it's (-) probably something to do with a step-dad or something like that [*	MC/DR	Tentative interpretation
78. 8	OK,] alright	MonPT	End of a unit of meaning
79. 9/8	My dad	MC/DR	Paraphrasing text
80. 7	This is a lot harder than the other one	MonTC	Text difficulty
81. 9	My	MC/DR	Paraphrasing text
82. 8	What was ? ((shows sheet to 9))	MonProb	Unfamiliar terms in text
83. 9	God, I don't	ActDiffglob	States failure to understand
84. 7	*		
85. 8	Er, I don't know what that means	ActDiffglob	States failure to understand
86. 9	It's son isn't it	MC/DR	Tentative interpretation
87. 8	Yeah	MC/DR	Concluding previous hypothesis is valid
88. 9	Son. So. My dad's, my dad's son has yeah	MC/DR	Paraphrasing text

89. 8	Isn't fille ? (1) What's fille again ?	MC/DR + ActDiffwp	Tentative interpretation + Use a dictionary
90. R	Fille is daughter		
91. 8	Daughter, right, [thank you	MC/DR	Restating text
92. 9	Right *] Yeah	MC/DR	Concluding previous hypothesis is valid
93. 8	My dad has a daughter (2) four, who's four years old. [She's called Maria	MC/DR	Paraphrasing text
94. 9	She's called Maria]	MC/DR	Paraphrasing text
95. 7	((sings)) Maria	MC/DR	Restating text
96. 8	She's (1) with (2)	MC/DR	Paraphrasing text
97. 7	Avec with sa nouvelle femme, a woman. (2) He is with another woman	MC/DR + MC/DR	Reading aloud + Paraphrasing text
98. 8	Oh yeah, and that was with a son with my mum and another man ((shows sheet to 7))	MC/DR	Jumping back to reconsider previous information
99. 7	Yeah	MC/DR	Concluding previous hypothesis is valid
100. 8	OK	MonPT	End of a unit of meaning
101. 7	Yeah, they're split up, they're split up	MC/DR	Concluding previous hypothesis is valid
102. 8	OK	MonPT	End of a unit of meaning

** using patterns to adjust interpretation is not one of P&A's coding categories*

Clearly here we have a much fuller discussion of meaning, but not one which proceeds automatically simply because the participants are substantially clearer about the lexis.

They do have to work at it, but the sharing of concentration by all three participants, and the flexibility with which they attack the text brings its own success.

We can note especially the way in which they explore unknown vocabulary amongst themselves before asking, (eg lines 55-6, 67-8), showing that 'not knowing' can be a positive strategy in a group context as long as it is discussed. Also worthy of comment is the way in which they hold meaning and recombine it with new elements in their tentative interpretations, (eg lines 77 and 98). This can involve being sensitive to the fact

that text often has discourse structures such as parallel statements, and we see an awareness of this also in lines 97-8. When this group paraphrases text they often link together well enough for it to seem that it is a single person working through the line of meaning construction, for example in lines 71-77 and 93-99. This contrasts sharply with lines 58-62 of the B1 transcript or with other examples to follow from A4's and A1's protocols. The essence of the difference, it appears, is that they do stop and recast frequently enough to gain a sense of emerging meaning, and this allows the macrostructure to be assembled much more effectively. It may be the combination of the same strategy types rather than the use of different strategy types here which is significant. Thus we see that an analysis of strategy use gives us part of a picture but we need to focus on other elements of group talk and dynamics to explain more fully why some gain meaning more effectively than others. Chapter Six will deal with this in much more detail.

But the particular strategic process which seems to lead to reasoned decision-making is exemplified by B3's next section, where they decide on the correct response to task statement 3.

Extract C - group B3 - TFI decision-making

198.	8	Maria and Guy	MC/DR	Paraphrasing text
199.	7	[Are both	MC/DR	Paraphrasing text
200.	8	Are] the	MC/DR	Paraphrasing text
201.	7	Live with their mum, no	MC/DR + MC/DR	Paraphrasing text + Adjusting tentative interpretation
202.	9	No	MC/DR	Concluding previously stated hypothesis invalid
203.	8	Are the same thing	MC/DR	Tentative interpretation

204.	7	It's got something, no their dad, no their mum, it's got something to do with their mum, mère	MC/DR	Tentative interpretation
205.	8	OK	MC/DR	Concluding previously stated hypothesis is valid
206.	9	[Where was that bit, we saw même earlier,	ActDiffglob + MC/DR	Formulates a question + Explicitly looking for related words in text
207.	7	Ont la même mère] What does même mean ?	MC/DR + MonProb	Restating text + Unfamiliar terms in text
208.	9	We saw it before but I can't remember where it was, was it with, where was it ((shows sheet to 8))	ActDiffglob	Pauses to scan for sources of difficulty
209.	8	Maybe it's like they've both	MC/DR	Tentative interpretation
210.	9	[It's there it's there look ((points))	MC/DR	Explicitly looking for related words in text
211.	8	the same mum.] Oh yeah	MC/DR	Adjusting initial ideas based on new information
212.	7	Guy and Maria both have the same mum	MC/DRcim	Confirming an inference
213.	8	Was that ?	ActDiffwp	Use a dictionary
214.	R	Yeah, um, même can actually mean more than one thing. In Number 3 that you're looking at même does mean the same. When you saw it up in the text it means something different there		
215.	7	Yeah so Maria and Guy have the [same mum	MC/DR	Restating text
216.	9/8	OK OK]	MC/DR	Concluding previously stated hypothesis is valid
217.	7	Is that true or false ?	ActDiffglob	Formulates a question
218.	9	Well was it his sister or his half sister, (1) cause then * was it	ActDiffglob	Formulates a question
219.	8	Um	MC/DR	Pausing to reflect
220.	9	Well he says that my dad has a s- a daughter [who is so	MC/DR	Tentative interpretation
221.	8	Oh with another] ((shows sheet to 7))	MC/DR	Tentative interpretation
222.	9	Well so they don't have the same mum do they ?	MC/DRcim	Drawing a conclusion
223.	7	So that's faux	MonPT	Achieving goal
224.	8	Yeah [that's false	MC/DR	Concluding previously stated hypothesis is valid

Here we see a continual cycle involving tentative interpretation, question formulation, the adjustment of ideas or testing of hypotheses made, and finally the drawing of a logical conclusion. Again all three participants are involved fairly equally, both in terms of numbers of turns and in terms of the quality of their strategic input. All three ask questions, all three make tentative interpretations, all three react to hypotheses as they are formulated. In this extract we see some very typical behaviour by this group. Subject 9 acts as the participant who looks back in the text to find evidence, through single vocabulary items or meaning at a more propositional level. Subject 8 is able to coordinate from the central seating position that she occupied in the group. Subject 7 tends to look ahead and react quickly to suggest readings to be evaluated by the other two participants.

We can contrast this with the decision-making by groups A1 and A4. A1 supplies the briefest example (at just 5 lines) of this part of the process of any of the eight groups on this particular statement. We have seen before the dangerous strategy for TFI tasks of simply looking for a match between statement and text. Here Subject 28 is advancing the proposal that because the word *même* is not in the text (although in fact it is !) that the statement should be counted as true. This is counter to the usual argument in this situation, where a parallel form is held to be true and a non-parallel form false. But Subject 26 opts for the safer middle line and suggests *pas clair*. Because the participants here do not really consider the meaning, and opt for the 'impossible to say' option simply on the basis that they cannot see a vocabulary item in the text, we have a much more restricted group of strategies in play. Evaluating the validity of an hypothesis is an

important strategy if the evaluation is based on a stated or at least perceived argument.

But the sequence here is not even slightly analogous to that shown in the B3 dialogue (eg lines 215-224) where the meaning of the relevant sentences is constructed through interpretation and questioning.

Extract D - Group A1 - TFI task decision-making

220	26	Maria et Guy ont la même mère	MC/DR	Reading aloud
221	28	I think that's vrai (4) Maria et Guy ont la même mère (2) doesn't say it in the text (7) doesn't say même there does it ?	MonPT + MC/DR + MC/DRcim	Achieving a goal + Reading aloud + Drawing a conclusion
222	26	Say pas clair	MC/DR	Concluding previous hypothesis invalid
223	28	(2) Yeah	MC/DR	Adjusting an interpretation
224	28/26	Pas clair	MC/DR	Maintaining an hypothesis

Group A4 on the other hand fail to see an incorrect translation of the task statement by Subject 10 (see further analysis of this issue in Chapter 6), and argue from this to a response which happens to be correct, although for entirely the wrong reason. There is more well-founded reasoning and some exploration of meaning at sentence level here, as well as signs from the subjects that they are looking carefully at the text. But it still lacks the focus of Group B3 and so the error remains uncorrected and the response given therefore wrong.

Extract E(i) - Group A4 - TFI task decision-making

146.	10	Maria	MC/DR	Paraphrasing text
147.	14	And Guy	MC/DR	Paraphrasing text
148.	10	and Guy, I think that's live with their mum (4)	MC/DR	Tentative interpretation
149.	14	Where's it say that ?	MonProb	Clarifying with peer

150.	10	* ((points))		
151.	14	*		
152.	10	It said there's a lot of other people who live there as well. Yeah look ((points))	MC/DR	Maintaining an hypothesis
153.	14	Yeah	MC/DR	Concluding previous hypothesis is valid
154.	10	Yeah, chez nous il y a - Oh no they live with their dad, look, chez nous	MC/DR + MC/DR	Reading aloud + Paraphrasing text
155.	14	Yeah	MC/DR	Concluding previous hypothesis is valid
156.	10	il y a mon père ((points))	MC/DR + MC/DRcim	Reading aloud + Drawing conclusion
157.	10	So that's false (3)	MonPT	Achieving a goal

Extract E(ii) - Group A4 - TFI task decision-making

222.	10	The third one was false wasn't it ?	MC/DR	Maintaining an hypothesis
223.	14	yeah, um fourth one is true	MC/DR	Concluding previous hypothesis is valid

In conclusion, we have seen that strategy use within groups does not necessarily inherently differ from strategy use by individuals. In some cases learners will differ because their behaviour within a group and that as an individual will not be the same, probably owing to affective rather than cognitive or metacognitive factors. A group does have the power to shape the strategy use of individuals within it if metacognitive discussion is overt, because clearly that needs to be addressed by the other group members. Similarly there is often a need to make meaning construction more public in the group context, and this inevitably affects how far members are overt about their readings of the text. But an important element in comprehension which has emerged through this analysis, the use of tentative interpretations, is not necessarily elicited by the group process any more than in the individual context. Restating and refining text readings are similarly not context-dependent. Reacting to readings is of course a feature of the group context, and clearly where a group has a habit of doing this the construction

will become more precise (although not necessarily more accurate). Decision-making is however probably affected by the group context in the majority of cases. It was barely necessary at any point for the researcher to ask for reasons for decisions in the group sessions, but was frequently so in the individual sessions. This shows that a natural process of accounting for statements is inherent in group-based tasks, and in turn this will affect the range and amount of clearly visible strategy use. We need to remember that individual strategy patterns are often hidden in the silences which accompany reading and thinking. But of course we can discern how effective they are through task outcomes. As we saw in the best v. worst comparison good readers can often voice their strategy use while reading and possibly that very articulation process actually enhances the reading. Poor readers will probably have less to say while reading and equally their performance remains unenhanced because of a lack of articulation of processes. Nevertheless the clear success of the B groups (when compared to individual performances) in the TFI task demonstrates that the use of a range of strategies, even if not more varied, is probably more effective in a group context as long as that context allows for productive variants of group talk. This concept is discussed in detail in Chapter Six along with more analysis of individual roles.

